Functional Analysis

Lou Yi

May 6, 2025

Contents

1 Preliminary

Recall basic knowledge on topological spaces, metric spaces, norms and inner products.

Notations and Conventions:

1. Let $C[a, b]$ denote the set of continuous functions on the interval a, b. This is space is complete under the sup norm but not complete under the norm

$$d_p(f, g) = \left(\int_a^b |f(x) - g(x)|^p dx \right)^{1/p}.$$

   The completion of the the d_p-norm is the $L^p([a, b])$.

2. Let $C^k[a, b]$, $k \in \mathbb{Z}_+$ denote the set of continuous functions with continuous derivative of order k.

3. Let $C^\infty[a, b]$ denote the set of all infinitely differentiable functions.

4. Let ℓ^p denote the set of all sequences $\mathbf{x} = \{x_i\}_{i=1}^\infty$ such that

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

   This space is complete under the p-norm.

5. We define the Hölder conjugate of $1 \leq p < \infty$ to be the number q such that

$$\frac{1}{p} + \frac{1}{q} = 1.$$

6. Let ℓ^∞ denote the set of all sequences $\mathbf{x} = \{x_i\}_{i=1}^\infty$ such that

$$\|x\|_\infty = \sup_i |x_i| < \infty.$$

   This space is complete under the ℓ^∞-norm.

Definition 1.1 (Isometry)**.** *Let (X, d_X) and (Y, d_Y) are metric spaces, then $T : X \to Y$ is called an **isometry** if $d_X(x_1, x_2) = d_Y(T(x_1), T(x_2))$. If further T is bijective, then we say X and Y are **isometric**.*

Theorem 1.2. *Let $1 \leq p < \infty$. Then ℓ^p is separable.*

Proof. Let

$$\mathbb{Q}_n = \{(x_1, \cdots, x_n, 0, 0, \cdots) : x_i \in \mathbb{Q}\}.$$

Then $\mathbb{Q}_n \subset \ell^p$ and $\mathbb{Q}_n$ is countable. Let

$$X = \bigcup_{n=1}^\infty \mathbb{Q}_n$$

Then X is countable subset of ℓ^p. It is easy to see that X is dense in ℓ^p. $\square$

Proposition 1.3. *ℓ^∞ is not separable.*

Proof. Let X be a dense subset of ℓ^∞. Then we consider the set of binary sequences which is a uncountable subset of ℓ^∞. Note the distance between any two distinct binary sequence is 1. But since X is dense in ℓ^∞, then for each ball of radius 1/3 centered at any binary sequence, exists $x \in X$ which belongs to the ball. By triangle inequality, elements of X cannot belong to more than one ball, which implies that X must be uncountable. $\square$

1.1 Zorn's Lemma

Definition 1.4 (Partial Order)**.** *A* ***partial order*** *on a set M is a binary relation denoted by $\leq$ such that*

1. *(Reflexivity) $a \leq a$ for every $a \in M$;*
2. *(Antisymmetry) If $a \leq b$ and $b \leq a$, then $a = b$;*
3. *(Transitivity) If $a \leq b$ and $b \leq c$, then $a \leq c$.*

The partial order becomes a ***total order*** *if every two elements of the set is comparable.*

An ***upper bound*** *of a subset W of a partial ordered set M is an element $u \in M$ such that*

$$x \leq u \quad \forall x \in W.$$

A ***maximal element*** *of M is an $m \in M$ such that if $m \leq x$, then $m = x$.*

Lemma 1.5 (Zorn's Lemma)**.** *Let $M \neq \emptyset$ be a partially ordered set. Suppose that every chain $C \subset M$ has an upper bound in M. Then M has at least one maximal element.*

1.2 Category Theorem

Definition 1.6 (Category)**.** *A subset M of a metric space X is said to be*

1. ***rare (or nowhere dense)*** *in X if its closure $\overline{M}$ has empty interior.*
2. ***meager (or of the first category)*** *in X if M is the union of countably many sets each of which is rare in X.*
3. ***nonmeager (or of the second category)*** *in X if M is not meager in X.*

Example: Every closed proper subspace of a vector space is nowhere dense. Suppose $Y \subsetneq X$ and Y has non-empty interior. So it contains $B(y,r)$. Then let $z \in X$, $z \neq 0$. Define $x := y + \frac{r}{2\|z\|}z$. Then we can show $x \in B(y,r)$. Then we have a contradiction, as Y is a subspace.

Theorem 1.7 (Baire's Category Theorem for Complete Metric Spaces)**.** *If a metric space $X \neq \emptyset$ is complete, it is nonmeager in itself. Hence if $X \neq \emptyset$ is complete and*

$$X = \bigcup_{k=1}^{\infty} A_k$$

where each A_k is closed, then at least one A_k contains a nonempty open subset.

Proof. Suppose the complete metric space $X \neq \emptyset$ is meager, then

$$X = \bigcup_{k=1}^{\infty} M_k \tag{1.1}$$

with each M_k is rare in X. We construct a Cauchy sequence (p_k) whose limit p is not in any of the M_k, thereby contradicting the representation (1.1).

By assumption, M_1 is rare in X, so by definition $\bar{M}_1$ does not contain a non-empty open set. But X does. This implies $\bar{M}_1 \neq X$, hence the complement of $\bar{M}_1$ is nonempty and open. We may thus choose a point p_1 in $\bar{M}_1^c$ and an open ball around it, say

$$B_1 = B(p_1, \epsilon_1) \subset \bar{M}_1^C,$$

By assumption, M_2 is rare in X, so that $\bar{M}_2$ does not contain a non-empty open set. Hence it does not contain the open ball $B(p_1, \frac{1}{2}\epsilon_1)$, this implies that $\bar{M}_2^c \cap B(p_1, \frac{1}{2}\epsilon)$ is not empty and open, so we can choose a point p_2 and ball $B(p_2, \epsilon_2)$ contained in the intersection with $\epsilon_2 < \frac{1}{2}\epsilon_1$. Doing this inductively, we obtain a sequence of nested balls

$$B_k = B(p_k, \epsilon_k)$$

with $\epsilon_k < \epsilon_1 2^{k-1}$, and $B_k \cap M_k = \emptyset$. Then (p_k) is Cauchy and converges, so must converge to some $p \in X$. Moreover, by triangle inequality, we must have that $p \in B_m$ for every m. Since $B_m \subset \bar{M}_m^c$, we see that $p \notin M_m$ for every m, so $p \notin \bigcup M_m = X$, which is a contradictions. □

Corollary 1.7.1. *The complement of a meager subset M of a complete metric space X is non-meager.*

Proof. Otherwise X is meager. □

Corollary 1.7.2. *If X is complete, then the countable union of (closed) nowhere dense sets has empty interior.*

Proof. If a set is nowhere dense, then its closure is also nowhere dense, so we can assume that all the nowhere dense set in the union is closed. Now suppose

$$Y = \bigcup_{i=1}^{\infty} A_n$$

where each A_n is nowhere dense. And $U \subset Y$ is open. Since open subsets of a complete metric space is completely metrizable and the new metric induce the same topology. Thus

$$U = \bigcup_{i=1}^{\infty} A_n \cap U$$

and each $A_n \cap U$ is nowhere dense. This contradicts the Baire Category Theorem. □

Lemma 1.8. *A subset M of a metric space X is rare if and only if $(\bar{M})^c$ is dense in X.*

Proof. Suppose X is a metric space. Then M is rare if and only if $\mathrm{int}(\overline{M}) = \emptyset$. Note that $\overline{\bar{M}^c} = X \setminus \mathrm{int}(\overline{M})$, then the lemma follows. □

Lemma 1.9. *A metric space X has the property that the countable union of closed nowhere dense sets has empty interior if and only if every meager set has empty interior.*

Proof. Suppose A_n is a countable collection of closed nowhere dense sets. Then

$$\bigcup_{n=1}^{\infty} A_n$$

is a meager set. If every meager set has empty interior, then the countable union of closed nowhere dense sets has empty interior. Conversely, suppose a meager set $\bigcup_{n=1}^{\infty} A_n$ has nonempty interior, then $\bigcup_{n=1}^{\infty} \bar{A_n}$ must have nonempty interior. So one of $\bar{A_n}$ is not nowhere dense. □

Lemma 1.10. *A metric space X has the property that every meager set has empty interior if and only if it has the property that the countable union of open dense subsets is dense.*

Proof. We know the first property is equivalent to saying that the countable union of closed nowhere dense sets has empty interior.

Now let $\{U_k\}$ be a countable collection of dense open subsets of a complete metric space X. Its intersection is dense if and only if the intersection intersects with any open sets U in X is nonempty. Let

$$G_k = X \setminus \{U_k\},$$

note each G_k is closed and nowhere dense. Then

$$\bigcup_{k=1}^{\infty} G_k \cap U \neq U \ \ \forall U$$

if and only if

$$\bigcap_{k=1}^{\infty} U_k \cap U \neq \emptyset \ \ \forall U$$

So the two properties are equivalent. □

Hence we have the following alternative formulation of Baire's Category Theorem stated below:

Theorem 1.11 (Complete Metric Spaces are Baire)**.** *Complete metric spaces are Baire spaces, that is the intersection of countably many dense open subsets is still dense.*

2 Normed Space

2.1 Normed Vector Space

Definition 2.1 (Hamel basis)**.** *If B is an independent spanning set vector space V, such a set B is called a **Hamel basis**.*

Definition 2.2 (Normed Linear Spaces)**.** *Given a vector space V, a **normed linear space** is $(V, \|\cdot\|)$, where $\|\cdot\| : V \to \mathbb{R}$ is a norm on V, that is:*

1. $\|v\| \geq 0$, $\forall v \in V$;
2. $\|v\| = 0$ *iff* $v = 0$;
3. $\|\alpha v\| = |\alpha| \cdot \|v\|$;
4. $\|v + w\| \leq \|v\| + \|w\|$.

Every norm induce a metric by $d(u, v) = \|u - v\|$.

Remark 2.2.1. *Let d be the metric induced by a norm $\|\cdot\| : V \to \mathbb{R}$. Then $\|\cdot\| : (V, d) \to (\mathbb{R}, |\cdot|)$ is continuous.*

Lemma 2.3. *Let X be a normed space and Y be a subspace of X, then the closure of Y, denoted $\bar{Y}$ is also a subspace of X.*

Proof. Clearly $0 \in X$. We show $x + ay$ if $x, y \in \bar{Y}$. Since $x, y \in \bar{Y}$, there is sequences $(x_n), (y_n) \subset Y$ such that $x_n \to x$ and $y_n \to y$, then the sequence $(x_n + ay_n) \subset Y$ is a sequence that converge to $x + ay$, so $x + ay \in \bar{Y}$. □

Definition 2.4 (Absolutely Convergent)**.** *If X is a normed space and (x_n) is a sequence of elements in X. We can define*

$$S_n = \sum_{k=1}^{n} x_k.$$

If there exists $s \in X$ such that

$$\lim_{n\to\infty} \|s_n - s\| = 0,$$

*then we write $s = \sum_{k=1}^{\infty} x_k$. If $\sum_{k=1}^{\infty} \|x_k\| < \infty$, we call the series **absolutely convergent**.*

Definition 2.5 (Schauder Basis)**.** *If X is a normed space and $(e_k)_{k=1}^{\infty}$ is a sequence of elements of X such that $\forall x \in X$, exists a unique sequence of scalar (α_n) such that*

$$\lim_{n\to\infty} \left\| \sum_{k=1}^{n} \alpha_k e_k - x \right\| = 0$$

*then we call (e_k) a **Schauder basis**. In this case, we write*

$$x = \sum_{k=1}^{\infty} \alpha_k e_k.$$

Example: The collection of vectors $e_n = (\delta_n)$ is a Schauder basis for ℓ^p $(1 \leq p < \infty)$ with p-norm.

Proposition 2.6. *If X is a normed space that has a Schauder basis, then X is separable.*

Proof. Let (e_k) be a Schauder basis. Notice that the set S of countable linear combination of (e_k) with coefficients in $\mathbb{Q}$ and converges in X is a countable dense subset of X. $\square$

2.2 Banach Space

Definition 2.7 (Banach Space)**.** *A normed space that is complete with respect to the metric induced by the norm is called a* ***Banach Space****.*

Theorem 2.8. *Let X be an infinite dimensional Banach space and suppose $B \subset X$ is a Hamel basis. Then B is uncountable.*

Proof. Suppose B is countable, denoted by $\{\beta_k\}_{k=1}^{\infty}$. Define $X_n = \operatorname{span}(\beta_1, \beta_2, \cdots, \beta_n)$. Then for any $x \in X$, $x \in X_n$ for some n. That is

$$X = \bigcup_{n=1}^{\infty} X_n.$$

Note that $\dim X_n = n$, so each X_n is closed. Moreover $X_n = \overline{X_n}$, so we know it is nowhere dense (a proper closed subspace is nowhere dense). This shows X is meager in itself, but X is complete, which contradicts the Baire Category Theorem. $\square$

Example: The vector space of the set of all polynomials cannot be normed so that it is complete.

Corollary 2.8.1. *A Banach space is finite dimensional iff every of its subspace is closed. In other words, infinite dimensional Banach space has non-closed subspaces.*

Proof. Otherwise take a countably infinite basis $\{x_1, \cdots, x_n, \cdots\}$ of X. Then consider

$$X_n = \operatorname{span}\{x_1, \cdots, x_n\}$$

and $Y = \bigcup X_n$. This space cannot be Banach, hence cannot be a closed subsapce of X. $\square$

Theorem 2.9. *A normed space X is a Banach space if and only if every absolutely convergent series converges.*

Proof. Clearly Banach implies absolutely convergent series converges. On the other hand, suppose we are given a Cauchy sequence $\{x_k\}_{k=1}^{\infty}$, then we can find a subsequence $\{x_{k_n}\}$ such that

$$\|x_{k_n} - x_{k_{n+1}}\| < \frac{1}{2^n}.$$

Thus $\{x_{k_n}\}$ converges absolutely. $\square$

Theorem 2.10. *A subspace Y of a Banach space X is complete iff its closed in X.*

Proof. Suppose Y is complete, let $y \in \bar{Y}$. Then exists a sequence $\{y_n\} \subset Y$ such that $y_n \to y$. Then $\{y_n\}$ converges in Y which shows that $y \in Y$.

Conversely, suppose Y is closed and let (y_n) be Cauchy in Y, since $Y \subset X$ and X is complete, then $y_n \to y$ for some $y \in X$. Further, by definition $y \in \bar{Y} = Y$. So Y is complete. $\square$

Lemma 2.11. *Let (X,d) and (Y,ρ) be metric spaces and suppose that $f : X \to Y$ is uniformly continuous. Then under f, the image of every Cauchy sequence is a Cauchy sequence. In fact, the converse holds.*

Theorem 2.12. *Let Y be a closed subspace of a normed linear space $(X, \|\cdot\|)$. Let X/Y denote the quotient space (elements of X/Y are additive cosets). For $x+Y \in X/Y$, we define the quotient norm $\|\cdot\|_*$ by*

$$\|x+Y\|_* = \inf_{y\in Y} \|x-y\| = \inf_{x^*} \|x^*\|,$$

where x^* ranges over all elements such that $x+Y = x^*+Y$. Then $\|\cdot\|_*$ is a norm on X/Y. Moreover, if X is a Banach space, then X/Y is also a Banach Space under the quotient norm.

Recall the following theorems:

Theorem 2.13 (Bolzano-Weierstrass Theorem)**.** *Any bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.*

Theorem 2.14 (Heine-Borel)**.** *A subset of $\mathbb{R}^n$ or $\mathbb{C}^n$ is compact if and only if it is closed and bounded. For general metric space, a subset of X is compact if and only if it is complete and totally bounded.*

Remark 2.14.1. *Recall that a metric space is compact if it is sequentially compact.*

Theorem 2.15. *If $f : K \to Y$ is continuous, and K is compact, then $f(K)$ is compact.*

Theorem 2.16. *If $f : K \to Y$ is continuous on the compact set K, then f is uniformly continuous.*

Theorem 2.17. *If K is compact and $f : K \to \mathbb{R}$ is continuous, then f attains its maximum and minimum in K.*

Lemma 2.18. *Let $x_1, \cdots, x_n$ be an independent list in $(X, \|\cdot\|)$. Then there exists constants $A, B > 0$ such that for any list of scalars $\alpha_1, \cdots, \alpha_n$,*

$$A\sum_{i=1}^{n} |a_i| \leq \|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \leq B\sum_{i=1}^{n} |a_i|.$$

Proof. Use Triangle Inequality we can easily get the second inequality by letting $B = \max\limits_{1\leq i\leq n} \|x_i\|$.

For the first inequality, define

$$S = \{(\alpha_1, \cdots, \alpha_n) \in \mathbb{R}^n \mid \sum_{i=1}^{n} |\alpha_i| = 1\}$$

which is compact (if the base field is $\mathbb{C}$, then replace $\mathbb{R}^n$ by $\mathbb{C}^n$). Define $f : S \to X$ by

$$f(\alpha_1, \cdots, \alpha_n) = \sum_{i=1}^{n} \alpha_i x_i.$$

We show that f is continuous:

$$\|f(\alpha_1, \cdots, \alpha_n) - f(\beta_1, \cdots, \beta_n)\| \leq \sum_{i=1}^{n} |\alpha_i - \beta_i|\|x_i\| \leq \max_i \|x_i\| \cdot \sum_{i=1}^{n} |\alpha_i - \beta_i|.$$

Hence $\|f(S)\|$ is compact since $\|\cdot\|$ is continuous, so $\|f(S)\|$ obtains a minimum (> 0) which we will denote by A (in fact we can obtain the maximum B in the same way). Then for any $(\alpha_1, \cdots, \alpha_n) \in S$, we have

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \geq A.$$

Lastly, the statement will be complete by a scaling argument, which is scaling by $1/\sum_{i=1}^{n} |\alpha_i|$. □

Corollary 2.18.1. *Let X be a finite-dimensional normed space and let $e_1, \cdots, e_n$ be any basis of X. Then there exists constants $A, B > 0$ such that $\forall x \in X$, if $x = \alpha_1 e_1 + \cdots + \alpha_n e_n$, then*

$$A \sum_{i=1}^{n} |a_i| \leq \|x\| \leq B \sum_{i=1}^{n} |a_i|.$$

Corollary 2.18.2. *All norms on a finite dimensional vector space are equivalent, that is, if $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms on a finite dimensional vector space X, then there exists two positive real number c and C such that for all $x \in X$,*

$$c\|x\|_2 \leq \|x\|_1 \leq C\|x\|_2.$$

Remark 2.18.1. *Equivalent norms induce the same topology and hence the same Cauchy sequences.*

Proposition 2.19. *Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on a linear space X. We say $\|\cdot\|_2$ is **stronger** than $\|\cdot\|_1$ if for any sequence $\{x_n\} \subset X$, $\|x_n\|_2 \to 0$ implies $\|x_n\|_1 \to 0$. Then $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$ if and only if there exists a constant $C > 0$ such that $\|x\|_1 \leq C\|x\|_2$ for all $x \in X$.*

Proof. The backward direction is straightforward. For the forward direction, suppose such C does not exists, then for any $k \in \mathbb{N}$, there exists $x_k \in X$ such that $\|x_k\|_2 = 1$, and $\|x_k\|_1 \geq k$. However $\{\frac{1}{k}x_k\}$ is a sequence that gives the contradiction. □

Corollary 2.19.1. *Every finite dimensional normed space over $\mathbb{R}$ or $\mathbb{C}$ is a Banach space that is isomorphic (hence homeomorphic) to $\mathbb{R}^n$ or $\mathbb{C}^n$. Hence a closed and bounded space subspace of a Banach space is compact.*

Corollary 2.19.2. *Every finite dimensional subspace of Y of a normed space X is closed in X.*

Lemma 2.20 (Riesz's Lemma)**.** *Let Y and Z be subspaces of normed space X (of any dimension), and suppose that Y is closed and is a proper subset of Z. Then for every real number $\theta \in (0, 1)$ there is a $z \in Z$ such that*

$$\|z\| = 1, \quad \|z - y\| \geq \theta \text{ for all } y \in Y.$$

Proof. We consider any $v \in Z \setminus Y$ and denote its distance from Y by a. Clearly $a > 0$ since Y is closed. We now take any $\theta \in (0, 1)$. By the definition of infimum, there is a $y_0 \in Y$ such that

$$a \leq \|v - y_0\| \leq \frac{a}{\theta}$$

Let $z = c(v - y_0)$, where

$$c = \frac{1}{\|v - y_0\|}.$$

Then $\|z\| = 1$, and we show that $\|z - y\| \geq \theta$ for every $y \in Y$. We have

$$\|z - y\| = \|c(v - y_0) - y\|$$

$$= c\|v - y_0 - c^{-1}y\|$$
$$= c\|v - y_1\|$$

where $y_1 = y_0 + c^{-1}y \in Y$. Hence $\|v - y_1\| \geq a$. So

$$\|z - y\| \geq ca = \frac{a}{\|v - y_0\|} \geq \frac{a}{a/\theta} = \theta$$

as desired. $\square$

Theorem 2.21. *Let X be a normed space over $\mathbb{R}$ or $\mathbb{C}$, then the unit ball is compact if and only if X is finite dimensional vector space.*

Proof. Finite dimensional, then isomorphic to $\mathbb{R}^n$ or $\mathbb{C}^n$, so trivial.

Now suppose $\dim X = \infty$ and assume towards a contradiction that $B = \{x : \|x\| \leq 1\}$ is compact. We choose any x_1 of norm 1. This generates a one dimensional subspace X_1 of X, which is closed and a proper subspace of X since $\dim X = \infty$. By Riesz's Lemma, there is an $x_2 \in X$ of norm 1 such that

$$\|x_2 - x_1\| \geq \theta = \frac{1}{2}.$$

Then the elements x_1, x_2 generate a two dimensional proper closed subspace X_2 of X. Again by Riesz's Lemma, we can pick x_3 that is distance at least $1/2$ away from x_1 and x_2. Doing this inductively, we would construct a sequence with no convergence subsequence, hence contradicting the compactness of B. $\square$

2.3 Linear Operator

Notation: $\mathscr{D}(T)$ denotes the domain of T; $\mathscr{R}(T)$ denotes the range of T; $\mathcal{N}(T)$ denotes the null space / kernel of T.

Definition 2.22 (Linear Operator)**.** *A linear operator T is an operator such that*

1. *the domain $\mathscr{D}(T)$ of T is a vector space and the range $\mathscr{R}(T)$ lies in a vector space over the same field.*
2. *for all $x, y \in \mathscr{D}(T)$ and scalars α,*

$$T(x + y) = Tx + Ty \quad T(\alpha x) = \alpha Tx.$$

Recall the following results from Linear Algebra:

Proposition 2.23. *Let T be a linear operator. Then*

1. *The range $\mathscr{R}(T)$ is a vector space;*
2. *If $\dim \mathscr{D}(T) = n < \infty$, then $\dim \mathscr{R}(T) \leq n$. In fact, in this case we have $\dim \mathscr{D}(T) = \dim \mathscr{R}(T) + \dim \mathcal{N}(T)$;*
3. *The null space $\mathcal{N}(T)$ is a vector space;*

4. *If T is bijective, then T^{-1} exists and is a linear operator. In particular we have a bijection between the basis of $\mathscr{D}(T)$ and $\mathscr{R}(T)$.*

5. $(ST)^{-1} = T^{-1}S^{-1}$

Definition 2.24 (Commutativity)**.** *Let V be a vector space and $S : X \to X$ and $T : X \to X$ be operators. Then we say S and T **commute** if $ST = TS$.*

Definition 2.25 (Bounded Linear Operator)**.** *Let X and Y be normed spaces and $T : \mathscr{D}(T) \to Y$ a linear operator, where $\mathscr{D}(T) \subset X$. The operator T is said to be **bounded** if there is a real number c such that for all $x \in \mathscr{D}(T)$,*

$$\|Tx\|_Y \leq c\|x\|_X.$$

*If this is the case, then we define the **operator norm** of the operator to be*

$$\|T\|_{op} = \sup_{\substack{x \in \mathscr{D}(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|}.$$

The following are easily to verify:

Lemma 2.26.

1. *The operator norm is a norm.*

2. *An alternative formula for the operator norm of T is*

$$\|T\| = \sup_{\substack{x \in \mathscr{D}(T) \\ \|x\|=1}} \|Tx\|.$$

3. *$\|Tx\| \leq \|T\|\|x\|$ for all $x \in \mathscr{D}(T)$. If T is bounded, then for any $x \in \mathscr{D}(T)$ with $\|x\| < 1$, then $\|Tx\| < \|T\|$.*

4. *If a normed space X is finite dimensional, then every linear operator on X is bounded.*

5. *$\|T_1 T_2\| \leq \|T_1\|\|T_2\|$, $\|T^n\| \leq \|T\|^n$.*

Theorem 2.27 (Continuity and boundedness)**.** *Let $T : \mathscr{D}(T) \to Y$ be a linear operator, where $\mathscr{D}(T) \subset X$ and X, Y are normed spaces. Then the following are equivalent:*

1. *T is bounded;*

2. *T maps bounded sets to bounded sets;*

3. *T is continuous;*

4. *T is sequentially continuous;*

5. *T is uniformly continuous and Lipschtiz continuous with Lipschtiz constant $\|T\|$;*

6. *T is continuous at a single point $x \in X$*

Corollary 2.27.1. *The null space $\mathcal{N}(T)$ of a bounded linear operator T is closed.*

Remark 2.27.1. *However, the range of a bounded linear operator is not closed in Y. Consider $T : \ell^\infty \to \ell^\infty$ defined by $(x_i) \mapsto (x_i/i)$. Then we consider the sequence*

$$x^{(n)} = (\sqrt{1}, \sqrt{2}, \cdots, \sqrt{n}, 0, 0, \cdots)$$

Then $Tx^{(n)} \to y = (1/\sqrt{i})_{i=1}^\infty$, however, y is not in the range of T. The same example also shows that the inverse $T^{-1} : \mathscr{R}(T) \to X$ of a bounded linear operator need not be bounded.

Proposition 2.28. *Let $T : \mathscr{D}(T) \to Y$ be a bounded linear operator, where $\mathscr{D}(T)$ lies in a normed space X and Y is a Banach space. Then T has an extension*

$$\tilde{T} : \overline{\mathscr{D}(T)} \to Y$$

where $\tilde{T}$ is a bounded linear opeartor of norm $\|\tilde{T}\| = \|T\|$.

Proof. The extension is unique by continuity. And by parsing to limit, we have $\|\tilde{T}\| \le \|T\|$ and clearly $\|\tilde{T}\| \ge \|T\|$. Hence we have equality. $\square$

Notation: Let X and Y be normed spaces (both real or complex), we use $L(X,Y)$ to denote the set of all linear operators from X to Y and we use $B(X,Y)$ to denote the set of all bounded linear operators from X into Y. Then $B(X,Y)$ is automatically a normed linear space as the following theorem states.

Theorem 2.29. *The vector space $B(X,Y)$ of all bounded linear operators from a normed space X into a normed space Y is a normed space with norm defined by*

$$\|T\| = \sup_{\substack{x\in X \\ x\neq 0}} \frac{\|Tx\|}{\|x\|}.$$

Theorem 2.30 (Completeness)**.** *If Y is a Banach space, then $B(X,Y)$ is a Banach space.*

Proof. Let $\{T_n\}$ be a Cauchy sequence in $B(X,Y)$, we show that (T_n) converges to an operator $T \in B(X,Y)$. Since $\{T_n\}$ is Cauchy, for every $\epsilon > 0$, there is an N such that

$$\|T_n - T_m\| < \epsilon \quad (m, n > N).$$

For all $x \in X$ and $m, n > N$ we have

$$\|T_n x - T_m x\| = \|(T_n - T_m)x\| \le \|T_n - T_m\|\|x\| < \epsilon\|x\|.$$

So $\{T_n(x)\}$ is a Cauchy sequence in Y for any $x \in X$, let its limit be y_x. Then we define $T(x) = y_x$. By linearity of limits, T is a linear operator. Now since $T_n x \to T_x$ for all x, then $\|T\| \le \sup_n \|T_n\|$, so $T \in B(X,Y)$. Lastly, we show $\|T_n - T\| \to 0$. Since $\|\cdot\|_Y$ is continuous, then

$$\|T_n x - Tx\| = \|T_n x - \lim_{m\to\infty} T_m x\| = \lim_{m\to\infty} \|T_n x - T_m x\| \le \epsilon\|x\|$$

where n is large enough. Hence $\|T_n - T\| \to 0$. $\square$

Remark 2.30.1. *In fact, we can show the converse: if $B(X,Y)$ is a Banach space, then Y is a Banach space; which implies that if Y is not a Banach space, then $B(X,Y)$ is not a Banach Space.*

2.4 Linear Functionals and Dual Space

Definition 2.31 (Linear functional)**.** *A* ***linear functional*** *f is a linear operator with domain in a vector space X and range in the scalar field $\mathbb{F}$ of X (We will always consider $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$).*

A ***bounded linear*** *functional f is a bounded linear operator*

Definition 2.32 (Dual Space)**.** *Let X be a normed vector space, then the* ***algebraic dual space*** *X^* consists of all linear functional. The* ***topological dual space*** *X' consists of all bounded linear functional. That is $X^* = L(X, \mathbb{F})$ and $X' = B(X, \mathbb{F})$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.*

Remark 2.32.1. *It is clear that when X is finite dimensional, then $X^* = X'$.*

Corollary 2.32.1. *If $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, then the dual space X' of a normed space X is a Banach Space.*

Proof. Follows from Theorem 2.30. □

Proposition 2.33. *Let X be an n-dimensional vector space and $E = \{e_1, \cdots, e_n\}$ be a basis for X. Then $F = \{f_1, \cdots, f_n\}$ where $f_i(e_j) = \delta_{ij}$ is a basis for X^*. In particular we have $\dim X^* = \dim X = n$.*

Lemma 2.34. *Let X be a finite dimensional vector space, if $x_0 \in X$ is such that $f(x_0) = 0$ for all $f \in X^*$, then $x_0 = 0$.*

Definition 2.35 (Embedding)**.** *Let X and Y be vector spaces, we say that X is* ***embeddable*** *if X is isomorphic with a subspace of a vector space Y.*

Definition 2.36 (Reflexive Space)**.** *There is a canonical natural embedding from X to X^{**} given by*

$$\phi : X \to X^{**}, \;\; x \mapsto \hat{x}$$

where $\hat{x}(f) = f(x)$ for any $f \in X^*$. If ϕ is also surjective, then we say that X is ***algebraically reflexive****. Moreover, if X is a normed space, and ϕ is an isometry (isometric linear isomorphism) from X to $(X')'$, then X is said to be* ***reflexive space****.*

Theorem 2.37. *A finite dimensional vector space is algebraically reflexive.*

Proof. Dimension argument. □

Proposition 2.38. *If the normed vector space over $\mathbb{R}$ or $\mathbb{C}$ is infinite dimensional, i.e. $\dim X = \infty$. Then $X' \neq X^*$.*

Proof. Since there is unbounded linear functional on X. □

Example:

1. The dual space of $\mathbb{R}^n$ is $\mathbb{R}^n$.

   Proof. We have $(\mathbb{R}^n)' = (\mathbb{R}^n)^*$. Then for every $f \in (\mathbb{R}^n)^*$ and $x = \sum_1^n \xi_k e_k$. Then

$$f(x) = \sum_1^n c_k \xi_k,$$

where $c_k = f(e_k)$. By Cauchy-Schwarz Inequality, we have

$$|f(x)| \le \sum |c_k \xi_k| \le (\sum c_k^2)^{1/2} (\sum \xi_k^2)^{1/2} = \|x\|_{L^2} \left(\sum \xi_k^2\right)^{1/2}.$$

Taking the supremum over all x of norm 1, we obtain

$$\|f\| \le \left(\sum \xi_k^2\right)^{1/2}.$$

When $x = (\xi_1, \cdots, \xi_n)$ equality is achieved. Hence we have

$$\|f\| = \left(\sum \xi_k^2\right)^{1/2}.$$

This proves that the norm of f is the Euclidean norm, and $\|f\| = \|c\|$, where $c = (\xi_1, \cdots, \xi_n) \in \mathbb{R}^n$. Hence the mapping of $(\mathbb{R}^n)'$ onto $\mathbb{R}^n$ defined by $f \mapsto c = (c_k)$, $c_k = f(e_k)$ is norm preserving and, since it is linear and bijective, it is an isomorphism. □

2. The dual space of ℓ^1, $(\ell^1)' = \ell^\infty$.

Proof. A Schauder basis for ℓ^1 is $\{e_k\}$, wehre $e_k = (\delta_{kj})$. Then every $x \in \ell^1$ has a unique representation

$$x = \sum_{k=1}^{\infty} \xi_k e_k.$$

Let $f \in (\ell^1)'$, since f is linear and bounded, then

$$f(x) = \sum_{k=1}^{\infty} \gamma_k \xi_k, \quad \gamma_k = f(e_k).$$

where the numbers $\gamma_k = f(e_k)$ are uniquely determined by f. Also $\|e_k\| = 1$ and

$$|\gamma_k| = |f(e_k)| \le \|f\| \|e_k\| = \|f\|,$$

hence $\sup_k |\gamma_k| \le \|f\|$, so $(\gamma_k) \in \ell^\infty$.

On the other hand, for every $b = (\beta_k) \in \ell^\infty$, we can obtain a corresponding bounded linear functional g on ℓ^1 by defining

$$g(x) = \sum_{k=1}^{\infty} \xi_k \beta_k$$

where $x = (\xi_k) \in \ell^1$. Then g is linear, and boundedness follows from

$$|g(x)| \le \sum |\xi_k \beta_k| \le (\sup_i |\gamma_i|) \sum |\xi_k| = \|x\| \sup_i |\gamma_i|.$$

Taking the supremum over all x of norm 1, we see that

$$\|f\| \le \sup_i |\gamma_i|.$$

This shows that

$$\|f\| = \sup_i |\gamma_i| = \|(\gamma_i)\|_{\ell^\infty}.$$

Hence this formula can be written as $\|f\| = \|c\|_{\ell^\infty}$, where $c = (\gamma_i) \in \ell^\infty$. It shows that the bijective linear mapping of (ℓ^1); onto ℓ^∞ defined by $f \mapsto c = (\gamma_i)$ is an isomorphism. $\square$

3. The dual space of ℓ^p is ℓ^q, here $1 < p < \infty$, and q is the Hölder conjugate of p.

Proof. A Schauder basis for ℓ^p is $\{e_k\}$ again. Then every $x \in \ell^p$ has unique representation

$$x = \sum_{k=1}^{\infty} \xi_k e_k.$$

Let $f \in (\ell^p)'$. Since f is linear and bounded, then

$$f(x) = \sum_{k=1}^{\infty} \gamma_k \xi_k, \quad \gamma_k = f(e_k).$$

where the numbers $\gamma_k = f(e_k)$ are uniquely determined by f. Let q be the Hölder conjugate of p, and consier $x_n = (\xi_k^{(n)})$ with

$$\xi_k^{(n)} = \begin{cases} |\gamma_k|^q/\gamma_k & \text{if } k \le n \text{ and } \gamma_k \neq 0 \\ 0 & \text{if } k > n \text{ or } \gamma_k = 0 \end{cases}.$$

Then we have

$$f(x_n) = \sum_{k=1}^{\infty} \xi_k^{(n)} \gamma_k = \sum_{k=1}^{n} |\gamma_k|^q.$$

Since $(q-1)p = q$, then we have

$$\begin{aligned} f(x_n) &\le \|f\| \|x_n\| \\ &= \|f\| \left(\sum |\xi_k^{(n)}|^p \right)^{1/p} \\ &= \|f\| \left(\sum |\gamma_k|^{(q-1)p} \right)^{1/p} \\ &= \|f\| \left(\sum |\gamma_k|^q \right)^{1/p} \end{aligned}$$

Where the last two sums is from 1 to n. Hence we conclude that

$$f(x_n) = \sum |\gamma_k|^q \le \|f\| \left(\sum |\gamma_k|^q \right)^{1/p}$$

by dividing $(\sum |\gamma_k|^q)^{1/p}$ on both sides, we conclude that

$$\|f\| \ge \left(\sum_{k=1}^{n} |\gamma_k|^q \right)^{1/q}.$$

By taking limit, we conclude that $\|(\gamma_k)\|_{\ell^q} < \|f\|$, so $(\gamma_k) \in \ell^q$.

Conversely, for any $b \in (\beta_k) \in \ell^q$, we get a corresponding bounded linear functional g on ℓ^p. In fact, we may define g on ℓ^p by setting

$$g(x) = \sum_{k=1}^{\infty} \xi_k \beta_k$$

where $x = (\xi_k) \in \ell^p$. Then g is linear and boundedness follows from Hölder's inequality. Hence $g \in (\ell^p)'$.

Lastly, we prove that the norm of f is the norm on the space ℓ^q. By Hölder's inequality, we have

$$|f(x)| = |\sum \xi_k \gamma_k| \leq (\sum |\xi_k|^p)^{1/p} (\sum |\gamma_k|^q)^{1/q} \leq \|x\| \|(\gamma_k)\|_{\ell^q}.$$

Since we have already established the inverse inequality, we conclude that

$$\|f\| = \|(\gamma_k)\|_{\ell^q}.$$

□

4. Let c_0 denote the set of sequences that converges to 0. Then $(c_0)' = \ell^1$.

   Proof. Note that c_0 has Schauder basis $\{e_k\}$, then a similar argument follows. □

5. The dual of ℓ^∞ is not ℓ^1.

   Proof. We show that there are bounded functionals on ℓ^∞, which are not of the form

   $$f(y) = \sum_k x_k y_k$$

   for some $x \in \ell^1$. Let $f : c \to \mathbb{R}$ (where $c \subseteq \ell^\infty$ denotes the set of convergent sequences) be given by $f(x) = \lim_n x_n$. Then f is bounded, as $|\lim_n x_n| \leq \sup_n |x_n| = \|x\|$. Let $g : \ell^\infty \to \mathbb{R}$ be a Hahn-Banach extension. If g were of the above mentioned form, we would have (with e_n the n-th unit sequence)

   $$x_n = g(e_n) = f(e_n) = 0$$

   hence $g = 0$. But $g \neq 0$, as for example $g(1, 1, \dots) = 1$. □

Theorem 2.39. *Let X be a normed linear space over $\mathbb{F}$ and suppose f be a linear functional on X. Then f is continuous if and only if its kernel is closed.*

Proof. Suppose f is trivial, then the statement clearly holds. Hence we assume f is not trivial, then we can clearly see by linearity, f is surjective.

If f is continuous, then clearly the kernel is closed. So we prove the converse.

Let $Y = \ker f$. Since Y is a closed subspace of X, then the hypothesis of the Riesz' Lemma is satisfied. Thus there exists $z \in X$ such that $\|z\| = 1$ and for any $y \in Y$, $\|z - y\| \geq \frac{1}{2}$. Since Y is a subspace, if $y \in Y$, then $-y \in Y$.

Hence
$$\|z - y\| = \|z + (-y)\| \geq \frac{1}{2}$$
for all $y \in Y$.

We check that $z \notin Y$, since $\|z + y\| \geq 1/2$ for all $y \in Y$. This shows $f(z) \neq 0$, hence we can divide by $f(z)$. Next, we compute the image of $x - [f(x)/f(z)]z$ under f:

$$\begin{aligned}
f\left(x - \frac{f(x)}{f(z)}z\right) &= f(x) - f\left(\frac{f(x)}{f(z)}z\right) \\
&= f(x) - \frac{f(x)}{f(z)}f(z) \\
&= f(x) - f(x) \\
&= 0
\end{aligned}$$

Hence $x - [f(x)/f(z)]z \in Y$.

Next suppose $x \in Y$, then $|f(x)| = 0$, so the inequality $|f(x)| \leq 2|f(x)|\|x\|$ trivially holds. When $x \notin Y$, $|f(x)| \neq 0$. Then as Y is a subspace, we have

$$x - [f(x)/f(z)]z \in Y \Longrightarrow \frac{f(z)}{f(x)}x - z \in Y.$$

We know for any $y \in Y$, $\|z + y\| \geq 1/2$, hence we have

$$\begin{aligned}
\|z + \frac{f(z)}{f(x)}x - z\| &\geq \frac{1}{2} \\
\|\frac{f(z)}{f(x)}x\| &\geq \frac{1}{2} \\
2\frac{|f(z)|}{|f(x)|}\|x\| &\geq 1 \\
2|f(z)|\|x\| &\geq |f(x)|
\end{aligned}$$

as desired.

Lastly to show f is continuous, it suffices to show that f is bounded. Let $x \in X$ be arbitrary, then we have

$$|f(x)| \leq 2|f(z)|\|x\|.$$

This shows that

$$\sup_{x \in X \setminus \{0\}} \frac{|f(x)|}{\|x\|} \leq 2|f(z)| < \infty.$$

Hence f is bounded, thus continuous. □

Corollary 2.39.1. *A linear map from a normed linear space into a finite dimensional normed linear space is continuous if and only if its kernel is closed.*

Proof. Let $f : X \to Y$ be a linear map from a normed linear space X into a finite dimensional normed linear space

Y. Since Y is a normed space, hence a metric space, then single is closed in Y. So suppose f is continuous, then $\ker f = f^{-1}(\{0\})$ being the preimage of a closed set is closed.

Let T be an isomorphism between Y and $\mathbb{F}^n$ that is also an homeomorphism. Then f is continuous if and only if Tf is continuous, as T is a homeomorphism. Moreover, as T is an isomorphism, then $\ker Tf = \ker f$. So instead of considering arbitrary finite dimensional normed linear spaces, we can just assume $Y = \mathbb{F}^n$ instead.

Since $\ker f$ is a closed subspace of X, then $X/\ker f$ is a normed vector space. Moreover, the canonical projection π of X onto $X/\ker f$ is continuous, since

$$\|\pi(x + \ker f)\| = \liminf_{y \in \ker f} \|x - y\| \leq \|x - 0\| = \|x\|.$$

Now we define $\tilde{T} : X/\ker f \to \mathbb{F}^n$ by $\tilde{T}(x + \ker f) = T(x)$, this is well-defined, since if $x + \ker f = x' + \ker f$, then $x - x' \in \ker f$, so $T(x) = T(x')$. It is not hard to see that $\tilde{T}$ and π are linear, and $T = \tilde{T} \circ \pi$. Hence if we can show that if $\tilde{T}$ is continuous, then T would be continuous.

By the first isomorphism theorem, $X/\ker f \cong f(X) \subset \mathbb{F}^n$, so $X/\ker f$ is a finite dimensional subspace. Let π_i be the canonical projection of $\mathbb{F}^n$ onto the i^{th} coordinate vector. Then by results from point-set topology, $\tilde{T}$ is continuous if and only if $\pi_i \circ \tilde{T}$ is continuous for $i = 1, \cdots, n$. Notice $\pi_i \circ \tilde{T}$ is a linear map from $X/\ker f$ to $\mathbb{F}$. This is continuous if and only if $\ker(\pi_i \circ \tilde{T})$ is continuous. Because $\ker(\pi_i \circ \tilde{T}) \subset X/\ker f$ is a finite dimensional normed space, then it is closed, as it is homeomorphic to $\mathbb{R}^k$ for some $k \in \mathbb{N}$. Hence $\ker(\pi_i \circ \tilde{T})$ is closed for any $i \in \{1, \cdots, n\}$. This shows f is continuous if $\ker f$ is closed. $\square$

3 Inner Product Spaces

3.1 Hilbert Space

Definition 3.1 (Inner Product)**.** *An **inner product** on vector space X is a mapping of $X \times X$ into the scalar field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ of X satisfying the following properties:*

1. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.

2. $\langle ax, y \rangle = a\langle x, y \rangle$ *where a is a scalar.*

3. $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

4. $\langle x, x \rangle \geq 0$, *and it is zero if and only if $x = 0$.*

Remark 3.1.1. *Every inner product induces a norm hence a metric.*

Remark 3.1.2. *The third property of an inner product is called **sesquilinear**, that is it is conjugate linear in the second factor. If $\mathbb{F} = \mathbb{R}$, then $\langle x, y \rangle = \langle y, x \rangle$, so the inner product is symmetric.*

Lemma 3.2. *Suppose $\langle \cdot, \cdot \rangle$ is an inner product, then*

- $\langle \alpha x + \beta y, z \rangle = \alpha\langle x, z \rangle + \beta\langle y, z \rangle$.
- $\langle x, \alpha y + \beta z \rangle = \bar{\alpha}\langle x, y \rangle + \bar{\beta}\langle x, z \rangle$.

Definition 3.3 (Hilbert Space)**.** *An **inner product space** (or pre-Hilbert space) is a vector space X with an inner product defined on X. A **Hilbert space** is a complete inner product space under the metric inducedby the inner product.*

Remark 3.3.1. *Hilbert Spaces are thus Banach Spaces, but not all Banach spaces are inner product spaces.*

Example:

1. The standard inner product on $\mathbb{R}^n$ induces the Euclidean metric. Moreover, this is a real Hilbert space.

2. The standard inner product on $\mathbb{C}^n$ induces the Euclidean metric. Moreover, this is a complex Hilbert space.

3. The inner product
$$\langle x, y \rangle = \int_a^b x(t)\overline{y(t)}dt$$
makes the space of $L^2([a, b])$ functions an inner product space. In fact, this is also a Hilbert space.

Lemma 3.4. *If X is a finite dimensional vector space, and $\{e_j\}$ is a basis for X. The any inner product on X is completely determined by its values $\gamma_{jk} = \langle e_j, e_k \rangle$. Conversely, if H is a **Hermitian Matrix** in the base field K, that is $H = \overline{H^T}$, then H induces an inner product on X.*

Proposition 3.5. *A normed space X is an inner product space if and only if it satisfies the parallelogram equality, that is for all $x, y \in X$,*
$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

If this is the case, then if X is a real inner product space, we have

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}.$$

If X is a complex inner product space, then

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2 + \|u+iv\|^2 i - \|u-iv\|^2 i}{4}.$$

Example: The space ℓ^p with $p \neq 2$ is not an inner product space, hence not a Hilbert space. Since if we take $x = (1, 1, 0, 0, \cdots)$ and $y = (1, -1, 0, 0, \dots)$, then

$$\|x\| = \|y\| = 2^{1/p} \quad \text{and} \quad \|x+y\| = \|x-y\| = 2.$$

Then we see that the parallelogram equality is not satisfied. Similarly, we can show that $C[a, b]$ with the sup norm is not an inner product space. Since for $x \equiv 1$ and $y(t) = \frac{t-a}{b-a}$, the parallelogram equality is not satisfied.

Definition 3.6 (Orthogonality). *An element x of an inner product space X is said to be* ***orthogonal*** *to an element $y \in X$ if*

$$\langle x, y\rangle = 0$$

and denoted $x \perp y$. If $A, B \subset X$. Then we write $A \perp B$ if $a \perp b$ for all $a \in A$ and all $b \in B$.

Theorem 3.7 (Pythagorean Theorem). *Suppose X is an inner product space and $x, y \in X$ are such that $x \perp y$, then*

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

Proof. By direct computation,

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 + \|y\|^2\end{aligned}$$

□

Remark 3.7.1. *The converse of this statement holds if X is a real inner product space. However, the converse may not be true in a complex inner product space, for example let $x = 1$ and $y = i$ in $\mathbb{C}^1$, then 1 is not orthogonal to i.*

Lemma 3.8. *Suppose $\{x_1, \cdots, x_m\}$ is a mutually orthogonal nonzero vectors, then this set is linearly independent.*

Proof. Multiple x_i, to show the coefficient of $x_i = 0$. □

Proposition 3.9 (Schwarz Inequality, Triangle Inequality). *An inner product and the corresponding norm satisfy the Schwarz inequality and the triangle inequality as follows:*

1. *(Schwarz Inequality) For $x, y \in X$*

$$|\langle x, y\rangle| \le \|x\|\|y\|$$

where the equality sign holds if and only if $\{x, y\}$ is a linearly dependent set.

2. *(Triangle Inequality) The norm induced by the inner product also satisfies*

$$\|x+y\| \leq \|x\| + \|y\|$$

with equality sign holds if and only if $y = 0$ or $x = cy$ for some real $c \geq 0$.

Remark 3.9.1. *Cauchy Schwarz inequality also holds if we only have a semi-inner product.*

Proof. If $y = 0$, then clearly we have Schwarz inequality. So assume $y \neq 0$, then for any scalar α, we have

$$\begin{aligned} 0 &\leq \|x - \alpha y\|^2 \\ &= \langle x - \alpha y, x - \alpha y \rangle \\ &= \langle x, x \rangle - \bar{\alpha}\langle x, y \rangle - \alpha[\langle y, x \rangle - \bar{\alpha}\langle y, y \rangle]. \end{aligned}$$

Let $\bar{\alpha} = \frac{\langle y,x \rangle}{\langle y,y \rangle}$. Then we have

$$0 \leq \langle x, x \rangle - \frac{\langle y, x \rangle}{\langle y, y \rangle} = \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2}$$

This gives the Schwarz Inequality. We see equality holds if and only if $x - \alpha y = 0$, so x and y linearly dependent.

Next we check Triangle's inequality. We have

$$\|x+y\|^2 = \langle x+y, x+y \rangle = \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2.$$

By Schwarz inequality,

$$|\langle x, y \rangle| = |\langle y, x \rangle| \leq \|x\|\|y\|.$$

So

$$\|x+y\|^2 \leq (\|x\| + \|y\|)^2$$

gives the desired triangle inequality. We see equality holds if and only if

$$\langle x, y \rangle + \langle y, x \rangle = \mathrm{Re}\langle x, y \rangle = 2\|x\|\|y\|.$$

But since $\|x\|\|y\| \geq |\langle x, y \rangle|$, we must have $y = 0$ or $x = cy$, where c is real. However, we may check that equality only holds if and only if $c \geq 0$ for the case $x = cy$. $\square$

Lemma 3.10. *In an inner product space, the following are equivalent:*

1. *$x \perp y$;*

2. *$\|x + \alpha y\| = \|x - \alpha y\|$ for all scalar α;*

3. *$\|x + \alpha y\| \geq \|x\|$ for all scalars.*

Proof. Suppose $x \perp y$, then we clearly have $\|x + \alpha y\| = \|x - \alpha y\|$ by Pythagorean Theorem. Now if $\|x + \alpha y\| = \|x - \alpha y\|$, then

$$2\|x + \alpha y\|^2 = \|x + \alpha y\|^2 + \|x - \alpha y\|^2 = 2\|x\|^2 + 2\|y\|^2 \geq 2\|x\|^2.$$

Lastly, if $\|x + \alpha y\| \geq \|x\|$ for all scalars α. Then we have

$$\|y\|^2 \geq -\frac{2}{\alpha^2}\text{Re}\langle x, \alpha y\rangle$$

for any α. If $\langle x, \alpha y\rangle \neq 0$, then we can take α such that the inequality does not hold. Hence we must have $x \perp y$. □

Lemma 3.11 (Continuity of Inner Product)**.** *If in an inner product space, $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$.*

Proof.

$$\begin{aligned}
|\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n, y_n\rangle - \langle x_n, y\rangle + |\langle x_n, y\rangle - \langle x, y\rangle| \\
&\leq |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\
&\leq \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\|.
\end{aligned}$$

Since $x_n \to x$, then $\|x_n\|$ is bounded. Since $y_n - y \to 0$ and $x_n - x \to 0$, we see that $\langle x_n, y_n\rangle \to \langle x, y\rangle$. □

Corollary 3.11.1. *In an inner product space H, if $y \perp x_n$ for all $n \in \mathbb{N}$ and $x_n \to x$, then $x \perp y$.*

Proof. Since the inner product is continuous. □

Definition 3.12 (Isomorphism of Inner Product Space)**.** *An **isomorphism** T of an inner product space X onto an inner product space $\tilde{X}$ over the same field is a bijective linear operator $T : X \to \tilde{X}$ which preserves the inner product, that is, for all $x, y \in X$,*

$$\langle Tx, Ty\rangle = \langle x, y\rangle.$$

*In this case, X and $\tilde{X}$ are called **isomorphic inner product spaces**.*

Definition 3.13 (Subspace)**.** *A **subspace** Y of an inner product space X is a vector subspace of X equipped with the inner product obtained from the restriction of the inner product on X.*

Similar to the completion of metric spaces, we have the following theorem:

Theorem 3.14 (Completion)**.** *For any inner product space X there exists a Hilbert space$H$ and an isomorphism A from X onto a dense subspace $W \subset H$. The space H is unique except for isomorphism.*

Theorem 3.15. *Let Y be a subspace of a Hilbert space H. Then:*

1. *Y is complete if and only if Y is closed in H.*
2. *If Y is finite dimensional, then Y is complete.*
3. *If H is separable, so is Y. More generally, every subset of a separable inner product space is separable.*

Proof. We prove the last statement. Since H is separable, and H is a metric space, it is second countable. Hence Y being a subspace of H is second countable. Then Y is separable. □

Proposition 3.16 (Zero Operator)**.** *Let $T : X \to X$ be a bounded linear operator on a complex inner product space X. If $\langle Tx, x\rangle = 0$ for all $x \in X$, then $T = 0$.*

Remark 3.16.1. *This is not true for real inner product space, counter example: rotation by $\pi/2$ on $\mathbb{R}^2$.*

Proof. We show for any $u, v \in T$, we have $\langle Tu, v\rangle = 0$, then $\langle Tu, Tu\rangle = 0$. By assumption,

$$\begin{aligned} 0 &= \langle T(u+av), u+av\rangle \\ &= \langle Tu, u\rangle + \bar{\alpha}\langle Tu, v\rangle + a\langle Tv, u\rangle + a\bar{a}\langle Tv, v\rangle \\ &= \bar{a}\langle Tu, v\rangle + a\langle Tv, u\rangle \end{aligned}$$

Then let $a = 1$ and $a = i$, we see

$$\langle Tu, v\rangle + \langle Tv, u\rangle = 0 \quad \text{and} \quad \langle Tu, v\rangle - \langle Tv, u\rangle = 0$$

which proves that $\langle Tu, v\rangle = 0$. □

3.2 Orthogonal Complements and Direct Sums

Theorem 3.17 (Minimizing Vector)**.** *Let X be an inner product space and $M \neq \emptyset$ a convex subset which is complete (in the metric induced by the inner product). Then for every given $x \in X$, there exists a unique $y \in M$ such that*

$$\delta = \inf_{\tilde{y}\in M} \|x - \tilde{y}\| = \|x - y\|.$$

Remark 3.17.1. *Convexity and completeness are both necessary for the existence of the y.*

Proof. Existence: by definition of an infimum, there is a sequence $\{y_n\} \subset M$ such that $\delta_n = \|x - y_n\| \to \delta$. We show that $\{y_n\}$ is Cauchy. Let $v_n := y_n - x$, then $\|v_n\| = \delta_n$ and

$$\|v_n + v_m\| = \|y_n + y_m - 2x\| = 2\left\|\frac{1}{2}(y_n + y_m) - x\right\| \geq 2\delta.$$

because M is convex, so $\frac{1}{2}(y_n + y_m) \in M$. Furthermore, we have $y_n - y_m = v_n - v_m$. Hence by the parallelogram equality

$$\begin{aligned} \|y_n - y_m\|^2 &= \|v_n - v_m\|^2 \\ &= -\|v_n + v_m\|^2 + 2(\|v_n\|^2 + \|v_m\|^2) \\ &\leq -(2\delta)^2 + 2(\delta_n^2 + \delta_m^2) \end{aligned}$$

This shows that $\{y_m\}$ is Cauchy. Since M is complete, then $\{y_n\}$ converges, say $y_n \to y \in M$. Since $y \in M$, we have $\|x - y\| \geq \delta$. On the other hand

$$\|x - y\| \leq \|x - y_n\| + \|y_n - y\| = \delta_n + \|y_n - y\| \to \delta$$

as $n \to \infty$. This shows that $\|x - y\| = \delta$.

Uniqueness: we assume that $y \in M$ and $y' \in M$ both satisfy

$$\|x - y\| = \delta \quad \text{and} \quad \|x - y_0\| = \delta$$

By parallelogram equality

$$\begin{aligned}\|y - y'\|^2 &= \|(y - x) - (y' - x)\|^2 \\ &= 2\|y - x\|^2 + 2\|y' - x\|^2 - \|(y - x) + (y' - x)\|^2 \\ &= 2\delta^2 + 2\delta^2 - 2^2\|\frac{1}{2}(y + y') - x\|^2.\end{aligned}$$

Since $\frac{1}{2}(y + y') \in M$, then the last term on is no less than $4\delta^2$. Hence we conclude that $\|y - y'\| = 0$, that is $y = y'$. □

Corollary 3.17.1. *Let X be an inner product space, and let $M \subset X$ be a convex complete subset. Let $\{x_n\} \subset M$ be such that $\|x_n\| \to d$, where $d = \inf_{x \in M} \|x\|$. Then $\{x_n\}$ converges in M, hence in X.*

Proof. The proof is exactly the same as the existence part of the Minimizing Vector Theorem, its just we take $x \in X$ to be the zero vector. □

Lemma 3.18. *Let Y be a complete subspace, and $x \in X$ fixed. Suppose $y \in Y$ is such that*

$$\|x - y\| = \inf_{\tilde{y} \in Y} \|x - \tilde{y}\|.$$

Then $z := x - y \perp Y$.

Proof. Suppose $z \perp Y$ does not hold, then there exists $y_1 \in Y$ such that

$$\langle z, y_1 \rangle = \beta \neq 0.$$

Clearly, $y_1 \neq 0$. Furthermore, for any scalar α, we have

$$\begin{aligned}\|z - \alpha y_1\|^2 &= \langle z - \alpha y_1, z - \alpha y_1 \rangle \\ &= \langle z, z \rangle - \bar{\alpha}\langle z, y_1 \rangle - \alpha[\langle y_1, z \rangle - \bar{\alpha}\langle y_1, y_1 \rangle] \\ &= \langle z, z \rangle - \bar{\alpha}\beta - \alpha[\bar{\beta} - \bar{\alpha}\langle y_1, y_2 \rangle].\end{aligned}$$

If we choose

$$\bar{\alpha} = \frac{\bar{\beta}}{\langle y_1, y_1 \rangle}$$

then the last term vanishes. So

$$\|z - \alpha y_1\|^2 = \|z\|^2 - \frac{|\beta|^2}{\langle y_1, y_2 \rangle} < \delta^2$$

where $\delta = \inf |x - \tilde{y}|$. But this is impossible since if we let $y_2 = y + \alpha y_1 \in Y$, then $\|x - y_2\| < \delta$. Thus we must have $\langle z, y_1 \rangle = 0$ for all $y_1 \in Y$. □

Definition 3.19 (Direct Sum)**.** *A vector space X is said to be the **direct sum** of two subspaces Y and Z of X, written*

$$X = Y \oplus Z,$$

*if each $x \in X$ has a unique representation $x = y + z$, $y \in Y$, $z \in Z$. If this is the case, then Z is called an **algebraic complement** of Y in X and vice versa. Y, Z are called a **complementary pair** of subspaces in X.*

Definition 3.20 (Orthogonal Complement)**.** *Let S be a subset of the inner product space X, then we define its* ***orthogonal complement****, $S^\perp$ by*

$$S^\perp = \{x \in X \mid x \perp S\}.$$

Theorem 3.21 (Direct Sum)**.** *Let Y be any closed subspace of a Hilbert space H. Then*

$$H = Y \oplus Z$$

where $Z = Y^\perp$.

Proof. Since H is complete and Y is closed, then Y is complete. Then for every $x \in H$, there exists $y \in Y$ such that

$$\|y - x\| = \inf_{\tilde{y} \in H} \|x - \tilde{y}\|$$

By Lemma (3.18), we also know that $y - x \perp Z$, hence $x = y + z$, where $y \in Y$ and $z \in Y^\perp$.

To prove the uniqueness of the representation, suppose

$$x = y + z = y_1 + z_1.$$

Then $y - y_1 = z_1 - z$, Since $z_1 - z \in Z = Y^\perp$, then $y - y_1 \in Y \cap Y^\perp = \{0\}$. This implies that $y = y_1$, so $z = z_1$. □

Corollary 3.21.1. *Let Y be a closed subspace of a Hilbert space X. Then X/Y in the quotient norm is isometrically isomorphic to $Y^\perp$.*

Proof. Direct verification, note that we can map $x \in Y^\perp$ to $x + Y$, then $\|x\| = \|x + Y\|$. □

Definition 3.22 (Orthogonal Projection)**.** *Let Y be any closed subspace of a Hilbert space H, we know*

$$H = Y \oplus Y^\perp.$$

Then for any $x \in H$, we define its ***orthogonal projection*** *on Y, to be the unique vector $y \in Y$ such that there exists $z \in Z$ satisfying $x = y + z$. The orthogonal projection gives a natural map, called the* ***projection operator****, P, where $P : H \to Y$, $x \mapsto Px = y$.*

Lemma 3.23 (Property of Projection Map)**.**

1. *P maps H onto Y.*
2. *P is the identity map on Y.*
3. *$Z = Y^\perp$ is the kernel of P.*
4. *P is idempotent, that is $P^2 = P$.*
5. *$Id - P$ is the projection map onto $Y^\perp$.*

Proposition 3.24 (Property of Orthogonal Complement)**.** *Let H be an inner product space and $S \subset H$, then*

1. *$S^\perp$ is a closed subspace.*

2. $S^\perp = (\overline{\operatorname{span} S})^\perp$.

3. $S^\perp \cap \overline{\operatorname{span} S} = \{0\}$, *hence* $S^\perp \cap \operatorname{span} S = \{0\}$.

4. $S \subset (S^\perp)^\perp$.

5. $\overline{\operatorname{span} S} \subset (S^\perp)^\perp$.

6. *If* $C \subset S$, *then* $S^\perp \subset C^\perp$.

7. $S^\perp = S^{\perp\perp\perp}$.

8. *If* S *is a vector subspace of* H, *then*

$$S^\perp = \{x \in H : \|x\| \leq \|x + v\| \ \forall v \in S\}.$$

where all the closure are taken in H *rather than its completion.*

Further, suppose H *is a Hilbert space. Then*

1. *If* W *is a subspace of* H, $(W^\perp)^\perp = \overline{W}$ *and* $W^\perp = \overline{W}^\perp$. *In particular, if* W *is closed if and only if* $W = W^{\perp\perp}$ *and shows that* $(W^\perp)^\perp$ *is the smallest closed closed subspace containing* W.

2. *The span of a subset* S *of* H *is dense in* H *if and only if* $S^\perp = \{0\}$.

Proof. $S^\perp$ is closed follows from Corollary (3.11.1). One can easily check that it is a subspace. It is clear that if $y \perp x$ for all $x \in S$, then by conjugate linearity in the second component of an inner product, we have $y \perp \operatorname{span} S$. Then by Corollary (3.11.1), we have $S^\perp = (\overline{\operatorname{span} S})^\perp$. Next, suppose $y \in S^\perp \cap \overline{\operatorname{span} S}$, then as $S^\perp = (\overline{\operatorname{span} S})^\perp$, we have $y \perp y$, so $y = 0$. It is easy to show that $S \subset (S^\perp)^\perp$ and $\overline{\operatorname{span} S} \subset (S^\perp)^\perp$. From the definition, we can easily see that $C \subset S$ implies $S^\perp \subset C^\perp$. We know $S^\perp \subset S^{\perp\perp\perp}$. Now since $S \subset S^{\perp\perp}$, then $S^\perp \supset (S^{\perp\perp})^\perp$. Lastly, by Lemma (3.10), we have property (8).

Next, if W is a subspace of H, then $\overline{W}$ is a closed subspace of H. Then $(W^\perp)^\perp = \overline{W}$ follows from the Direct Sum Decomposition (Theorem (3.21)). Since if $x \perp W^\perp$, then $x = y + z$, where $z \in W^\perp$ must be zero. We already know $W^\perp = \overline{W}^\perp$. Then clearly if W is closed, $W = W^{\perp\perp}$; and if $W = W^{\perp\perp}$, since the orthogonal complement is always closed, we have W is closed.
Lastly, if $\overline{\operatorname{span} S} = H$, then $S^\perp = \overline{\operatorname{span} S}^\perp = H^\perp = 0$. On the other hand, if $S^\perp = \{0\}$. Then $\overline{\operatorname{span} S}^\perp = \{0\}$. Suppose $x \notin \overline{\operatorname{span} S}$, then we know that $x - P_{\operatorname{span} S} x \neq 0$ and is orthogonal to $\overline{S}$, where P is the projection to $\overline{S}$. Hence we conclude that $\overline{S} = H$. □

3.3 Orthonormal Sets and Sequences

Definition 3.25 (Orthonormal Sets and Sequences)**.** *An* ***orthogonal set*** M *is an inner product space* X *is a subset* $M \subset X$ *whose elements are pairwise orthogonal. An* ***orthonormal set*** $M \subset X$ *is an orthogonal set in* X

whose elements have norm 1, that is, for all $x, y \in M$,

$$\langle x, y\rangle = \begin{cases} 0 & \text{if } x \neq y \\ 1 & \text{if } x = y. \end{cases}$$

If an orthogonal or orthonormal set M is countable, we can arrange it in a sequence $(x_n)_{n=1}^{\infty}$ and call it an ***orthogonal or orthonormal sequence*** *respectively.*

More generally, an indexed set, or family (x_α), $\alpha \in I$ is called ***orthogonal*** *if $x_\alpha \perp x_\beta$ for all $\alpha, \beta \in I$, $\alpha \neq \beta$. The family is called* ***orthonormal*** *if is orthogonal and all x_α have norm 1.*

Example: The sequence $\{e_n\}$ given by

$$e_0(t) = \frac{1}{\sqrt{2\pi}}, \quad e_n(t) = \frac{\cos nt}{\sqrt{\pi}}$$

is an orthonormal sequence of the space $C[0, 2\pi]$ endowed with the L^2 norm. Similarly, the sequence $\{\tilde{e}_n\}$ given by

$$\tilde{e}_n(t) = \frac{\sin nt}{\sqrt{\pi}}$$

is an orthonormal sequence.

Lemma 3.26 (Linear Independence)**.** *An orthonormal set is linearly independent even if the set if infinite.*

Lemma 3.27 (Representation Using Orthonormal Set)**.** *Suppose $E = \{e_1, \cdots, e_n\}$ are orthonormal and x is in the span of S, then*

$$x = \sum_{k=1}^{n} \langle x, e_k\rangle e_k$$

Moreover,

$$\|x\|^2 = \sum_{k=1}^{n} |\langle x, e_k\rangle|^2.$$

Proof. Firstly, we note that the representation is unique as E is linearly independent. Next, suppose

$$x = \sum_{k=1}^{n} a_k e_k,$$

then by taking inner product with e_j, we have

$$\langle x, e_j\rangle = a_j.$$

The second statement follows from Pythagorean Theorem. □

Corollary 3.27.1. *Let $E = \{e_1, \cdots, e_n\}$ be orthonormal and $x \in X$ be arbitrary. Define*

$$y := \sum_{k=1}^{n} \langle x, e_k\rangle e_k.$$

Then $x = y + z$, where $z = x - y$ satisfies that $z \perp y$, hence

$$\sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \leq \|x\|^2.$$

Remark 3.27.1. *This also show the projection of x onto $\operatorname{span}\{e_1, \cdots, e_n\}$ is given by y. Hence for all $\tilde{y} \in \operatorname{span}\{e_1, \cdots, e_n\}$, $\|x - y\|$ minimizes $\|x - \tilde{y}\|$.*

Proof. Note that in order to check $z \perp y$, just need to check $\langle z, e_k \rangle = 0$ for all $k \in \{1, \cdots, n\}$. Lastly, by Pythagorean Theorem, we conclude that

$$\|x\|^2 = \|y\|^2 + \|z\|^2 \geq \|y\|^2.$$

□

Theorem 3.28 (Bessel Inequality)**.** *Let $\{e_k\}_{k=1}^{\infty}$ be an orthonormal sequence in an inner product space X. Then for every $x \in X$*

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \leq \|x\|^2.$$

Remark 3.28.1. *The inner product $\langle x, e_k \rangle$ are called the **Fourier coefficients** of x with respect to the orthonormal sequence $\{e_k\}_{k=1}^{\infty}$.*

Remark 3.28.2. *Strict inequality can indeed occur, in fact, it happens iff $x \notin \overline{\operatorname{span}\{e_k\}_{k \in \mathbb{N}}}$.*

Proof. Since

$$\sum_{k=1}^{N} |\langle x, e_k \rangle|^2$$

is a monotone increasing bounded sequence (bounded by $\|x\|^2$), then its limit exists and is less than or equal to $\|x\|^2$. □

Corollary 3.28.1. *Let (e_k) be an orthonormal sequence in an inner product space X. Then for any $x, y \in X$,*

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle \langle y, e_k \rangle| \leq \|x\| \|y\|.$$

Proof. By Bessel's inequality, we have

$$\|x\|^2 \geq \sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2.$$

$$\|y\|^2 \geq \sum_{k=1}^{\infty} |\langle y, e_k \rangle|^2.$$

Then by Cauchy Schwarz Inequality, we have

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle| |\langle y, e_k \rangle| \leq \left(\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \right)^{1/2} \left(\sum_{k=1}^{\infty} |\langle y, e_k \rangle|^2 \right)^{1/2} \leq \|x\| \|y\|.$$

□

Proposition 3.29 (Gram-Schmidt Process). *Let $(x_n) \subset X$ be independent. Then there exists a set of orthonormal sequence (e_n) such that for all $k \in \mathbb{N}$.*

$$\operatorname{span}\{e_1, \cdots, e_k\} = \operatorname{span}\{x_1, \cdots, x_k\}.$$

Proof. We define e_i recursively. Let

$$e_1 = \frac{1}{\|x_1\|} x_1.$$

x_2 can be written as

$$x_2 = \langle x_2, e_1 \rangle e_1 + v_2$$

then let

$$e_2 = \frac{1}{\|v_2\|} v_2.$$

If we have already defined $e_1, \cdots, e_k$, then

$$x_{k+1} = \sum_{i=1}^{k} \langle x_{k+1}, e_i \rangle e_i + v_{k+1}$$

We define

$$e_{k+1} = \frac{1}{\|v_{k+1}\|} v_{k+1}.$$

□

Corollary 3.29.1. *A finite inner product space has a basis of orthonormal vectors.*

Theorem 3.30. *Let (e_k) be an orthonormal sequence in a Hilbert space H. Then*

1. *The series*

$$\sum_{k=1}^{\infty} \alpha_k e_k$$

converges if and only if

$$\sum_{k=1}^{\infty} |\alpha_k|^2$$

converges.

2. *If the series,*

$$\sum_{k=1}^{\infty} \alpha_k e_k = x,$$

then the coefficients α_k are the Fourier coefficients $\langle x, e_k \rangle$.

3. *For any $x \in H$, the series*

$$\sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$$

converges.

Proof.

1. let

$$s_n = \sum_{i=1}^{n} \alpha_i e_i \quad \text{and} \quad \sigma_n = \sum_{i=1}^{n} |\alpha_i|^2.$$

   Then, because of the orthonormality, for any m and $n > m$,

$$\begin{aligned} \|s_n - s_m\|^2 &= \|\alpha_{m+1} e_{m+1} + \cdots + \alpha e_n\|^2 \\ &= |\alpha_{m+1}|^2 + \cdots + |\alpha_n|^2 = \sigma_n - \sigma_m. \end{aligned}$$

   Hence (s_n) is Cauchy in H if and only if (σ_n) is Cauchy in $\mathbb{R}$. Since H and $\mathbb{R}$ are complete, then either they both converges or diverges.

2. Note that

$$\sum_{i=1}^{n} \alpha_i e_i = s_n = \sum_{i=1}^{n} \langle s_n, e_i \rangle e_i.$$

   So $\alpha_i = \langle s_n, e_i \rangle$ for $n \geq i$. Since $s_n \to x$, then by the continuity of inner product, we have $\langle x, e_i \rangle = \alpha_i$ as desired.

3. This is because by Bessel's Inequality, we have

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \leq \|x\|^2.$$

   Hence the statement follows from (1).

□

Corollary 3.30.1. *Let $\{e_j\}$ be an orthonormal sequence in a Hilbert space H and*

$$x = \sum_{j=1}^{\infty} a_j e_j \quad \textit{and} \quad y = \sum_{j=1}^{\infty} \beta_j e_j.$$

Then

$$\langle x, y \rangle = \sum_{j=1}^{\infty} \alpha_j \bar{\beta}_j.$$

In particular, we have

$$\langle x, y \rangle = \sum_{j=1}^{\infty} \langle x, e_k \rangle \overline{\langle y, e_k \rangle}.$$

Proof. We know $\alpha_j = \langle x, e_j \rangle$ and $\beta_j = \langle y, e_j \rangle$ by Theorem 3.30. Then by Corollary (3.28.1), we conclude that the series

$$\sum_{j=1}^{\infty} \alpha_j \overline{\beta}_j$$

converges absolutely. Now

$$\begin{aligned}\langle x, y\rangle &= \sum_{N\to\infty} \langle \sum_{j=1}^{N} \alpha_j e_j, \sum_{j=1}^{N} \beta_j e_j\rangle \\ &= \sum_{N\to\infty} \sum_{j=1}^{\infty} \alpha_j \bar{\beta}_j \\ &= \sum_{j=1}^{\infty} \alpha_j \bar{\beta}_j.\end{aligned}$$

□

Corollary 3.30.2. *Let $\{e_k\}$ be an orthonormal sequence in a Hilbert space H, then for every $x \in H$, the vector*

$$y = \sum_{k=1}^{\infty} \langle x, e_k\rangle e_k$$

exists in H and $x - y$ is orthogonal to every e_k.

Corollary 3.30.3. *Let $\{e_k\}$ be an orthonormal sequence in a Hilbert space H, and let $M = \operatorname{span}\{e_k\}$. Then for any $x \in H$, we have $x \in \bar{M}$ if and only if*

$$x = \sum_{k=1}^{\infty} \langle x, e_k\rangle e_k.$$

Lemma 3.31 (Fourier Coefficients)**.** *Any x in an inner product space X can have at most countably many nonzero Fourier coefficients $\langle x, e_\kappa\rangle$ with respect to an orthonormal family $(e_\kappa)_{\kappa\in I}$, in X.*

Remark 3.31.1. *By the fact that countable union of countable sets is countable, then for any subsequence of X, we can find a countable orthonormal subfamily of $(e_\kappa)_{\kappa\in I}$ that can represent everything element of the sequence.*

Proof. For each $m = 1, 2, \cdots$, the number of Fourier coefficients such that $|\langle x_i, e_\kappa\rangle| > \frac{1}{m}$ must be finite, otherwise, it would contradict the Bessel's Inequality. Hence this shows that the nonzero Fourier coefficients is at most the countable union of countable sets which is countable. □

Hence with any fixed $x \in H$, we can associate a series to x by

$$\sum_{\kappa\in I} \langle x, e_\kappa\rangle e_\kappa$$

even if I is uncountable. Since we can arrange the e_κ with $\langle x, e_\kappa\rangle \neq 0$ in a countable sequence. The convergence follows from Theorem (3.30). Yet, in order to show that the series indeed make sense, we need to show that the sum does not depend on the order in which those e_κ are arranged in a sequence.

Proposition 3.32. *In a Hilbert space, the series*

$$\sum_{\kappa\in I} \langle x, e_\kappa\rangle e_\kappa$$

is well-defined.

Proof. Let $\{w_m\}$ be a rearrangement of $\{e_n\}$, with bijection σ, that is $w_{\sigma(n)} = e_n$. We set

$$\alpha_n = \langle x, e_n \rangle \quad \text{and} \quad \beta_m = \langle x, w_m \rangle$$

and

$$x_1 = \sum_{n=1}^{\infty} a_n e_n \quad \text{and} \quad x_2 = \sum_{m=1}^{\infty} \beta_m w_m.$$

Then

$$\langle x, e_n \rangle = \alpha_n = \langle x_1, e_n \rangle \quad \text{and} \quad \langle x, w_m \rangle = \beta_m = \langle x_2, w_m \rangle.$$

Since $e_n = w_{\sigma(n)}$, we thus has

$$\begin{aligned} \langle x_1 - x_2, e_n \rangle &= \langle x_1, e_n \rangle - \langle x_2, w_{\sigma(n)} \rangle \\ &= \langle x, e_n \rangle - \langle x, w_{\sigma(n)} \rangle = 0 \end{aligned}$$

and similarly, we have $\langle x_1 - x_2, w_m \rangle = 0$. This implies

$$\begin{aligned} \|x_1 - x_2\|^2 &= \langle x_1 - x_2, \sum \alpha_n e_n - \sum \beta_m w_m \rangle \\ &= \sum \bar{\alpha_n} \langle x_1 - x_2, e_n \rangle - \sum \bar{\beta}_m \langle x_1, x_2, w_m \rangle = 0. \end{aligned}$$

Consequently, $x_1 = x_2$. □

3.4 Total Orthonormal Sets and Sequences

Definition 3.33 (Total Orthonormal Set)**.** *A* ***total set*** *or* ***fundamental set*** *in a normed space X is a subset $M \subset X$ whose span is dense in X. Accordingly, an orthonormal set (resp. sequence) in an inner product space X which is total in X is called* ***total orthonormal set*** *(resp.sequence) in X.*

Remark 3.33.1. *A total orthonormal family in X is sometimes called an orthonormal basis for X. However, it is important to note that it is not a basis in the sense of linear algebra.*

Theorem 3.34. *In every Hilbert space $H \neq \{0\}$, there exists a total orthonormal set.*

Proof. Zorn's Lemma. □

Theorem 3.35. *All total orthonormal sets in a given Hilbert space $H \neq \{0\}$ have the same cardinality, and this is known as the* ***Hilbert dimension*** *of H.*

Proof. WLOG, assume the two orthonormal sets $\{x_\alpha\}_{\alpha \in I}$ and $\{y_\beta\}_{\beta \in J}$ are not finite. Then for any $\alpha \in I$,

$$x_\alpha = \sum_{j=1}^{\infty} \langle x_\alpha, y_j \rangle y_j,$$

where $j \in J_\alpha \subset J$ is a countable subset. Then it is easy to verify that

$$J = \bigcup_{\alpha \in I} J_\alpha.$$

Since each $|J_\alpha| \leq |I|$, then

$$|J| = \left| \bigcup_{\alpha \in I} J_\alpha \right| \leq |I|$$

by cardinal arithmetic. □

Theorem 3.36. *Let M be a subset of an inner product space X. Then:*

1. *if M is total in X, then there does not exist a nonzero $x \in X$ which is orthogonal to every element of M. Equivalently, if $x \perp M$, then $x = 0$.*

2. *If X is complete, the above condition is also sufficient for the totality of M in X.*

Proof. Let H be the completion of X, then X is a dense subset in H. Since M is total, then $\operatorname{span} M$ is dense in X, hence in H. Thus if $x \perp M$, by Lemma (3.24), $x = 0$.

On the other hand, if X is a Hilbert space and M satisfies the condition, so that $M^\perp = \{0\}$, then again by Lemma (3.24), we conclude that M is total in X. □

Similarly, by interpreting the definition of total orthonormal set, we have

Theorem 3.37 (Parseval's identity)**.** *An orthonormal set M ins a Hilbert space H is total in H if and only if for all $x \in H$, the Parseval's Identity holds, that is*

$$\sum_k |\langle x, e_k \rangle|^2 = \|x\|^2.$$

Remark 3.37.1. *We note that the series $\sum_k |\langle x, e_k \rangle|^2$ is always well-defined by Theorem (3.32).*

Proof. Make use of Theorem (3.36) in both directions. □

Theorem 3.38 (Separable Hilbert Spaces)**.** *Let H be a Hilbert space. Then*

1. *If H is separable, every orthonormal set in H is countable.*

2. *If H contains an orthonormal sequence which is total in H, then H is separable.*

Proof. If H is separable, and B be any dense set in H and M any orthonormal set. Then any two distinct element x and y of M have distance $\sqrt{2}$ since

$$\|x - y\|^2 = \langle x - y, x - y \rangle = \langle x, x \rangle + \langle y, y \rangle = 2.$$

Hence spherical neighborhoods N_x of x and N_y of y of radius $\sqrt{2}/3$ are disjoint. Since B is dense in H, there is a $b \in B$ in N_x and a $\tilde{b} \in B$ in N_y with $b \neq \tilde{b}$. Hence if M were uncountable, we could have uncountably many such pairwise disjoint spherical neighborhoods, which implies B would be uncountable.

Now let $\{e_k\}$ be a total orthonormal sequence in H and A the set of all linear combinations

$$\gamma_1^{(n)} e_1 + \cdots + \gamma_n^{(n)} e_n \quad n = 1, 2, \cdots$$

where $\gamma_k^{(n)} = a_k^{(n)} + ib_k^{(n)}$ and $a_k^{(n)}$ and $b_k^{(n)}$ are rational. Clearly, A is countable (countable union of countable sets). We prove that A is dense in H by showing that for every $x \in H$ and $\epsilon > 0$, there is a $v \in A$ such that $\|x - v\| < \epsilon$. And this follows from Theorem (3.37). □

Example: (Non-separable Hilbert Space). Let $e_s(x) := e^{isx}$, $s \in \mathbb{R}$. Then $\langle e_s, e_\ell \rangle = 0$ iff $s \neq t$, where

$$\langle e_s, e_\ell \rangle = \lim_{m \to \infty} \frac{1}{2M} \int_{-M}^{M} e^{isx} e^{-itx} dx.$$

Then the completion of this space is the space of almost periodic functions. This space is not separable since we have an uncountable orthogonal basis.

Theorem 3.39 (Isomorphism and Hilbert Dimension)**.** *Two Hilbert spaces H and $\tilde{H}$, both real or both complex, are isomorphic if and only if they have the same Hilbert dimension.*

Proof. Suppose H is isomorphic with $\tilde{H}$ and $T : H \to \tilde{H}$ is an isomorphism. Then as T is inner product preserving, any orthonormal set in H will be mapped to an orthonormal set in $\tilde{H}$ by T. Since T is bijective and distance preserving, we thus conclude that T maps every total orthonormal set in H onto a total orthonormal set in $\tilde{H}$. Hence H and $\tilde{H}$ have the same Hilbert dimension.

Conversely, suppose that H and $\tilde{H}$ have the same Hilbert dimension. The case $H = \{0\}$ and $\tilde{H} = \{0\}$ is trivial. Let $H, \tilde{H} \neq \{0\}$ and any total orthonormal sets M in H and $\tilde{M}$ in $\tilde{H}$ have the same cardinality, so that we can index them by the same index set $\{\kappa\}$ and write $M = \{e_\kappa\}$ and $\tilde{M} = \{\tilde{e}_\kappa\}$. To show that H and $\tilde{H}$ are isomorphic, we construct an isomorphism of H onto $\tilde{H}$. For every $x \in H$, we have

$$x = \sum_\kappa \langle x, e_\kappa \rangle e_\kappa$$

where the right-hand side is a finite sum or an infinite series. Define

$$\tilde{x} = Tx = \sum_\kappa \langle x, e_k \rangle \tilde{e}_k$$

which converges since $\tilde{H}$ is a Hilbert space. The operator T is linear since the inner product is linear with respect to the first factor. T is isometric since

$$\|\tilde{x}\|^2 = \|Tx\|^2 = \sum_k |\langle x, e_k \rangle|^2 = \|x\|^2.$$

Then by polarizing identity, we see that T preserves the inner product. Hence T is injective. Finally, to show surjectivity, given any

$$\tilde{x} = \sum_\kappa \alpha_\kappa \tilde{e}_\kappa$$

in $\tilde{H}$, we have $\sum |\alpha_k|^2 < \infty$ by the Bessel inequality. Hence

$$\sum_\kappa \alpha_k e_k$$

converges to an $x \in H$, and $\alpha_k = \langle x, e_k \rangle$. We thus have $\tilde{x} = Tx$, so T is surjective. $\square$

3.5 Riesz's Theorem

Theorem 3.40 (Riesz's Theorem (Hilbert Spaces))**.** *Every bounded linear functional f on a Hilbert space H can be represented in terms of the inner product, namely*

$$f(x) = \langle x, z \rangle$$

where z is uniquely determined by f and moreover

$$\|z\| = \|f\|.$$

Proof. Firstly, we see if such z exist, then it is unique. Since we must have $\langle x, z_1 - z_2 \rangle = 0$ for all $x \in H$, thus $z_1 = z_2$.

Now suppose we have this representation, then $f(z) = \langle z, z \rangle = \|z\|^2 \leq \|f\|\|z\|$. This shows $\|z\| \leq \|f\|$. On the other hand, by the Schwarz inequality, we have

$$|f(x)| = |\langle x, z \rangle| \leq \|x\|\|z\|,$$

this implies $\|f\| \leq \|z\|$.

Lastly, we show the existence of z. If $f = 0$, then clearly we can take $z = 0$. So we assume $f \not\equiv 0$. Notice if such z exists, then $\langle x, z \rangle = 0$ for all $x \in \mathcal{N}(f)$. Hence it is natural we consider $\mathcal{N}(f)$ and $\mathcal{N}(f)^\perp$.

Since f is bounded, then $\mathcal{N}(f)$ is closed. Since $f \not\equiv 0$, $\mathcal{N}(f) \neq H$, so $\mathcal{N}(f)^\perp \neq \{0\}$ by the projection theorem. Then $\mathcal{N}(f)^\perp$ contains a $z_0 \neq 0$. We set

$$v = f(x)z_0 - f(z_0)x$$

where $x \in H$ is arbitrary. Apply f, we obtain

$$f(v) = f(x)f(z_0) - f(z_0)f(x) = 0.$$

This shows that $v \in \mathcal{N}(f)$. Since $z_0 \perp \mathcal{N}(f)$, we have

$$\begin{aligned} 0 = \langle v, z_0 \rangle &= \langle f(x)z_0 - f(z_0)x, z_0 \rangle \\ &= f(x)\langle z_0, z_0 \rangle - f(z_0)\langle x, z_0 \rangle. \end{aligned}$$

Since $\langle z_0, z_0 \rangle \neq 0$, then

$$f(x) = \frac{f(z_0)}{\langle z_0, z_0 \rangle}\langle x, z_0 \rangle.$$

So if we let

$$z := \frac{\overline{f(z_0)}}{\langle z_0, z_0 \rangle} z_0,$$

then for any $x \in H$, we have

$$f(x) = \langle x, z \rangle.$$

□

Remark 3.40.1. *Conversely, it is easy to show that $f(x) = \langle x, z \rangle$ defines a bounded linear functional f on X, of norm $\|z\|$. Hence we have a canonical isometry bijection from H to H' (this map is not linear, but conjugate linear). Similarly, we have a canonical isometry bijection from H' to H''.*

Example:

1. (Space $\mathbb{R}^n$), any linear functional f on $\mathbb{R}^n$ can be represented as taking dot product with an element of $\mathbb{R}^n$, that is for any $f \in X' = X^*$, there exists $z \in \mathbb{R}^n$ such that

$$f(x) = x \cdot z = \sum_{i=1}^{n} x_i z_i.$$

2. (Space ℓ^2) Every bounded linear functional f on ℓ^2 can be represented in the form

$$f(x) = \sum_{j=1}^{\infty} \xi_j \bar{\zeta}_j,$$

where $z = (\zeta_j) \in \ell^2$ is uniquely determined by f.

Corollary 3.40.1. *The dual space H' of a Hilbert space H is a Hilbert space with inner product $\langle \cdot, \cdot \rangle_1$ defined by*

$$\langle f_z, f_v \rangle_1 = \overline{\langle z, v \rangle} = \langle v, z \rangle,$$

where $f_z(x) = \langle x, z \rangle$.

Proof. We know any by Riesz representation theorem, any element of H' is given by f_z. Now clearly $\alpha f_z = f_{\bar{\alpha} z}$, and $f_z + f_v = f_{z+v}$. Hence $\langle v, z \rangle$ gives an inner product on H', moreover, it agrees with the norm since we have $\|f_z\| = \|z\|$. H' is complete since F' is complete. □

Proposition 3.41. *Let X be an inner product space such that the mapping $X \to X'$ given by $z \mapsto f_z$ is surjective, where $f_z : X \to \mathbb{F}$ is given by*

$$f_z(x) = \langle x, z \rangle.$$

Then X is a Hilbert space.

Proof. We have an isometric bijection between X and X'. Since $\mathbb{F}$ is complete, X' is complete. Since isometric bijections are homeomorphisms, then X is complete. □

Theorem 3.42 (Hilbert Spaces are Reflexive)**.** *Every Hilbert space is reflexive. That is to say the canonical natural embedding from H to $H^{''}$ given by*

$$\phi : H \to H^{''}, \quad x \mapsto \hat{x}$$

is an isometric linear isomorphism, where $\hat{x}(f) = f(x)$ for any $f \in H'$.

Proof. We know ϕ is injective, we show that ϕ is bijective and isometric. We know by Riesz Representation theorem, the elements of H' can be written as f_z, where

$$f_z(x) = \langle x, z \rangle \quad \forall x \in H.$$

Then by Riesz representation theorem again, any elements of H'' can be written as T_{f_z}, where

$$T_{f_z}(f_v) = \langle f_v, f_z \rangle = \overline{\langle v, z \rangle} = \langle z, v \rangle.$$

Then we note that $\phi(z) = T_{f_z}$, since

$$\hat{\phi}(f_v) = f_v(z) = \langle z, v \rangle.$$

Moreover, we have

$$\|T_{f_z}\| = \|f_z\| = \|z\|$$

Hence ϕ preserves norms. So ϕ is an isometric linear isomorphism. □

Definition 3.43 (Sesquilinear Form)**.** *Let X and Y be vector spaces over the same field $\mathbb{F}$. Then a **sesquilinear form** h on $X \times Y$ is a mapping*

$$h : X \times Y \to K$$

such that for all $x, x_1, x_2 \in X$ and $y, y_1, y_2 \in Y$ and all scalars α, β, we have

$$\begin{aligned} h(x_1 + x_2, y) &= h(x_1, y) + h(x_2, y) \\ h(x, y_1 + y_2) &= h(x, y_1) + h(x, y_2) \\ h(\alpha x, y) &= \alpha h(x, y) \\ h(x, \beta y) &= \overline{\beta} h(x, y). \end{aligned}$$

Hence h is linear in the first argument and conjugate linear in the second one.

If X and Y are normed spaces and it here is a real number c such that for all x, y,

$$|h(x, y)| \leq c\|x\|\|y\|,$$

*then h is said to be **bounded**. In this case, we define the **norm** of h to be*

$$\|h\| = \sup_{\substack{x \in X \setminus \{0\} \\ y \in Y \setminus \{0\}}} \frac{|h(x, y)|}{\|x\|\|y\|} = \sup_{\substack{\|x\|=1 \\ \|y\|=1}} |h(x, y)|$$

is called the norm of h.

Remark 3.43.1. *It is clear that $|h(x, y)| \leq \|h\|\|x\|\|y\|$ if h is bounded.*

Example: Any inner product $\langle \cdot, \cdot \rangle$ on an inner product space X is a bounded sesquilinear form with norm 1.

Theorem 3.44 (Riesz Representation). *Let H_1, H_2 be Hilbert spaces and*

$$h : H_1 \times H_2 \to \mathbb{F}$$

is a bounded sesquilinear form. Then h has a representation

$$h(x, y) = \langle Sx, y\rangle$$

where $S : H_1 \to H_2$ is a bounded linear operator. S is uniquely determined by h and has norm

$$\|S\| = \|h\|.$$

Proof. We consider $\overline{h(x, y)}$ which is linear in y. Let x be fixed, then the Riesz Representation for the Hilbert space states that there exists unique $z_x \in H_2$, such that for any $y \in H_2$, we have

$$\overline{h(x, y)} = \langle y, z_x\rangle.$$

Hence

$$h(x, y) = \langle z_x, y\rangle.$$

Then it is natural we define the operator $S : H_1 \to H_2$ by $Sx = z_x$. We check that S is linear.

$$\begin{aligned}\langle S(\alpha x_1 + \beta x_2), y\rangle &= h(\alpha x_1 + \beta x_2, y) \\ &= \alpha h(x_1, y) + \beta h(x_2, y) \\ &= \alpha\langle Sx_1, y\rangle + \beta\langle Sx_2, y\rangle \\ &= \langle \alpha Sx_1 + \beta Sx_2, y\rangle\end{aligned}$$

for all $y \in H_2$, hence $S(\alpha x_1 + \beta x_2) = \alpha Sx_1 + \beta Sx_2$.

We show that S is bounded.

$$\|h\| = \sup_{\substack{x\in X\setminus\{0\}\\ y\in Y\setminus\{0\}}} \frac{|\langle Sx, y\rangle|}{\|x\|\|y\|} \geq \sup_{\substack{x\neq 0\\ Sx\neq 0}} \frac{|\langle Sx, Sx\rangle|}{\|x\|\|Sx\|} = \sup_{x\neq 0} \frac{\|Sx\|}{\|x\|} = \|S\|.$$

This proves the boundedness of S, moreover, $\|h\| \geq \|S\|$. On the other hand, we show $\|h\| \leq \|S\|$. By Schwarz Inequality:

$$\|h\| = \sup_{\substack{x\in X\setminus\{0\}\\ y\in Y\setminus\{0\}}} \frac{|\langle Sx, y\rangle|}{\|x\|\|y\|} \leq \sup_{x\neq 0} \frac{\|Sx\|\|y\|}{\|x\|\|y\|} = \|S\|.$$

Lastly, we show S is unique. Suppose $T : H_1 \to H_2$ is such that for all $x \in H_1$ and $y \in H_2$, we have

$$h(x, y) = \langle Sx, y\rangle = \langle Tx, y\rangle,$$

then we see that for all $x \in H$, $Sx = Tx$, hence $S = T$. □

Definition 3.45 (Hermitian form). *Let X be a vector space over a field K. A* ***Hermitian sesquilinear form*** *h*

on $X \times X$ is a mapping $h : X \times X \to K$ such that for all $x, y, z \in X$ and $\alpha \in K$,

$$h(x + y, z) = h(x, z) + h(y, z)$$
$$h(\alpha x, y) = \alpha h(x, y)$$
$$h(x, y) = \overline{h(y, x)}.$$

In particular, we notice that if h is positive definite, then h gives an inner product on X.

3.6 Hilbert Adjoint Operators

Definition 3.46 (Hilbert-Adjoint operator)**.** *Let $T : H_1 \to H_2$ be a bounded linear operator, where H_1 and H_2 are Hilbert spaces. Then the Hilbert-adjoint operator T^* of T is the operator*

$$T^* : H_2 \to H_1$$

such that for all $x \in H_1$ and $y \in H_2$,

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

Theorem 3.47 (Existence of Adjoint Operator)**.** *The Hilbert-adjoint operator T^* of T exists, is unique and is a bounded linear operator with norm*

$$\|T^*\| = \|T\|.$$

Proof. The formula

$$h(y, x) = \langle y, Tx\rangle$$

defines a sesquilinear form on $H_2 \times H_1$ because the inner product is sesquilinear and T is linear. h is bounded, since by Schwarz inequality,

$$|h(y, x)| = |\langle y, Tx\rangle| \leq \|y\|\|Tx\| \leq \|T\|\|x\|\|y\|.$$

This also implies $\|h\| \leq \|T\|$. Moreover, we have $\|h\| \geq \|T\|$ from

$$\|h\| = \sup_{\substack{x \in X\setminus\{0\} \\ y \in Y\setminus\{0\}}} \frac{|\langle y, Tx\rangle|}{\|y\|\|x\|} \geq \sup_{\substack{x \neq 0 \\ Tx \neq 0}} \frac{|\langle Tx, Tx\rangle|}{\|Tx\|\|x\|} = \|T\|.$$

This shows $\|h\| = \|T\|$.

Then by the Riesz representation theorem for h, we have a unique bounded linear operator S such that

$$\langle y, Tx\rangle = h(y, x) = \langle Sy, x\rangle.$$

So if we set $T^* = S$, then $\|T^*\| = \|S\| = \|h\| = \|T\|$. Also $\langle y, Tx\rangle = \langle T^* y, x\rangle$, hence

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

□

Definition 3.48 (Positive Operator)**.** *Let $T : H \to H$ be a self-adjoint operator. Then T is* ***positive*** *if $\langle Ax, x\rangle \in \mathbb{R}$ and*

$$\langle Ax, x\rangle \geq 0$$

with equality if and only if $x = 0$.

Lemma 3.49 (Zero operator)**.** *Let X and Y be inner product spaces and $Q : X \to Y$ be a bounded linear operator. Then:*

1. *$Q = 0$ if and only if $\langle Qx, y\rangle = 0$ for all $x \in X$ and $y \in Y$.*
2. *If $Q : X \to X$, where X is complex, and $\langle Qx, x\rangle = 0$ for all $x \in X$, then $Q = 0$.*

Proof.

1. If $Q = 0$, then clearly $\langle Qx, y\rangle = 0$. Conversely, $\langle Qx, y\rangle = 0$ for all x, y implies $Qx = 0$ for all x, hence $Q = 0$.
2. By assumption, $\langle Qv, v\rangle = 0$ for every $v = ax + y \in X$. That is

$$\begin{aligned} 0 &= \langle Q(\alpha x + y), \alpha x + y\rangle \\ &= |\alpha|^2\langle Qx, x\rangle + \langle Qy, y\rangle + \alpha\langle Qx, y\rangle + \tilde{\alpha}\langle Qy, x\rangle. \end{aligned}$$

The first two terms on the right are zero by assumption. Let $\alpha = 1$ gives

$$\langle Qx, y\rangle + \langle Qy, x\rangle = 0.$$

Let $\alpha = i$, we have

$$\langle Qx, y\rangle - \langle Qy, x\rangle = 0.$$

Hence $\langle Qx, y\rangle = 0$, so $Q = 0$.

□

Theorem 3.50 (Properties of Adjoint Operators)**.** *Let H_1, H_2 be Hilbert spaces, $S : H_1 \to H_2$ and $T : H_1 \to H_2$ be bounded linear operators and α be a scalar. Then we have*

1. *$\langle T^*y, x\rangle = \langle y, Tx\rangle$ for all $x \in H_1$, $y \in H_2$.*
2. *$(S + T)^* = S^* + T^*$.*
3. *$(\alpha T)^* = \bar{\alpha}T^*$.*
4. *$(T^*)^* = T$, when defined.*
5. *$\|T^*T\| = \|TT^*\| = \|T\|^2$.*
6. *$T^*T = 0$ if and only if $T = 0$.*
7. *Suppose $H_2 = H_1$, then $(ST)^* = (T^*S^*)$*
8. *$0^* = 0$, $Id^* = Id$.*

9. *Suppose $H_2 = H_1$ and the inverse of T is also bounded, then*

$$(T^*)^{-1} = (T^{-1})^*.$$

10. *Suppose $(T_n) \to T$ is a sequence of bounded linear operators on a Hilbert space, then $T_n^* \to T^*$.*

Proof. Properties (1), (2), (3), (4) are easy to verify directly from definition. Now for (5), we see that $T^*T : H_1 \to H_1$ and $TT^* : H_2 \to H_2$. Then by the Schwarz inequality, we have

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle \leq \|T^*Tx\|\|x\| \leq \|T^*T\|\|x\|^2.$$

This shows $\|T\|^2 \leq \|T^*T\|$. But since $\|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2$, this shows $\|T^*T\| = \|T\|^2$. Replacing T by T^*, and as $(T^*)^* = T$, we have

$$\|T^*T\| = \|TT^*\| = \|T\|^2.$$

Then it directly follows that if $T^*T = 0$ iff $T = 0$. Property (7) and (8) are also easy to verify from the definition.

Let us verify (9). Since T is surjective, then T^* is injective, since

$$\langle x, T^*(y_1 - y_2)\rangle = \langle Tx, y_1 - y_2\rangle.$$

$T(H_1) = H_1$, so we must have $y_1 = y_2$ if $T^*(y_1 - y_2) = 0$. Next, we must have T^* is surjective by Proposition (3.51), so T^* admits an inverse. Lastly, we just need to show $(T^{-1})^*(T^*) = Id$ to show $(T^*)^{-1} = (T^{-1})^*$.

$$\langle x, (T^{-1})^*T^*y\rangle = \langle T^{-1}x, T^*y\rangle = \langle x, y\rangle.$$

This holds for all x, y, hence $(T^{-1})^*T^*y = y$.

Lastly, to show (10), suppose $\|T_n - T\| \to 0$, then $(T_n - T)^* = T_n^* - T^*$, hence

$$\|T_n - T\| = \|(T_n - T)^*\| = \|T_n^* - T^*\| \to 0.$$

□

Proposition 3.51. *Let H_1, H_2 be Hilbert spaces and $T : H_1 \to H_2$ be a bounded linear operator. In $M_1 \subset H_1$ and $M_2 \subset H_2$ are closed subspaces, then $T(M_1) \subset M_2$ if and only if $M_1^\perp \supset T^*(M_2^\perp)$.*

In particular, if we take $M_1 = \mathcal{N}(T)$, then

- $T^*(H_2) \subset \mathcal{N}(T)^\perp$.
- $[T(H_1)]^\perp \subset \mathcal{N}(T^*)$.
- $\mathcal{N}(T) = [T^*(H_2)]^\perp$.

Proof. Suppose $T(M_1) \subset M_2$, then if $x \in T^*(M_2^\perp)$, $x = T^*z$ for some $z \in M_2^\perp$, that is $\langle z, y\rangle = 0$ for any $y \in M_2$. So for any $\tilde{y} \in M_1$, we have

$$\langle x, \tilde{y}\rangle = \langle T^*z, \tilde{y}\rangle = \langle z, T\tilde{y}\rangle = 0.$$

Since $T\tilde{y} \in M_2$ by assumption. In particular, for this direction, we do not require M_1 and M_2 to be closed subspaces. Conversely, suppose $M_1^\perp \supset T^*(M_2^\perp)$. Let $y \in M_2 \perp$, then for any $x \in M_1$, we have

$$0 = \langle T^*(y), x\rangle = \langle y, Tx\rangle.$$

This show that $Tx \perp M_2^\perp$ for any $x \in M_1$. Hence $T(M_1) \subset M_2^{\perp\perp} = M_2$ by the fact that M_2 is a closed subspace.

Now if we take $M_1 = \mathcal{N}(T)$ which is a closed subspace of H_1. Then let $M_2 = \{0\}$, then $T(M_1) \subset M_2$, $M_2^\perp = H_2$, hence $T^*(H_1) \subset \mathcal{N}(T)^\perp$. Next, we show $[T(H_1)]^\perp \subset \mathcal{N}(T^*)$. Suppose $x \perp T(H_1)$, then for any $y \in H_1$,

$$0 = \langle Ty, x\rangle = \langle y, T^*x\rangle.$$

Hence this shows that T^*x must be zero, so $x \in \mathcal{N}(T^*)$.

Now since $T^*(H_2) \subset \mathcal{N}(T)^\perp$, and $\mathcal{N}(T)^{\perp\perp} = \mathcal{N}(T)$, we have $\mathcal{N}(T) \subset [T^*(H_2)]^\perp$. Similarly to above, we can show that $[T^*(H_2)]^\perp \subset \mathcal{N}(T)$, hence we have $\mathcal{N}(T) = [T^*(H_2)]^\perp$. □

Definition 3.52 (Self-Adjoint, Unitary and Normal Operators). *A bounded liner operator $T : H \to H$ on a Hilbert spaceH is said to be*

- ***self-adjoint*** *or* ***Hermitian*** *if $T^* = T$,*
- ***unitary*** *if T is bijective and $T^* = T^{-1}$.*
- ***normal*** *if $TT^* = T^*T$.*

Lemma 3.53. *If T is self-adjoint, then*

$$\langle Tx, y\rangle = \langle x, Ty\rangle.$$

If T is self-adjoint or unitary, T is normal. A normal operator need not to be self-adjoint nor unitary.

Proposition 3.54 (Matrix Representation of Adjoint). *If a basis for $\mathbb{C}^n$ is given and a linear operator on $\mathbb{C}^n$ is represented by a certain matrix, then its Hilbert-adjoint operator is represented by the complex conjugate transpose of that matrix.*

Remark 3.54.1. *Consequently, we say the representing matrix is Hermitian (unitary, normal, resp.) if T is Hermitian (unitary, normal, resp.) For the special case of real matrices, we say the matrix is real symmetric, if T is self-adjoint, and orthogonal if T is unitary. In summary we have the following:*

- *Hermitian : $\bar{A}^T = A$;*
- *skew-Hermitian: $\bar{A}^T = -A$;*
- *unitary: $\bar{A}^T = A^{-1}$;*
- *normal: $A\bar{A}^T = \bar{A}^T A$;*
- *real symmetric: $A^T = A$;*

- *real skew-symmetric:* $A^T = -A$;
- *orthogonal:* $A^T = A^{-1}$.

Theorem 3.55 (Self-adjointness). *Let $T : H \to H$ be a bounded linear operator on a Hilbert space H. Then:*

1. *If T is self-adjoint, $\langle Tx, x\rangle$ is real for all $x \in H$.*
2. *If H is complex and $\langle Tx, x\rangle$ is real for all $x \in H$, then the operator T is self-adjoint.*

Proof. If T is self-adjoint, then for all x,

$$\overline{\langle Tx, x\rangle} = \langle x, Tx\rangle = \langle Tx, x\rangle.$$

Hence $\langle Tx, x\rangle$ is real.

Conversely, if $\langle Tx, x\rangle$ is real for all x, then

$$\langle Tx, x\rangle = \overline{\langle Tx, x\rangle} = \overline{\langle x, T^*x\rangle} = \langle T^*x, x\rangle.$$

Hence $\langle (T - T^*)x, x\rangle = 0$ for all x, this shows $T - T^* = 0$ by Lemma (3.49). $\square$

Lemma 3.56 (Product of Self-Adjoint Operators). *The product of two bounded self-adjoint linear operators S and T on a Hilbert space H is self-adjoint if and only if the operators commute.*

Proof. Since $(ST)^* = T^*S^* = TS$, then $(ST) = (ST)^*$ if and only if $ST = TS$. $\square$

Corollary 3.56.1. *If $T : H \to H$ is a bounded self-adjoint linear operator, then so is T^n where n is a positive integer.*

Proposition 3.57 (Decomposition of Operators). *For any bounded linear operator T on H, the operators*

$$T_1 = \frac{1}{2}(T + T^*) \quad \textit{and} \quad T_2 = \frac{1}{2i}(T - T^*)$$

are self-adjoint, and

$$T = T_1 + iT_2 \quad \textit{and} \quad T^* = T_1 - iT_2.$$

Moreover, if S_1, S_2 are another two self-adjoint operators such that $T_1 + iT_2 = S_1 + iS_2$, then $S_1 = T_1$, $S_2 = T_2$.

Proof. Clearly T_1, T_2 are self adjoint, and $T = T_1 + iT_2$, $T^* = T_1 - iT_2$. We show uniqueness. Suppose $S_1 + iS_2 = T_1 + iT_2 = T$, then $S_1 - iS_2 = T_1 - iT_2 = T^*$. So $2S_1 = T$ and $2iS_2 = T - T^*$, which shows uniqueness. $\square$

Corollary 3.57.1. *T is normal if and only if T_1 and T_2 commutes.*

Proof. If T_1 and T_2 commutes, then clearly $TT^* = T^*T$. On the other hand, if $TT^* = T^*T$, then

$$(T_1 + iT_2)(T_1 - iT_2) = (T_1 - iT_2)(T_1 + iT_2),$$

that is $T_1T_2 = T_2T_1$. $\square$

Theorem 3.58 (Sequences of self-adjoint operators)**.** *Let (T_n) be a sequence of bounded self-adjoint linear operators $T_n : H \to H$ on a Hilbert space H. Suppose that (T_n) converges to T, then the limit operator T is a bounded self-adjoint linear operator on H.*

Proof. We show that $T^* = T$, i.e., $\|T - T^*\| = 0$. Since

$$\|T_n^* - T^*\| = \|(T_n - T)^*\| = \|T_n - T\|.$$

Then

$$\begin{aligned} \|T - T^*\| &\leq \|T - T_n\| + \|T_n - T_n^*\| + \|T_n^* - T^*\| \\ &= \|T - T_n\| + 0 + \|T_n - T\| \\ &= 2\|T_n - T\| \to 0. \end{aligned}$$

□

Theorem 3.59 (Unitary Operators)**.** *Let the operators $U : H \to H$ and $V : H \to H$ be unitary, where H is a Hilbert space. Then*

1. *U is an isometry, so $\|Ux\| = \|x\|$ for all $x \in H$;*
2. *$\|U\| = 1$, if $H \neq \{0\}$;*
3. *$U^{-1} = (U^*)$ is unitary;*
4. *UV is unitary;*
5. *A bounded linear operator T on a Hilbert space H is unitary if and only if T is isometric.*

Proof.

1. By assumption U is invertible. Suffices to prove $\|Ux\| = \|x\|$ since we have the polarizing identity. Now

$$\|Ux\|^2 = \langle Ux, Ux\rangle = \langle x, U^*Ux\rangle = \langle x, Ix\rangle = \|x\|^2.$$

2. Follows from (1).
3. Since U is bijective, so is U^{-1}. Then

$$(U^{-1})^* = (U^*)^* = U = (U^{-1})^{-1}.$$

4. UV is clearly bijective, and

$$(UV)^* = V^*U^* = V^{-1}U^{-1} = (UV)^{-1}.$$

5. One direction is clear. Now suppose T is an isometry, then it is invertible by assumption. It is also clear that it is bounded. Now since T preserves norms, then it preserves inner products, hence

$$\langle TT^*x, y\rangle = \langle Tx, Ty\rangle = \langle x, y\rangle.$$

So we conclude that $TT^* = I$.

□

Definition 3.60 (Unitary Equivalence)**.** *Let S and T be linear operators on a Hilbert space H. The operator S is said to be **unitarily equivalent** to T if there is a unitary operator U on H such that*

$$S = UTU^{-1} = UTU^{*}.$$

Remark 3.60.1. *Then it is easy to see that S is self-adjoint if and only if T is self adjoint.*

Proposition 3.61. *A bounded linear operator $T : H \to H$ on a complex Hilbert space H is normal if and only if $\|T^{*}x\| = \|Tx\|$ for all $x \in H$. Hence if T is normal, then $\|T^2\| = \|T\|^2$.*

Proof. Suppose T is normal, then

$$\|T^{*}x\| = \langle T^{*}x, T^{*}x\rangle = \langle TT^{*}x, x\rangle = \langle T^{*}Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\|.$$

Conversely, we can show if $\|T^{*}x\| = \|Tx\|$ for all $x \in H$, then

$$\langle (TT^{*} - T^{*}T)x, x\rangle = 0.$$

Then by Lemma (3.49), T is normal. Lastly, if T is normal, then clearly $\|T^2\| \leq \|T\|^2$. On the other hand,

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle TT^{*}x, x\rangle \leq \|T^{*}Tx\|\|x\| = \|T(Tx)\|\|x\| = \|T^2x\|\|x\|$$

shows that $\|T^2\| \geq \|T\|^2$. □

Proposition 3.62. *If $T_n : H \to H$ are normal linear operators on a Hilbert space, and $T_n \to T$, then T is a normal linear operator.*

Proof. We show $\|T^{*}T - TT^{*}\| = 0$,

$$\begin{aligned}
&\|T^{*}T - TT^{*}\| \\
&\leq \|T^{*}T - T_n^{*}T\| + \|T_n^{*}T - T_n^{*}T_n\| + \|T_n^{*}T_n - T_nT_n^{*}\| + \|T_nT_n^{*} - T_nT^{*}\| + \|T_nT^{*} - TT^{*}\| \\
&\leq \|T - T_n\|\|T\| + \|T_n\|\|T - T_n\| + 0 + \|T_n\|\|T - T_n\| + \|T_n - T\|\|T\|
\end{aligned}$$

Then as $n \to \infty$, the last line has limit 0. □

4 Fundamental Theorems for Normed and Banach Spaces

4.1 Hahn-Banach Theorem

Definition 4.1 (Sublinear and Positive-Homogeneous Functionals)**.** *Let X be a vector space, then a functional $p : X \to \mathbb{R}$ is **sublinear** if*

$$p(x + y) \leq p(x) + p(y)$$

*for all $x, y \in X$. p is **positive-homogeneous** if*

$$p(\alpha x) = \alpha p(x)$$

for all $\alpha \geq 0$ in $\mathbb{R}$ and $x \in X$.

Example: The norm function on any normed space is a sublinear and positive homogeneous functional. The norm of a linear functional is also sublinear and positive homogeneous. Moreover, we note that if p_1, p_2 are sublinear functionals on a vector space X and $c_1, c_2 > 0$, then $p = c_1 p_1 + c_2 p_2$ is sublinear on X.

Theorem 4.2 (Hahn-Bach Theorem)**.** *Let X be a real vector space and p a sublinear functional on X. Furthermore, let f be a linear functional which is defined on a subspace Z of X and satisfies*

$$f(x) \leq p(x) \quad \forall x \in Z.$$

Then f has a linear extension $\tilde{f}$ from Z to X satisfying

$$\tilde{f}(x) \leq p(x) \quad \forall x \in X.$$

Proof. We shall prove the following:

1. The set E of all linear extensions g of f satisfying $g(x) \leq p(x)$ on their domain $\mathscr{D}(g)$ can be partially ordered and Zorn's lemma yields a maximal element $\tilde{f}$ of E.

2. $\tilde{f}$ is defined on the entire space X.

For 1, let M denote the set of all linear functionals g such that $g|_Z = f$ and $g(x) \leq p(x)$. We define a partial order on M by letting $g \leq g'$ if $\mathscr{D}(g) \subset \mathscr{D}(g')$ and $g = g'$ on $\mathscr{D}(g)$. Then for any chain $C \subset M$, we see that the unary union of C is an upper bound of C that is in C. Hence by Zorn's lemma, M has a maximal element $\tilde{f}$. By the definition of E, this is a linear extension of f which satisfies $\tilde{f}(x) \leq p(x)$ on $\mathscr{D}(\tilde{f})$.

We show that $\mathscr{D}(\tilde{f})$ is all of X. Suppose not, then let $y \in X \setminus \mathscr{D}(\tilde{f})$ and consider the subspace Y of X spanned by $\mathscr{D}(\tilde{f})$ and y_1. Any $x \in Y_1$ can be written as

$$x = \alpha y_1 + z$$

where $z \in \mathscr{D}(\tilde{f})$. This representation is unique, since we have a direct sum. We define g on Y by

$$g_1(\alpha y_1 + z) = \alpha c + \tilde{f}(z),$$

then g is linear and an proper extension of $\tilde{f}$. We choose appropriate c such that $g \leq p$ on its domain.

We consider any $z_1, z_2 \in \mathscr{D}(\tilde{f})$, then we have

$$\begin{aligned}\tilde{f}(z_1) - \tilde{f}(z_2) &= \tilde{f}(z_1 - z_2)\\ &\le p(z_1 - z_2)\\ &\le p(y + z_1) + p(-z_2 - y).\end{aligned}$$

Then we have

$$-p(-z_2 - y) - \tilde{f}(z_2) \le p(y + z_1) - \tilde{f}(z_1).$$

Since z_1, z_2 are arbitrary, by taking sup on the left and inf on the right, we get that there exits $c \in \mathbb{R}$ such that

$$-p(-z_2 - y) - \tilde{f}(z_2) \le c \le p(y + z_1) - \tilde{f}(z_1)$$

for all $z_1, z_2 \in \mathscr{D}(\tilde{f})$. This is the c we choose.

Now suppose α is negative, then we get

$$-p\left(-y - \frac{1}{\alpha z}\right) - \tilde{f}\left(\frac{1}{\alpha}z\right) \le c,$$

so by multiplying $-a > 0$ on both sides, we have

$$\alpha p\left(-y - \frac{1}{\alpha z}\right) + \tilde{f}(z) \le -\alpha c$$

for all $z \in \mathscr{D}(\tilde{f})$. This shows that

$$g_1(x) = \alpha c + \tilde{f}(z) \le -\alpha p\left(-y - \frac{1}{\alpha}z\right) = p(\alpha y + z) = p(x).$$

For $\alpha = 0$, we have nothing to prove. Now if $\alpha > 0$, then we have

$$c \le p\left(\frac{1}{\alpha}z + y\right) - \tilde{f}\left(\frac{1}{\alpha}z\right)$$

for all $z \in \mathscr{D}(\tilde{f})$. Then similarly, we can also show that

$$g_1(\alpha y + z) = \alpha c + \tilde{f}(z) \le p(\alpha y + z).$$

□

Remark 4.2.1. *In particular, the extension we obtained also satisfies*

$$-p(-x) \le \tilde{f}(x).$$

Theorem 4.3 (Generalized Hahn-Banach Theorem)**.** *Let X be a real or complex vector space and p a real-valued functional on X which is subadditivity, that is for all $x, y \in X$,*

$$p(x + y) \le p(x) + p(y)$$

and for every scalar α *satisfies*

$$p(\alpha x) = |\alpha| p(x).$$

Furthermore, let f *be a linear functional which is defined on a subspace* Z *of* X *and satisfies*

$$|f(x)| \leq p(x) \quad \forall x \in Z.$$

Then f *has a linear extension* $\tilde{f}$ *from* Z *to* X *satisfying*

$$|\tilde{f}(x)| \leq p(x) \quad \forall x \in X.$$

Remark 4.3.1. *We note that if* p *satsifies the hypothesis in the theorem, then* $p(0) = 0$ *and* $p(x) \geq 0$*, hence* p *is a seminorm.*

Proof. Suppose X is real, then we know there exists $\tilde{f}(x) \leq p(x)$. Then

$$-\tilde{f}(x) = \tilde{f}(-x) \leq p(-x) = |-1| p(x) = p(x).$$

This shows $\tilde{f}(x) \geq -p(x)$, hence $|\tilde{f}(x)| \leq p(x)$ as desired.

Now let us consider the case where X is a complex vector space. f is complex-valued, we can write

$$f(x) = f_1(x) + i f_2(x)$$

with f_1, f_2 real valued. Since f is linear on Z and f_1 and f_2 are real valued, then

$$f_1(x) \leq |f(x)| \leq p(x) \quad \forall x \in Z_r,$$

where we temporarily treat Z as a real vector space Z_r. By the Hahn-Banach Theorem on real vector spaces, there exists a linear extension $\tilde{f}_1$ of f_1 from Z_r to X_r, such that

$$\tilde{f}_1(x) \leq p(x) \quad \forall x \in X_r.$$

Now for every $x \in Z$,

$$i[f_1(x) + i f_2(x)] = i f(x) = f(ix) = f_1(ix) + i f_2(ix).$$

The real parts on both sides must be equal, so

$$f_2(x) = -f_1(ix) \quad \forall x \in Z.$$

Hence if for all $x \in X$, we set

$$\tilde{f}(x) = \tilde{f}_1(x) - i \tilde{f}_1(ix).$$

We see that $\tilde{f}(x) = f(x)$ on Z. We verify $\tilde{f}$ is a linear functional on X:

$$\begin{aligned} \tilde{f}((a+ib)x) &= \tilde{f}_1(ax + ibx) - i\tilde{f} - 1(iax - bx) \\ &= a\tilde{f}_1(x) + b\tilde{f}_1(ix) - i[a\tilde{f}_1(ix) - b\tilde{f}_1(x)] \end{aligned}$$

$$
\begin{aligned}
&= (a+ib)[\tilde{f}_1(x) - i\tilde{f}_1(ix)] \\
&= (a+ib)\tilde{f}(x).
\end{aligned}
$$

We also show that $|\tilde{f}(x)| \leq p(x)$. If $\tilde{f}(x) = 0$, then since we see $p(x) \geq 0$, the inequality holds. Now let x be such that $\tilde{f}(x) \neq 0$, then we can write, using polar coordinates

$$\tilde{f}(x) = |\tilde{f}(x)|e^{i\theta}, \quad \text{thus} \quad |\tilde{f}(x)| = \tilde{f}(x)e^{-i\theta} = \tilde{f}(e^{-i\theta}x).$$

Since $|\tilde{f}(x)|$ is real, the last expression is real and thus equal to its real part. Hence we have

$$|\tilde{f}(x)| = \tilde{f}(e^{-i\theta}x) = \tilde{f}_1(e^{-i\theta}x) \leq p(e^{-i\theta}x) = |e^{-i\theta}|p(x) = p(x).$$

□

Theorem 4.4 (Hahn-Banach Theorem (Normed Spaces))**.** *Let f be a bounded linear functional on a subspace Z of a normed space X. Then there exists a bounded linear functional $\tilde{f}$ on X which is an extension of f to X and has the same norm*

$$\|\tilde{f}\|_X = \|f\|_Z.$$

Proof. If $Z = \{0\}$, then take $\tilde{f} = 0$. Otherwise, by the generalized Hahn-Banach theorem, since $p(x) = \|f\|_z\|x\|$ is a sublinear functional with $p(\alpha x) = |\alpha|p(x)$. Then there exists an extension $\tilde{f}$ on X such that

$$|\tilde{f}(x)| \leq p(x) = \|f\|_Z\|x\| \quad \forall x \in X.$$

This shows $\|\tilde{f}\|_X = \|f\|_Z$. □

Remark 4.4.1. *For the special case of the Hilbert space, suppose suppose f is a bounded linear functional on a closed subspace Z of X, then there exists $z \in Z$*

$$f(x) = \langle x, z\rangle \quad \forall x \in Z.$$

Hence we can define $\tilde{f}(x) = \langle x, z\rangle$ for all $x \in X$.

Theorem 4.5. *Let X be a normed space and $x_0 \neq 0$ be any element of X. Then there exists a bounded linear function $\tilde{f}$ on X such that*

$$\|\tilde{f}\| = 1, \quad \text{and} \quad \tilde{f}(x_0) = \|x_0\|.$$

Proof. Let $Z \subset X$ be the span of $\{x_0\}$. Define f on Z by $f(x) = f(\alpha x_0) = \alpha\|x_0\|$ and extend it to a linear functional on X. We see $\|\tilde{f}\| = \|f\| = 1$ and $\tilde{f}(x_0) = \|x_0\|$. □

Corollary 4.5.1 (Dual Representation of Norm)**.** *For every x in a normed space X, we have*

$$\|x\| = \sup_{\substack{f \in X' \\ f \neq 0}} \frac{|f(x)|}{\|f\|}.$$

Hence if x_0 is such that $f(x_0) = 0$ for all $f \in X'$, then $x_0 = 0$.

Proof. Clearly $\|x\|$ is at most the right hand side. Then by Theorem (4.5), we have the other inclusion. □

Theorem 4.6 (Riesz' Theorem for Functionals on $C(a,b)$)**.** *Every bounded linear functional f on $C[a,b]$ can be represented by a Riemann-Stieltjes integral*

$$f(x) = \int_a^b x(t)dw(t)$$

where w is of bounded variation on $[a,b]$ and has total variation

$$\mathrm{Var}(w) = \|f\|.$$

4.2 Adjoint Operator

Definition 4.7 (Adjoint Operator)**.** *Let $T : X \to Y$ be a bounded linear operator, where X and Y are normed spaces. Then the **adjoint operator** $T^\times : Y' \to X'$ of T is defined by*

$$(T^\times g)(x) = g(Tx) \quad \forall g \in Y'$$

where X' and Y' are the dual spaces of X and Y respectively.

Remark 4.7.1. *If T is represented by a matrix T_E, then the adjoint operator $T^\times$ is represented by the transpose of T_E.*

Theorem 4.8 (Norm of Adjoint Operator)**.** *The adjoint operator $T^\times$ is linear and bounded and*

$$\|T^\times\| = \|T\|.$$

Proof. The operator $T^\times$ is linear since its domain Y' is a vector space and we have

$$\begin{aligned}(T^\times(\alpha g_1 + \beta g_2))(x) &= (\alpha g_1 + \beta g_2)(Tx) \\ &= \alpha g_1(Tx) + \beta g_2(Tx) \\ &= \alpha(T^\times g_1)(x) + \beta(T^\times g_2)(x).\end{aligned}$$

Now let $f = T^\times g$, then

$$\|T^\times g\| = \|f\| \leq \|g\|\|T\|.$$

Taking the supremum over all $g \in Y'$ of norm one, we obtain the inequality

$$\|T^\times\| \leq \|T\|.$$

Conversely, by Theorem (4.5), for any nonzero $x_0 \in X$, there is a $g_0 \in Y'$ such that

$$\|g_0\| = 1 \quad \text{and} \quad g_0(Tx_0) = \|Tx_0\|.$$

Here, $g_0(Tx_0) = (T^\times g_0)(x_0)$ by the definition of the adjoint operator $T^\times$. Writing $f_0 = T^\times g_0$, we thus obtain

$$\|Tx_0\| = g_0(Tx_0) = f_0(x_0)$$

$$\begin{aligned} &\le \|f_0\|\|x_0\| \\ &= \|T^\times g_0\|\|x_0\| \\ &\le \|T^\times\|\|g_0\|\|x_0\|. \end{aligned}$$

Since $\|g_0\| = 1$, we thus have for every $x_0 \in X$,

$$\|Tx_0\| \le \|T^\times\|\|x_0\|.$$

But always

$$\|Tx_0\| \le \|T\|\|x_0\|,$$

and here $c = \|T\|$ is the smallest constant c such that $\|Tx_0\| \le c\|x_0\|$ holds for all $x_0 \in X$. Hence $\|T^\times\|$ cannot be smaller than $\|T\|$. $\square$

Lemma 4.9. *Suppose $S, T \in B(X, Y)$, and $\alpha \in \mathbb{F}$, then*

$$(S+T)^\times = S^\times + T^\times$$
$$(\alpha T)^\times = \alpha T^\times$$

If S^{-1} exists and $S^{-1} \in B(Y, X)$, then $(T^\times)^{-1}$ also exists, $(T^\times)^{-1} \in B(X', Y')$ and

$$(T^\times)^{-1} = (T^{-1})^\times.$$

Let X, Y, Z be normed spaces and $T \in B(X, Y)$ and $S \in B(Y, Z)$, then

$$(ST)^\times = T^\times S^\times.$$

Hence $(T^n)^\times = (T^\times)^n$ for all $n \in \mathbb{Z}$.

Proof. Let $g \in Y'$, then

$$(S+T)^\times(g) = g \circ (S+T) = g \circ S + g \circ T = S^\times(g) + T^\times(g) = (S^\times + T^\times)(g).$$

Similarly, we have $(\alpha T)^\times = \alpha T^\times$.

Next, if $S^{-1} \in B(Y, X)$, we show that $(S^{-1})^\times \circ S^\times = id$ and $S^\times \circ (S^{-1})^\times = id$.

$$((S^{-1})^\times \circ S^\times)(g) = S^{-1}(g \circ S) = g \circ (S \circ S^{-1}) = g$$

similarly, we have the other equality. Now since $S^{-1} \in B(Y, X)$, and $\|(S^{-1})^\times\| = \|S^{-1}\|$, we conclude that $(S^\times)^{-1} \in B(X', Y')$.

Lastly,

$$(ST)^\times(g) = g \circ (ST).$$

and

$$T^\times S^\times(g) = T^\times(g \circ S) = g \circ S \circ T = g \circ (ST).$$

□

Proposition 4.10 (Annihilator)**.** *Let X and Y be normed spaces, $T : X \to Y$ be a bounded linear operator and $M = \overline{\mathscr{R}(T)}$, then the annihilator of M is the nullspace of $T^\times$.*

Proof. Suppose g annihilates M, then

$$T^\times(g)(x) = g(T(x)) = 0 \quad \forall x \in X,$$

since $g(x) \in M$. On the other hand, if g is such that $(T^\times g)(x) = g(T(x)) = 0$ for all $x \in X$, then g clearly annihilates $\mathscr{R}(T)$. However, by continuity, we conclude that g annihilates $\mathscr{R}(T)$. □

To end off this section, we consider the relationship of this adjoint operator $T^\times$ and the Hilbert-adjoint operator T^*. In this case, we consider $X = H_1$ and $Y = H_2$, where H_1 and H_2 are two Hilbert spaces. If $T : H_1 \to H_2$ then $T^\times : H_2' \to H_1$. In particular, $T^\times$ is given by

$$T^\times g = f \quad \textbf{where } f(x) = g(T(x)).$$

By the Riesz representation theorem, we know there exists $x_0 \in H_1$ and $y_0 \in H_2$ such that

$$\begin{aligned} f(x) &= \langle x, x_0 \rangle \\ g(y) &= \langle y, y_0 \rangle \end{aligned}$$

which is uniquely determined by f and g. This defines operators:

$$\begin{aligned} A_1 : H_1' \to H_1 \quad &\textbf{by} \quad A_1 f = x_0, \\ A_2 : H_2' \to H_2 \quad &\textbf{by} \quad A_2 g = y_0. \end{aligned}$$

Moreover, we know A_1 and A_2 are bijective and isometric. Furthermore, this operation is conjugate linear, that is if we write $f_1(x) = \langle x, x_1 \rangle$ and $f_2(x) = \langle x, x_2 \rangle$, we have for all x and scalars α, β,

$$\begin{aligned} (\alpha f_1 + \beta f_2)(x) &= \alpha f_1(x) + \beta f_2(x) \\ &= \alpha \langle x, x_1 \rangle + \beta \langle x, x_2 \rangle \\ &= \langle x, \bar{\alpha} x_1 + \bar{\beta} x_2 \rangle. \end{aligned}$$

Hence $A_1(\alpha f_1 + \beta f_2) = \bar{\alpha} A_1 f_1 + \bar{\beta} A_1 f_2$. Similarly, A_2 is conjugate linear. Thus we have the following diagram:

$$\begin{tikzcd} H_1 \arrow[r, shift left, "T"] & H_2 \arrow[l, shift left, "T^*"] \\ H_1' \arrow[u, "A_1"] & H_2' \arrow[u, "A_2"'] \arrow[l, "T^\times"] \end{tikzcd}$$

Then if we define

$$T^* = A_1 T^\times A_2^{-1} : H_2 \to H_1, \quad T^*(y_0) = x_0 \tag{4.1}$$

we show that T^* is in fact the Hilbert-adjoint of T. In particular, we note that T^* given by the definition is linear and

$$\langle Tx, y_0\rangle = g(Tx) = f(x) = \langle x, x_0\rangle = \langle x, T^* y_0\rangle.$$

Hence e conclude the following theorem:

Theorem 4.11. *Formula (4.1) represents the Hilbert-adjoint operator T^* of a linear operator T on a Hilbert space in terms of the adjoint operator $T^\times$ of T.*

Remark 4.11.1. *$T^\times$ is defined on the dual of the space which contains the range of T, whereas T^* is defined directly on the space which contains the range of T. For $T^\times$, we have $(\alpha T)^\times = \alpha T^\times$ but for T^* we have $(\alpha T)^* = \bar{\alpha} T^*$. Then in the finite dimensional case, $T^\times$ is represented by the transpose of the matrix representing T, whereas T^* is represented by the complex conjugate transpose of the matrix representing T.*

4.3 Reflexive Space

Recall we have defined that

Definition 4.12 (Reflexive Space)**.** *There is a canonical natural embedding from X to X^{**} given by*

$$\phi : X \to X^{**}, \;\; x \mapsto \hat{x}$$

*where $\hat{x}(f) = f(x)$ for any $f \in X^*$. If ϕ is also surjective, then we say that X is **algebraically reflexive**. Moreover, if X is a normed space, and ϕ is an isometry (isometric linear isomorphism) from X to $(X')'$, then X is said to be **reflexive space**.*

Remark 4.12.1. *The range of the canonical embedding is always closed. Since we see later that the canonical embedding is always an isometry.*

Remark 4.12.2. *If X is reflexive, it is isomorphic to X'', however, the converse is not true. There are normed spaces X such that X and X'' are isomorphic but X is not reflexive.*

Remark 4.12.3. *Recall that we have shown every Hilbert space is reflexive.*

However, we left something unverified in the definition. That is $\hat{x}$ need to be bounded in order for ϕ to be a well-defined map mapping into X''. Hence this is exactly what we show:

Lemma 4.13. *For every $x \in X$, the functional $\hat{x} : X' \to \mathbb{F}$ given by $\hat{x}(f) = f(x)$ is a bounded linear functional on X', so $\hat{x} \in X''$. In fact,*

$$\|\hat{x}\| = \|x\|.$$

Proof. We know $\hat{x}$ is a linear map. Now definition, we have

$$\|\hat{x}\| = \sup_{\substack{f \in X' \\ f \neq 0}} \frac{|\hat{x}(f)|}{\|f\|} = \sup_{\substack{f \in X' \\ f \neq 0}} \frac{|f(x)|}{\|f\|} = \|x\|$$

where the last equality follows from the dual representation of the norm of $\|x\|$. $\square$

Remark 4.13.1. *This also shows that the canonical mapping always preserves norm and hence to check for isometry, suffices to check that the map is surjective. In fact, we have the following lemma.*

Lemma 4.14. *The canonical mapping, denoted by C, from X to X'', is an isomorphism of the normed space X onto the normed space $\mathscr{R}(C)$, the range of C. In particular, this shows that X is embeddable to $\mathscr{R}(C)$ as a normed space.*

Lemma 4.15 (Completeness)**.** *If a normed space X is reflexive, it is complete, hence a Banach Space.*

Proof. Since the X'' is complete. □

Lemma 4.16. *Every finite dimensional normed space is reflexive.*

Proof. Since on finite dimensional space, we have $X' = X^*$ and $X'' = X^{**}$. We also know that X is algebraically reflexive, hence the canonical mapping is surjective, so X is reflexive. □

Example:

1. ℓ^p with $1 < p < +\infty$ is reflexive. Moreover $L^p[a,b]$ with $1 < p < +\infty$ is reflexive. However, $C[a,b]$ is not reflexive since it is not complete.
2. We will show that ℓ^1 is not reflexive. In fact, this coincides with our prior knowledge, since we know that the dual of ℓ^1 is ℓ^∞ and the dual of ℓ^∞ is not ℓ^1.

Lemma 4.17. *Let Y be a proper closed subspace of a normed space X. Let $x_0 \in X \setminus Y$ be arbitary, and*

$$\delta = \inf_{\tilde{y} \in Y} \|\tilde{y} - x_0\|.$$

Then there exists an $\tilde{f} \in X'$ such that

$$\|\tilde{f}\| = 1, \quad \tilde{f}(y) = 0 \text{ for all } y \in Y, \quad \tilde{f}(x_0) = \delta.$$

Proof. We consider the subspace $Z \subset X$ spanned by Y and x_0, defined on Z a bounded linear functional f by

$$f(z) = f(y + \alpha x_0) = \alpha\delta \quad \text{for some } y \in Y.$$

Note Z is direct sum of Y and the span of $\{x_0\}$, hence f is well-defined on Z. Now can easily check that $f(y) = 0$ for all $y \in Y$, $f(x_0) = \delta$. Next, for any $z \in Z$,

$$\begin{aligned}
|f(z)| &= |\alpha|\delta \\
&= |\alpha| \inf_{\tilde{y} \in Y} \|\tilde{y} - x_0\| \\
&\le |\alpha| \| - \frac{1}{\alpha} y - x_0\| \\
&= \|y + \alpha x_0\| = \|z\|
\end{aligned}$$

Hence $\|f\| \le 1$. On the other hand, by the definition of infimum, Y contains a sequence (y_n) such that $\|y_n - x_0\| \to \delta$. Let $z_n = y_n - x_0$. Then we have $f(z_0) = -\delta$. So

$$\|f\| = \sup_{\substack{z \in Z \\ z \neq 0}} \frac{|f(z)|}{\|z\|} \ge \frac{|f(z_n)|}{\|z_n\|} = \frac{\delta}{\|z_n\|} \to 1.$$

Hence $\|f\| \geq 1$, so $\|f\| = 1$. Lastly, by Hahn-Banach theorem, f admits an extension $\tilde{f}$ to X such that the properties are still satisfied. $\square$

Corollary 4.17.1. *Different closed subspaces Y_1 and Y_2 of a normed space X have different annihilators.*

Corollary 4.17.2. *Let M be any subset of a normed space X. Then $x_0 \in X$ is in $\overline{\operatorname{span} M}$ if and only if $f(x_0) = 0$ for every $f \in X'$ such that $f|_M = 0$.*

Corollary 4.17.3. *A subset M of a normed space X is total in X if and only if every $f \in X'$ which is zero everywhere on M is zero everywhere on X.*

Theorem 4.18. *If the dual space X' of a normed space X is separable, then X is separable.*

Proof. Let S_1' denote the unit sphere in X', then we claim S_1' is separable. Let $\{x_n'\} \subset X'$ be a countable dense subset of X', then define

$$y_n' = \frac{x_n'}{\|x_n'\|}, \tag{4.2}$$

then $\{y_n'\}$ is a countable subset of S_1'. Now $\forall x' \in S_1^*$, $\exists x_{n_j}'$, $\|x_{n_j}' - x'\| \to 0$, then, $\|x_{n_j}'\| \to 1$, so

$$\|y_{n_j}' - x'\| \leq \left|1 - \frac{1}{\|x_{n_j}'\|}\right| \|x_{n_j}'\| + \|x_{n_j}' - x'\| \to 0. \tag{4.3}$$

Thus, $\{y_n'\}$ is a countable dense subset of S_1'.

Next, by definition of the norm of an operator, for each $n \in \mathbb{N}_{\geq 1}$, there exists $x_n \in X$, such that $\|x_n\| = 1$, $\langle y_n', x_n\rangle \geq 1/2$. Let $X_0 = \overline{\operatorname{span}\{x_1, x_2, \cdots,\}}$, we show $X_0 = X$. Suppose not, then using the Hahn-Banach Theorem, we can construct $y' \in S_1'$, such that $\langle y', x\rangle = 0$ for all $x \in X_0$, but then

$$\|y' - y_n'\| \geq |\langle y' - y_n', x_n\rangle| \geq \frac{1}{2}, \tag{4.4}$$

which is clearly contradicts that $\{y_n'\}$ is dense in S_1'. Thus $X_0 = X$, and X is separable.

$\square$

Corollary 4.18.1. *A separable normed space X with a nonseparable dual space X' cannot be reflexive.*

Corollary 4.18.2. *A Hilbert space H is separable if and only if H^* is separable.*

Theorem 4.19. *A Banach space X is reflexive if and only if its dual space X' is reflexive.*

Proof. Suppose X' is reflexive. Let $C : X \to X''$ and $C' : X' \to X'''$ be canonical injections. Suppose that $C(X) \neq X''$, then by the Hahn-Banach theorem, there exists $\zeta \in X'''$ such that $\zeta \neq 0$ and $\zeta \equiv 0$ on $C(X)$ ($C(X)$ is closed since C is isometric). Because X' is reflexive, there exists $\theta \in X'$ such that $\zeta = C'(\theta)$. Then for all $x \in X$,

$$0 = \langle \zeta, Cx\rangle = \langle C'\theta, Cx\rangle = \langle Cx, \theta\rangle = \langle \theta, x\rangle.$$

This shows that $\theta = 0$, hence $\zeta = 0$, which is a contradiction that $\zeta \neq 0$. Hence we conclude that X must be reflexive.

Conversely, suppose X is reflexive. We note that

$$C^{\times} \circ C' = id_{X'}.$$

Next, if an operator is invertible with bounded inverse, then so is its adjoint. So $C^{\times}$ is invertible with bounded inverse. Hence C' is invertible hence an isomorphism. $\square$

4.4 Uniform Boundedness Theorem

Theorem 4.20 (Uniform Boundedness Theorem)**.** *Let (T_n) be a sequence of bounded linear operators $T_n : X \to Y$ from a Banach space X into a normed space Y such that $(\|T_n x\|)$ is bounded for every $x \in X$, say,*

$$\|T_n x\| \leq c_x$$

for all $n = 1, \cdots$, where c_x is a real number only dependent on x. Then the sequence of the norms $\|T_n\|$ is bounded.

Remark 4.20.1. *The criterion of X being complete is essential, as otherwise we can easily construct a counter example.*

Proof. For every $k \in N$, let $A_k \subset X$ be the set of all x such that $\|T_n x\| \leq k$ for all n. Then A_k is closed, as it is the intersection of closed sets. Since $\|T_n x\| \leq c_x$, then we see that every $x \in X$ belongs to some A_k, hence

$$X = \bigcup_{k=1}^{\infty} A_k.$$

Since X is complete, Baire's theorem implies that some A_k contains an open ball, say

$$B_0 = B(x_0, r) \subset A_{k_0}.$$

Let $x \in X$ be arbitrary nonzero vector, we set

$$z = x_0 + \gamma x, \quad \gamma = \frac{r}{2\|x\|}.$$

Then $\|z - x_0\| < r$, so that $z \in B_0 \subset A_{k_0}$. Then $\|T_n z\| \leq k_0$ for all n. Also $\|T_n x_0\| \leq k_0$ since $x_0 \in B_0$. Since

$$x = \frac{1}{\gamma}(z - x_0).$$

This gives that for all n,

$$\|T_n x\| = \frac{1}{\gamma}\|T_n(z - x_0)\| \leq \frac{1}{\gamma}(\|T_n z\| + \|T_n x_0\|) \leq \frac{4}{r}\|x\| k_0.$$

Hence for all n, we have

$$\|T_n\| = \sup_{\|x\|=1} \|T_n x\| \leq \frac{4}{r} k_0.$$

$\square$

Example: (Space of polynomials) The normed space X of all polynomials with norm defined by

$$\|x\| = \max_j |\alpha_j| \qquad (\alpha_0, \alpha_1, \cdots \text{ the coefficients of } x)$$

is not complete.

Proof. We construct a sequence of bounded linear operators (T_n) on X which is pointwise bounded but the norm is not uniformly bounded, so that X cannot be complete.

We may write a polynomial $x \neq 0$ of degree N_x in the form

$$x(t) = \sum_{j=0}^{\infty} \alpha_j t^j \qquad (\alpha_j = 0 \text{ for } j > N_x).$$

(For $x = 0$ the degree is not defined in the usual discussion of degree, but this does not matter here.) As a sequence of operators on X we take the sequence of functionals T_n defined by

$$T_n 0 = 0, \qquad T_n x = \alpha_0 + \alpha_1 + \cdots + \alpha_{n-1}.$$

Each T_n is linear and bounded, since $\|T_n(x)\| \leq n\|x\|$. Furthermore, for each fixed $x \in X$, the sequence $(|f_n(x)|)$ is clearly bounded. However, we see that $\|T_n\|$ is not bounded, since we consider $x(t) = 1 + t + \cdots t^n$ which has norm one. But $\|T_n x\| = n\|x\|$, shows $\|T_n\|$ is not bounded. □

Corollary 4.20.1 (Banach Steinhause Theorem)**.** *Suppose that* $\lim_{n\to\infty} T_n x$ *exists for all* x*. Define* $T : X \to Y$ *by*

$$T_x = \lim_{n\to\infty} T_n x.$$

Then $T \in B(X, Y)$*. Moreover,*

$$\|T\| \leq \liminf_{n\to\infty} \|T_n\|.$$

Proof. Linearity is clear. Boundedness follows from the Uniform Boundedness Theorem. The last assertion follows since for all $x \in X$ with $\|x\| = 1$, we have

$$\|Tx\| = \liminf_{n\to\infty} \|T_n x\|.$$

□

Lemma 4.21. *If* (x_n) *is a sequence in a Banach Space* X *such that* $(f(x_n))$ *is bounded for all* $f \in X'$*, then* $(\|x_n\|)$ *is bounded.*

Proof. Consider $(\hat{x}_n)$, suffices to show $(\|\hat{x}_n\|)$ is bounded, since $\|\hat{x}_n\| = \|x_n\|$. Since we have for all $f \in X'$,

$$\|\hat{x}_n(f)\| = \|f(x_n)\|$$

is bounded, then by the uniform boundedness theorem, we must have $\|\hat{x}_n\|$ is bounded. □

Theorem 4.22. *If* X *and* Y *are Banach spaces and* $T_n \in B(X, Y)$ *for all* $n \in \mathbb{N}$*. Then the following statement are equivalent:*

1. $(\|T_n\|)$ *is bounded;*

2. $(\|T_n x\|)$ *is bounded for all* $x \in X$*;*

3. $(|g(T_n x)|)$ *is bounded for all* $x \in X$ *and all* $g \in Y'$*.*

Proof. By the uniform boundedness theorem, we know (1) and (2) are equivalent. Now suppose $\|T_n\|$ is bounded, then $|g(T_n x)| \leq \|g\|\|T_n\|\|x\|$ is bounded for each fixed x. On the other hand, suppose $(|g(T_n x)|)$ is bounded for all $x \in X$ and all $g \in Y'$, then by Lemma)4.21), we know $\|T_n x\|$ is bounded for all $x \in X$. $\square$

4.5 Different Types of Convergence

Definition 4.23 (Strong Convergence)**.** *A sequence* (x_n) *in a normed space* X *is said to be* ***strongly convergent*** *(or the convergent in the norm) if there is an* $x \in X$ *such that*

$$\lim_{n\to\infty} \|x_n\| = 0$$

and we write

$$\lim_{n\to\infty} x_n = x \quad \text{and} \quad x_n \to x.$$

Definition 4.24 (Weak Convergence)**.** *A sequence* $\{x_n\}$ *in a normed space* X *is said to be weakly convergent if there is an* $x \in X$ *such that for every* $f \in X$”*,*

$$\lim_{n\to\infty} f(x_n) = f(x).$$

and we denote

$$x_n \rightharpoonup x.$$

In particular, a sequence $\{x_n\}$ *in a Hilbert space* H *is said to be* ***converge weakly*** *to* x *if for all* $y \in H$*,*

$$\lim_{n\to\infty} \langle x_n, y\rangle = \langle x, y\rangle.$$

Lemma 4.25. *Let* X *and* Y *be normed space,* $T \in B(X,Y)$ *and* (x_n) *is a sequence in* X*. If* $x_n \rightharpoonup x$*, then* $Tx_n \rightharpoonup Tx$*.*

Proof. Since for any $f \in Y'$, $f \circ T \in X'$. $\square$

Lemma 4.26 (Weak Convergence)**.** *Let* (x_n) *be a weakly convergence sequence in a normed space* X *with* $x_n \rightharpoonup x$*. Then*

1. *The weak limit* x *of* (x_n) *is unique.*

2. *Every subsequence of* (x_n) *converges weakly to* x*.*

3. *The sequence* $(\|x_n\|)$ *is bounded.*

Suppose also (y_n) *converges weakly to* y*, and* $\alpha \in \mathbb{F}$*, then*

$$x_n + \alpha y_n \rightharpoonup x + \alpha y.$$

Proof. (1) and (2) are clear. (3) follows from Lemma (4.21). The last statement follows from the linearity of linear operators. □

Theorem 4.27 (Strong and Weak Convergence)**.** *Let (x_n) be a sequence in a normed space X, then*

1. *Strong convergence implies weak convergence with the same limit.*
2. *Weak convergence does not imply strong convergence.*
3. *If $\dim X < \infty$, then the two notion coincide.*

Proof. (1) is clear. Now for (2) we construct a counter example. Let (e_n) be an orthonormal sequence in a Hilbert space H. Let $f \in H'$, then we know $f(x) = \langle x, z \rangle$ for some z, hence $f(e_n) = \langle e_n, z \rangle$. Now the Bessel inequality tells us that

$$\sum_{n=1}^{\infty} |\langle e_n, z \rangle|^2 \leq \|z\|^2.$$

hence the series on the left converges, so that its terms must approach zero as $n \to \infty$. This implies

$$f(e_n) = \langle e_n, z \rangle \to 0.$$

Since $f \in H'$ was arbitrary, we see that $e_n \rightharpoonup 0$. However (e_n) does not converge strongly because

$$\|e_m - e_n\|^2 = \langle e_m - e_n, e_m - e_n \rangle = 2.$$

For (3) suppose that $x_n \rightharpoonup x$ and $\dim X = k$. Let $\{e_1, \cdots, e_k\}$ be any basis for X and say

$$x_n = \alpha_1^{(n)} e_1 + \cdots + a_k^{(n)} e_k.$$

Then by considering the projection to the span of the i^{th} basis e_i, we have that $\{\alpha_i^{(n)}\}$ must converge. Hence it x_n converges strongly since its finite dimensional. □

Example: If $x_n \in C[a, b]$, and $x_n \rightharpoonup x \in C[a, b]$, then (x_n) is pointwise convergent on $[a, b]$.

Lemma 4.28. *In a Hilbert space X, x_n converges to x if and only if x_n converges to x weakly and $\|x_n\|$ converges to $\|x\|$.*

Proof. The forward direction is trivial, for the converse, suffices to show that $\|x_n - x\| \to 0$ as $n \to \infty$. To this regard, we show $\|x_n - x\|^2 \to 0$ as $n \to \infty$. We have

$$\begin{aligned}
\|x_n - x\|^2 &= \langle x_n - x, x_n - x \rangle \\
&= \langle x_n, x_n \rangle + \langle x, x \rangle - \langle x_n, x \rangle - \langle x, x_n \rangle \\
&= \|x_n\|^2 + \|x\|^2 - 2\mathrm{Re}(\langle x_n, x \rangle).
\end{aligned}$$

Now since x_n converges weakly to x and $\|x_n\|$ converges to $\|x\|$, then $\langle x_n, x \rangle \to \langle x, x \rangle = \|x\|^2$ and $\|x_n\|^2 \to \|x\|$ as $n \to \infty$. Hence we have

$$\lim_{n \to \infty} \|x_n\|^2 = \|x\|^2 + \|x\|^2 - 2\mathrm{Re}(\|x\|^2) = 0.$$

So x_n converges to x. □

Lemma 4.29. *In a normed space X, $x_n \rightharpoonup x$ if and only if*

- *The sequence $(\|x_n\|)$ is bounded.*
- *For every element f of a total subset $M \subset X'$, $f(x_n) \to f(x)$.*

Proof. $\Rightarrow$: this direction is trivial.

$\Leftarrow$: suppose $(\|x_n\|)$ is bounded and for every element f of a total subset $M \subset X'$ we have $f(x_n) \to f(x)$. We show that $f(x_n) \to f(x)$.

Suppose $\|x_n\| \leq c$ and also that $\|x\| \leq c$. Since M is total in X', for every $f \in X'$ there is a sequence (f_j) in span M such that $f_j \to f$. Hence for any given $\epsilon > 0$ we can find a j such that

$$\|f_j - f\| \leq \frac{\epsilon}{3c}.$$

Moreover, since $f_j \in \operatorname{span} M$, then there is an N such that for all $n > N$, we have

$$|f_j(x_n) - f_j(x)| < \frac{\epsilon}{3}.$$

Then by triangle inequality, for $n > N$, we have

$$|f(x_n) - f(x)| \leq \epsilon.$$

□

Corollary 4.29.1. *For $1 < p < \infty$, in ℓ^p, $x_n \rightharpoonup x$ if and only if*

- *The sequence $(\|x_n\|)$ is bounded;*
- *For every fixed j we have $\xi_j^{(n)} \to x_j$ as $n \to \infty$; here $x_n = (\xi_j^{(n)})$ and $x = (\xi_j)$.*

Proposition 4.30. *If X is normed space and $x_n \rightharpoonup x$, then $x_0 \in \bar{Y}$ where $Y = \operatorname{span}(x_n)$. In particular, there is a sequence (y_m) of linear combinations of elements of (x_n) which converges strongly to x_0. Consequently, any closed subspace Y of a normed space X contains the limits of all weakly convergent sequences of its elements.*

Proof. Suppose $x_0 \notin \bar{Y}$, then we can construct a functional that annihilates $\bar{Y}$ but does not annihilate x_0 by Lemma (4.17). □

Proposition 4.31 (Weak Semi-lower Continuity of Norms)**.** *Let X be a normed space and $x_n \rightharpoonup x$, then*

$$\liminf_{n\to\infty} \|x_n\| \geq \|x\|.$$

Proof. By Hahn Banach's Theorem, there exists a functional f on X with

$$\|f\| = 1 \quad \text{and} \quad f(x) = \|x\|.$$

Then

$$\|f(x_n)\| \leq \|x_n\|.$$

But since $f(x_n) \to f(x) = \|x\|$, we must have

$$\liminf_{n\to\infty} |f(x_n)| \geq \|x\|.$$

Then

$$\liminf_{n\to\infty} \|x_n\| \geq \|x\|.$$

□

Definition 4.32 (Weak Completeness)**.** *Let X be a normed space, then a* ***weak Cauchy sequence*** *(x_n) of X is a sequence in X such that for every $f \in X'$, the sequence $(f(x_n))$ is Cauchy. We say that X is* ***weakly complete****, if each weak Cauchy sequence in X converges weakly in X.*

Remark 4.32.1. *Again by Lemma 4.21, weakly Cauchy sequences are bounded.*

Proposition 4.33. *If X is reflexive normed spaced, then X is weakly complete.*

Proof. Suppose (x_n) is weakly Cauchy in X, then consider $\hat{x}_n \in X''$ which is pointwise bounded, hence uniformly bounded. By Corollary (4.20.1), $\hat{x}_n$ converges pointwise to some $\hat{x} \in X''$. Since X is reflexive, then $\phi(x) = \hat{x}$ for some $x \in X$, where ϕ is the canonical embedding. One can verify easily that x is the weak limit of (x_n). □

Definition 4.34 (Convergence Of Sequences Of Operators)**.** *Let X and Y be normed spaces. A sequence (T_n) of operators $T_n \in B(X,Y)$ is said to:*

1. ***uniformly operator convergent*** *with uniform operator limit T, if (T_n) converges in the norm on $B(X,Y)$ to T;*
2. ***strongly operator convergent*** *with strong operator limit T if $(T_n x)$ converges strongly to Tx in Y for every $x \in X$;*
3. ***weakly operator convergent*** *with weak operator limit T, if $(T_n x)$ converges weakly to Tx in X for every $x \in X$.*

Remark 4.34.1. *It is clear from the definition that uniform convergent implies strong operator convergent implies weak operator convergent. Moreover the limit must also be the same. However, the converses do not necessarily hold. However, if Y is finite dimensional Banach space, then we see that the notion of strong operator convergent coincides with the notion of weak operator convergent since the notion of strong convergence and weak convergence coincides in finite dimensional normed spaces.*

Remark 4.34.2. *For the uniform operator limit of T_n, T must be bounded, as otherwise $\|T_n - T\|$ would not make sense. However, if the convergence is only strong or weak, T is linear but may be unbounded if X is not complete. On the other hand, If X is Banach space, then by Banach Steinhause Theorem, we know T must be bounded.*

Example: Let the space X be all sequences $x = (\xi_j)$ in ℓ^2 with only finitely many nonzero terms, taken with the metric on ℓ^2. This space is not complete. We consider a sequence of bounded linear operators T_n on X defined by

$$T_n x = (\xi_1, 2\xi_2, 3\xi_3, \cdots, n\xi_n, \xi_{n+1}, \xi_{n+2}, \cdots),$$

this sequence (T_n) converges strongly to the unbounded linear operator T defined by

$$Tx = (\eta_j), \quad \eta_j = j\xi_j.$$

Definition 4.35 (Strong And Weak Star Convergence Of Functionals)**.** *Let (f_n) be a sequence of bounded linear functionals on a normed space X. Then*

1. ***Strong convergence** of (f_n) means that there is an $f \in X'$ such that $\|f_n - f\| \to 0$. This is written as*

$$f_n \to f.$$

2. ***Weak star convergence** of (f_n) means that there is an $f \in X'$ such that $f_n(x) \to f(x)$ for all $x \in X$. This is written*

$$f_n \rightharpoonup^* f.$$

*The limit f are called **strong limit** and **weak star limit** of (f_n) respectively.*

Remark 4.35.1. *Note that for functionals, weak operator convergence and strong operator convergence coincide, hence together they are called weak star convergence ($\mathbb{F}'$ is just $\mathbb{F}$). In this case, the uniform operator convergence is called strong convergence.*

The following theorem can easily be proved:

Theorem 4.36. *A sequence (T_n) of operators $T_n \in B(X, Y)$, where X and Y are Banach spaces, is strongly operator convergent if and only if:*

1. *The sequence $(\|T_n\|)$ is bounded;*

2. *The sequence $(T_n x)$ is Cauchy in Y for every x in a total subset M of X.*

Then we have the following immediate corollary:

Corollary 4.36.1. *A sequence (f_n) of bounded linear functionals on a Banach space X is weak star convergent, the limit being a bounded linear functional on X, if and only if:*

1. *The sequence $(\|f_n\|)$ is bounded;*

2. *The sequence $(f_n(x))$ is Cauchy for every x in a total subset M of X.*

Example: Weakly Cauchy sequences are not necessarily weak convergent. Consider the sequence $y_n = e_1 + e_2 + \cdots + e_n \in c_0 \subset \ell^\infty$. This is weakly Cauchy but not weakly convergent.

4.6 Open Mapping Theorem

Definition 4.37 (Open Mapping)**.** *Let X and Y be metric spaces. Then $T : \mathscr{D}(T) \to Y$ with domain $\mathscr{D}(T) \subset X$ is called an **open mapping** if for every open set in $\mathscr{D}(T)$ the image is an open set in Y.*

Lemma 4.38. *A bounded linear operator T from a Banach space X onto (surjective) a Banach space Y has the property that the image $T(B_0)$ of the open unit ball $B_0 = B(0, 1) \subset X$ contains an open ball about $0 \in Y$.*

Remark 4.38.1. *Surjectivity is necessary, as otherwise the zero map is a counter example.*

Proof. Let $B_n := B(0, 2^{-n})$, $n \in \mathbb{N}$. Consider B_1, then for any $x \in X$, there exists $k > 0$ such that $x \in kB_1$. Then

$$X = \bigcup_{k=1}^{\infty} kB_1.$$

Since T is surjective,

$$Y = T(X) = T\left(\bigcup_{k=1}^{\infty} kB_1\right) = \bigcup_{k=1}^{\infty} kT(B_1) = \bigcup_{k=1}^{\infty} \overline{kT(B_1)}.$$

Since Y is a complete metric space, then by the Baire Category Theorem, at least one of $\overline{kT(B_1)}$ must contain an open ball. So $\overline{T(B_1)}$ must contain an open ball, which we will denote by $B(y_0, \epsilon)$. Then $B(0, \epsilon) = B(y_0, \epsilon) - y_0$. So

$$B(0, \epsilon) \subset \overline{T(B_1)} - y_0.$$

Next suppose $y \in \overline{T(B_1)} - y_0$. Then $y + y_0 \in \overline{T(B_1)}$ and $y_0 \in \overline{T(B_1)}$. Then we have

$$u_n \to y + y_0, \quad v_n \to y_0$$

with $u_n = Tw_n \subset T(B_1)$ and $v_n = Tz_n \subset T(B_1)$. Since

$$\|w_n - z_n\| \leq \|w_n\| + \|z_n\| < \frac{1}{2} + \frac{1}{2}$$

Then $w_n - z_n \in B_0$, and $T(w_n - z_n) = Tw_n - Tz_n = u_n - v_n \to y$. That is $y \in \overline{T(B_0)}$.

Since T is linear, $\overline{T(B_n)} = 2^{-n}\overline{T(B_0)}$. Let $V_n = B_Y\left(0, \frac{\epsilon}{2^n}\right) \subset Y$. Then by our previous analysis, we have $V_n \subset \overline{T(B_n)}$.

Next, we show that $V_1 \subset T(B_0)$. Let $y \in V_1$ be arbitrary, then $y \in \overline{T(B_1)}$. Let $v_1 \in T(B_1)$ such that $v_1 = Tx_1$, $x_1 \in B_1$ and $\|y - v_1\| \leq \frac{\epsilon}{4}$. Then since $\|y - v_1\| < \frac{\epsilon}{4}$, $y - Tx_1 \in V_2 \subset \overline{T(B_2)}$. So $\exists x_2 \in B_2$ such that $v_2 = Tx_2$ and $\|(y - Tx_1) - v_2\| < \frac{\epsilon}{8}$. Continue in this way, we could construct the sequence $\{x_k\}$ such that $x_k \in \overline{T(B_k)}$ and

$$\left\| y - \sum_{k=1}^{n} Tx_k \right\| < \frac{\epsilon}{2^{k+1}}.$$

Let $z_n = \sum_{k=1}^{n} x_k$. Then

$$\|z_n - z_m\| \leq \sum_{k=m+1}^{n} \|x_k\| < \sum_{k=m+1}^{\infty} \frac{1}{2^k}.$$

So z_n is Cauchy, so z_n converges to some $x \in X$ as X is a Banach space. Moreover, we have

$$\|x\| = \left\| \sum_{k=1}^{\infty} x_k \right\| \leq \sum_{k>1}^{\infty} \|x_k\| < \sum_{k=1}^{\infty} \frac{1}{2^k} = 1.$$

We have strict inequality in the middle since $\|x_1\| < \frac{1}{2} - \delta$ for some $\delta > 0$. This shows $x \in B_0$. Next by continuity, $Tz_n \to Tx$, so $y = Tx \in T(B_0)$. Since $y \in V_1$ is arbitrary, then $V_1 \subset T(B_0)$. Since V_1 is open ball centered at 0, this concludes the proof of the lemma. □

Theorem 4.39 (Open Mapping Theorem)**.** *A bounded linear operator T from a Banach space X onto (surjective) a Banach space Y is an open mapping.*

Proof. We prove that for every open set $A \subset X$ the image $T(A)$ is open in Y. Let $y \in Tx \in T(A)$, we show the set $T(A)$ contains an open all about $y = Tx$.

Let $y = Tx \in T(A)$. Since A is open, it contains an open ball with center x. hence $A - x$ contains an open ball with center 0. Let the radius of the ball be r and set $k = \frac{1}{r}$, so that $r = \frac{1}{k}$. Then $k(A - x)$ contaisn the open unit ball $B(0, 1)$. Then Lemma (6.26) implies $T(k(A - x)) = k[T(A) - Tx]$ contains an open ball about 0, and so does $T(A) - Tx$. hence $T(A)$ contains an open ball about $Tx = y$. Since $y \in T(A)$ was arbitrary, $T(A)$ is open. □

Theorem 4.40 (Bounded Inverse Theorem)**.** *If T is a bijective bounded linear operator from a Banach space X to a Banach space Y. Then T^{-1} is continuous and thus bounded. In this case, there are positive real numbers a and b such that $a\|x\| \le \|Tx\| \le b\|x\|$ for all $x \in X$.*

Proof. Suppose $T^{-1} : Y \to X$ is well-defined, it is continuous since T is an open map. Hence it is bounded. Then second part follows readily from the boundedness of T and T^{-1}. □

Example: Completeness is necessary in the Bounded inverse theorem, since we have the following counter example when X and Y is not complete (If we have a bijective bounded linear operator, then X not complete implies Y not complete). Let X be the normed space whose points are sequences of complex numbers $x = (\xi_j)$ with only finitely many nonzero terms and norm given by the sup norm. Let $T : X \to X$ be defined by

$$y = Tx = \left(\xi_1, \frac{1}{2}\xi_2, \frac{1}{3}\xi_3, \cdots\right).$$

Then T is linear with operator norm of being 1. However T^{-1} is unbounded.

Corollary 4.40.1. *Let X and Y be Banach spaces and $T : X \to Y$ an injective bounded linear operator. Then $T^{-1} : \mathscr{R}(T) \to X$ is bounded if and only if $\mathscr{R}(T)$ is closed in Y.*

Proof. The if part follows from the Bounded Inverse Theorem. Now suppose T^{-1} is bounded, then we can show using sequential limit that $\mathscr{R}(T)$ is closed. □

Corollary 4.40.2. *Let $X_1 = (X, \|\cdot\|_1)$ and $X_2 = (X, \|\cdot\|_2)$ be Banach spaces. If there is a constant c such that $\|x\|_1 \le c\|x\|_2$ for all $x \in X$, then there is a constant k such that $\|x\|_2 \le k\|x\|_1$ for all $x \in X$, so the two norms are in fact equivalent.*

Proof. Consider the map $T : X_1 \to X_2$ by $x \mapsto x$. □

4.7 Closed Graph Theorem

Definition 4.41 (Closed Linear Operator). *Let X and Y be normed spaces and $T : \mathscr{D}(T) \to Y$ a linear operator with domain $\mathscr{D}(T) \subset X$. Then T is called a* ***closed linear operator*** *if its graph*

$$\Gamma(T) = \{(x, y) \mid x \in \mathscr{D}(T),\ y = Tx\}$$

is closed in the normed space $X \times Y$ with norm

$$\|(x, y)\| = \|x\| + \|y\|$$

or just the product topology.

Remark 4.41.1. *In fact $\Gamma(T)$ equipped with the norm becomes a normed vector space.*

Remark 4.41.2. *From the definition, one can easily show that the sum of two closed linear operator or the scalar multiple of a closed linear operator is again a closed linear operator.*

Example: Let $X = C[0, 1]$ and

$$\begin{aligned} T : \mathscr{D}(T) &\to X \\ x &\mapsto x' \end{aligned}$$

where x' is the derivative of x and $\mathscr{D}(T)$ is the subsapce of functions $x \in X$ which have continuous derivatives equipped with the C^0-norm. Then T is not bounded, since we can consider the sequence $\{x^n\}_{n\in\mathbb{N}}$. However, T is closed, but $\mathscr{D}(T)$ is not closed.
Note that if we consider

$$T : C^1[0, 1] \to C[0, 1], \quad x \mapsto x'$$

directly, with $C^1[0, 1]$ equipped with the C^1-norm. Then T is indeed bounded.

Theorem 4.42 (Closed Graph Theorem). *Let X and Y be Banach spaces and $T : \mathscr{D}(T) \to Y$ a closed linear operator, where $\mathscr{D}(T) \subset X$. Then if $\mathscr{D}(T)$ is closed in X, the operator T is bounded.*

Remark 4.42.1. *As we see in the previous example, without the extra condition $\mathscr{D}(T)$ is closed. Then the closedness of an operator does not imply the boundedness of an operator. In fact, the converse doesn't hold either for general linear operators. For example, we can consider mapping an proper dense subset to a proper dense subset via the identity mapping.*

Proof. Since X and Y are complete, $X \times Y$ is complete. By assumption $\Gamma(T)$ and $\mathscr{D}(T)$ are closed, so $\Gamma(T)$ and $\mathscr{D}(T)$ are also complete.

Define $p : \Gamma(T) \to \mathscr{D}(T)$ by $(x, Tx) \mapsto x$. This is clearly a bijective and linear. Moreover p is bounded, since

$$\|p(x, Tx)\| = \|x\| \le \|x\| + \|Tx\| = \|(x, Tx)\|.$$

Therefore by the bounded inverse theorem, p^{-1} is is continuous/bounded. That is, there exists $b > 0$ such that $\|p^{-1}x\| \le b\|x\|$ for all $x \in X$. That is

$$\|(x, Tx)\| \le b\|x\|$$

But

$$\|Tx\| \leq \|Tx\| + \|x\| = \|(x, Tx)\| \leq b\|x\|$$

This shows T is bounded. □

Theorem 4.43. *Let $T : \mathscr{D}(T) \to Y$ be a linear operator, where $\mathscr{D}(T) \subset X$ and X and Y are normed spaces. Then T is closed if and only if it has the following property. If $x_n \to x$, where $x_n \in \mathscr{D}(T)$, and $Tx_n \to y$, then $x \in \mathscr{D}(T)$ and $Tx = y$.*

Proof. Suppose T is closed, then as $x_n \to x$, $Tx_n \to y$,

$$(x_n, Tx_n) \to (x, y).$$

So $(x, y) \in \mathscr{D}(T)$ with $Tx = y$. Conversely, suppose $x_n \to x$, and $Tx_n \to y$, then $(x, y) \in \Gamma(T)$ and $x \in \mathscr{D}(T)$ Take any limit point of $\Gamma(T)$, then we using sequential limit, we see it is in $\Gamma(T)$, this shows $\Gamma(T)$ is closed. □

Lemma 4.44. *Let $T : \mathscr{D}(T) \to Y$ be a bounded linear operator with domain $\mathscr{D}(T) \subset X$, where X and Y are normed spaces. Then:*

1. *if $\mathscr{D}(T)$ is a closed subset of X, then T is closed.*
2. *if T is closed and Y is complete, then $\mathscr{D}(T)$ is a closed subset of X.*

Proof.

1. If $x_n \to x$, then $Tx_n \to Tx$ by the boundedness of T. Now if $(x_n, Tx_n) \to (x, y)$, then we must have $x_n \to x$ and $Tx_n \to y$. Since T is closed, then $(x, y) = (x, Tx) \in \Gamma(T)$.
2. For $x \in \overline{\mathscr{D}(T)}$, there is a sequence (x_n) in $\mathscr{D}(T)$ such that $x_n \to x$. Since T is bounded

   $$\|Tx_n - Tx_m\| = \|T(x_n - x_m)\| \leq \|T\|\|x_n - x_M\|.$$

   This shows that (Tx_n) is also Cauchy, so $Tx_n \to y$ for some $y \in Y$, as Y is complete. Since T is closed, $x \in \mathscr{D}(T)$, as $(x, y) \in \mathscr{D}(T)$.

□

Lemma 4.45. *The inverse T^{-1} of a closed linear operator T is a closed linear operator.*

Proof. This follows from the definition of closed linear operator. □

Lemma 4.46. *The null space $\mathcal{N}(T)$ of a closed linear operator $T : X \to Y$ is a closed subspace of X.*

Proof. Let $x \in \overline{\mathcal{N}(T)}$ with $x_n \to x$, $x_n \in \mathcal{N}(T)$. Then $(x_n, Tx_n) = (x_n, 0) \to (x, 0) \in \Gamma(T)$. Since $\Gamma(T)$ is closed, then $x \in \mathcal{N}(T)$. □

5 Approximations

5.1 Banach Fixed Point Theorem

Theorem 5.1 (Banach Fixed Point Theorem)**.** *Consider a nonempty complete metric space $X = (X, d)$. If $T : X \to X$ is a contraction on X. Then T has precisely one fixed point.*

Corollary 5.1.1. *Under the condition of the previous theorem. Let $x_m = T^m x$, and x be the unique fixed point of T. Then the error estimates are as follows:*

- *The **prior estimate** is given by*
$$d(x_m, x) \leq \frac{\alpha^m}{1-\alpha} d(x_0, x_1).$$
- *The **posterior estimate** is given by*
$$d(x_m, x) \leq \frac{\alpha}{1-\alpha} d(x_{m-1}, x_m).$$

Theorem 5.2 (Contraction on a ball)**.** *Let T be a mapping of a complete metric space $X = (X, d)$ into itself. Suppose T is a contraction with contraction constant α on a closed ball $Y = \{x \,|\, d(x, x_0) \leq r\}$ and*

$$d(x_0, Tx_0) < (1-\alpha)r.$$

Then there exists a unique fixed point $\tilde{x}$ of T in Y such that $T^m x \to \tilde{x}$ for all $x \in Y$.

Proof. Since $d(x_0, Tx_0) < (1-\alpha)r$, then $T^m x \in Y$ if $x \in Y$. Hence the proof follows. $\square$

Theorem 5.3 (Linear Equations)**.** *If $x = Cx + b$ with $C = (c_{jk})_{\substack{1 \leq j \leq n \\ 1 \leq k \leq m}}$ and $b \in \mathbb{C}^n$ is a system of n linear equation in n unknowns $x = (x_1, \cdots, x_n)$ satisfies*

$$\sum_{k=1}^{n} |c_{jk}| < 1 \quad \forall j.$$

Then it has precisely one solution x. This solution can be obtained as the limit of the iterative sequence $(x^0, x^1, \cdots)$, where x^0 is arbitrary and

$$x^{m+1} = Cx^m + b.$$

The error bounds are

$$d(x^m, x) \leq \frac{\alpha}{1-\alpha} d(x^{m-1}, x^m) \leq \frac{\alpha^m}{1-\alpha} d(x^0, x^1).$$

Theorem 5.4 (Picard's Existence and Uniqueness Theorem for ODEs)**.** *Let f be continuous on a rectangle*

$$R = \{(t, x) \,:\, |t - t_0| \leq a,\, |x - x_0| \leq b\}$$

with $|f(t, x)| \leq c$ for all $(t, x) \in \mathbb{R}$. Suppose that f is k-Lipschitz on R with respect to x. Then the initial value problem

$$x' = f(t, x), \quad x(t_0) = x_0$$

has a unique solution. This solution exists on an interval $[t_0 - \beta, t_0 + \beta]$, where

$$\beta < \min\left\{a, \frac{b}{c}, \frac{1}{k}\right\}.$$

Theorem 5.5 (Fredholm Integral Equation). *Suppose k and v in the integral equation*

$$x(t) - \mu \int_a^b k(t,\tau)x(\tau)d\tau = v(t)$$

are continuous on $J \times J$ and $J = [a,b]$ respectively. Assume μ is such that

$$|\mu| < \frac{1}{c(b-a)}$$

with c being any upper bound of $|k(t,\tau)|$ on $J \times J$. Then the integral equation has a unique solution x on J. This function x is the limit of the iterative sequence $(x_0, x_1, \cdots)$, where x_0 is any continuous function on J and

$$x_{n+1}(t) = v(t) + \mu \int_a^b k(t,\tau)x_n(\tau)d\tau.$$

Theorem 5.6 (Volterra Integral Equation). *Suppose k and v in the integral equation*

$$x(t) - \mu \int_a^t k(t,\tau)x(\tau)d\tau = v(t)$$

are continuous on $a \leq \tau \leq t$, $a \leq t \leq b$ and a $[a,b]$ respectively. Then the integral equation has a unique solution x on $[a,b]$ for every μ.

Lemma 5.7 (Fixed Point). *Let $T : X \to X$ be a continuous mapping on a complete metric space $X = (X,d)$, and suppose that T^m is a contraction on X for some positive integer m. Then T has a unique fixed point.*

Proof. By assumption $B = T^m$ is a contraction on X, that is $d(Bx, By) \leq \alpha d(x,y)$ for some $x, y \in X$, with $\alpha < 1$. Hence for every $x_0 \in X$,

$$\begin{aligned} d(B^n T x_0, B^n x_0) &\leq \alpha d(B^{n-1} T x_0, B^{n-1} x_0) \\ &\leq \alpha^n d(Tx_0, x_0) \to 0 \end{aligned}$$

as $n \to \infty$. Then by the Banach's Fixed point theorem, B has a unique fixed point, call it x, and $B^n x_0 \to x$. Since the mapping T is continuous, this implies $B^n T x_0 = T B^n x_0 \to Tx$. hence

$$d(B^n T x_0, B^n x_0) \to d(Tx, x),$$

so we must have $d(Tx, x) = 0$. Lastly since every fixed point of T is also a fixed point of B, then T cannot have more than one fixed point. □

5.2 Approximation in Normed Spaces

Definition 5.8 (Best Approximation). *Let $X = (X, \|\cdot\|)$ be a normed space and Y be a fixed subspace of X. We let*

$$\delta = \delta(x, Y) = \inf_{y \in Y} \|x - y\|.$$

If there exists a $y_0 \in Y$ such that

$$\|x - y_0\| = \delta$$

*then y_0 is called a **best approximation** to x out of Y.*

Theorem 5.9 (Existence Theorem of Best Approximations)**.** *If Y is a finite dimensional subspace of a normed space $X = (X, \|\cdot\|)$, then for each $x \in X$ there exists a best approximation to x out of Y.*

Remark 5.9.1. *The finite dimensionality of Y is essential.*

Proof. Let $x \in X$ be fixed. Consider the closed ball

$$\tilde{B} = \{y \in Y : \|y\| \leq 2\|x\|\}.$$

Then $0 \in \tilde{B}$, so that for the distance from x to $\tilde{B}$, we obtain

$$\delta(x, \tilde{B}) = \inf_{\tilde{y} \in \tilde{B}} \|x - \tilde{y}\| \leq \|x - 0\| = \|x\|.$$

Now if $y \notin \tilde{B}$, then $\|y\| > 2\|x\|$ and

$$\|x - y\| \geq \|y\| - \|x\| > \|x\| \geq \delta(x, \tilde{B}).$$

This shows that $\delta(x, \tilde{B}) = \delta(x, Y) = \delta$ and this value cannot be assumed by a $y \in Y - \tilde{B}$. Next by the compactness of $\tilde{B}$ and continuity of norm, we conclude that there is a $y_0 \in \tilde{B}$ such that $\|x - y\|$ assumes a minimum at $y = y_0$. $\square$

Example: We consider $X = C[a, b]$ and Y being the finite dimensional subspace of the space $C[a, b]$ which is spanned by $\{1, x, x^2, \cdots, x^n\}$. Then Theorem (5.9) tells us that for any given continuous function f on $[a, b]$ there exists a polynomial p_n of degree at most n such that for every $y \in Y$,

$$\max_{x \in [a,b]} |f(x) - p_n(x)| \leq \max_{x \in [a,b]} |f(x) - y(x)|.$$

Lemma 5.10. *In a normed space $(X, \|\cdot\|)$, the set M of best approximations to a given point x out of a subspace Y of X is convex.*

Proof. Let δ denote the distance from x to Y. Suppose M has more than one point, namely y and z. Then by definition

$$\|x - y\| = \|x - z\| = \delta.$$

Let

$$w = \alpha y + (1 - \alpha)z, \quad 0 \leq \alpha \leq 1$$

we show $w \in M$. Since $\|x - w\| \geq \delta$, and

$$\begin{aligned} \|x - w\| &= \|\alpha(x - y) + (1 - \alpha)(x - z)\| \\ &\leq \alpha\|x - y\| + (1 - \alpha)\|x - z\| \\ &= \alpha\delta + (1 - \alpha)\delta = \delta. \end{aligned}$$

Hence $\|x - w\| = \delta$, i.e., M is convex. $\square$

Definition 5.11 (Strict Convexity)**.** *A **strictly convex norm** is a norm such that for all x, y of norm 1, and $x \neq y$, then*

$$\|x + y\| < 2.$$

Lemma 5.12. *Any Hilbert space with the norm generated by the inner product is strictly convex. The space $C[a, b]$ with the sup-norm is not strictly convex.*

Proof. For all $x \neq y$ in a Hilbert space with norm one. Let $\alpha = \|x - y\| > 0$, then by the parallelogram law, we have

$$\|x + y\|^2 = -\|x - y\|^2 + 2(\|x\|^2 + \|y\|^2) = -\alpha^2 + 4 < 4.$$

So $\|x + y\| < 2$.

On the other hand, let

$$f(x) \equiv 1 \quad g(x) = \frac{x - a}{b - a}.$$

Then one can show $\|f\| = \|g\| = 1$ and $\|f + g\| = 2$. □

Theorem 5.13 (Uniqueness Theorem for Best Approximation)**.** *In a strictly convex normed space X there is at most one best approximation to an $x \in X$ out of a given subspace. In particular, for every given x in a Hilbert space H and every given closed subspace Y of H, the unique best approximation to x out of Y is given by the projection Px of x onto Y.*

5.3 Uniform Approximation

Definition 5.14 (Haar Condition)**.** *A finitely dimensional subspace Y of the real space $C[a, b]$ is said to satisfy the **Haar condition** if for every $y \in Y$, $y \neq 0$, y has at most $n - 1$ zeros in $[a, b]$, where $n = \dim Y$.*

Remark 5.14.1. *The Haar condition is equivalent to the condition that for every basis $\{y_1, \cdots, y_n\} \subset Y$ and every n-tuple of distinct points $t_1, \cdots, t_n$ in the interval $J = [a, b]$,*

$$\begin{vmatrix} y_1(t_1) & y_1(t_2) & \cdots & y_1(t_n) \\ y_2(t_1) & y_2(t_2) & \cdots & y_2(t_n) \\ & & \ddots & \\ y_n(t_1) & y_n(t_2) & \cdots & y_n(t_n) \end{vmatrix} \neq 0.$$

This is the case since the row space must be dimension n.

Lemma 5.15. *Suppose a subspace Y of the real space $C[a, b]$ satisfies the Haar condition. If for a given $x \in C[a, b]$ and a $y \in Y$ the function $x - y$ has less than $n + 1$ extremal points, then y is not a best approximation to x out of Y, where $n = \dim Y$.*

Theorem 5.16. *Let Y be a finite dimensional subspace of the real space $C[a, b]$. Then the best approximation out of Y is unique for every $x \in C[a, b]$ if and only if Y satisfies the Haar condition.*

Theorem 5.17. *The best approximation to an x in the real space $C[a, b]$ out of Y_n is unique; here Y_n is the subsapce consisting of $y = 0$ and all polynomials of degree not exceeding a fixed given n.*

Definition 5.18 (Alternating Set). *Let $x \in C[a,b]$ and $y \in Y$, where Y is a subspace of the real space $C[a,b]$. A set of points $t_0, \cdots, t_k$ in $[a,b]$ where $t_0 < t_1 < \cdots < t_k$ is called an **alternating set** for $x - y$ if $x(t_j) - y(t_j)$ has alternately the values $+\|x-y\|$ and $-\|x-y\|$ at consecutive points t_j.*

Lemma 5.19. *Let Y be a subspace of the real space $C[a,b]$ satisfying the Haar condition. Given $x \in C[a,b]$, let $y \in Y$ be such that for $x - y$ there exists an alternating set of $n+1$ points, where $n = \dim Y$. Then y is the best uniform approximation to x out of Y.*

6 Spectral Theory

6.1 Basic Concepts

Definition 6.1 (Spectrum For Matrices)**.** *For $A \in \mathbb{F}^{n\times n}$, the collection of eigenvalues of A is called its **spectrum**, i.e., it is the set of all λ's in $\mathbb{F}$ such that there exists $x \neq 0$ and*

$$Ax = \lambda x.$$

Recall that the characteristic polynomial of A is given by $\det(A - \lambda I)$ and recall that the eigenvalues of a matrix is precisely the roots of the characteristic polynomial. Similarly, we can define the eigenvalues of a linear operator T mapping from a finite dimensional space to a finite dimensional vector space using matrix representation. Moreover, the eigenvalues of T is independent on the choice of bases.

Now for finite dimensional case, we note $\det(T - \lambda I) = 0$ iff $T - \lambda I$ is not invertible iff $T - \lambda I$ is not injective iff $Tx = \lambda x$ for some $x \neq 0$. However, this chain of equivalence fails in the infinite dimensional case as there are noninvertible operators T defined on the infinite dimensional vector spaces such that T is still injective (e.g., the right shift operator).

We note for general vector spaces, $T - \lambda I$ can fail to be invertible in several ways:

1. $T - \lambda I$ is not injective;
2. $T - \lambda I$ is injective, but not surjective and its range is dense.
3. $T - \lambda I$ is injective, but not surjective and its range is not dense.

Definition 6.2 (Spectrum)**.** *Let T be a linear operator. Suppose $T - \lambda I$ is not injective, then λ is said to belong to the **point spectrum** and say λ is an **eigenvalue**, and any nonzero vector in the nullspace of $T - \lambda I$ is called an **eigenvector**. If $T - \lambda I$ is injective, but not surjective and its range is dense, then we say λ belongs to the **continuous spectrum**. If $T - \lambda I$ is injective, but not surjective and its range is not dense, then we say λ belongs to the **residual spectrum**. We denote $\sigma(T)$ to be the collection of all the **spectrums** and denote $\rho(T)$ to be the complement of $\sigma(T)$, which is called the **resolvent set**. Any element of $\rho(T)$ is called a **resolvent** of T, and we also denote $(T - \lambda I)^{-1} = R_\lambda(T) = R_\lambda$.*

Remark 6.2.1. *We have $\rho(T) \sqcup \sigma_p(T) \sqcup \sigma_c(T) \sqcup \sigma_r(T) = \mathbb{C}$.*

Example: We consider the left shift operator L acting on ℓ^2. We calculates its point spectrum, we need to find $(L - \lambda I)x = 0$, then

$$(x_2, x_3, x_4, \cdots) - (\lambda x_1, \lambda x_2, \lambda x_3, \cdots) = 0$$

Then we note $x_{k+1} - \lambda x_k = 0$, that is the sequence x is given by

$$(1, \lambda, \lambda^2, \lambda^3, \cdots)$$

We note $x \in \ell^2$ if and only if $|\lambda| < 1$. Later we will show that the spectrum is always closed and contained in the radius centered at 0 with radius $\|T\|$. Hence this tells us that $\sigma(L)$ is the closed unit disc in $\mathbb{C}$. It turns out the continuous spectrum of L is the unit circle in $\mathbb{C}$, and the residual spectrum of L is the empty set.

Example: We can also consider the right shift operator R acting on ℓ^2. We note that the only possible x such that $(R-\lambda I)x=0$ is when $x=0$. Hence R has no point spectrum.

Example: Let us consider the space $L^2[0,1]$ and the map $m : L^2[0,1] \to L^2[0,1]$ by $mf(x) = xf(x)$. Suppose $xf(x) = mf(x) = \lambda f(x)$, then $f=0$ a.e., This shows m has no eigenvalues. Now if $\lambda \notin [0,1]$, then $(x-\lambda)^{-1}f(x) \in L^2$, so $m-\lambda I$ is invertible. If $\lambda \in [0,1]$, note that $c(x-\lambda)^{-1}$ is not in $L^2[0,1]$, so $m-\lambda I$ is not surjective, but the range is dense. Hence $[0,1]$ is the continuous spectrum.

Example: Let $X = C[0,1]$ and define $T : X \to X$ by $Tx = vx$, where $v \in X$ is fixed. Then we note that $\sigma(T) = v([0,1])$. Hence we can construct a linear operator T on X whsoe spectrum is the given interval $[a,b]$.

Lemma 6.3. *If Y is the eigenspace corresponding to an eigenvalue λ of an operator T, then the spectrum of $T|_Y$ is just $\{\lambda\}$.*

Proof. We note $T|_Y = \lambda I$. Hence $T - \lambda_0 I$ is invertible if and only if $\lambda_0 \neq 0$. □

We also recall the following proposition from linear algebra.

Proposition 6.4 (Linear Independence of Eigenvectors)**.** *Eigenvectors $x_1, \cdots, x_n$ corresponding to different eigenvalues $\lambda_1, \cdots, \lambda_n$ of a linear operator T on a vector space X constitute a linear independent set.*

6.2 Spectral Properties of Bounded Linear Operators

Theorem 6.5. *Let $T \in B(X,X)$, where X is a Banach space. If $\|T\| < 1$, then $(1-T)^{-1}$ exists and is bounded, moreover*

$$(1-T)^{-1} = \sum_{j=0}^{\infty} T^j = I + T + T^2 + \cdots.$$

Proof. We have $\|T^j\| \leq \|T\|^j$. So the series converges absolutely when $\|T\| < 1$. Since X is complete, so is $B(X,X)$. Then absolutely convergence implies convergence of the series. Let $S = \sum_{j=0}^{\infty} T^j$, we show $S = (I-T)^{-1}$.

$$(I-T)(1+T+\cdots+T^m) = I - T^{n+1}.$$

Let $n \to \infty$, then $T^{n+1} \to 0$ as $\|T\| < 1$. Thus we have

$$(I-T)S = S(I-T) = I.$$

□

Theorem 6.6 (Spectrum Closed)**.** *The resolvent set $\rho(T)$ of a bounded linear operator T on a complex Banach space X is open, hence the spectrum $\sigma(T)$ is closed.*

Proof. If $\rho(T) = \emptyset$, then it is open (later we will see this cannot actually happen). Let $\rho(T) \neq \emptyset$. For a fixed $\lambda_0 \in \rho(T)$ and any $\lambda \in \mathbb{C}$, we have

$$\begin{aligned} T - \lambda I &= T - \lambda_0 I - (\lambda - \lambda_0) I \\ &= (T - \lambda_0 I) \underbrace{\left[I - (\lambda - \lambda_0)(T - \lambda_0 I)^{-1} \right]}_{V}. \end{aligned}$$

Then $T_\lambda = T_{\lambda_0} V$. Since $\lambda_0 \in \rho(T)$ and T is bounded. Then $R_{\lambda_0} = T_{\lambda_0}^{-1} \in B(X, X)$ by the bounded inverse theorem. Furthermore, theorem (6.6) shows that V has a bounded inverse if

$$\|(\lambda - \lambda_0) R_{\lambda_0}\| < 1$$

that is, when

$$|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|}.$$

In this case, we must have

$$R_\lambda = T_\lambda^{-1} = (T_{\lambda_0} V)^{-1} = V^{-1} R_{\lambda_0}$$

we note R_λ is also bounded. Hence $\rho(T)$ is open. □

Corollary 6.6.1. *Let T be a bounded linear operator on a complex Banach space X. Then for every $\lambda \in \rho(T)$, the resolvent $R_\lambda(T)$ has the representation*

$$R_\lambda = \sum_{j=0}^{\infty} (\lambda - \lambda_0)^j R_{\lambda_0}^{j+1},$$

The series being absolutely convergent for every λ in the open disk given by

$$|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|}$$

in the complex plane. This disk is a subset of $\rho(T)$.

Theorem 6.7. *The spectrum $\sigma(T)$ of a bounded linear operator $T : X \to X$ on a complex Banach space X is compact and lies in the disk given by*

$$|\lambda| \leq \|T\|.$$

Hence the resolvent set $\rho(T)$ of T is not empty.

Proof. Let $\lambda \neq 0$ and $\kappa = \frac{1}{\lambda}$. From Theorem (6.6), we have

$$R_\lambda = (T - \lambda I)^{-1} = -\frac{1}{\lambda}(I - \kappa T)^{-1} = -\frac{1}{\lambda} \sum_{j=0}^{\infty} (\kappa T)^j = -\frac{1}{\lambda} \sum_{j=0}^{\infty} \left(\frac{1}{\lambda} T \right)^j.$$

The series converges for all λ such that

$$\left\| \frac{1}{\lambda} T \right\| = \frac{\|T\|}{|\lambda|} < 1 \Longrightarrow |\lambda| > \|T\|.$$

Then any such λ must be is in $\rho(T)$. Hence $\sigma(T) = \mathbb{C} \setminus \rho(T)$ must lie in the disk $|\lambda| \leq \|T\|$. By Theorem (6.7), we know $\sigma(T)$ is closed, hence compact. □

Corollary 6.7.1. *If $T \in B(X,X)$, then $\|R_\lambda(T)\| \to 0$ as $|\lambda| \to \infty$.*

Definition 6.8 (Spectral Radius)**.** *The **spectral radius** $r_\sigma(T)$ of an operator $T \in B(X,X)$ on a complex Banach space X is the radius*

$$r_\sigma(T) = \sup_{\lambda \in \sigma(T)} |\lambda|.$$

Remark 6.8.1. *By the previous theorem, we have that $\rho_\sigma(T) \leq \|T\|$.*

Theorem 6.9. *Let X be a complex Banach space, $T \in B(X,X)$ and $\lambda, \mu \in \rho(T)$. Then:*

1. *The resolvent R_λ of T satisfies the Hilbert relation or resolvent equation*

$$R_\mu - R_\lambda = (\mu - \lambda) R_\mu R_\lambda \tag{6.1}$$

   for $\lambda, \mu \in \rho(T)$.

2. *R_λ commutes with any $S \in B(X,X)$ which commutes with T.*

3. *We have*

$$R_\lambda R_\mu = R_\mu R_\lambda$$

   whenever $\lambda, \mu \in \rho(T)$.

Proof. (3) follows directly from (1). For (1), since we have $I = T_\lambda R_\lambda$ where $T_\lambda = T - \lambda I$ and $I = R_\mu T_\mu$, then

$$\begin{aligned} R_\mu - R_\lambda &= R_\mu(T_\lambda R_\lambda) - (R_\mu T_\mu) R_\lambda \\ &= R_\mu(T_\lambda - T_\mu) R_\lambda \\ &= (\mu - \lambda) R_\mu R_\lambda. \end{aligned}$$

For (2), by assumption, if $ST = TS$, then $ST_\lambda = T_\lambda S$. Using the fact that $I = T_\lambda R_\lambda = R_\lambda T_\lambda$, then

$$R_\lambda S = R_\lambda S T_\lambda R_\lambda = R_\lambda T_\lambda S R_\lambda = S R_\lambda.$$

□

Theorem 6.10 (Spectral Mapping Theorem for Polynomials)**.** *Let X be a complex Banach space, $T \in B(X,X)$ and*

$$f(\lambda) = \alpha_n \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_0, \quad (\alpha_n \neq 0).$$

Then

$$\sigma(f(T)) = f(\sigma(T)).$$

Proof. We assume that $\sigma(T) \neq \emptyset$ (this fact will be proved later). The case $n = 0$ is trivial; then $f(\sigma(T)) = \{\alpha_0\} = \sigma(f(T))$. Let $n > 0$, we first prove

$$\sigma(f(T)) \subset f(\sigma(T)).$$

For simplicity we write $S = f(T)$ and $S_\mu = f(T) - \mu I$, $\mu \in \mathbb{C}$.

If S_μ^{-1} exists, then the formula for S_μ shows that S_μ^{-1} is the resolvent operator of $f(T)$. We keep μ fixed. Since X is complex, the polynomial given by $s_\mu(\lambda) = f(\lambda) - \mu$ must factor completely into linear terms, say

$$s_\mu(\lambda) = f(\lambda) - \mu = a_n(\lambda - \gamma_1)(\lambda - \gamma_2)\cdots(\lambda - \gamma_n),$$

where $\gamma_1, \cdots, \gamma_n$ are the zeros of s_μ. So we have

$$S_\mu = \alpha_n(T - \gamma_1 I)(T - \gamma_2 I)\cdots(T - \gamma_n I).$$

If each γ_j is in $\rho(T)$, then each $T - \gamma_j I$ has a well-defined inverse, so S_μ is invertible with

$$S_\mu^{-1} = \frac{1}{\alpha_n}(T - \gamma_n I)^{-1}\cdots(T - \gamma_1 I)^{-1}.$$

hence we have $\mu \in f(\rho(T))$. This proves that if $\mu \in \sigma(f(T))$, then $\gamma_j \in \sigma(T)$ for some j, i.e., $\sigma(f(T)) \subset f(\sigma(T))$.

Next we prove $f(\sigma(T)) \subset \sigma(f(T))$. We do this by showing that

$$\kappa \in f(\sigma(T)) \Longrightarrow \kappa \in \sigma(f(T)).$$

Let $\kappa \in p(\sigma(T))$, then $\kappa = p(\beta)$ for some $\beta \in \sigma(T))$. There are two possibilities

1. $T - \beta I$ is not injective;
2. $T - \beta I$ is injective.

For the first case, we have $p(\beta) - \kappa = 0$. Hence β is a zero of the polynomial given by

$$s_\kappa(\lambda) = \rho(\lambda) - \kappa.$$

It follows that we can write

$$s_k(\lambda) = f(\lambda) - k = (\lambda - \beta)g(\lambda),$$

where $g(\lambda)$ denotes the product of the other $n-1$ linear factors and α_n. Corresponding to this representation, we have

$$S_\kappa = f(T) - \kappa I = (T - \beta I)g(T).$$

Since the factors of $g(T)$ all commutes with $T - \beta I$, we also have

$$S_\kappa = g(T)(T - \beta I)$$

Then the noninjectivity of $T_\beta I$ shows S_k is not injective. So $\kappa \in \sigma(f(T))$.

Now suppose $T - \beta I$ is injective, then the range of $T - \beta I$ must not be the entirety of X. So we have $\mathscr{R}(S_\kappa) \neq X$. This shows $\kappa \in \sigma(f(T))$. □

Corollary 6.10.1. *For any operator $T \in B(X, X)$ on a complex Banach space X,*

$$r_\sigma(\alpha T) = |\alpha| r\sigma(T), \quad \textit{and} \quad r_\sigma(T^k) = [r_\sigma(T)]^k$$

for any $k \in \mathbb{N}$.

6.3 Use of Complex Analysis in Spectral Theory

Definition 6.11 (Locally Holomorphy, Holomorphy)**.** *Let Λ be an open subset of $\mathbb{C}$ and X a Banach space. A vector valued function is of the form*

$$S : \Lambda \to B(X, X), \quad \lambda \mapsto S_\lambda.$$

*Then S is said to be **locally holomorphic** on Λ if for every $x \in X$ and $f \in X'$ the function h defined by*

$$h(\lambda) = f(S_\lambda x)$$

*is holomorphic at every $\lambda_0 \in \Lambda$. S is said to be **holomorphic** on Λ if S is locally holomorphic on Λ and Λ is a domain. S is said to be **holomorphic** at a point $\lambda_0 \in \mathbb{C}$ if S is holomorphic on some ϵ-neighbourhood of λ_0.*

Theorem 6.12 (Holomorphy of R_λ)**.** *The resolvent $R_\lambda(T)$ of a bounded linear operator $T : X \to X$ on a complex Banach space X is holomorphic at every point λ_0 of the resolvent set $\rho(T)$ of T. Hence it is locally holomorphic on $\rho(T)$.*

Proof. Recall that for every value $\lambda_0 \in \rho(T)$, by Corollary (6.7.1), we have

$$R_\lambda(T) = \sum_{j=0}^{\infty} R_{\lambda_0}(T)^{j+1}(\lambda - \lambda_0)^j$$

Which is absolutely convergent for λ sufficiently close to λ_0. Taking any $x \in X$ and $f \in X'$, and define h by

$$h(\lambda) = f(R_\lambda(T)(x)),$$

then we have

$$h(\lambda) = \sum_{j=0}^{\infty} c_j(\lambda - \lambda_0)^j$$

where

$$c_j = f(R_{\lambda_0}(T)^{j+1}x).$$

which is absolutely convergent when λ is close to λ_0. □

Theorem 6.13. *If $T \in B(X, X)$, where X is a complex Banach space, and $\lambda \in \rho(T)$, then*

$$\|R_\lambda(T)\| \geq \frac{1}{\delta(\lambda)}$$

where

$$\delta(\lambda) = \inf_{s \in \sigma(T)} |\lambda - s|$$

is the distance from λ to the spectrum $\sigma(T)$. Hence

$$\|R_\lambda(T)\| \to \infty \quad as \quad \delta(\lambda) \to 0.$$

Proof. Assume again $\sigma(T) \neq \emptyset$. For every $\lambda_0 \in \rho(T)$, we see that the distance from λ_0 to the spectrum must at least equal to

$$\frac{1}{\|R_{\lambda_0}\|}$$

since within this disk, we have an absolutely convergent power expansion for R_λ based at λ_0, This implies

$$\delta(\lambda_0) \geq \frac{1}{\|R_{\lambda_0}\|}.$$

□

Theorem 6.14 (Spectrum Non-empty)**.** *If $X \neq \{0\}$ is a complex Banach space and $T \in B(X,X)$, then $\sigma(T) \neq \emptyset$.*

Proof. If $T = 0$, then $\sigma(T) = \{0\} \neq \emptyset$. Let $T \neq 0$, so $\|T\| \neq 0$. Then we have

$$R_\lambda = -\frac{1}{\lambda}\sum_{j=0}^{\infty}\left(\frac{1}{\lambda}T\right)^j$$

whenever $\|\lambda\| \geq \|T\|$. Then

$$\|R_\lambda\| \leq \frac{1}{|\lambda|}\sum_{j=0}^{\infty}\left|\frac{1}{\lambda}T\right\|^j = \frac{1}{|\lambda| - \|T\|} \leq \frac{1}{\|T\|}$$

whenever $|\lambda| \geq 2\|T\|$. We show that the assumption $\sigma(T) = \emptyset$ leads to a contradiction. Suppose $\rho(T) = \mathbb{C}$, then R_λ is holomorphic. So if we fix any $x \in X$ and $f \in X'$, the function

$$h(\lambda) = f(R_\lambda x)$$

is holomorphic on $\mathbb{C}$. By then

$$|h(\lambda)| \leq \|f\|\|R_\lambda x\| \leq \|f\| \cdot \frac{1}{\|T\|}\|x\|$$

when $|\lambda| \geq 2\|T\|$. So h is bounded, hence is a constant by Liouville's Theorem. But then this shows for arbitrary $x \in X$ and $f \in X'$, h defined in this way is a constant function, hence R_λ must be independent of λ which is a contradiction. □

Theorem 6.15 (Spectral Radius)**.** *If T is a bounded linear operator on a complex Banach space, then for the spectral radius $r_\sigma(T)$ of T, we have*

$$r_\sigma(T) = \lim_{n\to\infty}\sqrt[n]{\|T\|^n}.$$

Remark 6.15.1. *The sequence $\{\sqrt[n]{\|T^n\|}\}_{n=1}^{\infty}$ need not be monotone.*

Proof. We have $\sigma(T^n) = [\sigma(T)]^n$, so

$$r_\sigma(T^n) = [r_\sigma(T)]^n.$$

We also know that

$$r_\sigma(T^n) \leq \|T^n\|.$$

So for all $n \in \mathbb{N}$,

$$r_\sigma(T) = \sqrt[n]{r_\sigma(T)^n} \leq \sqrt[n]{\|T^n\|}.$$

Hence

$$r_\sigma(T) \leq \liminf \sqrt[n]{\|T^n\|} \leq \limsup \sqrt[n]{\|T^n\|}.$$

We show the last expression equals to $r_\sigma(T)$. Let $\kappa = \frac{1}{\lambda}$, then we can write

$$R_\lambda = -\kappa \sum_{n=0}^{\infty} T^n \kappa^n$$

whenever $|\kappa| \leq r$ where r is the radius of absolutely convergence of this series. But then we know

$$r = \frac{1}{\limsup \sqrt[n]{\|T^n\|}}.$$

Hence this proves that whenever $|\lambda| \geq 3\frac{1}{r}$, R_λ is well-defined. In particular, this proves

$$r_\sigma(T) = \limsup \sqrt[n]{\|T^n\|}$$

□

Corollary 6.15.1. *Suppose T is nilpotent defined on a complex Banach space $X \neq \{0\}$, that is $T^m = 0$ for some positive integer m, then the spectrum of T is $\{0\}$.*

Corollary 6.15.2. *If X is a complex Banach space, $S, T \in B(X, X)$ and $ST = TS$, then*

$$r_\sigma(ST) \leq r_\sigma(S) r_\sigma(T).$$

Remark 6.15.2. *The condition $ST = TS$ is necessarily. Otherwise take two nilpotent matrix whose product is a non-nilpotent matrix.*

Corollary 6.15.3. *If T is a normal operator on a Hilbert space H, then $r_\sigma(T) = \|T\|$.*

Proof. Since T is normal, then $\|TT^*\| = \|T\|^2$ and T, T^* commutes. Hence

$$r_\sigma(TT^*) \leq r_\sigma(T) r_\sigma(T^*)$$

and

$$r_\sigma(TT^*) = \lim \sqrt[n]{\|TT^*\|} = \|T\|^2.$$

Now since $r_\sigma(T) \leq \|T\|$ and $r_\sigma(T^*) \leq \|T^*\| = \|T\|$. This shows that $r_\sigma(T) = \|T\|$. □

6.4 Banach Algebra

Definition 6.16 (Banach Algebra)**.** *An **algebra** A over a field K is a vector space A over K such that for each ordered pair of elements $x, y \in A$ a unique product $xy \in A$ is defined with the properties*

$$(xy)z = x(yz)$$
$$x(y+z) = xy + xz$$

$$(x+y)z = xz + yz$$
$$\alpha(xy) = (\alpha x)y = x(\alpha y)$$

for all $x, y, z \in A$ and scalars α.

If $K = \mathbb{R}$ or $\mathbb{C}$, then A is said to be real or complex, respectively.

A is said to be ***commutative*** *(or abelian) if the multiplication is commutative, that is, if for all $x, y \in A$,*

$$xy = yx.$$

A is called an algebra with ***identity*** *if A contains an element e such that for all $x \in A$,*

$$ex = xe = x.$$

Definition 6.17 (Normed Algebra, Banach Algebra)**.** *A normed algebra A is a normed space which is an algebra such that for all $x, y \in A$*

$$\|xy\| \leq \|x\|\|y\|$$

and if A has an identity e, with $\|e\| = 1$. A ***Banach algebra*** *is a normed algebra which is complete.*

Remark 6.17.1. *The fact that $\|xy\| \leq \|x\|\|y\|$ makes multiplication on A to be continuous respect to both factors.*

Examples:

1. The spaces $\mathbb{R}$ and $\mathbb{C}$ are naturally Banach algebras.
2. The $C[a, b]$ with the pointwise multiplication is a Banach algebra with identity ($e \equiv 1$).
3. The vector space X of all complex $n \times n$ matrices is a noncommutative algebra with identity I. We can endow a norm on this space to make it a Banach Algebra.
4. The space $B(X, X)$ on a complex Banach space $X \neq \{0\}$ is a Banach algebra with identity I. The multiplication is the composition of operators.

Definition 6.18 (Inverse)**.** *Let A be an algebra with identity. An $x \in A$ is said to be* ***invertible*** *if it has an* ***inverse*** *in A, that is, there exists x^{-1} such that*

$$x^{-1}x = xx^{-1} = e.$$

Definition 6.19 (Resolvent Set, Spectrum)**.** *Let A be a complex Banach algebra with identity. Then the* ***resolvent set*** *$\rho(x)$ of an $x \in A$ is the set of all λ in the complex plane such that $x - \lambda e$ is invertible. The spectrum is the complement of $\rho(x)$ in $\mathbb{C}$. Any $\lambda \in \sigma(x)$ is called a* ***spectral value*** *of x.*

Remark 6.19.1. *If $A = B(X, X)$, then this definition agrees with the notion of Resolvent and Spectrum of linear operators.*

Lemma 6.20 (Inverse)**.** *Let A be a complex Banach Algebra with identity e. If $x \in A$ satisfies $\|x\| < 1$, then $e - x$ is invertible, and*

$$(e-x)^{-1} = e + \sum_{j=1}^{\infty} x^j.$$

Lemma 6.21 (Invertible Elements)**.** *Let A be a complex Banach algebra with identity. Then the set G of all invertible elements of A is an open subset of A; hence the subset $M = A - G$ of all non-invertible elements of A is closed.*

Lemma 6.22. *Let A be a complex Banach algebra with identity e. Then for any $x \in A$, the spectrum $\sigma(x)$ is compact and the spectral radius satisfies*

$$r_\sigma(x) \le \|x\|.$$

Moreover,

$$\sigma(x) \neq \emptyset.$$

Definition 6.23. *Suppose A has no identity. Then we can always supply A with an identity in the following canonical fashion. Let $\tilde{A}$ be the set of all ordered pairs (x, α), where $x \in A$ and α is a scalar. Define*

$$\begin{aligned} (x,\alpha) + (y,\beta) &= (x+y, \alpha+\beta) \\ \beta(x,\alpha) &= (\beta x, \beta\alpha) \\ (x,\alpha)(y,\beta) &= (xy + \alpha y + \beta x, \alpha\beta) \\ \|(x,\alpha)\| &= \|x\| + |\alpha| \\ \tilde{e} &= (0,1). \end{aligned}$$

Then $\tilde{A}$ is a Banach algebra with identity $\tilde{e}$ if A is a Banach algebra. Furthermore, the mapping $x \mapsto (x,0)$ is an isomorphism of A onto a subspace of $\tilde{A}$ both regarded as normed spaces. The subspace has codimension 1. If we identity x with $(x,0)$, then $\tilde{A}$ is simply A plus the one-dimensional vector space generated by $\tilde{e}$.

6.5 Spectrum of The Adjoint Operator

Proposition 6.24. *Let X and Y be Banach spaces and $T \in B(X,Y)$, then T^* (same as $T^\times$) is injective if and only if T has dense range.*

Proof. T^* injective implies T has dense range: we prove the contrapositive of this statement. Suppose the range of T is not dense. Then let $y \in Y \setminus \overline{T(x)}$, then it is not hard to see that $M = \operatorname{span}\{y, \overline{T(x)}\} = \operatorname{span}\{y\} \oplus \overline{T(X)}$. We define a linear map $g : M \subset Y \to \mathbb{F}$ using universal property of linear maps. We first pick a basis $\{\alpha_i\}_{i\in I}$ of $\overline{T(x)}$, which must be linearly independent with y, then we set $g(y) = 1$ and $g(\alpha_i) = 0$ for any $i \in I$. By universal property this gives a linear functional defined on M, in particular, by this construction, we have $g(T(X)) = \{0\}$. We show that g is a bounded linear functional. Let $z \in M$ be arbitrary, then there exists unique $x \in \overline{T(X)}$ and $\alpha \in \mathbb{F}$ such that $z = \alpha y + x$. If $\alpha = 0$, then $g(z) = g(x) = 0$. Now suppose $\alpha \neq 0$. Then since $y \notin \overline{T(X)}$, let

$$d = \inf\{\|y - x\| \,:\, x \in \overline{T(X)}\}.$$

We must have $d > 0$, as otherwise $y \in \overline{T(X)}$ gives a contradiction. Then we have

$$d \le \left\| y - \left(-\frac{x}{|\alpha|}\right) \right\| = \frac{1}{|\alpha|} \left\| y + \frac{x}{\alpha} \right| .$$

Hence

$$|g(\alpha y + x)| = |\alpha| \le \frac{1}{d} \left\| y + \frac{x}{\alpha} \right| .$$

This in particular shows that g is bounded with norm at most $\frac{1}{d}$. Hence by the Hahn-Banach Theorem, we can extend g to a bounded linear function $\tilde{g} : Y \to \mathbb{F}$, in particular $\tilde{g} \in Y'$. We show that $\tilde{g}$ is in the nullspace of T^*. By construction, for any $x \in X$, we have

$$T^*(\tilde{g})(x) = \tilde{g}(T(x)) = g(T(x)) = 0$$

So $T^*(\tilde{g})$ is the zero map. Since $\tilde{g}$ is obviously not trivial. This shows that T^* is not injective. By contrapositive, we have that if T^* is injective, then T must have dense range.

T has dense range implies T^* is injective: suppose T has a dense range. Let $g \in Y'$ be in the kernel of T^*, we show that g must be trivial. Let $y \in Y$ be arbitrary, we show $g(y) = 0$ By assumption, $T^*(g) : X \to \mathbb{F}$ is the zero map, hence for all $x \in X$, we must have

$$(T^*(g))(x) = g(T(x)) = 0.$$

Since the range of T is dense in Y, then we can find a sequence $\{y_n\}_{n=1}^\infty \subset Y$ and $\{x_n\}_{n=1}^\infty \subset X$ such that $y_n = T(x_n)$ and $y_n \to y$ in Y as $n \to \infty$. Since g is bounded, it is sequential continuous, thus we have

$$g(y) = g(\lim_{n\to\infty} y_n) = g(\lim_{n\to\infty} T(x_n)) = \lim_{n\to\infty} g(T(x_n)) = \lim_{n\to\infty} 0 = 0.$$

Thus we conclude that g is the zero map, so T^* is injective. $\square$

Lemma 6.25. *A bounded linear operator $T : X \to Y$ between Banach spaces is bounded below if and only if T is one-to-one and has a closed range.*

Proof. Let $T : X \to Y$ be bounded. Suppose $\|Tx\| \geq \gamma\|x\|$ holds for all $x \in X$ and some $\gamma > 0$, then one can easily show that it is injective and have closed range. For the converse, one make use of the closed graph theorem. $\square$

Theorem 6.26 (Banach)**.** *A bounded operator $T : X \to Y$ between Banach spaces has closed range if and only if its adjoints operator $T^* : Y^* \to X^*$ has closed range.*

Proof. Suppose $\mathscr{R}(T)$ is closed, since the operator $T : X \to \mathscr{R}(T)$ is a surjective operator between Banach spaces, there exists some constant $c >)$ such that $cB_R \subset T(B_X)$ where B_R and B_X denote the open unit ball of $\mathscr{R}(T)$ and X respectively.

Let $\{x_n^*\} \subset \mathscr{R}(T^*)$ satisfying $x_n^* \to 0$. Pick a sequence $\{y_n^*\} \subset Y^*$ satisfying $x_n^* = T^*y_n^*$ for each n and let $y_n^*|_R$ be the restriction of y_n^* to $\mathscr{R}(T)$. From

$$\begin{aligned}
\|x_n^*\| &= \sup_{x\in B_X} |x_n^*(x)| \\
&= \sup_{x\in B_x} |(T^*y_n^*)(x)| \\
&= \sup_{x\in B_x} |y_n^*(Tx)| \\
&\geq c \sup_{y\in B_R} |y_n^*(y)| \\
&= c\|y_n^*|_R\|
\end{aligned}$$

and $\|x_n^*\| \to 0$, it follows that $\|y_n^*|_R\| \to 0$. Let z_n^* be a linear extension of $y_n^*|_R$ to all of Y such that $\|z_n^*\| = \|y_n^*|_R\|$. Then $T^*z_n^* = T^*y_n^*$ holds for all n and $\|z_n^*\| \to 0$. This shows the range of the operator T^* is closed.

For the converse, assume that $\mathscr{R}(T^*)$ is closed. Put $Z = \overline{\mathscr{R}(T)}$ and consider the operator $S : X \to Z$ defined by $Sx = Tx$ for each $x \in X$. Let $y^* \in Z^*$ be an arbitrary linear extension of z^* to all of Y. THen we have $S^*z^*(x) = x^*(Tx) = y^*(Tx) = T^*y^*(x)$ for all $x \in X$, i.e., $S^*z^* = T^*y^*$. This implies that $\mathscr{R}(S^*) = \mathscr{R}(T^*)$. In other words, replacing T by S if necessary, we can assume that the range of T is also dense. Under this assumption, we show $\mathscr{R}(T) = Y$.

Since $\overline{\mathscr{R}(T)} = Y$, then T^* is injective by Proposition (6.25). Since $\mathscr{R}(T^*)$ is assumed to be closed, by Lemma (6.26) there exists some constant $c > 0$ satisfying $c\|y^*\| \leq \|T^*y^*\|$ for all $y^* \in Y^*$. Next, we claim that we have $\{y \in Y : \|y\| < c\} \subset \overline{T(B_X)}$.

Let $y \in Y$ be such that $\|y\| < c$ and assume that $y \notin \overline{T(B_X)}$. Then there exists some $y^* \in Y^*$ by Lemma (4.17) such that

$$y^*(y) > \sup\{|y^*(Tx)| : x \in B_x\} = \|T^*y^*\| \geq c\|y^*\|.$$

Since $y^*(y) \leq \|y^*\| \cdot \|y\|$, we must get $\|y\| > c$ which is a contradiction. So $\{y \in Y : \|y\| < c\} \subset \overline{T(B_X)}$. Then If this is the case, then we get

$$\{y \in Y : \|y\| < \frac{c}{2}\} \subset T(B_X)$$

by mimicking the proof of the Open Mapping Theorem. This shows that $\mathscr{R}(T) = Y$. □

Proposition 6.27. *Suppose $T \in B(X, Y)$ where X and Y are Banach spaces. Then T is invertible if and only if T^* is invertible.*

Proof. We know one direction, that is if T is invertible then T^* is invertible with inverse $(T^{-1})^*$. Now suppose T^* is invertible, then T^* must have closed range, which is X^*, hence T has closed range by Theorem (6.27). Since T^* is injective, then T has dense range. Thus T is surjective. Now suppose T is not injective, then one can easily show that T^* cannot be surjective. Hence we establish that T must be bijective, hence by the bounded inverse theorem be invertible. □

Corollary 6.27.1. *Suppose $T \in B(X)$ where X is a Banach space. Then $\sigma(T) = \sigma(T^*)$.*

Proof. Notice that $(T - \lambda I)^* = T^* - \lambda I$. □

7 Compact Operators

7.1 Compact Linear Operator on Normed Spaces

Definition 7.1 (Compact Linear Operator). *Let X and Y be a normed spaces. An operator $T : X \to Y$ is called a **compact linear operator** if T is linear and if for every bounded subset M of X, the image $T(M)$ is relatively compact, that is the closure $\overline{T(M)}$ is compact.*

Remark 7.1.1. *In fact by the nature of linear operators, T is compact if and only if the image $T(M)$ of the unit ball $M \subset X$ is relatively compact in Y.*

Lemma 7.2. *Let X and Y be normed spaces, then*

1. *Every compact linear operator $T : X \to Y$ is bounded, hence continuous.*
2. *If $\dim X = \infty$, the identity operator $I : X \to X$ (which is continuous) is not compact.*

Proof. The unit sphere $B(0,1) \subset X$ is bounded. Since T is compact, $\overline{T(U)}$ is bounded, hence

$$\sup_{\|x\|=1} \|Tx\| < \infty$$

which shows that T is bounded.

Now if $\dim X = \infty$, then we know that the closed unit ball is bounded not not compact. This shows that I is not compact. □

It is clear that by parsing to the sequential limits, then we have the following equivalent definition of an operator being compact:

Theorem 7.3 (Compactness Criterion). *Let X and Y be normed spaces and $T : X \to Y$ a linear operator.Then T is compact if and only if it maps every bounded sequence (x_n) in X onto a sequence (Tx_n) in Y which has a convergent subsequence.*

Corollary 7.3.1. *The compact linear operators from X into Y form a vector subspace $C(X, Y)$ of $B(X, Y)$.*

The following theorem is also clear:

Theorem 7.4. *Let X and Y be normed spaces and $T : X \to Y$ be a linear operator. Then:*

- *if T is bounded and $\dim T(X) < \infty$, the operator T is compact.*
- *if $\dim X < \infty$, the operator T is compact.*

Theorem 7.5 (Sequence of compact linear Operators). *Let (T_n) be a sequence of compact linear operators from a normed space X into a Banach space Y. If (T_n) is uniformly operator convergent, that is $\|T_n - T\| \to 0$, then the limit operator T is compact.*

Proof. We make use of sequential criterion. Let (x_m) be a bounded sequence in X, then by inductively extract subsequence $(T_n x_{n,m})$ such that $(T_n x_{n,m})$ is a convergent sequence with $(x_{n,m})$ being a subsequence of $(x_{n-1,m})$. Then taking the diagonal sequence, $(x_{m,m})_{m\in\mathbb{N}}$, one can show by the uniform convergence of T_n, we have $(Tx_{m,m})$ is Cauchy hence convergent. □

Remark 7.5.1. *The theorem is false if we replace uniform operator convergence by strong operator convergence. We provide the following counter-example: consider $T_n : \ell^2 \to \ell^2$ by $T_n(x) = (\xi_1, \cdots, \xi_n, 0, 0, \cdots)$, where $x = (\xi_j) \in \ell^2$. Since T_n is linear and bounded, T_n is compact as its range is finite dimensional. Clearly $T_n x \to x = Id(x)$, but Id is not compact since $\dim \ell^2 = \infty$.*

Corollary 7.5.1. *If Y is a Banach space, then $C(X, Y)$ is a Banach subspace of $B(X, Y)$.*

Theorem 7.6. *Let X and Y be normed spaces and $T : X \to Y$ be a compact linear operator. Suppose that (x_n) in X is weakly convergent, say $x_n \rightharpoonup x$. Then (Tx_n) is strongly convergent in Y and has the limit $y = Tx$.*

Proof. We write $y_n = Tx_n$ and $y = Tx$. We first show that

$$y_n \rightharpoonup y.$$

Let g be any bounded linear functional on Y, we define a functional f on X by setting

$$f(z) = g(Tz) \quad \forall z \in X.$$

f is linear and bounded, and

$$|f(z)| = |g(Tz)| \leq \|g\|\|Tz\| \leq \|g\|\|T\|\|z\|.$$

Since $x_n \rightharpoonup x$, then $f(x_n) \to f(x)$, so $g(Tx_n) \to g(Tx)$, that is $g(y_n) \to g(y)$. Since g is arbitrary, then y_n converges to y weakly.

Now suppose that y_n does not converge to y in norm. Then (y_n) has a subsequence (y_{n_k}) such that

$$\|y_{n_k} - y\| \geq \eta$$

for some $\eta > 0$. But since (x_n) is weakly convergent, (x_{n_k}) is bounded, so (Tx_{n_k}) has a convergent subsequence which must converge to y. However, this is a contradiction to $\|y_{n_k} - y\| \geq \eta$. Thus we must have that $y_n \to y$. □

Lemma 7.7. *If $T : X \to Y$ is a continuous map between two metric spaces. Then the image of a relatively compact set $A \subset X$ is relatively compact.*

Proof. Since T is continuous, then $T(\bar{A}) \subset \overline{T(A)}$. But $T(\bar{A})$ is compact since A is relatively compact, hence closed. Then $T(\bar{A}) = \overline{T(A)}$, so $T(A)$ is relatively compact. □

We also recall the following topological facts:

Lemma 7.8 (Total Boundedness)**.** *Let B be a subset of a metric space. Then:*

1. *If B is relatively compact, B is totally bounded.*
2. *If B is totally bounded and X is complete, B is relatively compact.*
3. *If B is totally bounded, for every $\epsilon > 0$ it has a finite ϵ-net $M_\epsilon \subset B$.*
4. *If B is totally bounded, B is separable.*

Theorem 7.9. *The range $\mathscr{R}(T)$ of a compact linear operator $T : X \to Y$ is separable where X and Y are normed spaces.*

Proof. Consider the ball $B_n = B(0,n) \subset X$. Since T is a compact, the image $C_n = T(B_n)$ is relatively compact. C_n is separable by Lemma (7.8). Since

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

then

$$T(X) = \bigcup_{n=1}^{\infty} T(B_n) = \bigcup_{n=1}^{\infty} C_n.$$

Since C_n is separable, it has a countable dense subset D_n, and the union

$$D = \bigcup_{n=1}^{\infty} D_n$$

is countable. So $T(X)$ has a countable dense set D. □

Corollary 7.9.1. *There does not exist a surjective compact linear operator $T : \ell^{\infty} \to \ell^{\infty}$.*

Theorem 7.10 (Compact Extension)**.** *A compact linear operator $T : X \to Y$ from a normed space X into a Banach space Y has compact linear extension $\tilde{T} : \hat{X} \to Y$, where $\hat{X}$ is the completion of X.*

Proof. We may regard X as a subspace of $\hat{X}$. Since T is bounded, it has a bounded linear extension $\tilde{T} : \hat{X} \to Y$. One can show that $\tilde{T}$ is compact using the sequential criterion. □

Theorem 7.11 (Adjoint Operator)**.** *Let $T : X \to Y$ be a linear operator. If T is compact, so is its adjoint operator $T^{\times} : Y' \to X'$.*

Proof. We consider any subset B of Y' which is bounded by some constant $c > 0$. We show that $T^{\times}(B)$ is totally bounded, so that $T^{\times}(B)$ is relatively compact, since X' is complete.

We prove that for any fixed $\varepsilon_0 > 0$ the image $T^{\times}(B)$ has a finite ε_0 - net. Since T is compact, the image $T(U)$ of the unit ball

$$U = \{x \in X | \|x\| \leq 1\}$$

is relatively compact. Hence $T(U)$ is totally bounded and has a finite ε_1 net $M \subset T(U)$ for $T(U)$, where $\varepsilon_1 = \varepsilon_0/4c$. This means that U contains points $x_1, \cdots, x_n$ such that each $x \in U$ satisfies

$$\|Tx - Tx_j\| < \frac{\varepsilon_0}{4c} \quad \text{for some } j. \tag{7.1}$$

We define a linear operator $A : Y' \longrightarrow \mathbb{R}^n$ by

$$Ag = (g(Tx_1), g(Tx_2), \cdots, g(Tx_n)). \tag{7.2}$$

g and T are bounded. Hence A is continuous and as range is finite dimensional, A is compact. Since B is bounded, $A(B)$ is relatively compact, hence totally bounded.Then it contains a finite ε_2 net $\{Ag_1, \cdots, Ag_m\}$ for itself, where $\varepsilon_2 = \varepsilon_0/4$. This means that each $g \in B$ satisfies

$$\|Ag - Ag_k\|_0 < \frac{1}{4}\varepsilon_0 \quad \text{for some } k, \tag{7.3}$$

where $\|\cdot\|_0$ is the norm on $\mathbb{R}^n$. We shall show that $\{T^\times g_1, \cdots, T^\times g_m\}$ is the desired ε_0 - net for $T^\times(B)$; this will then complete the proof.

We note that that for every j and every $g \in B$ there is a k such that

$$\begin{aligned}
|g(Tx_j) - g_k(Tx_j)|^2 &\leq \sum_{j=1}^{n} |g(Tx_j) - g_k(Tx_j)|^2 && (7.4)\\
&= \|A(g - g_k)\|_0^2 && (7.5)\\
&< \left(\frac{1}{4}\varepsilon_0\right)^2. && (7.6)
\end{aligned}$$

Let $x \in U$ be arbitrary. Then there is a j for which (7.1) holds. Let $g \in B$ be arbitrary. Then there is a k such that (7.3) holds, and (7.6) holds for that k and every j. We thus obtain

$$\begin{aligned}
|g(Tx) - g_k(Tx)| &\leq |g(Tx) - g(Tx_j)| + |g(Tx_j) - g_k(Tx_j)| && (7.7)\\
&\quad + |g_k(Tx_j) - g_k(Tx)| && (7.8)\\
&< \|g\|\|Tx - Tx_j\| + \frac{\varepsilon_0}{4} + \|g_k\|\|Tx_j - Tx\| && (7.9)\\
&\leq c\frac{\varepsilon_0}{4c} + \frac{\varepsilon_0}{4} + c\frac{\varepsilon_0}{4c} < \varepsilon_0. && (7.10)
\end{aligned}$$

Since this holds for every $x \in U$ we get that

$$\begin{aligned}
\|T^\times g - T^\times g_k\| &= \sup_{\|x\|=1} |(T^\times(g - g_k))(x)|\\
&= \sup_{\|x\|=1} |g(Tx) - g_k(Tx)| < \varepsilon_0.
\end{aligned}$$

This shows that $\{T^\times g_1, \cdots, T^\times g_m\}$ is an ε_0 - net for $T^\times(B)$. □

Lemma 7.12. *The restriction of a compact operator is compact.*

Lemma 7.13. *If X and Y are normed spaces, then there exists $T \in B(X,Y)$ such that T is not compact but the restriction of T to an infinite dimensional subspace of X is compact.*

Proof. Consider the map $T : \ell^2 \to \ell^2$ by

$$(x_1, x_2, x_3, x_4, x_5, \cdots) \mapsto (x_1, 0, x_3, 0, x_5, 0, x_7, \cdots)$$

This map is clearly not compact. However, the restriction to the subspace whose odd coordinates are all zero is the zero map, which is clearly compact. □

Lemma 7.14. *Let $T : Y \to Z$ be a compact linear operator and $S : X \to Y$ a bounded linear operator on a normed space X. Then TS and ST are compact.*

Proof. Let $T : Y \to Z$ be a compact operator and $S : X \to Y$ be bounded operator. Let M be a bounded subset of X. Then there exists $r > 0$ such that

$$M \subset B_X(0, r)$$

where $B_X(0,r)$ denotes the ball in X centered at 0 with radius r induced by the norm on X. Since S is bounded, then for any $x \in B_X(0,r)$, we have

$$\|Sx\| \leq \|S\|\|x\| < \|S\|r.$$

Hence we have

$$S(M) \subset B_Y(0, \|S\|r)$$

which is a bounded subset of Y. Since T is compact, then $T(S(M))$ is relatively compact. Since $T(S(M)) = (T \circ S)(M)$ and M is chosen to be any arbitrary bounded sets in X. Then this shows that $T \circ S : X \to Z$ is a compact operator.

Now let $T : Y \to Z$ be a bounded operator and $S : X \to Y$ be a compact operator. We make use of the sequential criterion. Let $\{x_n\}_{n=1}^{\infty}$ be any bounded sequence in X. Then by the compactness of S, $\{Sx_n\}_{n=1}^{\infty}$ has a convergent subsequence $\{Sx_{n_k}\}_{k=1}^{\infty}$ with limit $y \in Y$. Since T is bounded, then there exists $C > 0$ such that

$$\|Tv\| \leq C\|v\|$$

for all $v \in Y$. Then

$$\|T(Sx_{n_k} - y)\| \leq C\|Sx_{n_k} - y\| \to 0$$

as $k \to \infty$. Hence $\{TSx_{n_k}\}_{k=1}^{\infty}$ is a convergent subsequence of $\{TSx_n\}_{n=1}^{\infty}$, that converges to $T(y)$. Hence $T \circ S$ is compact. □

Corollary 7.14.1. *Let H be a Hilbert space, $T : H \to H$ a bounded linear operator and T^* the Hilbert-adjoint operator of T. Then T^* is compact if and only if T is compact.*

Proof. This is because if T is compact, then T^* can be represented in terms of the adjoint of $T^\times$ (by Theorem 4.11) which is compact by Theorem (7.11). Next note that the $(T^*)^* = T$. □

Corollary 7.14.2. *Let H be a Hilbert space, $T : H \to H$ a bounded linear operator and T^* the Hilbert-adjoint operator of T. Then T is compact if and only if T^*T is compact which happens if and only if T^* is compact*

Proof. Suppose T is compact, then T^*T being the composition of a bounded operator and a compact operator is compact. If T^*T is compact, then let $\{x_n\}$ be a bounded sequence with bound M. Then there exists $\{x_{n_k}\}$ such that $\{T^*Tx_{n_k}\}$ converges. Then

$$\|Tx_{n_k} - Tx_{n_j}\|^2 = \langle T^*Tx_{n_k} - T^*Tx_{n_j}, x_{n_k} - x_{n_j}\rangle \leq 2M\|T^*Tx_{n_k} - T^*Tx_{n_j}\|.$$

So $\{Tx_{n_k}\}$ is Cauchy, hence must converge. □

7.2 Spectral Properties of Compact Linear Operators

Theorem 7.15. *The set of eigenvalues of a compact linear operator $T : X \to X$ on a normed space X is countable (perhaps finite or even empty), and the only possible point of accumulation is $\lambda = 0$.*

Remark 7.15.1. *For $\lambda = 0$. λ could be in any of $\sigma_p(T)$, $\sigma_c(T)$, $\sigma_r(T)$.*

Proof. Suffices to show that for every real $k > 0$, the set of all $\lambda \in \sigma_p(T)$ such that $|\lambda| \geq k$ is finite. Suppose the contrary for some $k_0 > 0$. Then there is a sequence (λ_n) of infinitely many distinct eigenvalues such that $|\lambda_n| \geq k_0$. Also $Tx_n = \lambda_n x_n$ for some $x_n \neq 0$. Then we know that set of all the x_n's is linearly independent. Let $M_n = \text{span}\{x_1, \cdots, x_n\}$. Then every $x \in M_n$ has a unique representation

$$x = \alpha_1 x_1 + \cdots + \alpha_n x_n.$$

We apply $T - \lambda_n I$ and use $Tx_j = \lambda_j x_j$:

$$(T - \lambda_n I)x = \alpha_1(\lambda_1 - \lambda_n)x_1 + \cdots + \alpha_{n-1}(\lambda_{n-1} - \lambda_n)x_{n-1}.$$

We see that x_n no longer occurs on the right. Hence

$$(T - \lambda_n I)x \in M_{n-1} \quad \text{for all } x \in M_n.$$

The M_n's are closed (since they are finite dimensional). By Riesz's lemma, there is a sequence (y_n) such that

$$y_n \in M_n, \quad \|y_n\| = 1, \quad \|y_n - x\| \geq \frac{1}{2} \text{ for all } x \in M_{n-1}.$$

We show that

$$\|Ty_n - Ty_m\| \geq \frac{1}{2}k_0 \quad (n > m)$$

so that (Ty_n) has no convergent subsequence because $k_0 > 0$. This contradicts the compactness of T since (y_n) is bounded.

We note that

$$Ty_n - Ty_m = \lambda_n y_n - \tilde{x} \quad \text{where} \quad \tilde{x} = \lambda_n y_n - Ty_n + Ty_m.$$

Let $m < n$. We show that $\tilde{x} \in M_{n-1}$. Since $m \leq n - 1$, we see that $y_m \in M_m \subset M_{n-1} = \text{span}\{x_1, \cdots, x_{n-1}\}$. Hence $Ty_m \in M_{n-1}$. We also have that

$$\lambda_n y_n - Ty_n = -(T - \lambda_n I)y_n \in M_{n-1}.$$

Together, this shows $\tilde{x} \in M_{n-1}$, this also gives that $x = \lambda_n^{-1}\tilde{x} \in M_{n-1}$. So

$$\|\lambda_n y_n - \tilde{x}\| = |\lambda_n|\|y_n - x\| \geq \frac{1}{2}|\lambda_n| \geq \frac{1}{2}k_0.$$

□

Theorem 7.16. *Let $T : X \to X$ be a compact linear operator on a normed space X. Then for every $\lambda \neq 0$, the null space $\mathcal{N}(T_\lambda)$ of $T_\lambda = T - \lambda I$ is finite dimensional.*

Proof. We show that the closed unit ball M in $\mathcal{N}(T_\lambda)$ is compact. Let (x_n) be in M. Then (x_n) is bounded ($\|x_n\| \leq 1$), and $T(x_n)$ has a convergent subsequence (Tx_{n_k}). Now $x_n \in M \subset \mathcal{N}(T_\lambda)$ implies $T_\lambda x_n = Tx_n - \lambda x_n = 0$, so that $x_n = \lambda^{-1}Tx_n$, since $\lambda \neq 0$. Consequently, $(x_{n_k}) = (\lambda^{-1}Tx_{n_k})$ also converges. The limit is in M since M is closed, hence M is compact. This shows $\dim \mathcal{N}(T_\lambda) < \infty$. □

Corollary 7.16.1. *For $n = 1, 2, \cdots,$*

$$\dim \mathcal{N}(T_\lambda^n) < \infty$$

and

$$\{0\} = \mathcal{N}(T_\lambda^0) \subset \mathcal{N}(T_\lambda) \subset \mathcal{N}(T_\lambda^2) \subset$$

Proof. We note $\dim \mathcal{N}(T_\lambda^n)$ is the $(-\lambda)^n$ eigenspace of a compact operator. Hence it must be finite dimensional as $(-\lambda)^n \neq 0$. □

Theorem 7.17. *Let $T : X \to X$ be a compact linear operator on a normed space X. Then for every $\lambda \neq 0$, the range of $T_\lambda = T - \lambda I$ is closed.*

Proof. Suppose that $T_\lambda(X)$ is not closed. Then there is $y \in \overline{T_\lambda(X)}$, $y \notin T_\lambda(X)$ and a sequence (x_n) in X such that

$$y_n = T_\lambda x_n \to y.$$

Since $T_\lambda(X)$ is a vector space, $0 \in T_\lambda(X)$. But $y \notin T_\lambda(X)$, so $y \neq 0$. This implies $y_n \neq 0$ and $x_n \notin \mathcal{N}(T_\lambda)$ for all sufficiently large n. WLOG, we assume this holds for all n. Since $\mathcal{N}(T_\lambda)$ is closed, the distance δ_n from x_n to $\mathcal{N}(T_\lambda)$ is positive. Then there is a sequence $(z_n) \in \mathcal{N}(T_\lambda)$ such that

$$a_n = \|x_n - z_n\| < 2\delta_n.$$

We show that

$$a_n = \|x_n - z_n\| \to \infty. \tag{7.11}$$

Suppose this is not the case, then $(x_n - z_n)$ has a bounded subsequence. Since T is compact, then $(Tx_n - Tz_n)$ has a convergent subsequence. Now from $T_\lambda = T - \lambda I$ and $\lambda \neq 0$, we have $I = \lambda^{-1}(T - T_\lambda)$. Using $T_\lambda z_n = 0$, we obtain

$$\begin{aligned} x_n - z_n &= \frac{1}{\lambda}(T - T_\lambda)(x_n - z_n) \\ &= \frac{1}{\lambda}[T(x_n - z_n) - T_\lambda x_n]. \end{aligned}$$

$(Tx_n - Tz_n)$ has a convergent subsequence and $(T_\lambda x_n)$ converges, hence $(x_n - z_n)$ has a convergent subsequence, $x_{n_k} - z_{n_k} \to v$. Since T is compact, T is continuous and so is T_λ. Hence we have

$$y = T_\lambda x_{n_k} = T_\lambda(x_{n_k} - z_{n_k}) \to T_\lambda v.$$

This contradicts $y \notin T_\lambda(x)$. Thus we must have (7.11) hold.

Now set

$$w_n = \frac{1}{a_n}(x_n - z_n)$$

we have $\|w_n\| = 1$. Since $a_n \to \infty$, whereas $T_\lambda z_n = 0$ and $(T_\lambda x_n)$ converges, it follows that

$$T_\lambda w_n = \frac{1}{a_n} T_\lambda x_n \to 0.$$

Using again $I = \lambda^{-1}(T - T_\lambda)$, we obtain

$$w_n = \frac{1}{\lambda}(Tw_n - T_\lambda w_n).$$

Since T is compact and (w_n) is bounded, (Tw_n) has a convergent subsequence. Furthermore, $(T_\lambda w_n)$ converges. Hence (w_n) has a convergent subsequence, say

$$w_{n_j} \to w.$$

It is clear that $T_\lambda w = 0$. Hence $w \in \mathcal{N}(T_\lambda)$. Since $z_n \in \mathcal{N}(T_\lambda)$, also

$$u_n = z_n + a_n w \in \mathcal{N}(T_\lambda).$$

Hence for the distance from x_n to u_n we must have

$$\|x_n - u_n\| \geq \delta_n.$$

However, we also have

$$\begin{aligned} \delta_n &\leq \|x_n - z_n - a_n w\| \\ &= \|a_n w_n - a_n w\| \\ &= a_n \|w_n - w\| \\ &< 2\delta_n \|w_n - w\|. \end{aligned}$$

This shows $\|w_n - w\| > \frac{1}{2}$ which is clearly a contradiction. □

Corollary 7.17.1. *Let $T : X \to X$ be a compact linear operator on a normed space X. Then for every $\lambda \neq 0$, the range of T_λ^n is closed for every $n = 0, 1, 2, \cdots$. Furthermore,*

$$X = T_\lambda^0(X) \supset T_\lambda(X) \supset T_\lambda^2(X) \supset \cdots.$$

Lemma 7.18. *Let $T : X \to X$ be a compact linear operator on a normed space X, and let $\lambda \neq 0$. Then there exists a smallest integer r (depending on λ) such that from $n = r$ onwards, the null spaces $\mathcal{N}(T_\lambda^n)$ are all equal. If $r > 0$, then the inclusions*

$$\mathcal{N}(T_\lambda^0) \subset \mathcal{N}(T_\lambda) \subset \cdots \subset \mathcal{N}(T_\lambda^r)$$

are all proper.

Proof. Let us denote $\mathcal{N}_n = \mathcal{N}(T_\lambda^n)$. We know $\mathcal{N}_m \subset \mathcal{N}_{m+1}$. Suppose that $\mathcal{N}_m$ does not equal to $\mathcal{N}_{m+1}$ for any m. Then $\mathcal{N}_n$ is a proper subspace of $\mathcal{N}_{n+1}$ for every n. Since these null spaces are closed, Riesz's lemma implies that there exists a sequence (y_n) such that

$$y_n \in \mathcal{N}_n, \quad \|y_n\| = 1, \quad \|y_n - x\| \geq \frac{1}{2} \text{ for all } x \in \mathcal{N}_{n-1}.$$

We show that

$$\|Ty_n - Ty_m\| \geq \frac{1}{2}|\lambda|$$

so that (Ty_n) has no convergent subsequence. This contradicts the compactness of T, since (y_n) is bounded. From $T_\lambda = T - \lambda I$ we have $T = T_\lambda + \lambda I$ and

$$Ty_n - Ty_m = \lambda y_n - \tilde{x} \quad \text{where} \quad \tilde{x} = T_\lambda y_m + \lambda y_m - T_\lambda y_n.$$

Let $m < n$. We show that $\tilde{x} \in \mathcal{N}_{n-1}$. Since $m \leq n-1$, we clearly have $\lambda_m \in \mathcal{N}_m \subset \mathcal{N}_{n-1}$. Also $y_m \in \mathcal{N}_m$ implies

$$0 = T_\lambda^m y_m = T_\lambda^{m-1}(T_\lambda y_m),$$

that is $T_\lambda y_m \in \mathcal{N}_{m-1} \subset \mathcal{N}_{n-1}$. Similarly, $y_n \in \mathcal{N}_n$ implies $T_\lambda y_n \in \mathcal{N}_{n-1}$. Together, $\tilde{x} \in \mathcal{N}_{n-1}$. Also $x = \lambda^{-1}\tilde{x} \in \mathcal{N}_{N-1}$, so we have

$$\|\lambda y_n - \tilde{x}\| = |\lambda|\|y_n - x\| \geq \frac{1}{2}|\lambda|.$$

Next we show that $\mathcal{N}_m = \mathcal{N}_{m+1}$ implies $\mathcal{N}_n = \mathcal{N}_{n+1}$ for all $n > m$. Suppose this does not hold. Then $\mathcal{N}_n$ is a proper subspace of $\mathcal{N}_{n+1}$ for some $n > m$. We consider an $x \in \mathcal{N}_{n+1} - \mathcal{N}_n$. By definition

$$T_\lambda^{n+1}x = 0 \quad \text{but} \quad T_\lambda^n x \neq 0.$$

Since $n > m$, we have $n - m > 0$. We set $z = T_\lambda^{n-m}x$. Then

$$T_\lambda^{m+1}z = T_\lambda^{n+1}x = 0 \quad \text{but} \quad T_\lambda^m z = T_\lambda^n x \neq 0.$$

□

Lemma 7.19. *Let $T : X \to X$ be a compact linear operator on a normed space X, and let $\lambda \neq 0$. The there exists a smallest integer q (depending on λ) such that from $n = q$ on , the ranges $T_\lambda^n(X)$ are all equal; and if $q > 0$, the inclusions*

$$T_\lambda^0(X) \supset T_\lambda(X) \supset \cdots \supset T_\lambda^q(x)$$

are all proper.

Proof. We denote $\mathscr{R}_n = T_\lambda^n(X)$. Suppose that $\mathscr{R}_S = \mathscr{R}_{s+1}$ for no s. Then $\mathscr{R}_{n+1}$ is a proper subspace of $\mathscr{R}_n$ for every n. Since these ranges are closed, then Riesz's Lemma implies the existence of a sequence (x_n) such that

$$x_n \in \mathscr{R}_n, \quad \|x_n\| = 1, \quad \|x_n - x\| \geq \frac{1}{2} \; \forall x \in \mathscr{R}_{n+1}.$$

Let $m < n$ Since $T = T_\lambda + \lambda I$, then

$$Tx_m - Tx_n = \lambda x_m - (-T_\lambda x_m + T_\lambda x_n + \lambda x_n).$$

On the right $\lambda x_m \in \mathscr{R}_m$, $x_m \in \mathscr{R}_m$, so that $T_\lambda x_m \in \mathscr{R}_{m+1}$ and since $n > m$, we also have $T_\lambda x_n + \lambda x_n \in \mathscr{R}_n \subset \mathscr{R}_{m+1}$. Hence

$$Tx_m - Tx_n = \lambda(x_m - x)$$

where $x \in \mathscr{R}_{m+1}$. Consequently,

$$\|Tx_m - Tx_n\| = |\lambda|\|x_m - x\| \geq \frac{1}{2}|\lambda| > 0.$$

This contradicts T being compact. Furthermore, suppose $\mathscr{R}_{q+1} = \mathscr{R}_q$ means that T_λ maps $\mathscr{R}_q$ onto itself. Hence repeated application of T_λ gives $\mathscr{R}_{n+1} = \mathscr{R}_n$ for every $n > q$. □

Theorem 7.20. *Let $T : X \to X$ be a compact linear operator on a normed space X, and let $\lambda \neq 0$. Then there exists a smallest integer $n = r = q$, r and q as in Lemma (7.19) and Lemma (7.18), such that Lemma (7.19) and*

Lemma (7.18) holds.

Proof. We show that $q \geq r$ and $r \geq q$.

We have $\mathscr{R}_{q+1} = \mathscr{R}_q$. This means that $T_\lambda(\mathscr{R}_q) = \mathscr{R}_q$. Hence $y \in \mathscr{R}_q$ implies $y = T_\lambda x$ for some $x \in \mathscr{R}_q$. We show $T_\lambda x = 0$, $x \in \mathscr{R}_q$ implies $x = 0$. Suppose this is not the case, Then $T_\lambda x_1 = 0$ for some nonzero $x_1 \in \mathscr{R}_q$. Now with $y = x_1$ gives $x_1 = T_\lambda x_2$ for some $x_2 \in \mathscr{R}_q$. Inductively we construct a sequence such that

$$0 \neq x_1 = T_\lambda x_2 = \cdots = T_\lambda^{n-1} x_n \quad \text{but} \quad 0 = T_\lambda x_1 = T_\lambda^n x_n.$$

Hence $x_n \notin \mathcal{N}_{n-1}$ but $x_n \in \mathcal{N}_n$. We have $\mathcal{N}_{n-1} \subset \mathcal{N}$ is proper for any n, which is a contradiction. But since $\mathcal{N}(T_\lambda) \cap \mathscr{R}_q = \{0\}$, then by a moment of thought, this shows that $\mathcal{N}_{q+1} = \mathcal{N}_q$, $q \geq r$.

The converse is proved in the same way. Suppose $q > r$. Then we can produce something that is in $\mathcal{N}_{r+1}$ but not $\mathcal{N}_r$. □

Theorem 7.21. *Let $T : X \to X$ be a compact linear operator on a Banach space X. Then every spectral value $\lambda \neq 0$ of T is an eigenvalue of T.*

Proof. If $\mathcal{N}(T_\lambda) \neq \{0\}$, then λ is an eigenvalue of T. Suppose that $\mathcal{N}(T_\lambda) = \{0\}$, where $\lambda \neq 0$. Then $T_\lambda x = 0$ implies that $x = 0$ and $T_\lambda^{-1} : T_\lambda(X) = X \to X$ exists. By the bounded inverse theorem, T_λ^{-1} is bounded. So $\lambda \in \rho(T)$. □

Corollary 7.21.1 (Direct Sum)**.** *Let X, T, λ and r be as in Theorem (7.20). Then*

$$X = \mathcal{N}(T_\lambda^r) \oplus T_\lambda^r(X).$$

7.3 Solving Operator Equations

Theorem 7.22. *Let $T : X \to X$ be a compact linear operator on a normed space X and $\lambda \neq 0$. Then*

$$Tx - \lambda x = y \tag{7.12}$$

has a solution x if and only if y is such that

$$f(y) = 0$$

for all $f \in X'$ satisfying

$$T^\times f - \lambda f = 0. \tag{7.13}$$

In particular, if (7.13) *has only the trivial solution, then* (7.12) *is solvable for any $y \in X$.*

Proof. Suppose (7.12) has a solution $x = x_0$, that is

$$y = Tx_0 - \lambda x_0 = T_\lambda x_0.$$

Let f be any solution of (7.13). Then we have

$$f(y) = f(Tx_0 - \lambda x_0) = f(Tx_0) - \lambda f(x_0).$$

Now $f(Tx_0) = (T^\times f)(x_0)$ by the definition of the adjoint operator. Hence

$$f(y) = (T^\times f)(x_0) - \lambda f(x_0) = 0.$$

Conversely, suppose y satisfies $f(y) = 0$ for all f solving (7.13). Suppose (7.12) has no solutions. Then $y = T_\lambda x$ for no x. Hence $y \notin T_\lambda(X)$. Since $T_\lambda(X)$ is closed, the distance δ from y to $T_\lambda(X)$ is positive. Then there exists an $\tilde{f} \in X'$ such that $\tilde{f}(y) = \delta$ and $\tilde{f}(z) = 0$ for every $z \in T_\lambda(X)$. Since $z \in T_\lambda(X)$, we have $z = T_\lambda x$ for some $x \in X$, so that $\tilde{f}(z) = 0$ becomes

$$\tilde{f}(T_\lambda x) = \tilde{f}(Tx) - \lambda(\tilde{f}(x)) = (T^\times \tilde{f})(x) - \lambda \tilde{f}(x) = 0.$$

This holds for every $x \in X$ since $z \in T_\lambda(X)$ was arbitrary. Hence $\tilde{f}$ is a solution of (7.13). By assumption, it $\tilde{f}(y) = 0$. But this contradicts $\tilde{f}(y) = \delta > 0$. Consequently, (7.12) must have a solution. □

Lemma 7.23. *Let $T : X \to X$ be a compact linear operator on a normed space and let $\lambda \neq 0$ be given. Then there exists a real number $c > 0$ which is independent of y in (7.12) and such that for every y for which (7.12) has a solution, at least one of these solutions, call it $\tilde{x}$ satisifes*

$$\|\tilde{x}\| \leq c\|T_\lambda \tilde{x}\|.$$

Proof. Let x_0 be solution of $T_\lambda x = y$. If x is any other solution of this equation, then difference $z = x - x_0$ satisfies the homogeneous equation

$$Tz - \lambda z = 0.$$

Hence every solution of (7.12) can be written as $x = x_0 + z$ where $z \in \mathcal{N}(T_\lambda)$. Conversely, for every $z \in \mathcal{N}(T_\lambda)$, the sum $x_0 + z$ is a solution of (7.12). For a fixed x_0, the norm of x depends on z. Let us write

$$p(z) = \|x_0 + z\| \quad \text{and} \quad k = \inf_{z \in \mathcal{N}(T_\lambda)} p(z).$$

By the definition of an infimum, $cN(T_\lambda)$ contains a sequence (z_n) such that

$$p(z_n) = \|x_0 + z_n\| \to k.$$

Since $(p(z_n))$ converges, it is bounded. Also (z_n) is bounded because

$$\|z_n\| \leq p(z_n) + \|x_0\|.$$

Since T is compact, (Tz_n) has a convergent subsequence. But $z_n \in \mathcal{N}(T_\lambda)$ means that $T_\lambda z_n = 0$, that is $Tz_n = \lambda z_n$, where $\lambda \neq 0$. Hence (z_n) has a convergent subsequence, say $z_{n_j} \to z_0$ where $z_0 \in \mathcal{N}(T_\lambda)$ since $\mathcal{N}(T_\lambda)$ is closed. This shows that the set of solutions contains a solution $\tilde{x} = x_0 + z_0$ of minimum norm.

Next we prove that there is a $c > 0$ such that $\|\tilde{x}\| \leq c\|y\|$. Suppose this is not the case, then there exists a sequence

(y_n) such that

$$\frac{\|\tilde{x}_n\|}{\|y_n\|} \to \infty.$$

Multiplication by an α shows that to αy_n, there corresponds $\alpha\tilde{x}_n$ as a solution of minimum norm. Hence we may assume that $\|\tilde{x}_n\| = 1$. This implies $\|y_n\| \to 0$. Since T is compact, and $(\tilde{x}_n)$ is bounded, $(T\tilde{x}_n)$ has a convergent subsequence, which converges to say $v_0 = \lambda\tilde{x_0}$. Since $y_n = T_\lambda\tilde{x}_n = T\tilde{x}_n - \lambda\tilde{x}_n$, we have $\lambda\tilde{x}_n = T\tilde{x}_n - y_n$. Since $\|y_n\| \to 0$, we have

$$\tilde{x}_{n_j} = \frac{1}{\lambda}(T\tilde{x}_{n_j} - y_{n_j}) \to \tilde{x}_0.$$

From this, since T is continuous, we have

$$T\tilde{x}_{n_j} \to T\tilde{x}_0.$$

So $T\tilde{x}_0 = \lambda\tilde{x}_0$. Since $T_\lambda\tilde{x}_n = y_n$, we see that $x = \tilde{x}_n - \tilde{x}_0$ satisfies $T_\lambda x = y_n$. Since $\tilde{x}_n$ is of minimum norm,

$$\|x\| = \|\tilde{x}_n - \tilde{x}_0\| \geq \|\tilde{x}_n\| = 1.$$

But this contradicts the convergence of $\|tildex_{n_j}$. □

Theorem 7.24. *Let $T : X \to X$ be a compact linear operator on a normed space X and let $\lambda \neq 0$. Then*

$$T^\times f - \lambda f = g \tag{7.14}$$

has a solution f if and only if g is such that $g(x) = 0$ for all $x \in X$ satisfying

$$Tx - \lambda x = 0 \tag{7.15}$$

Hence if (7.15) has only the trivial solution $x = 0$, then (7.14) with any given $g \in X'$ is solvable.

Proof. If (7.14) has a solution f and x satisfies (7.15), then $g(x) = 0$ for all such x since

$$g(x) = (T^\times f)(x) - \lambda f(x) = f(Tx - \lambda x) = f(0) = 0.$$

Conversely, assume that g satisfies $g(x) = 0$ for all x satisfies (7.15). We consider any $x \in X$ and set $y = T_\lambda x$. Then $y \in T_\lambda(X)$. We may define a functional f_0 on $T_\lambda(X)$ by

$$f_0(y) = f_0(T_\lambda x) = g(x).$$

This is well-defined.

f_0 is linear since T_λ and g are linear. We show that f_0 is bounded. Lemma (7.23) implies that for every $y \in T_\lambda(X)$ at least one of the corresponding x's satisfies

$$\|x\| \leq c\|y\|$$

where c does not depend on y. Boundedness of f_0 now follows from

$$|f_0(y)| = |g(x)| \leq \|g\|\|x\| \leq c\|g\|\|y\| = \tilde{c}\|y\|.$$

By the Hahn-Banach theorem, the functional f_0 has an extension f on X which is a bounded linear functional defined on all of X. And by definition, we have

$$f(T_x - \lambda x) = f(T_\lambda x) = f_0(T_\lambda x) = g(x).$$

Hence

$$(T^\times f)(x) - \lambda f(x) = f(Tx - \lambda x) = g.$$

$\square$

Theorem 7.25. *Let $T : X \to X$ be a compact linear operator on a normed space X and let $\lambda \neq 0$. Then*

1. *The equation*
$$Tx - \lambda x = y \tag{7.16}$$
has a solution x for every $y \in X$ if and only if the homogeneous equation
$$Tx - \lambda x = 0 \tag{7.17}$$
has only the trivial solution $x = 0$. In this case, the solution of (7.16) is unique, and T_λ has a bounded inverse.

2. *The Equation*
$$T^\times f - \lambda f = g \tag{7.18}$$
has a solution for every $g \in X'$ if and only if
$$T^\times f - \lambda f = 0. \tag{7.19}$$
has only the trivial solution $f = 0$. IN this case, the solution of (7.18) is unique.

Proof. We prove (1), (2) is a direct consequence of (1) since $T^\times$ is compact if T is compact.

$\Rightarrow$: Suppose $x_1 \neq 0$ is such that $Tx_1 - \lambda x_1 = 0$. Then since (7.16) with any y is solvable, then there exists $T_\lambda x_2 = x_1$. Inductive, for any $k = 2, 3, \cdots$,, there exists

$$0 \neq x_1 = T_\lambda x_2 = T_\lambda^2 x_3 = \cdots = T_\lambda^{k-1} x_k$$

and

$$0 = T_\lambda x_1 = T_\lambda^k x_k.$$

Hence $x_k \in \mathcal{N}(T_\lambda^k)$ but $x_k \notin \mathcal{N}(T_\lambda^{k-1})$. This means that the nullspace is always increasing, which is a contradictions.

$\Leftarrow$: suppose $x = 0$ is the only solution of (7.17). Then (7.18) with any g is solvable by Theorem (7.24). Now $T^\times$ is compact, so that we can apply the first part of this proof to conclude that $f \equiv 0$ is the only solution to $T^\times f - \lambda f = 0$. Thus $Tx - \lambda x = y$ is solvable for any y by Theorem (7.22). The uniqueness of such solution is clear. The boundedness of the inverse of T_λ follows from Lemma (7.23). $\square$

Corollary 7.25.1. *Let $T : X \to X$ be a compact linear operator on a normed space X. Then if T has a nonzero spectral values. Every one of them must be an eigenvalue of T.*

Proof. If T_λ is not injective, then $\lambda \in \sigma_p(T)$ by definition. Let $\lambda \neq 0$, and T_λ is injective. Then $T_\lambda x = 0$ implies $x = 0$. Hence the homogeneous equation only has the trivial solution, then R_λ exists and is bounded. □

Next, we recall the following lemma from Linear Algebra:

Lemma 7.26. *Given a linearly independent set $\{f_1, \cdots, f_m\}$ in the dual space X' of a normed space X, there are elements $z_1, \cdots, z_m$ in X such that*

$$f_j(z_k) = \delta_{jk} = \begin{cases} 0 & j \neq k \\ 1 & j = k \end{cases}.$$

Consequently, we have the following Theorem:

Theorem 7.27. *Let $T : X \to X$ be a compact linear operator on a normed space X, and let $\lambda \neq 0$. Then the equations*

$$Tx - \lambda x = 0 \quad \textit{and} \quad T^\times f - \lambda f = 0$$

has the same number of linearly independent solutions.

7.4 Fredholm Alternative

Definition 7.28 (Fredholm Alternative)**.** *A bounded linear operator $A : X \to X$ on a normed space X is said to satisfy the* ***Fredholm Alternative*** *if A is such that either one of the following holds:*

- *The nonhomogeneous equation*

  $$Ax = y, \quad A^\times f = g$$

  have solutions x and f, respectively for every given $y \in X$ and $g \in X'$ and the solution is unique. The corresponding homogeneous equation

  $$Ax = 0, \quad A^\times f = 0$$

  have only the trivial solution $x = 0$ and $f = 0$ respectively.

- *The homogeneous equations*

  $$Ax = 0, \quad A^\times f = 0$$

  have the same number of linearly independent solutions

  $$x_1, \cdots, x_n \quad \textit{and} \quad f_1, \cdots, f_n$$

  respectively with $n \geq 1$. The homogeneous equations

  $$Ax = y, \quad A^\times f = g$$

  are not solvable for all y and g, respectively; they have a solution if and only if y and g are such that

  $$f_k(y) = 0, \quad g(x_k) = 0$$

  for $k = 1, 2, \cdots, n$, respectively.

Corollary 7.28.1. *Let $T : X \to X$ be a compact linear operator on a normed space X, and let $\lambda \neq 0$. Then $T_\lambda = T - \lambda I$ satisfies the Fredholm alternative.*

Theorem 7.29 (Fredholm Alternative for Integral Equations)**.** *Given the integral equation*

$$x(s) - \mu \int_a^b k(s,t)x(t)dt = \tilde{y}(s). \tag{7.20}$$

If k is such that $T : X \to X$ given by

$$(Tx)(s) = \int_a^b k(s,t)x(t)dt$$

is a compact linear operator on a normed space X. Then the Fredholm alternative holds for T_λ; thus either (7.20) has a unique solution for all $\tilde{y} \in X$ or the homogeneous equation corresponding to (7.20) has finitely many linearly independent nontrivial solutions.

Proof. Equation (7.20) can be written as

$$x - \mu T x = \tilde{y}.$$

Let $\mu = \frac{1}{\lambda}$ and $\tilde{y}(s) = -y(s)/\lambda$, then it becomes

$$Tx - \lambda x = y, \quad \lambda \neq 0.$$

□

Assume that X is a complex Banach space. If $|\lambda| \geq \|T\|$, then we know that $\lambda \in \rho(T)$, this implies $R_\lambda(T)$ exists and we have

$$R_\lambda(T) = -\lambda^{-1}(I + \lambda^{-1}T + \lambda^{-2}T^2 + \cdots).$$

Consequently, for solutions $x = R_\lambda(T)y$, we have the representation

$$x = -\frac{1}{\lambda}\left(y + \frac{1}{\lambda}Ty + \frac{1}{\lambda^2}T^2y + \cdots\right),$$

which is called a **Neumann series**.

On the other hand, suppose $\lambda \in \sigma(T)$ is nonzero. Then we know λ is an eigenvalue of T. This shows that the homogeneous equation has at least one linearly independent solutions. This corresponds to the second case of Definition (7.28).

Lemma 7.30. *Let $J = [a,b]$ be any compact interval and suppose that k is continuous on $J \times J$. Then the operator $T : X \to X$ defined by $T : X \to X$,*

$$(Tx)(s) = \int_a^b k(s,t)x(t)dt$$

where $X = C[a,b]$ is a compact linear operator.

Proof. This follows from the Arzela Ascoli's Theorem. □

8 Spectral Theory of Self-Adjoint Operators

8.1 Spectral Properties of Bounded Self-Adjoint Operators

Recall that if $T : H \to H$ is an operator defined on a Hilbert space H, then its adjoint $T^* : H \to H$ is such that

$$\langle Tx, y\rangle = \langle x, Ty\rangle.$$

We say that T is self-adjoint if $T = T^*$. We point out that if $T = T^*$ on all of H, that is

$$\langle Tx, y\rangle = \langle x, Ty\rangle$$

Then T must be bounded. To prove this point, we prove that T has closed graph. Suppose $x_n \to x$ and $Tx_n \to y$, then

$$\langle y, z\rangle = \lim_n \langle Tx_n, z\rangle = \lim_n \langle x_n, Tz\rangle = \langle x, Tz\rangle = \langle Tx, z\rangle$$

for any $z \in H$. This shows that $Tx = y$, therefore the graph is closed.

Theorem 8.1. *Let $T : H \to H$ be a bounded self-adjoint linear operator on a complex Hilbert space H. Then*

1. *All the eigenvalues of T (if they exist) are real.*
2. *Eigenvectors corresponding to different eigenvalues of T are orthogonal.*

Proof. 1. Let λ be any eigenvalue of T and x a corresponding eigenvector. Then $x \neq 0$ and $Tx = \lambda x$. Using the self - adjointness of T, we obtain

$$\begin{aligned}\lambda\langle x, x\rangle &= \langle \lambda x, x\rangle = \langle Tx, x\rangle \\ &= \langle x, Tx\rangle = \langle x, \lambda x\rangle = \overline{\lambda}\langle x, x\rangle.\end{aligned}$$

Here $\langle x, x\rangle = \|x\|^2 \neq 0$ since $x \neq 0$, and division by $\langle x, x\rangle$ gives $\lambda = \overline{\lambda}$. Hence λ is real.

2. Let λ and μ be eigenvalues of T, and let x and y be corresponding eigenvectors. Then $Tx = \lambda x$ and $Ty = \mu y$. Since T is self - adjoint and μ is real,

$$\begin{aligned}\lambda\langle x, y\rangle &= \langle \lambda x, y\rangle = \langle Tx, y\rangle \\ &= \langle x, Ty\rangle = \langle x, \mu y\rangle = \mu\langle x, y\rangle.\end{aligned}$$

Since $\lambda \neq \mu$, we must have $\langle x, y\rangle = 0$, which means orthogonality of x and y.

□

Theorem 8.2. *Let $T : H \to H$ be a bounded self-adjoint linear operator on a complex Hilbert space H. Then a number λ belongs to the resolvent set $\rho(T)$ of T if and only if there exists a $c > 0$ such that for every $x \in H$,*

$$\|T_\lambda x\| \geq c\|x\|. \tag{8.1}$$

Proof. Suppose $\lambda \in \rho(T)$, then $R_\lambda = T_\lambda^{-1}$ is a bounded operator by the bounded inverse theorem. Hence we get the only if direction.

Now suppose there exists a $c > 0$ such that (8.1) holds. Then one can verify easily that T_λ is injective. We show $T_\lambda(H) = H$. Suppose $x_0 \perp \overline{T_\lambda(H)}$, then

$$0 = \langle T_\lambda x, x_0\rangle = \langle Tx, x_0\rangle - \lambda\langle x, x_0\rangle$$

for all $x \in H$. Since T is self-adjoint, we thus obtain

$$\langle x, Tx_0\rangle = \langle Tx, x_0\rangle = \langle x, \overline{\lambda} x_0\rangle.$$

This shows $Tx_0 = \overline{\lambda}x_0$. By Theorem (8.1), if $x_0 \neq 0$, then $\lambda = \overline{\lambda}$. However, this contradicts

$$= 0\|T_\lambda x_0\| \geq c\|x_0\| > 0$$

Hence we conclude that $\overline{T_\lambda(H)}^{\perp} = \{0\}$. This shows $\overline{T_\lambda(H)} = H$. Next, it is easy to show that $T_\lambda(H)$ is closed. Since a convergent in $T_\lambda(H)$ implies the Cauchyness of the preimage sequence, which has a limit. This finishes the proof of the theorem. □

Theorem 8.3 (Spectrum of bounded self-adjoint operators)**.** *The spectrum $\sigma(T)$ of a bounded self-adjoint linear operator $T : H \to H$ on a complex Hilbert space H is real.*

Proof. We show that $\lambda = \alpha + i\beta$ with $\beta \neq 0$ must belong to $\rho(T)$. For every $x \neq 0$ in H, we have

$$\langle T_\lambda x, x\rangle = \langle Tx, x\rangle - \lambda\langle x, x\rangle$$

and since $\langle x, x\rangle$ and $\langle Tx, x\rangle$ are real, we have

$$\overline{T_\lambda x, x\rangle} = \langle Tx, x\rangle - \overline{\lambda}\langle x, x, \rangle.$$

Subtract the previous two equalities, we have

$$\overline{\langle T_\lambda x, x\rangle} - \langle T_\lambda x, x\rangle = (\lambda - \overline{\lambda})\langle x, x\rangle = 2i\beta\|x\|^2.$$

The left hand side is $-2i\Im\langle T_\lambda x, x\rangle$. So

$$|\beta|\|x\|^2 = |\Im\langle T_\lambda x, x, \rangle| \leq |\langle T_\lambda x, x\rangle| \leq \|T_\lambda x\|\|x\|.$$

Then by theorem (8.2), this shows $\lambda \in \rho(T)$. □

Example: Let $T : \ell^2 \to \ell^2$ be defined by

$$y = (\eta_i) = Tx, \quad x = (\xi_i), \quad \eta_i = \lambda_i \xi_i,$$

where (λ_i) is a bounded sequence on $\mathbb{R}$. Then $\sigma(T)$ is the closure of the set of eigenvalues of T, which is given by $\{\lambda_i\}$.

Theorem 8.4. *The spectrum $\sigma(T)$ of a bounded self-adjoint linear operator $T : H \to H$ on a complex Hilbert space H lies in the closed interval $[m, M]$ on the real axis, where*

$$m = \inf_{\|x\|=1} \langle Tx, x\rangle, \quad M = \sup_{\|x\|=1} \langle Tx, x\rangle.$$

In particular, $\sigma(T)$ is compact.

Proof. We already know that $\sigma(T)$ lies on the real axis. We show that any real $\lambda = M + c$ with $c > 0$ belongs to the resolvent set $\rho(T)$. For every $x \neq 0$, and $v = \|x\|^{-1}x$ we have $x = \|x\|v$ and

$$\langle Tx, x\rangle = \|x\|^2\langle Tv, v\rangle \leq \|x\|^2 \sup_{\|v'\|=1} \langle Tv', v'\rangle = \langle x, x, \rangle M.$$

Hence $-\langle Tx, x\rangle \geq -\langle x, x\rangle M$, and by the Schwarz inequality, we obtain

$$\begin{aligned} \|T_\lambda x\|\|x\| &\geq -\langle T_\lambda x, x\rangle \\ &= -\langle Tx, x\rangle + \lambda\langle x, x\rangle \\ &\geq (-M + \lambda)\langle x, x\rangle \\ &= c\|x\|^2, \end{aligned}$$

where $c = \lambda - M > 0$. This shows $\|T_\lambda x\| \geq c\|x\|$ hence we have $\lambda \in \rho(T)$ by Theorem (8.2). Similarly, we can prove for the case $\lambda < m$. □

Proposition 8.5. *For any bounded self-adjoint linear operator T on a complex Hilbert space H, we have*

$$\|T\| = \max(|m|, |M|) = \sup_{\|x\|=1} |\langle Tx, x\rangle|.$$

Proof. By Schwarz Inequality, it is clear that $|m|, |M|$ is no greater than $\|T\|$. Hence suffices to show $\|T\| \leq K$, where $K = \max(|m|, |M|)$. If T is the zero operator, then it is trivial. Otherwise, there exists z of norm 1 and $Tz \neq 0$. Let $v = \|Tz\|^{1/2}z$ and $w = \|Tz\|^{-1/2}Tz$. Then

$$\|v\|^2 = \|w\|^2 = \|Tz\|.$$

We now set $y_1 = v + w$ and $y_2 = v - w$. Then one can verify that

$$\langle Ty_1, y_1\rangle - \langle Ty_2, y_2\rangle = 4\|Tz\|^2.$$

Now for every $y \neq 0$ and $x = \|y\|^{-1}y$ we have $y = \|y\|x$ and

$$\langle Ty, y\rangle = \|y\|^2|\langle Tx, x\rangle| \leq \|y\|^2 K$$

Hence we have

$$\begin{aligned} |\langle Ty_1, y_1\rangle - \langle Ty_2, y_2\rangle| &\leq |\langle Ty_1, y_1\rangle| + |\langle Ty_2, y_2\rangle| \\ &\leq K(\|y_1\|^2 + \|y_2\|^2) \\ &= 2K(\|v\|^2 + \|w\|^2) \end{aligned}$$

$$= 4K\|Tz\|.$$

Hence we see that $4\|Tz\|^2 \leq 4K\|Tz\|$, so $\|Tz\| \leq K$ and $\|T\| \leq K$. □

Corollary 8.5.1. *m and M are spectral values of T.*

Proof. After a translation, we may assume $0 \leq m \leq M$. Then by Proposition (8.5), we have

$$M = \sup_{\|x\|=1} \langle Tx, x\rangle = \|T\|.$$

By the definition of a supremum, there is a sequence (x_n) such that

$$\|x_n\| = 1, \quad \langle Tx_n, x_n\rangle = M - \delta_n, \quad \delta_n \geq 0, \quad \delta_n \to 0.$$

Then $\|Tx_n\| \leq \|T\|\|x_n\| = \|T\| = M$, and since T is self-adjoint,

$$\begin{aligned}\|Tx_n - Mx_n\|^2 &= \langle Tx_n - Mx_n, Tx_n - Mx_n\rangle \\ &= \|Tx_n\|^2 - 2M\langle Tx_n, x_n\rangle + M^2\|x_n\|^2 \\ &\leq M^2 - 2M(M - \delta_n) + M^2 = 2M\delta_n \to 0.\end{aligned}$$

Thus there is no positive c such that

$$\|T_M x_n\| = \|Tx_n - Mx_n\| \geq c = c\|x_n\|.$$

□

Corollary 8.5.2. *A compact self-adjoint linear operator $T : H \to H$ on a complex Hilbert space $H \neq \{0\}$ has at least one eigenvalue.*

Proof. Suppose $\lambda \in \sigma(T)$ is nonzero, then by the compactness of T, we know that λ is an eigenvalue. Otherwise, if not such λ exists. Then we must have that $m = M = 0$, i.e., it is the zero operator. Hence 0 is an eigenvalue. □

Theorem 8.6. *The residual spectrum $\sigma_r(T)$ of a bounded-self adjoint linear operator $T : H \to H$ on a complex Hilbert space H is empty.*

Proof. Suppose $\sigma_r(T) \neq \emptyset$ and $\lambda \in \sigma_r(T)$. THen the $T_\lambda(H)$ is not dense in H. Hence by the projection theorem, there is a $y \neq 0$ in H which is orthogonal to $T_\lambda(H)$, so

$$\langle T_\lambda x, y\rangle = 0 \quad \forall x \in H.$$

Since λ is real and T is self-adjoint, we thus obtain

$$\langle x, T_\lambda y\rangle = 0 \quad \forall x \in H.$$

Let $x = T_\lambda y$, we thus have $\|T_\lambda y\|^2 = 0$, so $T_\lambda y = Ty - \lambda y = 0$. However, this shows that λ is instead an eigenvalue. □

8.2 Positive Operators

Recall that a positive operator $T : H \to H$ is one that is bounded, self-adjoint and

$$\langle Tx, x\rangle \geq 0 \quad \forall x \in H.$$

We also write $T_1 \leq T_2$ for two self-adjoint operators iff $0 \leq T_2 - T_1$.

Lemma 8.7. *If two bounded self-adjoint linear operators S and T on a Hilbert space H are positive and commutes, then their product ST is positive.*

Proof. We show $\langle STx, x\rangle \geq 0$ for all $x \in H$ as it is clear that ST is positive. If $S = 0$, then this is clear. Suppose $S \neq 0$. We define

$$S_1 = \frac{1}{\|S\|}S, \quad S_{n+1} = S_n - S_n^2.$$

It is clear that S_n's are self-adjoint. We prove by induction that

$$0 \leq S_n \leq I.$$

It is clear that when $n = 1$, the inequality holds. Now suppose $0 \leq S_k \leq I$ for some $k \in \mathbb{N}$, then

$$\langle S_k^2(I - S_k)x, x\rangle = \langle (I - S_k)S_kx, S_kx\rangle = \langle (I - S_k)y, y\rangle \geq 0$$

where we set $y = S_kx$. This implies

$$S_k^2(I - S_k).$$

Similarly, we can show that

$$S_k(I - S_k)^2 \geq 0.$$

Then

$$0 \leq S_k^2(I - S_K) + S_K(I - S_k)^2 = S_k - S_k^2 = S_{k+1}.$$

On the other hand, $I - S_k \geq 0$ and $S_k^2 \geq 0$, so

$$0 \leq I - S_k + S_k^2 = I - S_{k+1} \Longrightarrow S_{k+1} \leq I.$$

We now show that $\langle STx, x\rangle \geq 0$ for all $x \in H$. By definition, we have

$$S_1 = S_1^2 + S_2^2 + \cdots + S_n^2 + S_{n+1}.$$

Since $S_{n+1} \geq 0$, this implies

$$S_1^2 + \cdots + S_n^2 = S_1 - S_{n+1} \leq S_1.$$

Thus

$$\sum_{j=1}^{n} \|S_jx\|^2 \leq \langle S_1x, x\rangle.$$

Since n is arbitrary, the infinite series

$$\sum_{j=1}^{\infty} \|S_j x\|^2$$

converges. This shows $\|S_n x\| \to 0$ and $S_n x \to 0$. Then

$$\left(\sum_{j=1}^{n} S_j^2\right) x = (S_1 - S_{n+1})x \to S_1 x$$

as $n \to \infty$.

Next, notice that all the S_j's commutes with T. Using $S = \|S\|S_1$, and $T \geq 0$, and the continuity of the inner product, we thus obtain that for every $x \in H$, and $y_j = S_j x$,

$$\begin{aligned}
\langle STx, x\rangle &= \|S\|\langle TS_1 x, x\rangle \\
&= \|S\| \lim_{n\to\infty} \langle TS_j^2 x, x\rangle \\
&= \|S\| \lim_{n\to\infty} \langle Ty_j, y_j\rangle \geq 0.
\end{aligned}$$

This shows $\langle STx, x\rangle \geq 0$. □

Lemma 8.8. *The following are true:*

1. *if $S \leq T$ and $S \geq T$, then $S = T$.*
2. *If $T_1 \leq T_2$, then $T_1 + T \leq T_2 + T$.*
3. *If $T_1 \leq T_2$, then $\alpha T_1 \leq \alpha T_2$, where $\alpha \geq 0$.*
4. *A, B, T are bounded self-adjoint linear operators. If $T \geq 0$ and commutes with A and B, then*

$$A \leq B \Longrightarrow AT \leq BT.$$

5. *TT^* and T^*T are self-adjoint and positive. In particular, TT^* and T^*T have nonnegative spectra.*
6. *Suppose T is a bounded self-adjoint linear operator T on a complex Hilbert space. Then it is positive if and only if its spectrum consists of nonnegative real values only.*
7. *If $T : H \to H$ and $S : H \to H$ are bounded. If T is compact and $S^*S \leq T^*T$, then S is compact.*

Definition 8.9 (Monotone Sequence)**.** *A **monotone sequence** (T_n) of self-adjoint linear operators T_n on a Hilbert space H is a sequence (T_n) which is either monotone increasing, that is,*

$$T_1 \leq T_2 \leq T_3 \leq \cdots$$

or monotone decreasing.

Theorem 8.10. *Let (T_n) be a sequence of bounded self-adjoint linear operators on a complex Hilbert space H such that*

$$T_1 \leq T_2 \leq \cdots \leq T_n \leq \cdots \leq K$$

where K is a bounded self-adjoint linear operator on H. Suppose that any T_j commutes with K and with every T_m. Then (T_n) is strongly operator convergent and the limit operator T is linear, bounded and self-adjoint and satisfies $T \leq K$.

Proof. Consider $S_n = K - T_n$, then S_n is self-adjoint. We have

$$S_m^2 - S_n S_m = (S_m - S_n)S_m = (T_n - T_m)(K - T_m).$$

Let $m < n$. Then $T_n - T_m$ and $K - T_m$ are positive. By Lemma (8.7), $S_m^2 - S_n S_m \geq 0$, so $S_m^2 \geq S_n S_m$ for $m < n$. Similarly,

$$S_n S_m m - S_n^2 = (K - T_n)(T_n - T_m(\geq 0$$

so $S_n S_m \geq S_n^2$. Together, we have

$$S_m^2 \geq S_n S_m \geq S_n^2.$$

By definition, we thus have

$$\langle S_m^2 x, x \rangle \geq \langle S_n^2 x, x \rangle \geq 0.$$

So the sequence $(\langle S_n^2 x, x \rangle)$ converges. The rest follows from the uniform boundedness theorem. It is clear that $T \leq K$ is self-adjoint and positive. □

Proposition 8.11. *Let T be a positive operator, then $(I + T)^{-1}$ exists.*

Proof. We know the inverse exists since $-1 \in \rho(T)$. □

Definition 8.12 (Positive Square Root)**.** *Let $T : H \to H$ be a positive bounded self-adjoint linear operator on a complex Hilbert space H. Then a bounded self-adjoint linear operator A is called a* ***square root*** *of T if*

$$A^2 = T.$$

If in addition, $A \geq 0$, then A is called a ***positive square root*** *of T and is denoted by*

$$A = T^{1/2}.$$

Theorem 8.13 (Positive Square Root)**.** *Every positive bounded self-adjoint linear operator $T : H \to H$ on a complex Hilbert space has a positive square root A which is unique. This operator A commutes with every bounded linear operator on H which commutes with T.*

Proof. If $T = 0$, we take $A = T^{1/2} = 0$. So assume $T \neq 0$. Then by the Schwarz inequality, we have

$$\langle Tx, x \rangle \leq \|Tx\| \|x\| \leq \|T\| \|x\|^2.$$

Dividing by $\|T\| \neq 0$ and setting $Q = (1/\|T\|)T$, we obtain

$$\langle Qx, x \rangle \leq \|x\|^2 = \langle Ix, x \rangle.$$

That is $Q \leq I$. Assume that Q has a unique positive square root, $B = Q^{1/2}$, we have $B^2 = Q$ and wee see that a square root of $T = \|T\|Q$ is $\|T\|^{1/2}B$. Also it is not difficult to see that the uniqueness of $Q^{1/2}$ implies the uniqueness of the positive square root of T.

So WLOG, we assume $T \leq I$. Define $A_0 = 0$ and

$$A_{n+1} = A_n + \frac{1}{2}(T - A_n^2), \quad n = 0, 1, 2, \cdots .$$

Then A_ni s a polynomial in T. Hence A_n's are self-adjoint and all commute with every operator that T commutes with. We now prove the following:

- $A_n \leq I$;
- $A_n \leq A_{n+1}$;
- $A_n x \to Ax$, where $A = T^{1/2}$;
- If $ST = TS$, and S is a bounded linear operator on H, then $AS = SA$.

The proofs are as follows:

- It is clear that $A_0 \leq I$. Let $n > 0$, Since $I - A_{n-1}$ is self-adjoint, $(I - A_{n-1})^2 \geq 0$. Also $T \leq I$ implies $I - T \geq 0$. So we have

$$\begin{aligned} 0 &\leq \frac{1}{2}(I - A_{n-1})^2 + \frac{1}{2}(I - T) \\ &= I - A_{n-1} - \frac{1}{2}(T - A_{n-1}^2) \\ &= I - A_n. \end{aligned}$$

- It is clear that $A_0 \leq A_1 = \frac{1}{2}T$. We show that $A_{n-1} \leq A_n$ implies $A_n \leq A_{n+1}$. We note that

$$\begin{aligned} A_{n+1} - A_n &= A_n + \frac{1}{2}(T - A_n^2) - a_{n-1} - \frac{1}{2}(T - A_{n-1}^2) \\ &= (A_n - A_{n-1})[I - \frac{1}{2}(A_n + A_{n-1})] \\ &\geq 0 \end{aligned}$$

- How we note that (A_n) is a monotone sequence with $A_n \leq I$. Then it converges to a bounded self-adjoint linear operator A such that $A_n x \to Ax$ for all $x \in H$. Since $(A_n x)$ converges, we have

$$A_{n+1}x - A_n x = \frac{1}{2}(Tx - A_n^2 x) \to 0.$$

  Hence $Tx - A^2 x = 0$ for all x. It is clear that $A \geq 0$.

- Since $ST = TS$ implies $A_n S = SA_n$, then $A_n Sx = SA_n x$ for all $x \in H$. Letting $n \to \infty$, we have $AS = SA$.

Lastly, we prove the uniqueness of the positive square root. Suppose A and B are two positive square roots of T. Then $A^2 = B^2 = T$. Also $BT = BB^2 = B^2 B + TB$, so that $AB = BA$. Let $x \in H$ be arbitrary and $y = (A - B)x$. Then $\langle Ay, y \rangle \geq 0$ and $\langle By, y \rangle 0$. So

$$\langle Ay, y \rangle + \langle By, y \rangle = \langle (A + B)y, y \rangle = \langle (A + B)(A - B)x, y \rangle = \langle (A^2 - B^2)x, y \rangle = 0$$

This shows $\langle Ay, y\rangle = \langle By, y\rangle 0$. Since A is positive, then it has a positive square root C, we have

$$0 = \langle Ay, y\rangle = \langle C^2 y, y\rangle = \langle Cy, Cy\rangle = \|Cy\|^2$$

and $Cy = 0$. Also $Ay = C^2 y = C(Cy) = 0$. Similarly, $By = 0$. Hence $(A - B)y = 0$. Using $y = (A - B)x$, we thus have for all $x \in H$,

$$\|Ax - Bx\|^2 = \langle (A - B)^2 x, x\rangle = \langle (A - B)y, x\rangle = 0.$$

This shows that $Ax - Bx = 0$ for all $x \in H$ and proves $A = B$. □

Corollary 8.13.1. $\|T^{1/2}\|^2 = \|T\|$.

Proof. By Proposition (8.5), we have

$$\|T\| = \sup_{\|x\|=1} |\langle Tx, x\rangle| = \sup_{\|x\|=1} \langle T^{1/2}x, T^{1/2}x\rangle = \sup_{\|x\|=1} \|T^{1/2}x\|^2 = \|T^{1/2}\|^2$$

□

Corollary 8.13.2. *Let $T : H \to H$ be bounded positive self-adjoint linear operator on a complex Hilbert space. Then for any $x, y \in H$, we have*

$$|\langle Tx, y\rangle| \leq \langle Tx, x\rangle^{1/2} \langle Ty, y\rangle^{1/2}.$$

Proof. By Schwarz Inequality, we have

$$\begin{aligned}
|\langle Tx, y\rangle| &\leq \langle T^{1/2}x, T^{1/2}y\rangle| \\
&\leq \|T^{1/2}(x)\|^{1/2} \|T^{1/2}(y)\|^{1/2} \\
&= \langle Tx, x\rangle^{1/2} \langle Ty, y\rangle^{1/2}.
\end{aligned}$$

□

8.3 Projection Operators

Recall that given any closed subspace Y of a Hilbert space H, its orthogonal complement $Y^\perp$ exists and

$$H = Y \oplus Y^\perp.$$

Then the projection operator

$$P : H \to H, \quad x \mapsto Px = y$$

where y is the unique vector in Y such that $x = y + z$, $z \in Y^\perp$.

Lemma 8.14. *A bounded linear operator $P : H \to H$ on a Hilbert space H is a projection if and only if P is a self-adjoint and idempotent, that is $P^2 = P$.*

Remark 8.14.1. *An idempotent operator itself may not gaurantee to be a projection operator. E.g. consider*

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Proof. One direction is clear. We prove the converse. Suppose that $P^2 = P = P^*$ and denote $P(H)$ by Y. Then for every $x \in H$,

$$x = Px + (I - P)x.$$

Orthogonality $Y = P(H) \perp (I - P)(H)$ follows from

$$\langle Px, (I - P)v\rangle = \langle x, P(I - P)v\rangle = \langle x, Pv - P^2v\rangle = \langle x, 0\rangle = 0.$$

Y is the null space $\mathcal{N}(I - P)$ of $I - P$, because $Y \subset \mathcal{N}(I - P)$ can be seen from

$$(I - P)Px = Px - P^2x = 0,$$

and $Y \supset \mathcal{N}(I - P)$ follows if we note that $(I - P)x = 0$ implies $x = Px$. Hence Y is closed. Finally $P|_Y$ is the identity operator on Y since writing $y = Px$, we have $Py = P^2x = Px = y$. □

The following proposition is easy to verify.

Proposition 8.15. *For any projection P on a Hilbert space H, the following are true:*

1. $\langle Px, x\rangle = \|Px\|^2$;
2. $P \geq 0$;
3. $\|P\| \leq 1$ *and* $\|P\| = 1$ *if* $P(H) \neq \{0\}$.
4. $S^{-1}PS$ *is a projection if* S *is unitary.*

Proposition 8.16.

1. *Let P_1, P_2 be projections. Then $P = P_1P_2$ is a projection on H if and only if the projections P_1 and P_2 commutes. Then P projects H onto $Y = Y_1 \cap Y_2$, where $Y_j = P_j(H)$.*
2. *Two closed subspaces Y and V of H are orthogonal if and only if the corresponding projections satisfy $P_YP_V = 0$.*

Proof. 1. Suppose that $P_1P_2 = P_2P_1$. Then P is self-adjoint. It is easy to see that $P^2 = P$. Hence P is a projection and for every $x \in H$, we have

$$Px = P_1(P_2x) = P_2(P_1x).$$

So $Px \in Y_1 \cap Y_2$ and it is clear that $P(H) = Y_1 \cap Y_2$.
Conversely, if $P = P_1P_2$ is a projection defined on H, then P is self-adjoint and $P_1P_2 = P_2P_1$ follows from Lemma (3.56).

2. If $Y \perp V$, then $Y \cap V = \{0\}$ and $P_Y P_V x = 0$ for all $x \in H$. So $P_Y P_V = 0$. On the other hand, if $P_Y P_V = 0$, then for every $y \in Y$ and $v \in V$, we obtain

$$\langle y, v\rangle = \langle P_Y y, P_V v\rangle = \langle y, P_Y P_V v\rangle = \langle y, 0\rangle = 0.$$

Hence $Y \perp V$.

□

Proposition 8.17. *Let P_1 and P_2 be projections on a Hilbert space H. Then*

1. *The sum $P = P_1 + P_2$ is a projection on H if and only if $Y_1 = P_1(H)$ and $Y_2 = P_2(H)$ are orthogonal.*

2. *If $P = P_1 + P_2$ is a projection, P projects H onto $Y = Y_1 \oplus Y_2$.*

Remark 8.17.1. *This can easily be extended to the sum of n projections.*

Proof. (2) is clear if we prove (1). If $P = P_1 + P_2$ is a projection, then

$$P_1 + P_2 = (P_1 + P_2)^2 = P_1^2 + P_1P_2 + P_2P_1 + P_2^2 = P_1 + P1P_2 + P2P_1 + P_2.$$

Hence we have

$$P_1P_2 + P_2P_1 = 0.$$

Multiplying by P_2 from the left, we get

$$P_2P_1P_2 + P_2P_1 = 0.$$

Multiplying $P_2$from the right, we have

$$2P_2P_1P_2 = 0,$$

this shows $P_2P_1 = 0$, and $Y_1 \perp Y_2$.

Conversely, if $Y_1 \perp Y_2$, then $P_1P_2 = P_2P_1 = 0$, so $P^2 = P$. □

Theorem 8.18. *Let P_1 and P_2 be projections defined on a Hilbert space H. Denote by $Y_1 = P_1(H)$ and $Y_2 = P_2(H)$ the subspaces onto which H is projected by P_1 and P_2, and let $\mathcal{N}(P_1)$ and $\mathcal{N}(P_2)$ be the null spaces of these projections. Then the following are equivalent:*

1. *$P_2P_1 = P_1P_2 = P_1$;*

2. *$Y_1 \subset Y_2$;*

3. *$\mathcal{N}(P_1) \supset \mathcal{N}(P_2)$;*

4. *$\|P_1x\| \le \|P_2x\|$ for all $x \in H$;*

5. *$P_1 \le P_2$.*

Theorem 8.19. *Let P_1 and P_2 be projections on a Hilbert space H. Then*

1. *The difference $P = P_2 - P_1$ is a projection on H if and only if $Y_1 \subset Y_2$ where $Y_j = P_j(H)$;*

2. *If $P = P_2 - P_1$ is a projection, P projects H onto Y, where Y is the orthogonal complement of Y_1 in Y_2.*

Proof. We only prove (1). If $P = P_2 - P_1$ is a projection, then

$$P_2 - P_1 = (P_2 - P_1)^2 = P_2^2 - P_2P_1 - P_1P_2 + P_1^2$$

On the right, $P_2^2 = P_2$ and $P_1^2 = P_1$, so

$$P_1P_2 + P_2P_1 = 2P_1.$$

Then $P_2P_1P_2 = P_2P_1$, $P_2P_1P_2 = P_1P_2$, so

$$P_2P_1 = P_1P_2 = P_1.$$

This shows $Y_1 \subset Y_2$.

Conversely, if $Y_1 \subset Y_2$, going backwards, it is clear that P is idempotent. P is self-adjoint since P_1, P_2 are self-adjoint. Hence P is a projection. □

Lemma 8.20. *If (P_n) is a sequence of projections defined on a Hilbert space H and $P_n \to P$, then P is a projection defined on H.*

Proof. The limit of self-adjoint operator is self-adjoint. So suffices to prove P is idempotent. But since $P_n \to P$, then $P_nx \to Px$ and $P_nx = P_n^2x \to P^2x$, hence $P^2x = Px$. □

Theorem 8.21 (Monotone Increasing Sequence)**.** *Let (P_n) be a monotone increasing sequence of projections P_n defined on a Hilbert space. Then*

1. *(P_n) is strongly operator convergent, say $P_nx \to Px$ for every $x \in H$, then the limit operator P is a projection defined on H.*
2. *P projections H onto*
$$P(H) = \overline{\bigcup_{n=1}^{\infty} P_n(H)}.$$
3. *P has the null space*
$$\mathcal{N}(P) = \bigcap_{n=1}^{\infty} \mathcal{N}(P_n).$$

8.4 Spectral Family

Definition 8.22 (Spectral Family)**.** *A real* ***spectral family*** *(or real decomposition of unity) is a one-parameter family $\mathscr{E} = (E_\lambda)_{\lambda\in\mathbb{R}}$ of projections E_λ defined on a Hilbert space H which depends on a real parameter λ and is such that*

1. *For any $\lambda < \mu$, $E_\lambda \leq E_\mu$, hence $E_\lambda E_\mu = E_\mu E_\lambda = E_\lambda$.*
2. $$\lim_{\lambda\to-\infty} E_\lambda x = 0.$$
3. $$\lim_{\lambda\to+\infty} E_{|}lambdax = x.$$

4.

$$E_{\lambda+0} = \lim_{\mu\to\lambda^+} E_\mu x = E_\lambda x.$$

$\mathscr{E}$ is called a ***spectral family on an interval*** *$[a,b]$ if*

$$E_\lambda = 0 \text{ for } \lambda < a, \quad E_\lambda = I \text{ for } \lambda \geq b.$$

Definition 8.23 (Positive Part)**.** *Given a bounded self-adjoint linear operator $T : H \to H$ on a complex Hilbert space H. We denote B_λ to be the positive square root of $T_\lambda^2 = (T - \lambda I)^2$. We define the* ***positive part*** *of T_λ to be the operator*

$$T_\lambda^+ = \frac{1}{2}(B_\lambda + T_\lambda).$$

We define the ***negative part*** *of T_λ to be the operator*

$$T_\lambda^- = \frac{1}{2}(B - \lambda - T_\lambda).$$

If $\lambda = 0$, then we simply denote B, T^+ and B^-.

The spectral family of $\mathscr{E}$ of T is then defined by $\mathscr{E} = (E_\lambda)_{\lambda\in\mathbb{R}}$, where E_λ is the projection of H onto the null space $\mathcal{N}(T_\lambda^+)$ of T_λ^+. If $\lambda = 0$, then we simply denote the projection operator by $E : H \to Y = \mathcal{N}(T^+)$.

Remark 8.23.1. *It is clear from the definition that a bounded self-adjoint operator is positive iff $T^+ = T$ and $T^- = 0$.*

Lemma 8.24. *The operators defined above have the following properties:*

1. *B, T^+ and T^- are bounded and self-adjoint;*
2. *B, T^+ and T^- commute with every bounded linear operator that T commutes with; in particular*

$$BT = TB, \quad T^+T = TT^+, \quad T^-T = TT^-, \quad T^+T^- = T^-T^+.$$

3. *E commutes with every bounded self-adjoint linear operator that T commutes with; in particular,*

$$ET = TE, \quad EB = BE.$$

4. *The following are ture:*
   - *$T = T^+ - T^-$ and $B = T^+ + T^-$.*
   - *$T^+T^- = 0$ and $T^-T^+ = 0$.*
   - *$T^+E = ET^+ = 0$ and $T^-E = ET^- = T^-$.*
   - *$TE = -T^-$ and $T(I - E) = T^+$.*
   - *$T^+ \geq 0$ and $T^- \geq 0$.*

Proof.

1. T and B are both bounded and self-adjoint.

2. Suppose that $TS = ST$. Then $T^2S = TST = ST^2$, and $BS = SB$ follows from Theorem (8.13). Hence

$$T^+S = \frac{1}{2}(BS + TS) = \frac{1}{2}(SB + ST) = ST^+.$$

   Similarly, T^- will commute with any bounded self-adjoint operator that commutes with T.

3. For every $x \in H$, we have $y = Ex \in Y = \mathcal{N}(T^+)$. Hence $T^+y = 0$ and $ST^+y = S0 = 0$. From $TS = ST$ and (2), we have $ST^+ = T^+S$, so

$$T^+SEx = T^+Sy = ST^+y = 0.$$

   Hence $SEx \in Y$. Since E projects H onto Y¡ we thus have $ESEx = SEx$ for every $x \in H$, that is $ESE = SE$. Thus

$$ES = E^*S^* = (SE)^* = (ESE)^* = E^*S^*E^* = ESE = SE.$$

4. Direct verification.

□

Lemma 8.25. *Lemma (8.24) remains true if we replace*

$$T, B, T^+, T^-, E \quad \textit{by} \quad T_\lambda, B_\lambda, T_\lambda^+, T_\lambda^-, E_\lambda,$$

respectively, where $\lambda \in \mathbb{R}$. Moreover, for any real $\kappa, \lambda, \mu, \nu, \tau$, the following operators all commute:

$$T_\kappa, B_\lambda, T_\mu^+, T_\nu^-, E_\tau.$$

Theorem 8.26 (Spectral Family)**.** *Let $T : H \to H$ be a bounded self-adjoint linear operator on a complex Hilbert space H. Furthermore, let E_λ be the projection of H onto the null space $Y_\lambda = \mathcal{N}(T_\lambda^+)$ of the positive part T_λ^+ of $T_\lambda = T - \lambda I$. Then $\mathscr{E} = (E_\lambda)_{\lambda \in \mathbb{R}}$ is a spectral family on the intervals $[m, M] \subset \mathbb{R}$, where*

$$m = \inf_{\|x\|=1} \langle Tx, x \rangle, \quad M = \sup_{\|x\|=1} \langle Tx, x \rangle.$$

Proof. We prove the following:

1. $\lambda < \mu$ implies $E_\lambda \le E_\mu$;
2. $\lambda < m$ implies $E_\lambda = 0$;
3. $\lambda \ge M$ implies $E_\lambda = I$;
4. $\mu \to \lambda^+$ implies $E_\mu x \to E_\lambda x$.

The proof are as follows:

1. Let $\lambda < \mu$. We have $T_\lambda = T_\lambda^+ - T_\lambda^- \le T_\lambda^+$ because $-T^- \le 0$. Hence

$$T_\lambda^+ - T_\mu \ge T_\lambda - T_\mu = (\mu - \lambda)I \ge 0.$$

   $T_\lambda^+ - T_\mu$ is self-adjoint and commutes with T_μ^+, and $T_\mu^+ \ge 0$. Thus

$$T_\mu^+(T_\lambda^+ - T_\mu) = T_\mu^+(T_\lambda^+ - T_\mu^+ + T_\mu^-) \ge 0.$$

Since $T_\mu^+ T_\mu^- = 0$. Then $T_\mu^+ T_\lambda^+ \geq (T_\mu^+)^2$. This shows that for any $x \in H$,

$$\langle T_\mu^+ T_\lambda^+ x, x\rangle \geq \langle (T_\mu^+)^2 x, x\rangle \geq 0.$$

This shows that $T_\lambda^+ x = 0$ implies $T_\mu^+ x = 0$. Hence $\mathcal{N}(T_\lambda^+) \subset \mathcal{N}(T_\mu^+)$, that is $E_\lambda \leq E_\mu$.

2. Let $\lambda < m$. Suppose towards a contradiction that $E_\lambda \neq 0$. Then $E_\lambda z \neq 0$ for some z. We set $x = E_\lambda z$. Then $E_\lambda x = E_\lambda^2 z = E_\lambda z = x$, and we may assume $\|x\| = 1$ without the loss of generality. It follows that

$$\begin{aligned}\langle T_\lambda E_\lambda x, x\rangle &= \langle T_\lambda x, x\rangle \\ &= \langle Tx, x\rangle - \lambda \\ &\geq \inf_{\|\tilde{x}\|=1} \langle T\tilde{x}, \tilde{x}\rangle - \lambda \\ &= m - \lambda > 0.\end{aligned}$$

But this contradicts $T_\lambda E_\lambda = -T_\lambda^- \leq 0$.

3. Suppose that $\lambda > M$ but $E_\lambda \neq I$, so that $I - E_|lambda \neq 0$. Then $(I - E_\lambda)x = x$ for some x of norm $\|x\| = 1$. Hence

$$\begin{aligned}\langle T_\lambda (I - E_\lambda)x, x\rangle &= \langle T_\lambda x, x\rangle \\ &= \langle Tx, x\rangle - \lambda \\ &\leq \sup_{\|\tilde{x}\|=1} \langle T\tilde{x}, \tilde{x}\rangle - \lambda \\ &= M - \lambda < 0\end{aligned}$$

But this contradicts $T_\lambda(I - E_\lambda) = T_\lambda^+ \geq 0$. $E_M = I$ follows from the right continuity which we will prove now.

4. Let $\Delta = (\lambda, \mu]$. We associate the operator $E(\Delta)$ to Δ by

$$E(\Delta) = E_\mu - E_\lambda.$$

Since $\lambda < \mu$, we have $E_\lambda \leq E_\mu$. Hence $E_\lambda(H) \subset E_\mu(H)$ and $E(\Delta)$ is a projection by Theorem (8.19). Also $E(\Delta) \geq 0$. Then

$$E_\mu E(\Delta) = E_\mu^2 - E_\mu E_\lambda = E_\mu - E_\lambda = E(\Delta)$$

and

$$(I - E_\lambda)E(\Delta) = E(\Delta) - E_\lambda(E_\mu - E_\lambda) = E(\Delta).$$

Since $E(\Delta)$, T_μ^- and T_λ^+ are positive and commutes. The products $T_\mu^- E(\Delta)$ and $T_\lambda^+ E(\Delta)$ are positive. Hence

$$T_\mu E(\Delta) = T_\mu E_\mu E(\Delta) = -T_\mu^{-1} E(\Delta) \leq 0$$
$$T_\lambda E(\Delta) = T_\lambda (I - E_\lambda) E(\Delta) = T_\lambda^+ E(\Delta) \geq 0.$$

This implies $TE(\Delta) \le \mu E(\Delta)$ and $TE(\Delta) \ge \lambda E(\Delta)$. Together,

$$\lambda E(\Delta) \le TE(\Delta) \le \mu E(\Delta). \tag{8.2}$$

We keep λ fixed and let $\mu \to \lambda$ from the right in a monotone fashion. Then $E(\Delta)x \to P(\lambda)x$. Since $E(\Delta)$ is idempotent, $P(\lambda)$ is a projection. Also $\lambda P(\lambda) = TP(\lambda)$, that is $T_\lambda P(\lambda) = 0$. From this, we have

$$T_\lambda^+ P(\lambda) = T_\lambda (I - E_\lambda) P(\lambda) = (I - E_\lambda) T_\lambda P(\lambda) = 0.$$

Hence $T_\lambda^+ P(\lambda)x = 0$ for all $x \in H$. This shows that $P(\lambda)x \in \mathcal{N}(T_\lambda^+)$. By definition, E_λ projects H onto $\mathcal{N}(T_\lambda^+)$. Consequently, we have $E_\lambda P(\lambda)x = P(\lambda)x$, so $E_\lambda P(\lambda) = P(\lambda)$. On the other hand, if we let $\mu \to 0^+$, in $(I - E_\lambda)E(\Delta) = E(\Delta)$, we have

$$(I - E_\lambda)P(\lambda) = P(\lambda).$$

We conclude that $P(\lambda) = 0$, i.e., $\mathscr{E}$ is continuous from the right.

□

8.5 Spectral Theorem

Theorem 8.27 (Spectral Theorem)**.** *Let $T : H \to H$ be a bounded self-adjoint linear operator on a complex Hilbert space H. Then:*

1. *T has the spectral representation*

$$T = \int_{m-0}^{M} \lambda dE_\lambda$$

where $\mathscr{E} = (E_\lambda)$ is the spectral family associated with T; the integral is to be understood in the sense of uniform operator convergence for all $x, y \in H$,

$$\langle Tx, y\rangle = \int_{m-0}^{M} \lambda dw(\lambda),$$

where

$$w(\lambda) = \langle E_\lambda x, y\rangle,$$

and the integral is an ordinary Riemann-Stieltjes integral.

2. *More generally, if p is a polynomial in λ with real cooeficients, say*

$$p(\lambda) = \alpha_n \lambda^N + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_0,$$

then the operator $p(T)$ has the spectral representation

$$p(T) = \int_{m-0}^{M} p(\lambda) dE_\lambda$$

and for all $x, y \in H$,

$$\langle p(T)x, y\rangle = \int_{m-0}^{M} p(\lambda) dw(\lambda).$$

Remark 8.27.1. *$m-0$ is written to indicate that one must take into account a contribution at $\lambda = m$ which occurs if $E_m \neq 0$, thus using any $a < m$, we can write*

$$\int_a^M \lambda dE_\lambda = \int_{m-0}^M \lambda dE_|lambda = mE_m + \int_m^M \lambda dE_\lambda.$$

Proof. We choose a sequence $(\mathcal{P}_n)$ of partitions of $(a, b]$ (where $a < m$ and $M < b$) into the intervals

$$\Delta_{nj} = (\lambda_{nj}, \mu_{nj}], \quad j = 1, 2, \cdots, n$$

of length $\ell(\Delta_{nj}) = \mu_{nj} - \lambda_{nj}$. Note that $\mu_{nj} = \lambda_{n,j+1}$ for $j = 1, \cdots, n-1$. We assume the sequence $(\mathcal{P}_n)$ be such that

$$\eta(\mathcal{P}_n) = \max_j \ell(\Delta_{nj}) \to 0.$$

Then by Inequality (8.2), we have

$$\lambda_{nj} E(\Delta_{nj}) \leq TE(\Delta_{nj}) \leq \mu_{nj} E(\Delta_{nj}).$$

By summation over j from 1 to n we obtain for every n

$$\sum_{j=1}^n \lambda_{nj} E(\Delta_{nj}) \leq \sum_{j=1}^n TE(\Delta_{nj}) \leq \sum_{j=1}^n \mu_{nj} E(\Delta_{nj}).$$

Since $\mu_{nj} = \lambda_{n,j+1}$ for $j = 1, \cdots, n-1$, then we have

$$T \sum_{j=1}^n E(\Delta_{nj}) = T \sum_{j=1}^n (E_{\mu_{n_j}} - E_{\lambda_{n_j}}) = T(I - 0) = T.$$

Then letting $\eta(\mathcal{P}_\backslash) \to 0$, we have

$$\sum_{j=1}^n \mu_{nj} E(\Delta_{nj}) - \sum_{j=1}^n \lambda_{nj} E(\Delta_{nj}) = \sum_{j=1}^n (\mu_{nj} - \lambda_{nj}) E(\Delta_{nj}) \to 0.$$

Hence

$$\left\| T - \sum_{j=1}^n \lambda_{nj} E(\Delta_{nj}) \right\| \to 0.$$

Next we prove the theorem for polynomials. By linearity, suffices to prove for the case $p(\lambda) = \lambda^r$, where $r \in \mathbb{N}$. For any $\kappa < \lambda \leq \mu < \nu$, we have

$$\begin{aligned}(E_\lambda - E_\kappa)(E_\mu - E_\nu) &= E_\lambda E_\mu - E_\lambda E_\nu - E_\kappa E_\mu + E_\kappa E_\nu \\ &= E_\lambda - E_\lambda - E_\kappa + E_\kappa = 0.\end{aligned}$$

This shows that $E(\Delta_{nj})E(\Delta_{nk}) = 0$ for $j \neq k$. Also since $E(\Delta_{nj})$ is a projection, $E(\Delta_{nj})^s = E(\Delta_{nj})$ for every

$s = 1, 2, \cdots$. Consequently, we have

$$\left[\sum_{j=1}^{n} \lambda_{nj} E(\Delta_{nj})\right]^r = \sum_{j=1}^{n} \lambda_{nj}^r E(\Delta_{nj}).$$

Hence

$$\left\|T^r - \sum_{j=1}^{n} \lambda_{nj}^r E(\Delta_{nj})\right| = \left\|T^r - \left[\sum_{j=1}^{n} \lambda_{nj} E(\Delta_{nj})\right]^r\right\| \to 0.$$

□

Lemma 8.28 (Properties of $p(T)$). *Let $T : H \to H$ be a bounded self-adjoint linear operator on a complex Hilbert space H. Then:*

1. *$p(T)$ is self-adjoint.*
2. *If $p(\lambda) = \alpha p_1(\lambda) + \beta p_2(\lambda)$, then $p(T) = \alpha p_1(T) + \beta p_2(T)$.*
3. *If $p(\lambda) = p_1(\lambda) p_2(\lambda)$, then $p(T) = p_1(T) p_2(T)$.*
4. *If $p(\lambda) \geq 0$ for all $\lambda \in [m, M]$, then $p(T) \geq 0$.*
5. *If $p_1(\lambda) \leq p_2(\lambda)$ for all $\lambda \in [m, M]$, then $p_1(T) \leq p_2(T)$.*
6. *$\|p(T)\| \leq \max_{\lambda \in J} |p(\lambda)|$, where $J = [m, M]$.*
7. *If a bounded linear operator commutes with T, it also commutes with $p(T)$.*

Proof. We prove (4), (5) and (6).

For (4), since p has real cooeficients, complex zeros must come in conjugate pairs. Then we can factorize p as

$$p(\lambda) = \alpha \prod_j (\lambda - \beta_j) \prod_k (\gamma_k - \lambda) \prod_\ell [(\lambda - \mu_\ell)^2 + \nu_\ell^2]$$

where $\beta_j \leq m$ and $\gamma_k \geq M$. Then it is clear that $\alpha > 0$ if $p \neq 0$ (otherwise, it is trivial). Then by replacing λ by T, each of the factors is a positive operator, and clearly they commutes. Hence we establish (4). And now (5) immediately follows from (4) by considering $p = p_2 - p_1$.

It remains to prove (6). Let k denote the maximum of $|p(\lambda)|$ on J. Then $0 \leq p(\lambda)^2 \leq k^2$ for $\lambda \in J$. Hence

$$p(T)^2 \leq k^2 I$$

by (5). So

$$\langle p(T)x, p(T)x\rangle = \langle p(T)^2 x, x\rangle \leq k^2 \langle x, x\rangle.$$

This shows that $\|P(T)\| \leq k$. □

Next, we want to define $f(T)$ for general continuous function f. By the Weierstrass's Theorem, for any continuous function on $[m, M]$, there exists a sequence of polynomials (p_n) with real coefficients such that

$$p_n(\lambda) \to f(\lambda)$$

uniformly on $[m, M]$. By Lemma (8.28), we have

$$\|p_n(T) - p_r(T)\| \leq \max_{\lambda \in J} |p_n(\lambda) - p_r(\lambda)|.$$

Thus $p_n(T)$ is Cauchy, hence we define the limit to be $f(T)$. Note that this limit is independent of the choice of the approximation sequence. Suppose we have two sequences of polynomials approximating p, then join them alternatingly, we note that the new sequence formed is again a Cauchy sequence converging to the same limit.

Theorem 8.29 (Spectral Theorem Generalized)**.** *Let $T : H \to H$ be a bounded self-adjoint linear operator on a complex Hilbert space H and f a continuous real-valued function on a complex Hilbert space H and f a continuous real-valued function on $[m, M]$. Then $f(T)$ has the spectral representation*

$$f(T) = \int_{m-0}^{M} f(\lambda) dE_\lambda$$

where $\mathscr{E} = (E_\lambda)$ is the spectral family associated with T; the integral is to be understood in the sense of uniform operator convergence, and for all $x, y \in H$,

$$\langle f(T)x, y \rangle = \int_{m-0}^{M} f(\lambda) dw(\lambda)$$

where

$$w(\lambda) = \langle E_\lambda x, y \rangle.$$

Proof. This is true since we can approximate continuous functions by polynomials. □

Lemma 8.30 (Properties of $f(T)$)**.** *Let $T : H \to H$ be a bounded self-adjoint linear operator on a complex Hilbert space H. Then:*

1. *$f(T)$ is self-adjoint.*
2. *If $f(\lambda) = \alpha f_1(\lambda) + \beta f_2(\lambda)$, then $f(T) = \alpha f_1(T) + \beta f_2(T)$.*
3. *If $f(\lambda) = f_1(\lambda) f_2(\lambda)$, then $f(T) = f_1(T) f_2(T)$.*
4. *If $f(\lambda) \geq 0$ for all $\lambda \in [m, M]$, then $f(T) \geq 0$.*
5. *If $f_1(\lambda) \leq f_2(\lambda)$ for all $\lambda \in [m, M]$, then $f_1(T) \leq f_2(T)$.*
6. *$\|f(T)\| \leq \max_{\lambda \in J} |f(\lambda)|$, where $J = [m, M]$.*
7. *If a bounded linear operator commutes with T, it also commutes with $f(T)$.*

Theorem 8.31. *Let $T : H \to H$ be a bounded self-adjoint linear operator on a complex Hilbert space H and $\mathscr{E} = (E_\lambda)$ the corresponding spectral family. Then $\lambda \mapsto E_\lambda$ has a discontinuity at any $\lambda = \lambda_0$ (That is $E_{\lambda_0} \neq E_{\lambda_0 - 0}$) if and only if λ_0 is an eigenvalue of T. IN this case, the corresponding eigenspace is*

$$\mathcal{N}(T - \lambda_0 I) = (E_{\lambda_0} - E_{\lambda_0 - 0})(H).$$

Proof. λ_0 is an eigenvalue of T iff $\mathcal{N}(T - \lambda_0 I) \neq \{0\}$, so the first statement of the theorem follows immediately from

$$\mathcal{N}(T - \lambda_0 I) = (E_{\lambda_0} - E_{\lambda_0 - 0})(H).$$

We write $F_0 = E_{\lambda_0} - E_{\lambda_0 - 0}$. We first show that $F_0(H) \subset \mathcal{N}(T - \lambda_0 I)$. Recall Equation (8.2), let $\lambda = \lambda_0 - \frac{1}{n}$ and $\mu = \lambda_0$, we get

$$\left(\lambda_0 - \frac{1}{n}\right) E(\Delta_0) \leq TE(\Delta_0) \leq \lambda_0 E(\Delta_0).$$

Let $N \to \infty$, then $E(\Delta_0) \to F_0$, so we have

$$\lambda_0 F_0 \leq TF_0 \leq \lambda_0 F_0.$$

Hence $TF_0 = \lambda_0 F$, that is $(T - \lambda_0 I)F_0 = 0$, this shows $F_0(H) \subset \mathcal{N}(T - \lambda_0 I)$.

Next we show that $F_0(H) \supset \mathcal{N}(T - \lambda_0 I)$. Let $x \in \mathcal{N}(T - \lambda_0 I)$, we show that $F_0 x = x$. If $\lambda_0 \not\in [m, M]$, then $\lambda_0 \in \rho(T)$. Hence in this case, $\mathcal{N}(T - \lambda_0 I) = \{0\} \subset F_0(H)$. Let $\lambda_0 \in [m, M]$. By assumption, $(T - \lambda_0 I)x = 0$, then $(T - \lambda_0 I)^2 x = 0$. Hence by the spectral theorem,

$$\int_a^b (\lambda - \lambda_0)^2 dw(\lambda) = 0, \quad w(\lambda) = \langle E_\lambda x, x, \rangle$$

where $a < m$ and $b > m$. Here $(\lambda - \lambda_0)^2 \geq 0$ and $\lambda \mapsto \langle E_\lambda x, x \rangle$ is monotone increasing. Hence the integral over any subinterval of positive length must be zero. In particular, for every $\epsilon > 0$ we must have

$$0 = \int_a^{\lambda_0 - \epsilon} (\lambda - \lambda_0)^2 dw(\lambda) \geq \epsilon^2 \int_a^{\lambda_0 - \epsilon} dw(\lambda) = \epsilon^2 \langle E_{\lambda_0 - \epsilon} x, x \rangle$$

and

$$0 = \int_{\lambda_0 + \epsilon}^b (\lambda - \lambda_0)^2 dw(\lambda) \geq \epsilon^2 \int_{\lambda_0 + \epsilon}^b dw(\lambda) = \epsilon^2 \langle Ix, x \rangle - \epsilon^2 \langle E_{\lambda_0 + \epsilon} x, x \rangle.$$

Since $\epsilon > 0$, then we have

$$\langle E_{\lambda_0 - \epsilon} x, x \rangle = 0 \Longrightarrow E_{\lambda_0 - \epsilon} x = 0$$

and

$$\langle x - E_{\lambda_0 + \epsilon} x, x \rangle = 0 \Longrightarrow x - E_{\lambda_0 + \epsilon} x = 0.$$

We may thus write

$$x = (E_{\lambda_0 + \epsilon} - E_{\lambda_0 - \epsilon})x.$$

Let $\epsilon \to 0$, by the right continuity, we have $x = F_0 x$. □

Theorem 8.32. *Let $T : H \to H$ be a bounded self-adjoint linear operator on a complex Hilbert space H and*

$\mathscr{E} = (E_\lambda)$ *the corresponding spectral family. Then a real* λ_0 *belongs to the resolvent set* $\rho(T)$ *of* T *if and only if there is a* $\gamma > 0$ *such that* $\mathscr{E} = (E_\lambda)$ *is a constant on the interval* $[\lambda_0 - \gamma, \lambda_0 + \gamma]$.

Proof. We make use of the fact that $\lambda_0 \in \rho(T)$ if and only if

$$\|(T - \lambda_0 I)x\| \geq \gamma \|x\|.$$

Suppose that λ_0 is real and such that $\mathscr{E}$ is constant on $J = [\lambda_0 - \gamma, \lambda_0 + \gamma]$ for some $\gamma > 0$. Then

$$\|(T - \lambda_0 I)x\|^2 = \langle (T - \lambda_0 I)^2 x, x\rangle = \int_{m-0}^{M} (\lambda - \lambda_0)^2 d\langle E_\lambda x, x\rangle.$$

Since $\mathscr{E}$ is constant on J, integration over J yields the value zero, and for $\lambda \notin J$, we have $(\lambda - \lambda_0)^2 \geq \gamma^2$. Hence we have

$$\|(T - \lambda_0 I)x\|^2 \geq \gamma^2 \int_{m-0}^{M} d\langle E_\lambda x, x\rangle = \gamma^2 \langle x, x, \rangle.$$

Conversely, suppose that $\lambda_0 \in \rho(T)$. Then there exists $\gamma > 0$ such that

$$\|(T - \lambda_0 I)x\| \geq \gamma \|x\|.$$

So

$$\int_{m-0}^{M} (\lambda - \lambda_0)^2 d\langle E_\lambda x, x\rangle \geq \gamma^2 \int_{m-0}^{M} d\langle E_\lambda x, x\rangle. \tag{8.3}$$

We show that we obtain a contradiction if we assume that $\mathscr{E}$ is not constant on the interval J. Then we can find positive $\eta < \gamma$ such that $E_{\lambda_0+\eta} - E_{\lambda_0-\eta} \neq 0$. Hence there is a $y \in H$ such that

$$x = (E_{\lambda_0+\eta} - E_{\lambda_0-\eta})y \neq 0.$$

Then

$$E_\lambda x = E_\lambda (E_{\lambda_0+\eta} - E_{\lambda_0-\eta})y.$$

But then substitute this into (8.3), shows this is 0 if $\lambda < \lambda_0 - \eta$ and $(E_{\lambda_0+\eta} - E_{\lambda_0-\eta})y$ when $\lambda > \lambda_0 + \eta$. Let $K = [\lambda_0 - \eta, \lambda_0 + \eta]$. If $\lambda \in k$, then

$$\langle E_\lambda x, x\rangle = \langle (E_\lambda - E_{\lambda_0-\eta} y, y\rangle.$$

So

$$\int_{\lambda_0-\eta}^{\lambda_0+\eta} (\lambda - \lambda_0)^2 d\langle E_\lambda y, y\rangle \geq \gamma^2 \int_{\lambda_0-\eta}^{\lambda_0+\eta} d\langle E_\lambda y, y\rangle.$$

However this is a contradiction. □

Corollary 8.32.1. *Let* $T : H \to H$ *be a bounded self-adjoint linear operator on a complex Hilbert space* H *and* $\mathscr{E} = (E_\lambda)$ *the corresponding spectral family. Then a real* λ_0 *belongs to the continuous spectrum* $\sigma_c(T)$ *of* T *if and only if* $\mathscr{E}$ *is continuous at* λ_0 *and is not constant in any neighbourhood of* λ_0 *on* $\mathbb{R}$.

Proof. Since $\sigma_r(T) = \emptyset$. □

9 Spectral Theory of Unbounded Operators

9.1 Unbounded Linear Operators

Throughout this chapter, we consider linear operators $T : \mathscr{D}(T) \to H$ whose domain $\mathscr{D}(T)$ lies in a complex Hilbert space H and admit that such an operator T may be unbounded. We know that if T is self-adjoint and defined on H, then it is necessarily bounded. So $\mathscr{D}(T) = H$ is impossible for unbounded linear operators such that

$$\langle Tx, y \rangle = \langle x, Ty \rangle$$

for all $x, y \in H$.

We shall use the following notation for convenience sake. If T is an extension of S, then we denote $S \subset T$. We call an extension T of S a proper extension if $\mathscr{D}(S)$ is a proper subset of $\mathscr{D}(T)$.

Definition 9.1 (Densely Defined Operators)**.** *We say an operator T is **densely defined** in H if $\mathscr{D}(T)$ is dense in H.*

Given an unbounded operator T, we want to define T^* by such that

$$\langle Tx, y \rangle = \langle x, y^* \rangle, \quad \text{where} \quad y^* = T^* y.$$

However, the uniqueness of y^* is guaranteed if and only if T is densely defined. Indeed, if $\mathscr{D}(T)$ is not dense in H, then $\overline{D(T)} \neq H$, and its orthogonal complex in H contains a nonzero y_1. Then $y_1 \perp \mathscr{D}(T)$, hence

$$\langle x, y^* \rangle = \langle x, y^* + y_1 \rangle$$

which shows non-uniqueness. We agree to call T an operator on H if $\mathscr{D}(T)$ is all of H and an operator in H if $\mathscr{D}(T)$ lies in H.

Definition 9.2 (Hilbert Adjoint Operator)**.** *Let $T : \mathscr{D}(T) \to H$ be a densely defined linear operator in a complex Hilbert space H. Then the **Hilbert-adjoint operator** $T^* : \mathscr{D}(T^*) \to H$ of T is defined as follows. The domain $\mathscr{D}(T^*)$ of T^* consists of all $y \in H$ such that there is a $y^* \in H$ satisfying*

$$\langle Tx, y \rangle = \langle x, y^* \rangle$$

for all $x \in \mathscr{D}(T)$. For each such $y \in \mathscr{D}(T^*)$ the Hilbert-adjoint operator T^* is defined in terms of that y^* by

$$y^* = T^* y.$$

Remark 9.2.1. *T^* is indeed a linear operator.*

We also define composition and sum of linear operators (may be unbounded) on the maximal set where they make sense.

Lemma 9.3. *The following are easy to verify:*

1. $(T_1T_2)T_3 = T_1(T_2T_3)$;
2. $(T_1 + T_2)T_3 = T_1T_3 + T_2T_3$;
3. $T_1(T_2 + T_3) \supset T_1T_2 + T_1T_3$;
4. $(\alpha T)^* = \bar{\alpha}T^*$;
5. $(S + T)^* \supset S^* + T^*$.

Theorem 9.4. *Let $S : \mathscr{D}(S) \to H$ and $T : \mathscr{D}(T) \to H$ be linear operators which are densely defined in a complex Hilbert space H. Then:*

1. *If $S \subset T$, then $T^* \subset S^*$.*
2. *If $\mathscr{D}(T^*)$ is dense in H, then $T \subset T^{**}$.*

Proof.

1. By the definition of T^*, we have
$$\langle Tx, y\rangle = \langle x, T^*y\rangle$$
for all $x \in \mathscr{D}(T)$ and all $y \in \mathscr{D}(T^*)$. Since $S \subset T$, this implies
$$\langle Sx, y\rangle = \langle x, T^*y\rangle$$
for all $x \in \mathscr{D}(S)$ and $y \in \mathscr{D}(T^*)$. By the definition of S^*, we have
$$\langle x, T^*\rangle = \langle Sx, y\rangle = \langle x, S^*y\rangle$$
for all $x \in \mathscr{D}(S)$ and $y \in D(T^*)$.

2. Taking complex conjugate, we have
$$\langle T^*y, x\rangle = \langle y, Tx\rangle$$
for all $y \in \mathscr{D}(T^*)$ and all $x \in \mathscr{D}(T)$. Since $\mathscr{D}(T^*)$ is dense in H, the operator T^{**} exists and by definition
$$\langle T^*y, x\rangle = \langle y, T^{**}x\rangle$$
for all $y \in \mathscr{D}(T^*)$ and all $x \in \mathscr{D}(T^{**})$. Hence we see that $T \subset T^{**}$.

□

Proposition 9.5. *Let $T : \mathscr{D}(T) \to H$ be linear operators which are densely defined in a complex Hilbert space H be such that T is injective and its range $\mathscr{R}(T)$ is dense in H. Then T^* is injective and*
$$(T^*)^{-1} = (T^{-1})^*.$$

Proof. T^* exists since T is densely defined in H. Also T^{-1} exists since T is injective. $(T^{-1})^*$ exists since $\mathscr{D}(T^{-1}) = \mathscr{R}(T)$ is dense in H. We must show that $(T^*)^{-1}$ exists and satisfies $(T^*)^{-1} = (T^{-1})^*$.

Let $y \in \mathscr{D}(T^*)$. Then for all $x \in \mathscr{D}(T^{-1})$ we have $T^{-1}x \in \mathscr{D}(T)$ and

$$\langle T^{-1}x, T^*y\rangle = \langle TT^{-1}x, y\rangle = \langle x, y\rangle.$$

On the other hand, by the definition of the Hilbert-adjoint operator of T^{-1}, we have

$$\langle T^{-1}x, T^*y\rangle = \langle x, (T-1)^*T^*y\rangle$$

for all $x \in \mathscr{D}(T^{-1})$. This shows that $T^*y \in \mathscr{D}((T^{-1})^*)$. Furthermore, we conclude that

$$(T^{-1})^*T^*y = y$$

for any $y \in \mathscr{D}(T^*)$. We see that $T^*y = 0$ implies $Y = 0$. Hence $(T^*)^{-1} : \mathscr{R}(T^*) \to \mathscr{D}(T)^*$. Furthermore, since $(T^*)^{-1}T^*$ is the identity operator on $\mathscr{D}(T^*)$, this shows that $(T^*)^{-1} \subset (T^{-1})^*$. For the other inclusions, let $x \in \mathscr{D}(T)$ and $y \in \mathscr{D}((T^{-1})^*)$. Then $Tx \in \mathscr{R}(T) = \mathscr{D}(T^{-1})$ and

$$\langle Tx, (T^{-1})^*y\rangle = \langle T^{-1}Tx, y\rangle = \langle x, y\rangle.$$

On the other hand, by the definition of the Hilbert-adjoint operator of T, we have

$$\langle Tx, (T^{-1})^*y\rangle = \langle x, T^*(T^{-1})^*y\rangle$$

for all $x \in \mathscr{D}(T)$. Hence we conclude that $(T^{-1})^*y \in \mathscr{D}(T^*)$ and $T^*(T^{-1})^*y = y$ for any $y \in \mathscr{D}((T^{-1})^*)$. □

Lemma 9.6. *If S and T are such that ST is densely defined in H, then*

$$(ST)^* \supset T^*S^*,$$

and if S is defined on all of H and is bounded, then

$$(ST)^* = T^*S^*.$$

Definition 9.7 (Symmetric Linear Operator and Self-Adjoint Operator)**.** *Let $T : \mathscr{D}(T) \to H$ be a linear operator which is densely defined in a complex Hilbert space H. Then T is called a **symmetric linear operator** if and only if for all $x, y \in \mathscr{D}(T)$,*

$$\langle Tx, y\rangle = \langle x, Ty\rangle.$$

*T is called **self-adjoint** if*

$$T = T^*.$$

The following lemma is clear from the definition:

Lemma 9.8. *A densely defined linear operator T in a complex Hilbert space H is symmetric if and only if*

$$T \subset T^*.$$

Lemma 9.9. *Let H be a complex Hilbert space and $T : \mathscr{D}(T) \to H$ linear and dense defined in H. Then T is symmetric if and only if $\langle Tx, x\rangle$ is real for all $x \in \mathscr{D}(T)$.*

Proof. Suppose T is symmetric, then

$$\langle Tx, x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle}$$

this shows $\langle Tx, x\rangle \in \mathbb{R}$. Conversely, suppose $\langle Tx, x\rangle \in \mathbb{R}$ for all $x \in \mathscr{D}(T)$, then we see that T is symmetric. □

Lemma 9.10. *If T is symmetric then T^{**} is also symmetric.*

Proof. Since T is symmetric, $T \subset T^*$, hence $T^* \supset T^{**}$. But since $T^* \subset T^{***}$, then we have $T^{**} \subset T^{***}$, so T^{**} is symmetric. □

9.2 Closed Linear Operators and Closures

Recall the definition of a closed linear operator. It is defined irrespective whether an operator is bounded or unbounded.

Recall the following statements stated in the proposition.

Proposition 9.11. *Let $T : \mathscr{D}(T) \to H$ be a linear operator and $\mathscr{D}(T) \subset H$ is a subset in a complex Hilbert space. Then*

1. *T is closed if and only if whenever $x_n \to x$, $x_n \in \mathscr{D}(T)$ and $Tx_n \to y$, then $x \in \mathscr{D}(T)$ and $Tx = y$.*
2. *If T is closed and $\mathscr{D}(T)$ is closed. Then T is bounded (by closed graph theorem).*
3. *Let T be bounded. Then T is closed if and only if $\mathscr{D}(T)$ is closed.*

Theorem 9.12. *The Hilbert-adjoint operator T^* is closed.*

Proof. Let $(y_n) \in \mathscr{D}(T^*)$ be such that $y_n \to y_0$ and $T^*y_n \to z_0$. We show that $y_0 \in \mathscr{D}(T^*)$ and $z_0 = T^*y_0$. By definition, we have that for every $y \in \mathscr{D}(T)$,

$$\langle Ty, y_n\rangle = \langle y, T^*y_n\rangle.$$

Since the inner product is continuous, by letting $n \to \infty$, we have

$$\langle Ty, y_0\rangle = \langle y, z_0\rangle$$

for every $y \in \mathscr{D}(T)$. Hence $y_0 \in \mathscr{D}(T^*)$ and $z_0 = T^*y_0$. □

Definition 9.13 (Closable Operators, Closure)**.** *If a linear operator T has an extension T_1 which is a closed linear operator, then T is said to be* ***closable*** *and T_1 is called a* ***closed linear extension of*** *T. A closed linear extension $\bar{T}$ of a closable linear operator T is said to be minimal if every closed linear extension T_1 of T is a closed linear extension of $\bar{T}$. This minimal extension $\bar{T}$ of T is called the* ***closure of*** *T.*

Remark 9.13.1. *We note that the intersection of closed operators is closed. Hence we conclude that the closure is the intersection of all closed linear extensions, hence closure is unique.*

Theorem 9.14. *Let $\mathscr{D}(T) \to H$ be a linear operator where H is a complex Hilbert space and $\mathscr{D}(T)$ is dense in H. Then if T is symmetric, its closure $\bar{T}$ exists and is unique.*

Proof. Suffices to prove existence. We define $\bar{T}$ by first defining the domain $M = \mathscr{D}(\bar{T})$ and then $\bar{T}$ itself. Then we show that $\bar{T}$ is indeed the closure of T.

Let M be a set of all $x \in H$ for which there is a sequence (x_n) in $\mathscr{D}(T)$ and a $y \in H$ such that $x_n \to x$ and $Tx_n \to y$. Then it is not difficult to see that M is a vector space. Clearly $\mathscr{D}(T) \subset M$. On M we define $\bar{T}$ by setting

$$y = \bar{T}x$$

with y being the limit of Tx_n. Then we first show that this indeed gives a well-defined function.

Suppose x is arbitrary, and $(\tilde{x}_n)$ is another sequence such that $\tilde{x}_n \to x$ and $T\tilde{x}_n \to \tilde{y}$. Then by linearity of T, we have $Tx_n - T\tilde{x}_n = T(x_n - \tilde{x}_n)$. Since T is symmetric, then we have that for every $v \in \mathscr{D}(T)$,

$$\langle v, Tx_n - T\tilde{x_n} \rangle = \langle Tv, x_n - \tilde{x}_n \rangle.$$

Let $n \to \infty$, then we get

$$\langle v, y - \tilde{y} \rangle = \langle Tv, x - x \rangle = 0.$$

This shows that $y - \tilde{y} \perp \mathscr{D}(T)$. Since $\mathscr{D}(T)$ is dense in H, then $y = \tilde{y}$ hence $\bar{T}$ is well-defined. Moreover, by taking the constant sequence, we see that $\bar{T}$ is an extension of T.

Clearly $\bar{T}$ is also linear. WE show that $\bar{T}$ is in fact symmetric as well. Let $x, z \in \mathscr{D}(\bar{T})$ and $(x_n), (z_n) \subset \mathscr{D}(T)$ are such that $x_n \to x$, $Tx_n \to \bar{T}x$ and $z_n \to z$, $Tz_n \to \bar{T}z$. Since T is symmetric, $\langle z_n, Tx_n \rangle = \langle Tz_n, x_n \rangle$. Let $n \to \infty$, we obtain $\langle z, \bar{T}x \rangle = \langle \bar{T}z, x \rangle$. Hence $\bar{T}$ is symmetric.

Finally we show that $\bar{T}$ is closed and is the closure of T. It is clear that the closure of T must extends $\bar{T}$ defined above. Next by triangle inequality and parsing to sequences, one can easily show that $\bar{T}$ is closed. $\square$

Corollary 9.14.1. *Let $\mathscr{D}(T) \to H$ be a linear operator where H is a complex Hilbert space and $\mathscr{D}(T)$ is dense in H. Then if T is symmetric, then it is closable, and its closure is also symmetric.*

Theorem 9.15. *For a symmetric linear operator T. We have*

$$(\bar{T})^* = T^*.$$

Proof. Since $T \subset \bar{T}$, then $(\bar{T})^* \subset T^*$. Hence $\mathscr{D}((\bar{T})^*) \subset \mathscr{D}(T^*)$. We just need to show that if $y \in \mathscr{D}(T^*)$ implies $y \in \mathscr{D}((\bar{T})^*)$.

Let $y \in (T^*)$. Then for every $x \in \mathscr{D}(\bar{T})$, we show that

$$\langle \bar{T}x, y \rangle = \langle x, T^* y \rangle.$$

However, this is follows from the continuity of inner products. $\square$

Corollary 9.15.1. *If T is a symmetric linear operator, then T^{**} is a closed symmetric linear extension of T.*

9.3 Spectral Properties of General Self-adjoint operators

Lemma 9.16. *Let $T : \mathscr{D}(T) \to H$ be a linear operator whose Hilbert-adjoint operator T^* exists. If $\lambda \in \sigma_r(T)$, then $\overline{\lambda} \in \sigma_p(T^*)$.*

Proof. Consider

$$\langle Tx, y\rangle = \langle x, T^*y\rangle$$

where $y \perp \mathscr{R}(T)$. Then we have no choice but to make $T^*y = 0$. □

Theorem 9.17. *Let $T : \mathscr{D}(T) \to H$ be a self-adjoint linear operator which is densely defined in a complex Hilbert space H. Then a number λ belongs to the resolvent set $\rho(T)$ of T if and only if there exists a $c > 0$ such that for every $x \in \mathscr{D}(T)$*

$$\|T_\lambda x\| \geq c\|x\|$$

where $T_\lambda = T - \lambda I$.

Remark 9.17.1. *This is equivalent to say that $\lambda \in \sigma(T)$ if and only if we can find a sequence (x_n) in $\mathscr{D}(T)$ such that $\|x_n\| = 1$ and*

$$(T - \lambda I)x_n \to 0.$$

Theorem 9.18. *The spectrum $\sigma(T)$ of a self-adjoint linear operator $T : \mathscr{D}(T) \to H$ is real and closed where H is a complex Hilbert space, and $\mathscr{D}(T)$ is dense in H.*

Proof. We prove closedness of $\sigma(T)$. We show that the resolvent set $\rho(T)$ is open. Let $\lambda_0 \in \rho(T)$ and we show that every λ sufficiently closed to λ_0 also belongs to $\rho(T)$.

By the triangle inequality,

$$\begin{aligned}\|Tx - \lambda_0 x\| &= \|Tx - \lambda x + (\lambda - \lambda_0)x\| \\ &\leq \|Tx - \lambda x\| + |\lambda - \lambda_0|\|x\|.\end{aligned}$$

This can be written as

$$\|Tx - \lambda x\| \geq \|Tx - \lambda_0 x\| - |\lambda - \lambda_0|\|x\|.$$

Since $\lambda_0 \in \rho(T)$, then there is a $c > 0$ such that for all $x \in \mathscr{D}(T)$, $\|Tx - \lambda_0 x\| \geq c\|x\|$. Hence when $|\lambda - \lambda_0| \leq c/2$, we have

$$\|Tx - \lambda x\| \geq \frac{c}{2}\|x\|$$

this shows that $\rho(T)$ is open. □

Corollary 9.18.1. *The residual spectrum of a self-adjoint operator is empty, even if the operator is unbounded.*

Proof. Since $T = T^*$, and by Lemma (9.16), we have that if $\lambda \in \sigma_r(T)$, then $\bar{\lambda} \in \sigma_p(T)$. However, since $\lambda \in \mathbb{R}$, then this cannot happen. □

9.4 Spectral Representation of Self-Adjoint Operators

Proposition 9.19. *If $U : H \to H$ is a unitary linear operator on a complex Hilbert space $H \neq \{0\}$, then the spectrum $\sigma(U)$ is a closed subset of the unit circle; thus $|\lambda| = 1$ for every $\lambda \in \sigma(U)$.*

Proof. We know that $\|U\| = 1$, so $|\lambda| \leq 1$ for any $\lambda \in \sigma(U)$. Also $0 \in \rho(U)$ since for $\lambda = 0$, the resolvent operator of U is $U^{-1} = U^*$. The operator U^{-1} is unitary and $\|U^{-1}\| = 1$ implies that for every λ satisfying $|\lambda| < 1/\|U^{-1}\| = 1$ belongs to $\rho(U)$. Hence the spectrum of U must lie on the unit circle. □

Lemma 9.20. *Let*

$$h(\lambda) = \sum_{n=0}^{\infty} \alpha_n \lambda^n$$

be absolutely convergent for all λ such that $|\lambda| \leq k$. Suppose that $S \in B(H, H)$ is self-adjoint and has norm $\|S\| \leq k$ where H is a complex Hilbert space. Then

$$h(S) = \sum_{n=0}^{\infty} \alpha_n S^n$$

is a bounded self-adjoint linear operator and

$$\|h(S)\| \leq \sum_{n=0}^{\infty} |\alpha_n| k^n.$$

If a bounded linear operator commutes with S, it also commutes with $h(S)$.

Proof. The absolute convergence of the sequence implies the uniform convergence of the operators. We know that the limit of bounded self-adjoint operators over complex Hilbert space is self-adjoint. The rest of the statement in the lemma follows easily. □

Lemma 9.21 (Wecken's Lemma)**.** *Let W and A be bounded self-adjoint linear operators on a complex Hilbert space H. Suppose that $WA = AW$ and $W^2 = A^2$. Let P be the projection of H onto the null space $\mathcal{N}(W - A)$. Then*

1. *If a bounded linear operator commutes with $W - A$, it also commutes with P.*
2. *$Wx = 0$ implies $Px = x$.*
3. *We have $W = (2P - I)A$.*

Proof.

1. Suppose that B commutes with $W - A$. Since we have $Px \in \mathcal{N}(W - A)$ for every $x \in H$, we thus obtain

   $$(W - A)BPx = B(W - A)Px = 0.$$

   This shows that $BPx \in \mathcal{N}(W - A)$ and implies $P(BPx) = BPx$, that is

   $$PBP = BP.$$

   We show that $PBP = PB$. Since $W - A$ is self-adjoint, then

   $$(W - A)B^* = [B(W - A)]^* = [(W - A)B]^* = B^*(W - A).$$

This shows that $W - A$ commutes with B^*. Hence with the same reasons before, we have $PB^*P = B^*P$. Now since the projection operators are self-adjoints, then we have

$$PBP = (PB^*P)^* = (B^*P^*) = PB.$$

2. Let $Wx = 0$. Since A and W are self-adjoint with $A^2 = W^2$, we have

$$\|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^2x, x\rangle = \langle W^2x, x\rangle = \|Wx\|^2 = 0.$$

So $Ax = 0$. Hence $(W - A)x = 0$. This shows that $x \in \mathcal{N}(W - A)$, so $Px = x$.

3. From the assumption $W^2 = A^2$ and $WA = AW$, we have

$$(W - A)(W + A) = W^2 - A^2 = 0.$$

Hence $(W + A)x \in \mathcal{N}(W - A)$ for every $x \in H$. Since P projects H onto $\mathcal{N}(W - A)$, we thus obtain

$$P(W + A)x = (W + A)x$$

for every $x \in H$. Hence $P(W + A) = W + A$.

By part (1), we have $P(W - A) = (W - A)P = 0$. Hence

$$2PA = P(W + A) - P(W - A) = W + A.$$

This shows that $2PA - A = W$, so $W = (2P - I)A$.

□

Theorem 9.22 (Spectral Theorem for Unitary Operators)**.** *Let $U : H \to H$ be unitary operator on a complex Hilbert space $H \neq \{0\}$. Then there exists a spectral family $\mathscr{E} = (E_\theta)$ on $[-\pi, \pi]$ such that*

$$U = \int_{-\pi}^{\pi} e^{i\theta} dE_\theta = \int_{-\pi}^{\pi} (\cos\theta + i\sin\theta) dE_\theta.$$

More generally, for every continuous function f defined on the unit circle

$$f(U) = \int_{-\pi}^{\pi} f(e^{i\theta}) dE_\theta$$

where the integral is to be understood in the sense of uniform operator convergence and for all $x, y \in H$,

$$\langle f(U)x, y\rangle = \int_{-\pi}^{\pi} f(e^{i\theta}) dw(theta),$$

where $w(\theta) = \langle E_\theta x, y\rangle$.

Now for any self-adjoint operator $T : \mathscr{D}(T) \to H$ on a complex Hilbert space H, where $\mathscr{D}(T)$ is dense in H and T

may be unbounded. We associate T with the operator

$$U = (T - iI)(T + iI)^{-1}$$

which is called the **Cayley Transform** of T. This operator is well-defined, since $-i$ is not in the spectrum of T. Note that the linear fractional transformation transforms

$$u = \frac{t - i}{t + i}.$$

Lemma 9.23. *The Cayley transform*

$$U = (T - iI)(T + iI)^{-1}$$

of a self-adjoint linear operator $T : \mathscr{D}(T) \to H$ exists on H and is a unitary operator, where $H \neq \{0\}$ is a complex Hilbert space.

Proof. Since T is self-adjoint, then $\sigma(T)$ is real. Hence i and $-i$ belong to the resolvent set $\rho(T)$. Consequently, the inverses $(T+iI)^{-1}$ and $(T-iI)^{-1}$ exist on a dense subset of H and are bounded operators. Since $T = T^*$, then T is closed, thus we see that the inverses are defined on all of H. Hence U is well-defined and an automorphism on H. We show that U is an isometry. Let $x \in H$, and set $y = (T + iI)^{-1}x$, then

$$\begin{aligned}
\|Ux\|^2 &= \|(T - iI)y\|^2 \\
&= \langle Ty - iy, Ty - iy\rangle \\
&= \langle Ty, Ty\rangle + i\langle Ty, y\rangle - i\langle y, Ty\rangle + \langle iy, iy\rangle \\
&= \langle Ty + iy, Ty + iy\rangle \\
&= \|(T + iI)y\|^2 \\
&= \|(T + iI)(T + iI)^{-1}x\|^2 \\
&= \|x\|^2
\end{aligned}$$

□

Lemma 9.24. *If U is the Cayley transform of T, then*

$$T = i(I + U)(I - U)^{-1}.$$

Furthermore, 1 is not an eigenvalue of U.

Theorem 9.25 (Spectral Theorem for Self-adjoint linear operators)**.** *Let $T : \mathscr{D}(T) \to H$ be a self-adjoint linear operator where $H \neq \{0\}$ is a complex Hilbert space and $\mathscr{D}(T)$ is dense in H. Let U be the Cayley transform of T and (E_θ) the spectral family in the spectral representation (9.22) of $-U$. Then for all $x \in \mathscr{D}(T)$,*

$$\begin{aligned}
\langle Tx, x\rangle &= \int_{-\pi}^{\pi} \tan\frac{\theta}{2} dw(\theta) \quad w(\theta) = \langle E_\theta x, x\rangle \\
&= \int_{-\infty}^{+\infty} \lambda dv(\lambda) \quad v(\lambda) = \langle F - \lambda x, x\rangle
\end{aligned}$$

where $F_\lambda = E_{2\arctan\lambda}$.