Lecture notes for MA4211: Functional Analysis

WANG Dong

1 Normed spaces and Banach spaces

1.1 Basic definitions

In the study of metric spaces in Mathematical Analysis III, the Euclidean n-space is an prototype. But the Euclidean n-space is not only a metric space, but also a vector space.

Definition 1. A *vector space* (or *linear space*) over a field $K = \mathbb{R}$ or $\mathbb{C}$ is a nonempty set X of elements $x, y, \dots$ (called *elements*) together with two algebraic operations, the vector addition and multiplication of vectors by scalars (elements in K) such that the vector addition satisfies

- $x + y = y + x$,
- $x + (y + z) = (x + y) + z$,
- there exists $0 \in X$ such that $0 + x = x$ for all $x \in X$,
- for any $x \in X$ there exists a $-x \in X$ such that $x + (-x) = 0$,

and the multiplication by scalars satisfies

- $\alpha(\beta x) = (\alpha\beta)x$,
- $1x = x$,
- $\alpha(x + y) = \alpha x + \alpha y$,
- $(\alpha + \beta)x = \alpha x + \beta x$.

Remark 1. Here the fancy term "field K" simply means $\mathbb{R}$ or $\mathbb{C}$. We use this term only to save space. Depending on $K = \mathbb{R}$ or $\mathbb{C}$, the vector space is called a real vector space or complex vector space.

A vector space with a "compatible" metric is called a normed space, as it is formally defined by a norm, as follows.

Definition 2. A real or complex *normed space* X is a vector space with a *norm* defined on it, where the norm $\|\cdot\|$ is a real-valued function on X such that

- $\|x\| \geq 0$,
- $\|x\| = 0$ if and only if $x = 0$,

- $\|\alpha x\| = |\alpha|\|x\|$,
- $\|x + y\| \le \|x\| + \|y\|$.

A normed space has a natural metric structure with the distance d defined as $d(x, y) = \|x - y\|$.

Definition 3. A *Banach space* is a normed space whose metric is complete.

Example 1. $C([a, b]) = \{\text{continuous function on } [a, b]\}$ is a real normed space if we consider real-valued continuous functions on $[a, b]$ and let the vector addition be $(x + y)(t) = x(t) + y(t)$, the multiplication by scalars be $(\alpha x)(t) = \alpha x(t)$ and the norm be $\|x\| = \max_{t\in[a,b]}|x(t)|$. The metric structure of this normed space is $d(x, y) = \max_{t\in[a,b]}|x(t) - y(t)|$, and we know that it is a real Banach space since the completeness is proved in Mathematical Analysis III. If we consider complex-valued functions on $[a, b]$, similarly $C([a, b])$ is a complex normed space, and it is an exercise that it is also a complex Banach space.

From Mathematical Analysis III we know that an incomplete metric space R has a unique (up to isometry) completion $\hat{R}$, such that the incomplete metric space is isometric to a subspace W of the completion $\hat{R}$. Thus given a normed space, we can construct its completion. But the completion may not be a Banach space, since it may not be a vector space. The following theorem shows that the completion of a normed space is a Banach space.

Theorem 1. *Let $X = (X, \|\cdot\|)$ be a normed space. Then there is a Banach space $\hat{X}$ and an isometry A from X onto a subspace W of $\hat{X}$ which is dense in $\hat{X}$. The space $\hat{X}$ is unique, except for isometries.*

Proof. First we have a complete metric space $\hat{X} = (\hat{X}, \hat{d})$ and an isometry $A : X \to W = A(X)$ where W is dense in $\hat{X}$. Denote the metric in the normed space $(X, \|\cdot\|)$ by d, then for any $\hat{x}, \hat{y} \in W$ such that $\hat{x} = Ax, \hat{y} = Ay$, $\hat{d}(\hat{x}, \hat{y}) = d(x, y)$. By an abuse of notation, we denote a "norm" on W by $\|\cdot\|$ such that $\|\hat{x}\| = \|x\|$ if $\hat{x} = Ax$. It suffices to show that we can define vector addition and multiplication by scalar on $\hat{X}$ and a norm on $\hat{X}$ such that the metric determined by the vector space and the norm agrees with the metric $\hat{d}$. Ideally the norm is an extension of the $\|\cdot\|$ on W.

To define the vector addition of any $\hat{x}, \hat{y} \in \hat{X}$, we take sequences $\{\hat{x}_n\}, \{\hat{y}_n\} \subseteq W$ such that $\hat{x}_n \to \hat{x}$, $\hat{y}_n \to \hat{y}$. Since $\{\hat{x}_n\}$ and $\{\hat{y}_n\}$ are Cauchy sequences and

$$\begin{aligned}\hat{d}((\hat{x}_m + \hat{y}_m), (\hat{x}_n + \hat{y}_n)) &= \|(\hat{x}_m - \hat{x}_n) + (\hat{y}_m - \hat{y}_n)\| \\ &\le \|\hat{x}_m - \hat{x}_n\| + \|\hat{y}_m - \hat{y}_n\| \\ &= \hat{d}(\hat{x}_m, \hat{x}_n) + \hat{d}(\hat{y}_m, \hat{y}_n),\end{aligned}$$

the sequence $\{\hat{x}_n + \hat{y}_n\}$ is also a Cauchy sequence, and then we define

$$\hat{x} + \hat{y} = \lim_{n\to\infty} (\hat{x}_n + \hat{y}_n).$$

This definition does not depend on the choice of $\{\hat{x}_n\}$ and $\{\hat{y}_n\}$. Suppose we choose another pair of $\{\hat{x}'_n\}, \{\hat{y}'_n\} \subseteq W$ such that $\hat{x}'_n \to \hat{x}$ and $\hat{y}'_n \to \hat{y}$, then

$$\begin{aligned}\hat{d}((\hat{x}_n + \hat{y}_n), (\hat{x}'_n + \hat{y}'_n)) &= \|(\hat{x}_n - \hat{x}'_n) + (\hat{y}_n - \hat{y}'_n)\| \\ &= \|(\hat{x}_n - \hat{x}) + (\hat{x} - \hat{x}'_n) + (\hat{y}_n - \hat{y}) + (\hat{y} - \hat{y}'_n)\| \\ &\leq \hat{d}(\hat{x}_n, \hat{x}) + \hat{d}(\hat{x}'_n, \hat{x}) + \hat{d}(\hat{y}_n, \hat{y}) + \hat{d}(\hat{y}'_n, \hat{y}),\end{aligned}$$

and we have that $\{\hat{x}_n + \hat{y}_n\}$ and $\{\hat{x}'_n + \hat{y}'_n\}$ converge to the same point since they are equivalent Cauchy sequences.

Similarly, for any $\hat{x} \in \hat{X}$ and $\alpha \in K$, we pick up a sequence $\{\hat{x}_n\} \subseteq W$ such that $\hat{x}_n \to \hat{x}$. The sequence $\{\alpha\hat{x}_n\}$ is a Cauchy sequence since

$$\hat{d}(\alpha\hat{x}_m, \alpha\hat{x}_n) = \|\alpha\hat{x}_m - \alpha\hat{x}_n\| = |\alpha|\|\hat{x}_m - \hat{x}_n\| = |\alpha|\hat{d}(\hat{x}_m, \hat{x}_n).$$

Hence we define

$$\alpha\hat{x} = \lim_{n\to\infty} \alpha\hat{x}_n.$$

Similar to the well-definedness of $\hat{x} + \hat{y}$, this definition does not depend on the choice of $\{\hat{x}_n\}$.

The vector addition and multiplication by scalars defined above satisfy all the required relations. We prove one for example. Suppose $\{\hat{x}_n\}, \{\hat{y}_n\} \subseteq W$ and $\hat{x}_n \to \hat{x}$, $\hat{y}_n \to \hat{y}$, then

$$\begin{aligned}\alpha(\hat{x} + \hat{y}) &= \alpha \lim_{n\to\infty} (\hat{x}_n + \hat{y}_n) = \lim_{n\to\infty} \alpha(\hat{x}_n + \hat{y}_n) = \lim_{n\to\infty} \alpha\hat{x}_n + \alpha\hat{y}_n = \lim_{n\to\infty} \alpha\hat{x}_n + \lim_{n\to\infty} \alpha\hat{y}_n \\ &= \alpha\hat{x} + \alpha\hat{y}.\end{aligned}$$

Next we need to define a norm $\|\cdot\|_1$ in $\hat{X}$. The natural choice is

$$\|\hat{x}\|_1 = \hat{d}(\hat{0}, \hat{x}) \quad \text{where} \quad \hat{0} = A(0) \in W.$$

We need to show that the norm $\|\cdot\|_1$ satisfies all the requirements for a norm and the relation that

$$\hat{d}(\hat{x}, \hat{y}) = \|\hat{x} - \hat{y}\|_1. \tag{1}$$

We verify the relation (1), and the method works for the verification of other conditions. If $\hat{x}, \hat{y} \in W$, then by the isometry of W and X, we have $\|\hat{x}\|_1 = \|x\|$ and $\|\hat{y}\|_1 = \|y\|$ if $\hat{x} = Ax$ and $\hat{y} = Ay$. Then the relation (1) for $\|\cdot\|_1$ is equivalent to the corresponding relation for $\|\cdot\|$, so it is automatically satisfied. If $\hat{x}, \hat{y}$ are not in W, we take $\{\hat{x}_n\}, \{\hat{y}_n\} \subseteq W$ such that $\hat{x}_n \to \hat{x}$ and $\hat{y}_n \to \hat{y}$. Then

$$\hat{d}(\hat{x}, \hat{y}) = \lim_{n\to\infty} \hat{d}(\hat{x}_n, \hat{y}_n) = \lim_{n\to\infty} \|\hat{x}_n - \hat{y}_n\| = \lim_{d\to\infty} \hat{d}(\hat{0}, \hat{x}_n - \hat{y}_n) = \hat{d}(\hat{0}, \hat{x} - \hat{y}) = \|\hat{x} - \hat{y}\|_1. \tag{2}$$

□

1.2 Finite dimensional normed spaces

In functional analysis, we concentrate on infinite dimensional spaces. Here we use finite dimensional spaces as a warm-up. We will get some results about finite dimensional normed spaces that are intuitively convincing, but will later be shown to be false in the infinite dimensional setting.

Recall that in linear algebra, we say a finite set of vectors $x_1, x_2, \ldots, x_n$ are *linearly independent* if the only scalars $\alpha_1, \ldots, \alpha_n$ that make $\alpha_1 x_1 + \cdots \alpha_n x_n = 0$ are $\alpha_1 = \alpha_2 = \cdots \alpha_n = 0$.

Definition 4. A set of vectors $M \subseteq X$ is a *Hamel basis* for X if any finite set of vectors in M are linearly independent, and any $x \in X$ is a linear combination of finitely many vectors in M, that is, $x = \alpha_1 x_1 + \ldots + \alpha_n x_n$ where $x_1, \cdots, x_n \in M$.

If the vector space X has a finite Hamel basis, it is said to be a finite dimensional space, otherwise it is a infinite dimensional space.

Remark 2. Here we do not rule out the possibility that some infinite dimensional spaces have no Hamel basis at all. Actually all vector spaces have Hamel bases, but we do not prove it now.

If a vector space X is of finite dimension, then it has a finite basis $\{e_1, \ldots, e_n\}$, consisting of n elements, and any point $x \in X$ is expressed as

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n$$

in a unique way. The number n of elements in the basis is called the *dimension* of the vector space. It is an linear algebraic problem (as an exercise) to show that the dimension of a finite dimensional vector space is well defined, that is, if $\{e'_1, \ldots, e'_m\}$ is another basis of X, then $m = n$.

Below we give a technical lemma that is valid for all normed spaces but particularly useful for finite dimensional ones.

Lemma 1. *Let $x_1, \ldots, x_n$ be a linearly independent set of vectors in a normed space X. Then there is a positive number $c > 0$ such that for any choice of scalars $\alpha_1, \ldots, \alpha_n$,*

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \geq c(|\alpha_1| + \cdots + |\alpha_n|). \tag{3}$$

Proof. The case $\alpha_1 = \cdots = \alpha_n = 0$ is trivial. We observe that the inequality (3) is homogeneous in the sense that if the inequality holds for $\alpha_1, \ldots, \alpha_n$, then if we change α_i into $r\alpha_i$ for the same $r \neq 0$, then the inequality holds for the same value of c. Thus without loss of generality, we assume $|\alpha_1| + \cdots + |\alpha_n| = 1$.

Suppose to the contrary that there exist a sequence $\{y_k\} \supseteq X$ such that

$$y_k = \alpha_1^{(k)} x_1 + \cdots + \alpha_n^{(k)} x_n \quad \text{where} \quad \sum_{i=1}^{n} |\alpha_i^{(k)}| = 1$$

and $\|y_k\| \to 0$. Since $|\alpha_i^{(k)}| \leq 1$ for all $i = 1, \ldots, n$, by the Bolzano-Weierstrass theorem, a subsequence

$$\{(\alpha_1^{(k_m)}, \ldots, \alpha_n^{(k_m)})\} \subseteq \{(\alpha_1^{(k)}, \ldots, \alpha_n^{(k)})\} \subseteq K^n$$

converges to $(\alpha_1, \ldots, \alpha_n)$. (The classical Bolzano-Weierstrass theorem is stated in $\mathbb{R}^n$. But by identifying $\mathbb{C}$ as $\mathbb{R}^2$ and then $\mathbb{C}^n$ as $\mathbb{R}^{2n}$, we see that it is valid also for $\mathbb{C}^n$.)

This convergence implies that $|\alpha_1| + \cdots + |\alpha_n| = 1$. Then the subsequence $\{y_{k_m}\} \subseteq \{y_k\}$ converges to $y = \alpha_1 x_1 + \cdots + \alpha_n x_n$. To see it, we check that

$$\begin{aligned} d(y_m, y) = \|y_m - y\| &= \|(\alpha_1^{(k_m)} - \alpha_1)x_1 + \cdots + (\alpha_n^{(k_m)} - \alpha_n)x_n\| \\ &\le |\alpha_1^{(k_m)} - \alpha_1|\|x_1\| + \cdots + |\alpha_n^{(k_m)} - \alpha_n|\|x_n\|. \end{aligned}$$

Thus we have

$$\|y\| = d(0, y) = \lim_{n\to\infty} (0, y_{k_m}) = \lim_{n\to\infty} \|y_{k_m}\| = 0,$$

and so that $y = 0$, contradictory to the linear independence of $x_1, \ldots, x_n$. $\square$

The next theorem shows that all normed spaces of dimension n are topologically equivalent to the Euclidean n-space or its complex counterpart, the unitary n-space. In this module, we define the Euclidean n-space as the normed space $(X, \|\cdot\|)$ where

$$X = \mathbb{R}^n = \{x = (x_1, \ldots, x_n) \mid x_i \in \mathbb{R}\}, \quad \text{and} \quad \|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

The *unitary n-space* is defined as the normed space $(X, \|\cdot\|)$ where

$$X = \mathbb{C}^n = \{x = (x_1, \ldots, x_n) \mid x_i \in \mathbb{C}\}, \quad \text{and} \quad \|x\| = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}.$$

We know from Mathematical Analysis III that the Euclidean n-space, as a metric space, has many special properties. As a metric space, the unitary n-space is isomorphic to the Euclidean $2n$-space. (The construction of the isometry is an exercise.)

Definition 5. A norm $\|\cdot\|$ on a vector space X is called *equivalent* to another norm $\|\cdot\|_0$ on X if there are positive numbers $a < b$ such that for all $x \in X$

$$a\|x\|_0 \le \|x\| \le b\|x\|_0. \tag{4}$$

Note that if two norms $\|\cdot\|$ and $\|\cdot\|_0$ are equivalent, then the metrics d and d_0 defined by them are *strongly equivalent*, that is, for any $x, y \in X$, $ad_0(x, y) \le d(x, y) \le bd_0(x, y)$, and then the topologies defined by these metrics are the same.

Theorem 2. *On a finite dimensional vector space, any two norms are equivalent.*

Proof. Let $\{e_1, \ldots, e_n\}$ be a basis of X, and we define a norm $\|\cdot\|_2$ as

$$\|\alpha_1 e_1 + \cdots + \alpha_n e_n\|_2 = \sqrt{|\alpha_1|^2 + \cdots + |\alpha_n|^2}. \tag{5}$$

Since the equivalence of norms is *transitive*, that is, if both $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent to $\|\cdot\|_2$, then $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent to each other, we only need to show that any norm $\|\cdot\|$ is equivalent to $\|\cdot\|_2$ defined above.

By Lemma 1, there is a $c > 0$ such that

$$\|x\| = \|\alpha_1 e_1 + \cdots + \alpha_n e_n\| \ge c(|\alpha_1| + \cdots + |\alpha_n|) \ge c\sqrt{|\alpha_1|^2 + \cdots + |\alpha_n|^2} = c\|x\|_2.$$

On the other hand,

$$\begin{aligned} \|x\| &= \|\alpha_1 e_1 + \cdots + \alpha_n e_n\| \\ &\le |\alpha_1|\|e_1\| + \cdots + |\alpha_n|\|e_n\| \\ &\le \max_{i=1}^{n} \|e_i\|(|\alpha_1| + \cdots + |\alpha_n|) \\ &\le \max_{i=1}^{n} \|e_i\|\sqrt{n}\sqrt{|\alpha_1|^2 + \cdots + |\alpha_n|^2} = \max_{i=1}^{n} \|e_i\|\sqrt{n}\|x\|_2. \end{aligned}$$

Thus the $\|\cdot\|$ and $\|\cdot\|_2$ are equivalent with $a = c$ and $b = \sqrt{n}\max_{i=1}^{n}\|e_i\|$. $\square$

Remark 3. In the case that $K = \mathbb{R}$, if we identify the vector space X with $\mathbb{R}^n$ with

$$\alpha_1 e_1 + \cdots + \alpha_n e_n \mapsto (\alpha_1, \ldots, \alpha_n), \tag{6}$$

then the norm $\|\cdot\|_2$ defines the *Euclidean n-space* on X. So any n-dimensional real normed space is equivalent as a normed space, strongly equivalent as a metric space, and then homeomorphic as a topological space, to the Euclidean n-space. In the case that $K = \mathbb{C}$, any n-dimensional real normed space is equivalent as a normed space, strongly equivalent as a metric space, and then homeomorphic as a topological space, to the unitary n-space.

Corollary 1. *Let $(X, \|\cdot\|)$ be a finite dimensional normed space. Then*

(a) $(X, \|\cdot\|)$ is a Banach space.

(b) Any closed and bounded subset $M \subseteq X$ is compact. (Here M is bounded if $\sup_{x\in M}\|x\| < \infty$).

Proof. We use the completeness of the Euclidean n-space and the unitary n-space, and the Bolzano-Weierstrass theorem for the Euclidean n-space and the unitary n-space.

Let $e_1, \ldots, e_n$ be a basis of X, and we define another norm $\|\cdot\|_2$ on X by (5), so that the space $(X, \|\cdot\|_2)$ is (isomorphic to) the Euclidean n-space or the unitary n-space.

Suppose $\{x_n\}$ is a Cauchy sequence in $(X, \|\cdot\|)$. Then by Theorem 2, it is also a Cauchy sequence in $(X, \|\cdot\|_2)$. So it converges to a point $x \in X$ with respect to the metric defined by $\|\cdot\|_2$, that is, $\|x_n - x\|_2 \to 0$. Using Theorem 2 again, we have that $\|x_n - x\| \to 0$ and the sequence converges to x with respect to the metric defined by $\|\cdot\|$. Thus we show that the metric defined by $\|\cdot\|$ on X is complete and prove part (a).

Suppose M is a subset of X that is bounded with respect to $\|\cdot\|$, and is closed with respect to (the topology defined by) $\|\cdot\|$ (via the metric defined by the norm). Then by Theorem 2, M is also bounded with respect to $\|\cdot\|_2$. To show that it is also closed with respect to $\|\cdot\|_2$, we take any $x \in X \setminus M$. By the closedness with respect to $\|\cdot\|$, there is $\epsilon > 0$ such that the ball

$$B_\epsilon = \{y \in X \mid \|y - x\| < \epsilon\}$$

is disjoint to M. Then by Theorem 2, with a properly chosen $b > 0$ as in (4), we have that the ball

$$\{y \in X \mid \|y - x\|_2 < \epsilon\} \subseteq B_\epsilon$$

is also disjoint to M. So M is also closed with respect to $\|\cdot\|_2$. Using the Bolzano-Weierstrass theorem, M is compact with respect to the topology defined by $\|\cdot\|_2$. Since $\|\cdot\|$ defines the same topology on X, we prove that M is compact in the original normed space and finish part (b). $\square$

1.3 Bounded linear operators

The mapping from a linear space to another is called an *operator*. In this module, we are mostly concerned with linear operators.

Definition 6. A *linear operator* T is an operator such that

- The domain $D(T)$ of T is a vector space and the range $R(T)$ lies in another vector space, called the target space, over the same field ($\mathbb{R}$ or $\mathbb{C}$).

- For all $x, y \in D(T)$ and scalar α,

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

For a linear operator T, we also define the *null space* $N(T)$ as

$$N(T) = \{x \in D(T) \mid Tx = 0\}.$$

Note that $0 \in N(T)$ since $T(0) = T(0 \cdot 0) = 0T(0) = 0$.

Remark 4. Usually we consider a linear space X as the domain of the linear operator T, that is, $D(T) = X$, and a target space Y such that $R(T) \subseteq Y$. Then T is an operator from X *into* Y, and we write $T : X \to Y$. T is *surjective*, or is from X *onto* Y, if $R(T) = Y$.

Suppose X is a linear space, we say $M \subseteq X$ is a *subspace* of X if for any $x, y \in M$ and scalar α, β, $\alpha x + \beta y \in M$. Then Y is a linear space over the same field.

Theorem 3. *Let $T : X \to Y$ be a linear operator. Then*

(a) $R(T)$ is a subspace of Y.

(b) $N(T)$ is a subspace of X.

Proof. For part (a), we take any $y_1, y_2 \in R(T)$ and two scalars α, β. Suppose $y_1 = Tx_1$ and $y_2 = Tx_2$, then

$$\alpha y_1 + \beta y_2 = \alpha Tx_1 + \beta Tx_2 = T(\alpha x_1) + T(\beta x_2) = T(\alpha x_1 + \beta x_2) \in R(T).$$

For part (b), we take any $x_1, x_2 \in N(T)$ and two scalars α, β. Then

$$T(\alpha x_1 + \beta x_2) = \alpha Tx_1 + \beta Tx_2 = \alpha \cdot 0 + \beta \cdot 0 = 0.$$

□

The linear operator $T : X \to Y$ is *injective*, or *one-to-one*, if $Tx_1 = x_2$ implies $x_1 = x_2$. If T is both injective and surjective, then it has an *inverse*, denoted as T^{-1}, such that $T^{-1}Tx = x$ for all $x \in X$ and $TT^{-1}y = y$ for all $y \in Y$.

Theorem 4. *Let $T : X \to Y$ be a linear operator, Then*

(a) T is injective if and only if $Tx = 0$ implies $x = 0$.

(b) If T^{-1} exists, then it is a linear operator $T^{-1} : Y \to X$.

Proof. For part (a), we need only to show the "if" part. Given $x_1, x_2 \in X$, such that $Tx_1 = Tx_2$, then $T(x_1 - x_2) = Tx_1 - Tx_2 = 0$, so $x_1 - x_2 = 0$, and we have $x_1 = x_2$.

For part (b), we need to show that for all $y_1, y_2 \in Y$ and scalars α, β, $T^{-1}(\alpha y_1 + \beta y_2) = \alpha T^{-1}y_1 + \beta T^{-1}y_2$. This is because

$$T(T^{-1}(\alpha y_1 + \beta y_2)) = \alpha y_1 + \beta y_2 = \alpha TT^{-1}y_1 + \beta TT^{-1}y_2 = T(\alpha T^{-1}y_1 + \beta T^{-1}y_2),$$

and the injectivity of T. □

The following lemma is well known for bijective (i.e., both surjective and injective) linear operators between finite dimensional vector spaces, but we now consider it in the more general setting.

Lemma 2. *Let $T : X \to Y$ and $S : Y \to Z$ be bijective linear operators, where X, Y, Z are vector spaces over the same field K. Then the inverse $(ST)^{-1} : Z \to X$ of the product ST exists and*

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

Proof. Since both S and T are bijective, the composition ST is bijective, and then ST has an inverse. For any $z \in Z$, suppose $(ST)^{-1}z = x \in X$, it is equivalent to $(ST)x = S(Tx) = z$. Denote $y = Tx$, then $Sy = z$. So

$$(T^{-1}S^{-1})z = T^{-1}(S^{-1}z) = T^{-1}y = x.$$

By the arbitrarity of z, we show that $(ST)^{-1} = T^{-1}S^{-1}$. □

Now we consider linear operators between normed spaces.

Definition 7. Let X and Y be normed spaces and $T : X \to Y$ be a linear operator. Then operator T is said to be *bounded* if there is a positive real number c such that for all $x \in X$

$$\|Tx\| \le c\|x\|.$$

If $T : X \to Y$ is a bounded operator, then we define the *norm* of the operator T as

$$\|T\| = \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|Tx\|}{\|x\|} = \sup_{\substack{x\in X\\ \|x\|=1}} \|Tx\| = \sup_{\substack{x\in X\\ \|x\|=c}} \frac{\|Tx\|}{c} = \sup_{\substack{x\in X\\ \|x\|\le c}} \frac{\|Tx\|}{c} = \sup_{\substack{x\in X\\ \|x\|<c}} \frac{\|Tx\|}{c}, \quad \text{for all } c > 0, \tag{7}$$

where the equality of the five suprema is due to that $\|T(\alpha x)\|/\|\alpha x\| = \|Tx\|/\|x\|$.

Since operators between normed spaces can be viewed as mappings between metric spaces, we say an operator is *continuous* if it is a continuous mapping with respect to the metrics defined by the norms. More precisely, $T : X \to Y$ is *continuous at* $x_0 \in X$ if for every $\epsilon > 0$, there is a $\delta > 0$ such that

$$\|Tx - Tx_0\| < \epsilon \quad \text{for all } x \in X \text{ with } \|x - x_0\| < \delta. \tag{8}$$

For linear operators, continuity and boundedness are equivalent, as shown in the following theorem.

Theorem 5. *Let $T : X \to Y$ be a linear operator. Then*

(a) T is continuous if and only if it is bounded.

(b) If T is continuous at a single point, then it is continuous.

Proof. By (7), if T is bounded, then for any $\epsilon > 0$, we take $\delta < \epsilon/\|T\|$, and have that (8) is satisfied for $x_0 = 0$ and then T is continuous at $0 \in X$. Conversely, if T is continuous at 0, then we take a pair of ϵ, δ to make (8) hold, and then by (7)

$$\|T\| = \sup_{\substack{x\in X\\ \|x\|<\delta}} \frac{\|Tx\|}{\delta} \le \frac{\epsilon}{\delta}.$$

So the boundedness of T is equivalent to the continuity of T at 0.

On the other hand, the continuity of T at any point $x_0 \in X$ is equivalent to the continuity at 0. For any $x \in X$, denote $x' = x - x_0$, then the condition (8) is equivalent to

$$\|Tx'\| < \epsilon \quad \text{for all } x' \in X \text{ with } \|x'\| < \delta.$$

Therefore we prove the theorem. □

It is clear that if X and Y are vector spaces over the same field K, the set of all the linear operators from X to Y, denoted as $L(X, Y)$, form a vector space over K, where the vector addition and scalar multiplication are defined as

$$(T_1 + T_2)x = T_1x + T_2x, \quad (\alpha T)x = \alpha(Tx), \quad \text{for all } x \in X. \tag{9}$$

Below we show that the set of all the bounded linear operators from X to Y, denoted as $B(X, Y)$, form a normed space over K, where the vector addition and scalar multiplication are defined by (9) and the norm is the operator norm defined in (7).

Theorem 6. *Let X and Y be normed spaces over the same field, and $B(X, Y)$ be the set of all bounded linear operators from X into Y. Then $B(X, Y)$ is a subspace of the vector space $L(X, Y)$, and it is a normed space with the norm given by* (7)*.*

Proof. To show that $B(X, Y)$ is a subspace of $L(X, Y)$, we show that if $T, S \in B(X, Y)$ and α is a scalar, then $\alpha T \in B(X, Y)$ and $T + S \in B(X, Y)$. First,

$$\sup_{\substack{x \in X \\ \|x\|=1}} \|(\alpha T)x\| = \sup_{\substack{x \in X \\ \|x\|=1}} \|\alpha(Tx)\| = |\alpha| \sup_{\substack{x \in X \\ \|x\|=1}} \|Tx\|, \tag{10}$$

so $\|\alpha T\|$ is well defined and $\alpha T \in B(X, Y)$. Next

$$\sup_{\substack{x \in X \\ \|x\|=1}} \|(T + S)x\| = \sup_{\substack{x \in X \\ \|x\|=1}} \|Tx + Sx\| \leq \sup_{\substack{x \in X \\ \|x\|=1}} \|Tx\| + \sup_{\substack{x \in X \\ \|x\|=1}} \|Sx\|, \tag{11}$$

so $\|T + S\|$ is well defined and $T + S \in B(X, Y)$. Furthermore, by (10) and (11), we have that

$$\|\alpha T\| = |\alpha| \|T\|, \quad \|S + T\| \leq \|S\| + \|T\|.$$

To show that $B(X, Y)$ is a normed space, we need only that

(i) $\|T\| \geq 0$ for all $T \in B(X, Y)$. (This is obvious).

(ii) $\|T\| = 0$ if and only if $T = 0$, the zero operator, defined as $0(x) = x$ for all $x \in X$.

To verify condition (ii), we observe that $\|0\| = 0$ by (7), and if $T \neq 0$, there is $0 \neq x \in X$ such that $Tx = y \neq 0$. Then by (7),

$$\|T\| \geq \frac{\|Tx\|}{\|x\|} = \frac{\|y\|}{\|x\|} > 0.$$

and thus finish the proof. □

Naturally, if X and Y are both Banach spaces, we expect $B(X, Y)$ to be a Banach space. Actually there is a stronger result.

Theorem 7. *If Y is a Banach space and X is a normed space (not necessarily Banach space), then $B(X, Y)$ is a Banach space.*

Proof. Let $\{T_n\}$ be a Cauchy sequence in $B(X, Y)$, that is, for any $\epsilon > 0$, there is a N_ϵ such that $\|T_m - T_n\| < \epsilon$ for all $m, n > N_\epsilon$. Then for any $x \in X$, $\{T_n x\}$ is a Cauchy sequence, since

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

So by the completeness of Y, $\{T_n x\}$ has a unique limit, and we denote it as Tx.

First we show that the operator $T : x \to Tx$ is linear. For any $x_1, x_2 \in X$ and $\alpha, \beta \in K$,

$$T(\alpha x_1 + \beta x_2) = \lim_{n\to\infty} T_n(\alpha x_1 + \beta x_2) = \alpha \lim_{n\to\infty} T_n x_1 + \beta \lim_{n\to\infty} T_n x_2 = \alpha T x_1 + \beta T x_2,$$

and we get the linearity. Next we show that the operator is a bounded linear operator. Let $\epsilon > 0$. for any $x \in X$,

$$\|Tx\| \le \|Tx - T_n x\| + \|T_n x - T_{N_\epsilon+1} x\| + \|T_{N_\epsilon+1} x\|.$$

As $n \to \infty$, the first term vanishes, the second is bounded by $\epsilon\|x\|$ and the third bounded by $\|T_{N_\epsilon+1} x\|$, so $\|Tx\| \le (\|T_{N_\epsilon+1}\| + \epsilon)\|x\|$, and we get that as a linear operator, T is bounded, with the norm $\|T\| \le \|T_{N_\epsilon+1}\| + \epsilon$.

Thus $T \in B(X, Y)$. We claim that T is the limit of T_n in the normed space $B(X, Y)$, and then prove the theorem. To see the convergence, let $\epsilon > 0$ be arbitrary, we have for all $n > N_\epsilon$

$$\|(T - T_n)x\| \le \|Tx - T_m x\| + \|T_m x - T_n x\|.$$

As $m \to \infty$, the first term vanishes and the second is bounded by $\epsilon\|x\|$, so we conclude that $\|(T - T_n)\| \le \epsilon$, and so T_n converges to T in $B(X, Y)$. □

Note that K is a vector space over K itself. (The dimension is 1, and any nonzero element of K gives a basis). It is also a normed space, and then a Banach space by Corollary 1, with the norm defined by the absolute value of either real numbers or complex numbers. The linear operators from a vector space X (over K) to K are of special importance, and are called *linear functionals*. Similarly, A *bounded linear functional* $f : X \to K$ is a bounded operator whose target space is K. Thus a linear functional is bounded if and only if there is a positive real number c such that

$$|f(x)| \le c\|x\|$$

for all $x \in X$, and the norm of a bounded linear functional is defined by (7), where T denotes a bounded linear functional and $\|Tx\| = |Tx|$.

Example 2. The space $\ell^2 = \{(\alpha_1, \alpha_2, \dots) \mid \sum_{i=1}^{\infty} |\alpha_i|^2 < \infty\}$ is a vector space, over either $\mathbb{R}$ or $\mathbb{C}$, with

$$(\alpha_1, \alpha_2, \dots) + (\beta_1, \beta_2, \dots) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots), \quad \alpha(\alpha_1, \alpha_2, \dots) = (\alpha\alpha_1, \alpha\alpha_2, \dots).$$

It is also a normed space, with the norm

$$\|(\alpha_1, \alpha_2, \dots)\|_2 = \sqrt{\sum_{i=1}^{\infty} |\alpha_i|^2}.$$

Now consider the (bounded) linear functionals on ℓ^2. Suppose f is a linear functional on ℓ^2, then for each $e_i = (0, \dots, 0, 1, 0, \dots)$ such that the i-th coordinate is 1 and others 0, $f(e_i)$ is well defined, and we denote it as $\bar{\xi}_i$. (Note the complex conjugate. We do not explain it now, but it will turn out to be important later.) If f is a bounded linear operator, then f is uniquely determined by the infinite dimensional vector $(\xi_1, \xi_2, \dots)$. To see it, we take an arbitrary $a = (\alpha_1, \alpha_2, \dots)$, and let $a^{(n)} = (\alpha_1, \alpha_2, \dots, \alpha_n, 0, 0, \dots)$, the truncation of a. Then in ℓ^2 (as a metric space), $a^{(n)} \to a$, that is, $\|a^{(n)} - a\|_2 \to 0$. Thus

$$\lim_{n\to\infty} |f(a^{(n)}) - f(a)| = \lim_{n\to\infty} |f(a^{(n)} - a)| \leq \lim_{n\to\infty} \|f\| \cdot \|a^{(n)} - a\|_2 = 0,$$

and $f(a)$ is the limit of $f(a^{(n)})$, whose existence is guaranteed by the completeness of ℓ^2. In this sense, we use the infinite vector $(\xi_1, \xi_2, \dots)$ to represent f.

On the other hand, not all infinite vectors $(\xi_1, \xi_2, \dots)$ define uniquely a valid (bounded) linear operator on ℓ^2. For example, let f be represented by $(1, 1, 1, \dots)$, then $f(a)$, where $a = (1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots)$, is not well defined taking $f(a) = \lim_{n\to\infty} f(a^{(n)})$.

We claim that $(\xi_1, \xi_2, \dots)$ defines a bounded linear operator f if and only if

$$\sum_{i=1}^{\infty} |\xi_i|^2 < \infty. \tag{12}$$

Suppose $(\xi_1, \xi_2, \dots)$ satisfies (12), then for all $a = (\alpha_1, \alpha_2, \dots)$, by the Cauchy-Schwarz inequality (called Schwarz inequality in MA3209),

$$\begin{aligned}\sum_{i=1}^{\infty} |\bar{\xi}_i \alpha_i| = \lim_{n\to\infty} \sum_{i=1}^{n} |\bar{\xi}_i \alpha_i| \leq \lim_{n\to\infty} \sqrt{\sum_{i=1}^{n} |\xi_i|^2} \sqrt{\sum_{i=1}^{n} |\alpha_i|^2} &= \sqrt{\sum_{i=1}^{\infty} |\xi_i|^2} \sqrt{\sum_{i=1}^{\infty} |\alpha_i|^2} \\ &= \sqrt{\sum_{i=1}^{\infty} |\xi_i|^2} \|a\|_2,\end{aligned}$$

and the equal sign is attained if $a = (\xi_1, \xi_2, \dots)$. (Exercise: Prove the complex form of the Cauchy-Schwarz inequality.) So the bounded operator f can be defined for all $a \in \ell^2$ as

$$f(a) = \sum_{i=1}^{\infty} \bar{\xi}_i \alpha_i,$$

and we have

$$\|f\| = \sqrt{\sum_{i=1}^{\infty} |\xi_i|^2}.$$

On the other hand, if (12) is not satisfied, then we show that even if there exists a linear operator f satisfying $f(e_i) = \bar{\xi}_i$, it cannot be bounded. Let $x^{(n)} = (\xi_1, \xi_2, \dots, \xi_n, 0, \dots)$, the truncation of $(\xi_1, \xi_2, \dots)$, then

$$\lim_{n\to\infty} \frac{|f(x^{(n)})|}{\|x^{(n)}\|_2} = \lim_{n\to\infty} \frac{\sum_{i=1}^n \bar{\xi}_i \xi_i}{\sqrt{\sum_{i=1}^n |\xi_i|^2}} = \lim_{n\to\infty} \sqrt{\sum_{i=1}^n |\xi_i|^2} = \infty.$$

In conclusion, the normed space of bounded linear functionals of ℓ^2, is again ℓ^2.

Definition 8. Let X be a normed space. Then the normed space of bounded linear functionals of X, with the norm given by (7), is called the *dual space*, and is denoted by X'.

With the dual space X' defined, it is natural to consider the *second dual space* $X'' = (X')'$. There is a *canonical mapping* (actually an operator) $C : X \to X''$ such that for any $x \in X$ and $f \in X'$

$$(Cx)f = f(x).$$

It is not difficult to show that C is linear (exercise). Also C is bounded. From

$$\|Cx\| = \sup_{\substack{f\in X'\\ f\neq 0}} \frac{|(Cx)f|}{\|f\|} = \sup_{\substack{f\in X'\\ f\neq 0}} \frac{|f(x)|}{\|f\|} \le \sup_{\substack{f\in X'\\ f\neq 0}} \frac{\|f\| \cdot \|x\|}{\|f\|} = \|x\|, \tag{13}$$

we have that $\|C\| \le 1$. Later we will show that $\|C\| = 1$, and even more is true: C is an isometric mapping in the sense that $\|Cx\| = \|x\|$ for all $x \in X$.

However, C is not necessarily surjective. If the canonical mapping $C : X \to X''$ is isometric and surjective, that is, X'' is isomorphic to X by the canonical mapping, we say that the normed space X is *reflexive*.

Example 2 implies that the ℓ^2 space is reflexive. By Theorem 7, X needs to be a Banach space to be reflexive, but this is not a sufficient condition.

1.4 Bounded linear functionals and operators of finite dimensional normed spaces

As our first example, the Euclidean n-space and the unitary n-space are simplified finite dimensional counterparts of the ℓ^2-space over $\mathbb{R}$ and $\mathbb{C}$, so they have the same property:

Example 3. The bounded linear functionals on the Euclidean n-space or the unitary n-space are the normed space $\{f = (\xi_1, \dots, \xi_n) \mid \xi_i \in \mathbb{R} \text{ or } \mathbb{C}\}$ with the norm

$$\|f\|_2 = \sqrt{|\xi_1|^2 + \cdots + |\xi_n|^2}, \tag{14}$$

such that the functional is defined as

$$f(a) = \bar{\xi}_1 \alpha_1 + \cdots + \bar{\xi}_n \alpha_n, \quad \text{for all } a = (\alpha_1, \dots, \alpha_n) \in \mathbb{R} \text{ or } \mathbb{C}.$$

Now consider finite dimensional normed spaces with general norm. Let $(X, \|\cdot\|)$ be a finite dimensional normed space with basis $\{e_1, \dots, e_n\}$. Suppose f is a linear functional of X (not necessarily bounded) such that

$$f(e_i) = \bar{\xi}_i, \quad \text{for } i = 1, \dots, n,$$

then f is uniquely determined since by linearity

$$f(\alpha_1 e_1 + \cdots + \alpha_n e_n) = \bar{\xi}_1 \alpha_1 + \cdots + \bar{\xi}_n \alpha_n.$$

Define the norm $\|\cdot\|_2$ on X in the way similar to that in (5) in the proof of Theorem 2, such that $\|\alpha_1 e_1 + \cdots + \alpha_n e_n\|_2 = (|\alpha_1|^2 + \cdots + |\alpha_n|^2)^{1/2}$. By Theorem 2, $\|\cdot\|$ and $\|\cdot\|_2$ are equivalent and there are $a, b > 0$ such that $a\|x\|_2 \le \|x\| \le b\|x\|_2$ for all $x \in X$. Then for any $x = \alpha_1 e_1 + \cdots \alpha_n e_n \in X$,

$$\frac{|f(x)|}{\|x\|} = \frac{|\sum_{i=1}^n \bar{\xi}_i \alpha_i|}{\|x\|} \le \frac{1}{a} \frac{\sum_{i=1}^n |\bar{\xi}_i \alpha_i|}{\left(\sum_{i=1}^n |\alpha_i|^2\right)^{1/2}} \le \frac{1}{a} \sqrt{\sum_{i=1}^n |\xi_i|^2},$$

where in the last step we use Cauchy-Schwarz inequality. So f is a bounded linear functional on X with $\|f\| \le a^{-1}\|f\|_2$ where $\|f\|_2$ is expressed by (14).

Thus X' is also a normed space of dimension n. Repeat this argument, we have that X'' is again a normed space of dimension n. Below we show that the canonical mapping is isometric. It is clear that an isometric mapping is injective. From linear algebra, an injective linear operator from one n-dimensional space to another n-dimensional space is surjective. (If you forget it, prove it!) So any finite dimensional space is reflexive.

Since we have already had inequality (13), we need only to show that for any $x \in X$, $\|Cx\| \ge \|x\|$, and it sufficient to show that for any $0 \ne x \in X$ such that $\|x\| = 1$, there is an $0 \ne f \in X'$ with $\|f\| = 1$ such that $|(Cx)f| = |f(x)| = \|f\| \cdot \|x\| = 1$. Without loss of generality, we consider only $x = e_1$ and assume $\|e_1\| = 1$.

Let $X_1 \subseteq X_2 \subseteq \cdots \subseteq X_n = X$ be a sequence of nested subspaces of X such that X_i is *spanned* by $\{e_1, \dots, e_i\}$, that is, any $x \in X_i$ can be written as $\alpha_1 e_1 + \cdots + \alpha_i e_i$. Then each $(X_i, \|\cdot\|)$ has its own dual space X_i'. We construct inductively $f_1 \in X_1', \dots, f_n \in X_n' = X'$ such that

- $f_1(e_1) = 1$.
- $\|f_i\| = 1$ for all $i = 1, \dots, n$.
- f_i is the *restriction* of f_{i+1} in the sense that for any $x \in X_i \subseteq X_{i+1}$, $f_i(x) = f_{i+1}(x)$.

Then $f_n \in X'$ satisfies $\|f_n\| = 1$ and $f_n(e_1) = 1$. Thus $\|Ce_1\| \ge \|e_1\| = 1$ and we prove that the canonical mapping is isometric and hence X is reflexive.

The construction of f_1 is obvious: Let f_1 be the unique linear functional such that $f_1(e_1) = 1$. Suppose we have $f_i \in X_i'$ and $\|f_i\| = 1$, we show that there is $f_{i+1} \in X_{i+1}'$ with $\|f_{i+1}\| = 1$. Since any $x \in X_{i+1}$ can be uniquely written as $x = x' + \alpha e_{i+1}$ with $x' \in X_i$, the construction of f_{i+1} is equivalent to choosing a value $\xi \in K$ such that $f_{i+1}(e_{i+1}) = \xi$.

First we consider the case $K = \mathbb{R}$. It is obvious that $\|f_{i+1}\| \ge \|f_i\|$, so the condition that $\|f_{i+1}\| = 1$ is equivalent to $\|f_{i+1}\| \le 1$, or more concretely, for any $x = x' + \alpha e_{n+1}$ and $x' \in X_i$

$$|f_{i+1}(x)| = |f_i(x') + \alpha f_{i+1}(e_{i+1})| \le \|x\| = \|x' + \alpha e_{i+1}\|.$$

It sufficient to consider the special case $\alpha = 1$, and then we write the inequality $|f_i(x') + f_{i+1}(e_{i+1})| \le \|x' + e_{i+1}\|$ as

$$-\|x' + e_{i+1}\| - f_i(x') \le f_{i+1}(e_{i+1}) \le \|x' + e_{i+1}\| - f_i(x').$$

If the inequality

$$\sup_{x' \in X_i} -\|x' + e_{i+1}\| - f_i(x') \le \inf_{x' \in X_i} \|x' + e_{i+1}\| - f_i(x') \tag{15}$$

holds, then simply choose the value of $f_{i+1}(e_{i+1})$ as a real number between the two sides of (15), and the construction of f_{i+1} is done (in the $K = \mathbb{R}$ case). Below we show by contradiction that the inequality (15) always holds. Otherwise, there are $x', y' \in X_i$ such that

$$- \|x' + e_{i+1}\| - f_i(x') > \|y' + e_{i+1}\| - f_i(y'), \tag{16}$$

or equivalently,

$$f_i(y' - x') > \|y' + e_{i+1}\| + \|x' + e_{i+1}\|.$$

But since $\|f_i\| = 1$,

$$f_i(y' - x') \le \|y' - x'\| = \|(y' + e_{i+1}) - (x' + e_{i+1})\| \le \|y' + e_{i+1}\| + \|x' + e_{i+1}\|,$$

we derive the contradiction to (16).

Since the construction of f_{i+1} from f_i above uses the property of real numbers in an essential way, (do you have a reminiscence of the Dedekind cut?) it is not easily applied to the $K = \mathbb{C}$ case. Below we use the result for real normed space to solve the problem for the complex normed space.

First we explain the relation between a complex vector/normed space and the real space underlying it. An n-dimensional complex vector space X with basis $\{e_1, \dots, e_n\}$ can be viewed as a $2n$- dimensional real vector space $X_{\mathbb{R}}$ with basis $\{r(e_1), i(e_1), \dots, r(e_n), i(e_n)\}$, such that any $x = \alpha_1 e_1 + \ldots + \alpha_n e_n \in X$ corresponds to $R(x) \in X_{\mathbb{R}}$ as

$$R(x) = \Re(\alpha_1) r(e_1) + \Im(\alpha_1) i(e_1) + \cdots + \Re(\alpha_n) r(e_n) + \Im(\alpha_n) i(e_n).$$

It is obvious that $R : X \to X_{\mathbb{R}}$ is bijective. If $(X, \|\cdot\|)$ is a normed space, we give $X_{\mathbb{R}}$ a norm defined as

$$\|R(x)\| = \|x\|.$$

A (complex) linear functional f on X corresponds to a pair of (real) linear functionals $f_{\Re}$ and $f_{\Im}$ on $X_{\mathbb{R}}$, called the real part and the imaginary part of f respectively, such that

$$f_{\Re}(r(e_i)) = \Re f(e_i), \quad f_{\Re}(i(e_i)) = -\Im f(e_i), \quad f_{\Im}(r(e_i)) = \Im f(e_i), \quad f_{\Im}(i(e_i)) = \Re f(e_i),$$

or equivalently,

$$f_{\Re}(R(x)) = \Re f(x), \quad f_{\Im}(R(x)) = \Im f(x).$$

Conversely, if there are a pair of linear functionals $f_{\Re}, f_{\Im}$ on $X_{\mathbb{R}}$ such that $f_{\Re}(r(e_i)) = f_{\Im}(i(e_i))$ and $f_{\Re}(i(e_i)) = -f_{\Im}(r(e_i))$ for all $i = 1, \dots, n$, then they correspond to a functional f on X such that

$$f(e_i) = f_{\Re}(r(e_i)) - i f_{\Re}(i(e_i)) = -f_{\Im}(i(e_i)) + i f_{\Im}(r(e_i)),$$

or equivalently,

$$f(x) = f_{\Re}(R(x)) + if_{\Im}(R(x)).$$

Note that we can recover f by $f_{\Re}$ (or $f_{\Im}$) alone.

Suppose $\|f\| = r$, then there exists $x \in X$, such that $\|x\| = 1$ and $f(x) = re^{i\theta}$. (Here we use without proof that the superum in the definition can be attained. Prove it by the Bolzano-Weierstrass theorem.) Let $x_1 = e^{-i\theta}x$ and $x_2 = ie^{-i\theta}$. We have $\|R(x)\| = \|x_1\| = \|R(x_2)\| = \|x_2\| = \|x\| = 1$. On the other hand, $f(x_1) = r$, $f(x_2) = ri$, so $f_{\Re}(x_1) = f_{\Im}(x_2) = r$. Thus $\|f_{\Re}\| \geq r$ and $\|f_{\Im}\| \geq r$. Since it is not hard to see that $\|f_{\Re}\| \leq \|f\|$ and $\|f_{\Im}\| \leq \|f\|$, we conclude that $\|f_{\Re}\| = \|f_{\Im}\| = \|f\|$.

Now come back to the construction of f_{i+1} from f_i in the case $K = \mathbb{C}$. The linear functional f_i corresponds to a pair of linear functionals $f_{i,\Re}$ and $f_{i,\Im}$ on $X_{i,\mathbb{R}}$ such that $\|f_{i,\Re}\| = \|f_{i,\Im}\| = 1$. Since $X_{i+1,\mathbb{R}}$ is two dimensional higher than its subspace $X_{i,\mathbb{R}}$, we play twice the construction of f_{i+1} from f_i in the case $K = \mathbb{R}$, and get a linear functional $\tilde{f}_{i+1,\Re}$ on $X_{i+1,\mathbb{R}}$ such that $\tilde{f}_{i+1,\Re}(R(x)) = f_{i,\Re}(R(x))$ for all $x \in X_i$, and $\|f_{i+1,\Re}\| = 1$. Then we define f_{i+1} by

$$f_{i+1}(e_{i+1}) = \tilde{f}_{i+1,\Re}(r(e_{i+1})) - i\tilde{f}_{i+1,\Re}(i(e_{i+1})).$$

It is clear that the real part of f_{i+1} is exactly $\tilde{f}_{i+1,\mathbb{R}}$. So $\|f_{i+1}\| = \|\tilde{f}_{i+1,\Re}\| = 1$. Therefore we finish the construction of f_{i+1} from f_i in the $K = \mathbb{C}$ case.

A detour Below we consider a generalization of the result that will be used later. On a vector space X (either on $\mathbb{R}$ or $\mathbb{C}$), a function $p : X \to \mathbb{R}$ is called a *sublinear functional* if it is *subadditive*, that is,

$$p(x + y) \leq p(x) + p(y) \quad \text{for all } x, y \in X,$$

and *positive-homogeneous*, that is,

$$p(\alpha x) = \alpha p(x) \quad \text{for all } \alpha \geq 0 \text{ in } \mathbb{R} \text{ and } x \in X.$$

Note that if $(X, \|\cdot\|)$ is a normed space, then the norm $\|\cdot\| : X \to \mathbb{R}$ is a special sublinear functional. For a complex vector space, a condition stronger than positive-homogeneity, but often more convenient, is the equality

$$p(\alpha x) = |\alpha| p(x) \quad \text{for all } \alpha \in \mathbb{C}. \tag{17}$$

The method of construction of f_{i+1} from f_i can be generalized to (a special case of) the celebrated Hahn-Banach theorem.

Theorem 8 (Hahn-Banach. Extension of a linear functional by one dimension)**.** *1. Let X be a vector space over $\mathbb{R}$, and p be a sublinear functional on X. Suppose Z is a subspace of X, $y \in X \setminus Z$, such that any vector $x \in X$ can be expressed in a unique way $x = z + \alpha y$, $\alpha \in \mathbb{R}$. Furthermore, let f be a linear functional which is defined on Z and satisfies*

$$f(x) \leq p(x) \quad \textit{for all } x \in Z.$$

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

$$\tilde{f}(x) \leq p(x) \quad \textit{for all } x \in X, \tag{18}$$

that is, $\tilde{f}$ is a linear functional on X, satisfies (50) *on X and $\tilde{f}(x) = f(x)$ for every $x \in Z$.*

2. Let X be a vector space over $\mathbb{C}$, and p be a mapping from X to $\mathbb{R}$ that satisfies subadditivity and condition (17)*. Suppose Z is a subspace of X, $y \in X \setminus Z$, such that any vector $x \in X$ can be expressed in a unique way $x = z + \alpha y$, $\alpha \in \mathbb{C}$. Furthermore, let f be a linear functional which is defined on Z and satisfies*

$$|f(x)| \leq p(x) \quad \textit{for all } x \in Z.$$

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

$$|\tilde{f}(x)| \leq p(x) \quad \textit{for all } x \in X. \tag{19}$$

The proof of this theorem is left as a homework problem.

Now we turn back to finite dimensional normed spaces and consider the operators on them. We state only one result:

Theorem 9. *Let $(X, \|\cdot\|)$ be a finite dimensional normed space and $(Y, \|\cdot\|)$ be any normed space. Suppose $T : X \to Y$ is a linear operator. Then T is a bounded linear operator.*

This is a generalization of the result that any linear functional on a finite dimensional normed space is bounded. The proof is left as an exercise.

1.5 Examples of infinite dimensional normed spaces

Example 4 (Two norms on a vector space may not be equivalent). Let

$$X = \{x(t) \in C[0,1] \mid x'(t) \text{ exists and } x'(t) \in C[0,1]\}.$$

Define two normed spaces $(X, \|\cdot\|_0)$ and $(X, \|\cdot\|)$ as

$$\|x(t)\|_0 = \max_{t\in[0,1]} |x(t)|, \quad \|x(t)\|_= \max_{t\in[0,1]} |x(t)| + \max_{t\in[0,1]} |x'(t)|.$$

Then $(X, \|\cdot\|_0)$ is a subspace of the normed space $C[0,1]$. (A *subspace* of a normed space is understood as a normed space that is the subspace in the sense of vector space, with the norm induced by the original normed space.) $(X, \|\cdot\|)$ is the normed space $C'[0,1]$ defined in Problem 2.8.6 in the textbook. The two norms are not equivalent. Consider the sequence of functions $\{x_n(t) = \sin(n\pi t)\}$. We have

$$\|x_n(t)\|_0 = 1, \quad \|x_n(t)\| = n + 1,$$

and get the conclusion.

Example 5 (A normed space may not be a Banach space). The normed space $(X, \|\cdot\|_0)$ defined in Example 4 is not a Banach space. Here we give a "soft" argument to show it. By the Weierstrass approximation theorem, for any continuous function $f(t) \in C[0,1]$ there is a sequence of polynomials $\{p_n\}$ such that

$$\max_{t\in[0,1]} |f(t) - p_n(t)| \to 0, \quad \text{as } n \to \infty.$$

Let $f(t)$ be defined as

$$f(t) = \begin{cases} 0 & \text{for } t \in [0, \frac{1}{2}], \\ t - \frac{1}{2} & \text{for } t \in (\frac{1}{2}, 1], \end{cases}$$

and $\{p_n\} \subseteq X$ be the corresponding sequence of polynomials. It is clear that $\{p_n\}$ form a Cauchy sequence in $(X, \|\cdot\|_0)$, but they do not converge to any point in X. Suppose to the contray, $\tilde{f} \in X$ is the limit of $\{p_n\}$, with respect to $\|\cdot\|_0$. Then there is a small $\epsilon > 0$ such that

$$\left| \frac{\tilde{f}(t) - \tilde{f}(\frac{1}{2})}{t - \frac{1}{2}} - \tilde{f}'(\frac{1}{2}) \right| < \frac{1}{2} \quad \text{for all } t \in [\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon]. \tag{20}$$

However, we have that $\tilde{f}(\frac{1}{2}) = \tilde{f}(\frac{1}{2} - \epsilon) = 0$ and $\tilde{f}(\frac{1}{2} + \epsilon) = \epsilon$. Then there is no possible value of $\tilde{f}(\frac{1}{2})$ to make (20) hold, since $(\tilde{f}(t) - \tilde{f}\frac{1}{2})/(t - \frac{1}{2})$ is 0 if $t = \frac{1}{2} - \epsilon$ and is 1 if $t = \frac{1}{2} + \epsilon$.

Actually, the completion of $(X, \|\cdot\|_0)$ is exactly $C[0, 1]$.

Example 6 (The unit closed ball is not compact in an infinite dimensional normed space).

Lemma 3 (Riesz). *Let Y and Z be subspaces of a normed space X, and suppose Y is closed and is a proper subset of Z. Then for any real number $\theta \in (0, 1)$, there is a $z \in Z$ such that*

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

Proof. Take any $x \in Z \setminus Y$. Since Y is closed, $\inf_{y \in Y} \|x - y\| > 0$ since otherwise $x \in \overline{Y} = Y$. Suppose $\inf_{y \in Y} \|x - y\| = a$. Then there is a $y_\theta \in Y$ such that $\|x - y_\theta\| \leq \theta^{-1} a$. Let

$$z = \frac{x - y_\theta}{\|x - y_\theta\|} = \frac{\theta}{a}(x - y_\theta).$$

Then $\|z\| = 1$, and for any $y \in Y$,

$$\|z - y\| = \|\frac{\theta}{a}(x - y_0) - y\| = \frac{\theta}{a}\|x - (y_0 + \frac{a}{\theta} y)\| \geq \frac{\theta}{a} a = \theta.$$

□

Then as a corollary, we have that the unit closed ball $B[0, 1] = \{x \in X \mid \|x\| \leq 1\}$ in any infinite dimensional normed space is not compact. To see it, we construct a sequence $\{x_n\} \subseteq B[0, 1]$ such that $\|x_m - x_n\| \geq 1/2$ for all $m \neq n$. Then it has no convergent subsequence, and hence $B[0, 1]$ is not (sequentially) compact.

Let $x_1 = 0$, and we choose other x_n by induction. For any $k > 0$, let Y_k be the finite dimensional subspace of X spanned by $x_1, \ldots, x_k$, and Y_k is a proper subspace of X. Later we show that Y_k is a closed subspace in X. Then we apply Lemma 3 and choose $x_{k+1} \in X \setminus Y$ such that $\|x_{k+1}\| = 1$ and $\|x - y\| \geq 1/2$ for all $y \in Y_k$. Hence $\|x_{k+1} - x_i\| \geq 1/2$ for all $i = 1, \ldots, k$. Thus we obtain the desired sequence. The closedness of Y_k is by the following lemma.

Lemma 4. *Any finite dimensional subspace $(Y, \|\cdot\|)$ of a normed space $(X, \|\cdot\|)$ is closed.*

Proof. Recall a result for metric space: Suppose (Z, d) is a metric space, $M \subseteq Z$ and (M, d) is the subspace. If (M, d) is a complete metric space, then M is a closed set in Z. (Proof: If $z \in Z$ is a limit of a sequence $\{m_n\} \subseteq M$, then $\{m_n\}$ is a Cauchy sequence in both (Z, d) and (M, d). Since (M, d) is complete, $\{m_n\}$ has a limit $z' \in M$. Then $z' = z$ since it is also the limit of $\{m_n\}$ in (Z, d).)

Since we have shown that the finite dimensional normed space $(Y, \|\cdot\|)$ is complete in Corollary 1(a), we have that it is also a closed set in X. □

Example 7 (A Banach space may not be reflexive). Consider the Banach spaces (although we do not prove that they are Banach, it is not difficult to check)

$$\begin{aligned}
\ell^p &= (\{(\alpha_1, \alpha_2, \dots) \mid \sum_{i=1}^{\infty} |\alpha_i|^p < \infty\} \quad \|(\alpha_1, \alpha_2, \dots)\|_p = \left(\sum_{i=1}^{\infty} |\alpha_i|^p\right)^{\frac{1}{p}}, \quad (1 \le p < \infty) \\
\ell^\infty &= (\{(\alpha_1, \alpha_2, \dots) \mid \sup_{i=1}^{\infty} |\alpha_i| < \infty\} \quad \|(\alpha_1, \alpha_2, \dots)\|_\infty = \sup_{i=1}^{\infty} |\alpha_i|, \\
c_0 &= \text{subspace of } \ell^\infty \text{ such that } \lim_{i \to \infty} \alpha_i = 0.
\end{aligned}$$

From a tutorial problem, the Banach space c_0 has its dual space ℓ^1. A homework problem will show that the dual space of ℓ^1 is ℓ^∞, that is bigger than c_0. (The canonical mapping can be defined as the inclusion of c_0 in ℓ^∞.)

2 Inner product spaces and Hilbert spaces

2.1 Definitions and properties

Now we revisit the archetypes of real and complex norm spaces, the Euclidean n space and the unitary n-space. In these spaces, not only is the norm of a vector defined, but the norm is defined by another structure, the inner product of a pair of vectors. Thus they are also archetypes of real and complex inner product spaces.

Definition 9. 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. (Complete in the metric defined by the inner product as below.) Here an inner product on X is a mapping of $X \times X$ into the scalar field K of X; that is, whith every pair of vectors x and y there is associated a scalar which is written $\langle x, y\rangle$ and is called the inner product of x and y, such that for all vectors x, y, z and scalars

$$\begin{aligned} \langle x+y, z\rangle &= \langle x, z\rangle + \langle y, z\rangle, \\ \langle \alpha x, y\rangle &= \alpha\langle x, y\rangle, \\ \langle x, y\rangle &= \overline{\langle y, x\rangle}, \\ \langle x, x\rangle &\geq 0 \quad \text{and} \quad \langle x, x\rangle = 0 \Leftrightarrow x = 0. \end{aligned}$$

Since an inner product on X defines a norm on X given by

$$\|x\| = \sqrt{\langle x, x\rangle}$$

and then a metric given by

$$d(x, y) = \|x - y\| = \sqrt{\langle x-y, x-y\rangle},$$

an inner product is a normed space and a Hilbert space is a Banach space.

If the scalar field is $\mathbb{R}$, then we have $\langle x, y\rangle = \langle y, x\rangle$. If $K = \mathbb{C}$, it is not hard to derive from the definition of inner product that

$$\begin{aligned} \langle \alpha x + \beta y, z\rangle &= \alpha\langle x, z\rangle + \beta\langle y, z\rangle, \\ \langle x, \alpha y\rangle &= \overline{\langle \alpha y, x\rangle} = \overline{\alpha\langle y, x\rangle} = \bar{\alpha}\overline{\langle y, x\rangle} = \bar{\alpha}\langle x, y\rangle, \\ \langle x, \alpha y + \beta z\rangle &= \bar{\alpha}\langle x, y\rangle + \bar{\beta}\langle y, z\rangle. \end{aligned}$$

A not so ovbious property of inner product is the *parallelogram equality*

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

Actually, a normed space $(X, \|\cdot\|)$ is an inner product space if and only if the norm satisfies the parallelogram equality. The "only if" part is shown above, and the "if" part follows that if $\|\cdot\|$ satisfies the equality, then the mapping $I : X \times X \to K$ defined by the *polarization identity*

$$I(x, y) = \begin{cases} \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) & \text{if } K = \mathbb{R}, \\ \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) + \frac{i}{4}(\|x+iy\|^2 - \|x-iy\|^2) & \text{if } K = \mathbb{C}, \end{cases}$$

satisfies all the conditions required for an innter product, and as an inner product it defines the norm $\|\cdot\|$. (Exercise)

Example 8 (ℓ^2-space). The inner product defined on the (real or complex) ℓ^2-space is

$$\langle(\alpha_1, \alpha_2, \dots), (\beta_1, \beta_2, \dots)\rangle = \sum_{i=1}^{\infty} \alpha_i \overline{\beta}_i.$$

It is straightforward to check that the inner product is defined by the norm on ℓ^2-space via the polarization identity.

Before giving the next example, we consider the *completion* of an inner product space. Let $(X, \langle,\rangle)$ be an inner product space, with the norm $\|\cdot\|$ and metric d defined by the inner product. Then there is a unique complete metric space $(\hat{X}, \hat{d})$ that is the completion of (X, d), in the sense that there is an isometry $A : X \to \hat{X}$ and $A(X)$ is dense in $\hat{X}$. Theorem 1 shows that $\hat{X}$ is a vector space and there is a norm $\|\cdot\|_1$ defined on $\hat{X}$ such that $(\hat{X}, \|\cdot\|_1)$ is a normed space (and then a Banach space) and $\|Ax\|_1 = \|x\|$ for all $x \in X$. Below we show that $\|\cdot\|_1$ also satisfies the parallelogram equality. We need

Lemma 5. *In a normed space, the norm is a continuous mapping to* $\mathbb{R}$.

The proof is a simple exercise. Then for any $\hat{x}, \hat{y} \in \hat{X}$, there are $\{Ax_n\} \to \hat{x}$ and $\{Ay_n\} \to \hat{y}$. We have

$$\begin{aligned}
\|\hat{x}+\hat{y}\|_1^2 + \|\hat{x}-\hat{y}\|_1^2 &= \lim_{n\to\infty} \|Ax_n + Ay_n\|_1^2 + \|Ax_n - Ay_n\|_1^2 \\
&= \lim_{n\to\infty} \|x_n + y_n\|^2 + \|x_n - y_n\|^2 \\
&= \lim_{n\to\infty} 2(\|x_n\|^2 + \|y_n\|^2) \\
&= \lim_{n\to\infty} 2(\|Ax_n\|_1^2 + \|Ay_n\|_1^2) \\
&= 2(\|\hat{x}\|_1^2 + \|\hat{y}\|_1^2).
\end{aligned}$$

Thus we have that on $\hat{X}$ the inner product

$$\langle x, y\rangle_1 = \begin{cases} \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) & \text{if } K = \mathbb{R}, \\ \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) + \frac{i}{4}(\|x+iy\|^2 - \|x-iy\|^2) & \text{if } K = \mathbb{C}, \end{cases}$$

is well defined so that $(\hat{X}, \langle,\rangle)$ is a Hilbert space, and it is easy to check that

$$\langle Ax, Ay\rangle_1 = \langle x, y\rangle, \quad \text{for all } x, y \in X.$$

Therefore we prove that

Theorem 10. *For any inner product space X there exists a Hilbert space H and an isomorphism A from X onto a dense subspace $W \subseteq H$, that is, $A : X \to W$ is bijective and $\langle Ax, Ay\rangle = \langle x, y\rangle$ for all $x, y \in X$. The space H is unique except for isomorphisms.*

A byproduct of the proof is

Corollary 2. *In an inner product space $(X, \langle,\rangle)$, $\langle x, y\rangle$ is a continuous mapping from $X \times X$ to K in the sense that if $\{x_n\} \to x$, $\{y_n\} \to y$, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$.*

The proof is by Lemma 5 and the polarization identity.

Example 9 ($L^2[a,b]$). First we have the metric space

$$C^2[a,b] = \left(\{\text{continuous functions on } [a,b]\}, d(x(t), y(y)) = \left(\int_a^b |x(t) - y(t)|^2 dt \right)^{\frac{1}{2}} \right).$$

(In Mathematical Analysis III, we only consider real-valued functions for $C^2[a,b]$, but the definition makes sense also for complex-valued functions.) This metric space is induced by the real/complex normed space

$$C^2[a,b] = \left(\{\text{continuous functions on } [a,b]\}, \|x(t)\| = \left(\int_a^b |x(t)|^2 dt \right)^{\frac{1}{2}} \right).$$

We can verify that the norm on $C^2[a,b]$ satisfies the parallelogram equality, so it is also a real/complex inner product space, such that

$$C^2[a,b] = \left(\{\text{continuous functions on } [a,b]\}, \langle x(t), y(t) \rangle = \left(\int_a^b |x(t)\bar{y}(t)| dt \right)^{\frac{1}{2}} \right).$$

Note that $C^2[a,b]$ is not a complete metric space. We denote its completion by $L^2[a,b]$. By Theorem 1, $L^2[a,b]$ is also a normed space (and then a Banach space), such that if $\hat{x} \in L^2[a,b]$ and $\{x_n\} \to \hat{x}$ where $\{x_n\} \subseteq C^2[a,b]$, then

$$\|\hat{x}\| = \hat{d}(\hat{x}, 0) = \lim_{n\to\infty} d(x_n, 0) = \lim_{n\to\infty} \|x_n\|.$$

Similarly we check that the norm on $L^2[a,b]$ satisfies the parallelogram equality, so $L^2[a,b]$ is an inner product space (and then a Hilbert space) with the inner product given by the polarization identity.

Lemma 6 (Schwarz inequality and triangle inequality)**.** *In an inner product space $(X, \langle , \rangle)$, any $x, y \in X$ satisfy the Schwarz inequality*

$$|\langle x, y \rangle| \leq \|x\|\|y\|, \tag{21}$$

where the equality sign holds if and only if x, y are dependent, and the triangle inequality

$$\|x + y\| \leq \|x\| + \|y\| \tag{22}$$

where the equality sign holds if and only if $y = 0$ or $x = cy$ where c is real and nonnegative.

Proof. The case that either x or y is 0 is simple. Hence we assume that both x and y are nonzero below.

In the case that x and y are linearly dependent, we denote

$$e_1 = \frac{x}{\|x\|}, \tag{23}$$

such that $\|e_1\| = 1$ and $x = \|x\|e_1$. We also suppose $y = \alpha e_1$, and then

$$\|y\| = |\alpha|, \quad \|x + y\| = |\|x\| + \alpha|, \quad |\langle x, y \rangle| = |\alpha|\|x\|.$$

Hence we verify that (21) takes the equality sign and (22) also holds, while the equality sign in (22) holds if and only if α is real and nonnegative.

In the case that x and y are linearly independent, we denote e_1 by (23), and furthermore

$$e_2 = \frac{y - \langle y, e_1\rangle e_1}{\|y - \langle y, e_1\rangle e_1\|}.$$

Note that $\|e_2\| = 1$ and

$$\langle e_1, e_2\rangle = \frac{1}{\|y - \langle y, e_1\rangle e_1\|}(\langle y, e_1\rangle - \overline{\langle e_1, y\rangle}\langle e_1, e_1\rangle) = 0.$$

Write $x = \|x\|e_1$ and $y = \alpha e_1 + \beta e_2$ ($\beta \neq 0$). We have

$$\|y\| = \sqrt{|\alpha|^2 + |\beta|^2}, \quad \|x + y\| = \sqrt{|\|x\| + \alpha|^2 + |\beta|^2}, \quad |\langle x, y\rangle| = |\alpha|\|x\|.$$

After some work, we check that both inequalities (21) and (22) holds strictly. □

The vectors e_1, e_2 are *orthogonal* to each other since $\langle e_1, e_2\rangle = 0$. We denote the orthogonality as $e_1 \perp e_2$.

Since in an inner product space, the norm satisfies the parallelogram equality, inner product spaces have special properties not available for general normed spaces.

Theorem 11. *Let $(X, \langle,\rangle)$ be an inner product space and $M \neq \emptyset$ a convex subset. (That is, for all $x, y \in M$ and $\alpha \in (0, 1)$, $\alpha x + (1 - \alpha)y \in M$.) Suppose (M, d) is a complete metric space where d is the metric induced by the inner product $\langle,\rangle$. Then for any given $x \in X$, there exists a unique $y \in M$ such that*

$$\delta = \inf_{\tilde{y}\in M} \|x - \tilde{y}\| = \|x - y\|. \tag{24}$$

Proof. By the assumption, there is $\{y_n\} \subseteq M$ such that

$$\|x - y_n\| = \delta_n, \quad \lim_{n\to\infty} \delta_n = \delta.$$

Below we show that $\{y_n\}$ is a Cauchy sequence. For any $\epsilon > 0$, there is an N_ϵ such that for all $n > N_\epsilon$, $\|x - y_n\| < \delta + \epsilon$. Then for any $m, n > N_\epsilon$, by the parallelogram equality,

$$\|(x - y_m) + (x - y_n)\|^2 + \|(x - y_m) - (x - y_n)\|^2 = 2(\|x - y_m\|^2 + \|x - y_n\|^2),$$

or equivalently,

$$\|x - \frac{y_m + y_n}{2}\|^2 + \frac{1}{4}\|y_n - y_m\|^2 = \frac{1}{2}(\|x - y_m\|^2 + \|x - y_n\|^2).$$

Since $\|x - (y_m + y_n)/2\| \geq \delta$, $\|x - y_m\| < \delta + \epsilon$ and $\|x - y_n\| < \delta + \epsilon$, we have

$$\|y_n - y_m\| < \sqrt{\frac{\delta\epsilon}{2} + \frac{\epsilon^2}{4}}.$$

Thus we conclude that $\{y_n\}$ is Cauchy and then it has a limt $y \in M$. By assumption, $\|x - y\| \geq \delta$, and on the other hand, by

$$\|x - y\| \leq \|x - y_n\| + \|y_n - y\|$$

and that $\|x-y_n\| \to \delta$ and $\|y_n - y\| \to 0$ as $n \to \infty$, we have $\|x-y\| \leq \delta$. So $\|x-y\| = \delta$.

To show that the vector $y \in M$ satisfying (24) is unique, we suppose $y' \in M$ also satisfies (24). Then $(y+y')/2 \in M$, and by the parallelogram equality

$$\|x - \frac{y+y'}{2}\|^2 + \frac{1}{4}\|y-y'\| = \frac{1}{2}(\|x-y\|^2 + \|x-y'\|) = \delta.$$

We conclude that $y' = y$, otherwise we get the contradiction $\|x - \frac{y+y'}{2}\| < \delta$. $\square$

A consequence of Theorem 11 is that suppose Y is a subspace of X as an inner product space, (that is, Y is a subspace of X as a vector space, and Y has an inner product induced by that of X,) and Y is complete as a metric space (or equivalently, Y is a Hilbert space), then for any $x \in X$, there is a $y \in Y$ such that

$$\delta = \inf_{\tilde{y} \in Y} \|x - \tilde{y}\| = \|x - y\|. \tag{25}$$

Furthermore, we have

Lemma 7. *Let X be an inner product space and Y be a subspace of X. Suppose Y is complete as a metric space. For any $x \in X$, let $y \in Y$ be the vector such that* (25) *holds. Then $z = x - y$ is* orthogonal *to all $\tilde{y} \in Y$, that is, $\langle z, \tilde{y} \rangle = 0$ for all $\tilde{y} \in Y$, or denoted as $z \perp Y$. Conversely, if $z = x - y$ with $y \in Y$ is orthogonal to all $\tilde{y} \in Y$, then y is the vector that makes* (25) *holds.*

Proof. Let z satisfies the orthogonality condition. Then for any $\tilde{y} \in Y$,

$$\|x - \tilde{y}\|^2 = \|x - y\|^2 + \|y - \tilde{y}\|^2 \geq \|x - y\|^2,$$

and we have that y satisfies (25).

On the other hand, suppose there is $y' \in Y$ such that

$$\langle z, y' \rangle = re^{i\theta} \neq 0, \quad r > 0,$$

then let $y'' = e^{i\theta} y'$, and

$$\langle z, y'' \rangle = r > 0.$$

For $\epsilon > 0$, we have

$$\|x - (y - \epsilon y'')\|^2 = \|z - \epsilon y''\|^2 = \|z\|^2 - \langle z, \epsilon y'' \rangle - \langle \epsilon y'', z \rangle + \langle \epsilon y'' \rangle = \|z\|^2 - 2\epsilon r + \epsilon^2 \|y''\|^2.$$

If ϵ is a small enough positive number, then $-2\epsilon r + \epsilon^2 \|y''\|^2 < 0$ and then $\|x - (y - \epsilon y'')\| < \|x - y\|$ and x cannot satisfy (25). $\square$

Now let X be a Hilbert space, and Y be a closed subspace of X (as an inner product space). Then Y is also a Hilbert space (why?). Then for any $x \in X$, there is a unique $y \in Y$ such that (25) holds. We call y the *orthogonal projection* of x on Y, and denote it as Px. We understand P as an operator from X to Y. It is an exercise to show that P is a bounded linear operator.

For any subspace $Z \subseteq X$, define the *orthogonal complement* by

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

If Y is a closed subset of X, we have that

Lemma 8. *The orthogonal complement $Y^\perp$ of a closed subspace Y of a Hilbert space X is the null space $N(P)$ of the orthogonal projection of X onto Y, that is,*

$$Y^\perp = \{x \in X \mid Px = 0\}.$$

This lemma is a direct corollary of Lemma 7. Another consequence of Lemma 7 is the following result.

Theorem 12. *Let Y be a closed subspace of a Hilbert space X. Then $X = Y \oplus Y^\perp$, that is, any $x \in X$ can be written in a unique way as $x = y + z$, where $y \in Y$ and $z \in Y^\perp$.*

Proof. For any $x \in X$, by the orthogonal projection P onto Y, we write

$$x = Px + (x - Px).$$

By Lemma 7, $x - Px \in Y^\perp$. On the other hand, if we have

$$x = y + z, \quad y \in Y \text{ and } z \in Y^\perp,$$

then since $z = x - y \perp Y$, using Lemma 7 in another direction, we have that $y = Px$. □

Since $Y^\perp$ is also a subspace of X, we can consider its orthogonal complement

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

It is clear that $Y \subseteq Y^{\perp\perp}$, since any $y \in Y$ satisfies $y \perp z$ for all $z \in Y^\perp$. On the other hand, we have

Lemma 9. *If Y is a closed subspace of a Hilbert space X, then $Y = Y^{\perp\perp}$.*

Proof. Suppose $\tilde{y} \in Y^{\perp\perp}$, we use Theorem 12 to write

$$\tilde{y} = y + z,$$

where $y \in Y$ and $z \in Y^\perp$. Then since $\tilde{y} \perp z \in Y^\perp$, we have

$$\langle \tilde{y}, z\rangle = \langle y, z\rangle + \langle z, z\rangle = 0,$$

and then $\|z\| = 0$, or equivalently, $z = 0$. So $\tilde{y} = y \in Y$. □

The following lemma will be used later.

Lemma 10. *For any subset M of a Hilbert space X, the span of M is dense in X if and only if $M^\perp = 0$. (Here $M^\perp = \{x \in X \mid x \perp y \text{ for all } y \in M\}$.*

Proof. Suppose $\operatorname{span} M$ is dense in X, and $0 \neq x \in X$. Then there is $x' \in \operatorname{span} M$ such that $\|x' - x\| \leq \|x\|/2$. We have

$$\langle x, x'\rangle = \langle x, x\rangle + \langle x, x' - x\rangle \geq \|x\|^2 - \frac{1}{2}\|x\|^2 = \frac{1}{2}\|x\|^2 > 0.$$

Hence $x' \notin (\operatorname{span} M)^\perp$, and we conclude that $(\operatorname{span} M)^\perp = \{0\}$.

On the other hand, suppose $\operatorname{span} M$ is not dense in X. Let $Y = \overline{\operatorname{span} M}$. We have that Y is a vector space. To see it, we check

- If $y \in Y$ such that $\{y_n\} \to y$ where $\{y_n\} \subseteq \operatorname{span} M$, then $\{\alpha y_n\} \to \alpha y$ and then $\alpha y \in Y$.
- If $y, y' \in Y$ such that $\{y_n\} \to y$ and $\{y'_n\} \to y'$ where $\{y_n\}, \{y'_n\} \subseteq Y$, then $\{y_n + y'_n\} \to y + y'$ and then $y + y' \in Y$.

By assumption, Y is a closed and proper subspace of X. So by Riesz's Lemma 3, there is $x \in X$ such that $\|x\| = 1$ and $\|x - y\| > 1/2$ for all $y \in Y$. By Theorem 12, x has a decomposition $x = y + z$ where $y \in Y$ and $z \in Y^\perp$. It is clear that $z \neq 0$, and so $Y^\perp \neq \{0\}$. □

2.2 Orthonormal basis

In an inner product space, we say a set M sa an *orthonormal set* if any $x \in M$ has norm $\|x\| = 1$ and any pair of distinct $x, y \in M$ are orthogonal to each other $\langle x, y\rangle = 0$.

First we consider finite orthonormal sets. Suppose $\{e_1, \dots, e_n\}$ are orthonormal in an inner product space X. They span a finite dimensional subspace, denoted as X_n. It is easy to compute the norms and inner products of vectors in X_n:

$$\begin{aligned}
&\langle \alpha_1 e_1 + \cdots + \alpha_n e_n, \beta_1 e_1 + \cdots + \beta_n e_n\rangle \\
&= \alpha_1 \langle e_1, \beta_1 e_1 + \cdots + \beta_n e_n\rangle + \ldots + \alpha_n \langle e_n, \beta_1 e_1 + \cdots + \beta_n e_n\rangle \\
&= \alpha_1 \bar{\beta}_1 + \cdots + \alpha_n \bar{\beta}_n,
\end{aligned}$$

$$\|\alpha_1 e_1 + \ldots + \alpha_n e_n\| = \sqrt{|\alpha_1|^2 + \cdots + |\alpha_n|^2}.$$

By these formulas, we have that $e_1, \dots, e_n$ are linearly independent.

For any $x = \alpha_1 e_1 + \ldots \alpha_n e_n \in X_n$, it is clear that $\alpha_i = \langle x, e_i\rangle$, and we have the identity

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

Generally for $x \in X$, we have the decomposition $x = P_{X_n} x + z$ where $z \in X_n^\perp$, and

$$\sum_{i=1}^n |\langle x, e_i\rangle|^2 = \sum_{i=1}^n |\langle P_{X_n} x, e_i\rangle|^2 = \|P_{X_n} x\|^2 \leq \|x\|^2,$$

where the last inequality is because

$$\begin{aligned}
\|x\|^2 &= \langle P_{X_n} x + z, P_{X_n} x + z\rangle = \langle P_{X_n} x, P_{X_n} x\rangle + \langle P_{X_n} x, z\rangle + \langle z, P_{X_n} x\rangle + \langle z, z\rangle \\
&= \|P_{X_n} x\|^2 + \|z\|^2.
\end{aligned}$$

Now consider an *orthonormal sequence* $\{e_n\}_{n=1}^\infty$, such that

$$\langle e_i, e_j\rangle = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

We have the following result

Theorem 13 (Bessel inequality)**.** *Let $\{e_n\}$ be an orthonormal sequence in an inner product space X. Then for any $x \in X$*

$$\sum_{i=1}^{\infty} |\langle x, e_i\rangle|^2 \le \|x\|^2.$$

Proof. Note that $\sum_{i=1}^{n} |\langle x, e_i\rangle|^2 \le \|x\|^2$ for all n, and then use the monotone convergence theorem. □

Suppose $\{x_1, x_2, \dots\}$ be a sequence of linearly independent bectors in an inner product space. We can find an orthonormal sequence $\{e_1, e_2, \dots\}$ such that for any n

$$\operatorname{span}\{e_1, \dots, e_n\} = \operatorname{span}\{x_1, \dots, x_n\}. \tag{26}$$

An abstract construction of $\{e_n\}$ is as follows. Denote $X_n = \operatorname{span}\{x_1, \dots, x_n\}$ for all n. Then X_n is a closed subspace of X. Let $e_1 = \|x_1\|^{-1} x_1$, and for any $n > 1$, let

$$x_n = P_{X_{n-1}} x_n + z_n$$

be the decomposition of x_n with respect to X_{n-1}, such that $z_n \in X_{n-1}^{\perp}$. Since $x_1, \dots, x_n$ are linearly independent, $z_n \neq 0$, and we define $e_n = \|z_n\|^{-1} z_n$. It is not hard to check that $\{e_n\}$ are orthonormal and they satisfiy (26).

To make the construction practical, we need to have a convenient way to compute $P_{X_{n-1}}(x_n)$. Suppose we have already had $e_1, \dots, e_{n-1}$, then we check that

$$P_{X_{n-1}}(x_n) = \langle x_n, e_1\rangle e_1 + \dots + \langle x_n, e_{n-1}\rangle e_{n-1}.$$

Since any $x \in X_{n-1}$ can be written as $\alpha_1 e_1 + \dots + \alpha_{n-1} e_{n-1}$, we only need to check that

$$\langle \alpha_1 e_1 + \dots + \alpha_{n-1} e_{n-1}, z_n\rangle = A - B = 0,$$

where

$$\begin{aligned} A &= \langle \alpha_1 e_1 + \dots + \alpha_{n-1} e_{n-1}, x_n\rangle, \\ B &= \langle \alpha_1 e_1 + \dots + \alpha_{n-1} e_{n-1}, \langle x_n, e_1\rangle e_1 + \dots + \langle x_n, e_{n-1}\rangle e_{n-1}\rangle. \end{aligned}$$

By the orthonormality of $\{e_i\}$, we compute that $A = B = \alpha_1 \overline{\langle x_n, e_1\rangle} + \dots + \alpha_{n-1} \overline{\langle x_n, e_{n-1}\rangle}$.

The method of constructing an orthonormal sequence from an arbitrary linearly independent sequence is called the *Gram-Schmidt process.*

Example 10. In the inner product space $C^2[0,1]$ with scalar field $K = \mathbb{C}$, consider the sequence $\{e_n(t)\}_{n=-\infty}^{\infty}$, such that

$$e_n(t) = e^{2n\pi i t}.$$

This is an orthonormal sequence, because

$$\langle e_j, e_k\rangle = \int_0^1 e^{2j\pi i t} \overline{e^{2k\pi i t}} dt = \int_0^1 e^{2j\pi i t} e^{-2k\pi i t} dt = \int_0^1 e^{2(j-k)\pi i t} dt = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{otherwise.} \end{cases}$$

Here we note that it is convenient to take $K = \mathbb{C}$. If $K = \mathbb{R}$, we have a similar but less elegant result.

Example 11. In the inner product space $C^2[0,1]$ with scalar field $K = \mathbb{R}$, consider the sequence $\{e_n(t)\}_{n=-\infty}^{\infty}$ where

$$e_n(t) = \begin{cases} 1 & \text{if } n = 0, \\ \sqrt{2}\cos(2n\pi t) & n > 0, \\ \sqrt{2}\sin(-2n\pi t) & n < 0. \end{cases}$$

Then

$$\langle e_j, e_k \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{otherwise.} \end{cases}$$

We only check that for $j, k > 0$,

$$\begin{aligned} \langle e_j, e_k \rangle &= \int_0^1 2\cos(2j\pi t)\cos(2k\pi t)dt = \int_0^1 [\cos((j-k)\pi t) + \cos((j+k)\pi t)]dt \\ &= \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Example 12 (Legendre polynomials). Consider the inner product space $C^2[-1,1]$. We construct an orthonormal sequence by performing the Gram-Schmidt process to the monomials $\{1, x, x^2, \dots\}$. So e_k $(k = 0, 1, \dots)$ is a polynomial of degree k. It is easy to find that $e_0 = \frac{1}{\sqrt{2}}$, $e_1 = \sqrt{\frac{3}{2}}x$, but the calculation becomes more tedious as k becomes large. Luckily, there is a simple formula for all e_k.

Consider the *Legendre polynomials*

$$P_k(t) = \frac{1}{2^k k!}\frac{d^k}{dt^k}[(t^2-1)^k].$$

It is clear that P_k is a polynomial of degree k. They are orthogonal in $C[-1,1]$. To check it, we suppose $j \le k$, and compute

$$\begin{aligned} \langle P_j, P_k \rangle = \int_{-1}^1 P_j(t)P_k(t)dt &= \frac{1}{2^{j+k}j!k!}\int_{-1}^1 \frac{d^j}{dt^j}[(t^2-1)^j]\frac{d^k}{dt^k}[(t^2-1)^k]dt \\ &= \frac{-1}{2^{j+k}j!k!}\int_{-1}^1 \frac{d^{j+1}}{dt^{j+1}}[(t^2-1)^j]\frac{d^{k-1}}{dt^{k-1}}[(t^2-1)^k]dt, \end{aligned}$$

where we use the integration by parts and that

$$\frac{d^j}{dt^j}[(t^2-1)^j]\underbrace{\frac{d^{k-1}}{dt^{k-1}}[(t^2-1)^k]}_{=0 \text{ at } t=\pm 1}$$

vanishes at ± 1. Use this trick repeatedly, we have

$$\langle P_j, P_k \rangle = \frac{(-1)^k}{2^{j+k}j!k!}\int_{-1}^1 \frac{d^{j+k}}{dt^{j+k}}[(t^2-1)^j](t^2-1)^k dt.$$

If $j < k$, then by counting the degree of $\frac{d^{j+k}}{dt^{j+k}}[(t^2-1)^j]$, we have that

$$\frac{d^{j+k}}{dt^{j+k}}[(t^2-1)^j] = 0$$

and then $\langle P_j, P_k \rangle = 0$. If $j = k$, then

$$\frac{d^{j+k}}{dt^{j+k}}[(t^2-1)^j] = (2j)!$$

and then

$$\begin{aligned}
\langle P_j, P_k \rangle &= \frac{(-1)^j(2j)!}{2^{2j}(j!)^2}\int_{-1}^{1}(t^2-1)^j dt \\
&= \frac{(-1)^j(2j)!}{2^{2j}(j!)^2}(-1)^j \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^{2j}\theta d\sin\theta \\
&= \frac{(2j)!}{2^{2j}(j!)^2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^{2j+1}\theta d\theta \\
&= \frac{(2j)!}{2^{2j}(j!)^2}\cdot 2\frac{2}{3}\frac{4}{5}\cdots\frac{2j}{2j+1} \\
&= \frac{2}{2j+1}.
\end{aligned}$$

Thus we have the orthonormal sequence $\{e_k = \sqrt{\frac{2k+1}{2}}P_k(t)\}_{k=0}^{\infty}$.

Let $\{e_n\}_{n=1}^{\infty}$ be an orthonormal sequence in a Hilbert space H. We consider the series $\sum_{k=1}^{\infty}\alpha_k e_k$ where α_k are scalars. The meaning and well-definedness of the series is given in the following theorem.

Theorem 14. *Let $\{e_n\}_{n=1}^{\infty}$ be an orthonormal sequence in a Hilbert space H.*

(a) The sequence $\{s_n\}$ where

$$s_n = \sum_{k=1}^{n}\alpha_k e_k \tag{27}$$

converges if and only if $\sum_{k=1}^{\infty}|\alpha_k|^2 < \infty$. If $\{s_n\}$ converges, we denote its limit as $\sum_{k=1}^{\infty}\alpha_k e_k$.

(b) If $\{s_n\}$ defined by (27) converges to $x \in H$, then $\alpha_k = \langle x, e_k\rangle$. In other words, for this x,

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

(c) For any $x \in X$, denote $\alpha_k = \langle x, e_k\rangle$, then the sequence $\{s_n\}$ defined by (27) converges, or equivalently, $\sum_{k=1}^{\infty}\alpha_k e_k$ exists, though it may not be equal to x.

Proof. (a) If $\sum_{k=1}^{\infty}|\alpha_k| < \infty$, then $\{s_n\}$ is a Cauchy sequence, since for any $n > m$,

$$\|s_n - s_m\|^2 = \|\sum_{k=m+1}^{n} \alpha_k e_k\|^2 = \sum_{k=m+1}^{n} |\alpha_k|^2.$$

By the same reason, if $\sum_{k=1}^{\infty}|\alpha_k| = \infty$, then $\{s_n\}$ is not a Cauchy sequence. Thus in a Hilbert space where a sequence converges if and only if it is Cauchy, $\sum_{k=1}^{\infty} \alpha_k e_k$, the limit of s_n, converges if and only if $\sum_{k=1}^{\infty}|\alpha_k|^2 < \infty$.

(b) By the continuity of the inner product,

$$\langle x, e_k \rangle = \lim_{n\to\infty} \langle s_n, s_k \rangle = \alpha_k.$$

(c) By the Bessel inequality,

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

□

Definition 10. A *total set* in a normed space X is a subset $M \subseteq X$ whose span is everywhere dense in X. A *total orthonormal set*, or an *orthonormal basis* in an inner product space X is a total set that is also orthonormal.

Note that in a finite dimensional inner product space, an orhonormal basis is a special kind of (Hamel) basis. But in an infinite dimensional inner product space, an orthonormal basis is not a Hamel basis, as we can see in example 13. It is called a basis in the sense of *Schauder basis* (whose definition can be found in the textbook or `http://en.wikipedia.org/wiki/Schauder_basis`).

The totality of an arbitrary set is characterized as follows.

Theorem 15. *Let M be a subset of an inner product space X. Then*

(a) If M is total in X, then there does not exist a nonzero $x \in X$ which is orthogonal to every element of M, that is,

$$x \perp M \Rightarrow x = 0.$$

(b) If X is complete, then the condition above is also sufficient for the totality of M in X.

Proof. (a) We just need to show that if there is an $x \neq 0$ such that $x \perp M$, then M is not total, that is, $\overline{\operatorname{span} M} \neq X$. Suppose otherwise $\overline{\operatorname{span} M} = X$, then for any $x \neq 0$, there is $\alpha_1 m_1 + \cdots + \alpha_k m_k \in \operatorname{span} M$ where $m_1, \cdots, m_k \in M$, such that

$$\|x - (\alpha_1 m_1 + \cdots + \alpha_k m_k)\| < \frac{\|x\|}{2}. \tag{28}$$

On the other hand, by the orthogonality,

$$\langle x, (\alpha_1 m_1 + \cdots + \alpha_k m_k)\rangle = \bar{\alpha}_1\langle x, m_1\rangle + \cdots + \bar{\alpha}_k\langle x, m_k\rangle = 0,$$

so

$$\|x - (\alpha_1 m_1 + \cdots + \alpha_k m_k)\|^2 = \|x\|^2 + \|\alpha_1 m_1 + \cdots + \alpha_k m_k\|^2 \geq \|x\|^2,$$

which is contradictory to (28).

(b) We just need to show that if M is total, then there is such an $x \neq 0$ that is orthogonal to M, given that X is complete. Note that $\overline{\operatorname{span} M}$ is a closed subspace of X, and by the assumption it is not X itself. Let $z \in X \setminus \overline{\operatorname{span} M}$ and use the orthogonal projection onto $\overline{\operatorname{span} M}$ to decompose z into

$$z = x + y, \quad \text{where } y \in \overline{\operatorname{span} M} \text{ and } x \in (\overline{\operatorname{span} M})^{\perp}.$$

Then the x satisfies our requirement.

□

In any Hilbert space, there exists an orthonormal basis, as claimed without proof in the textbook. The proof requires Zorn's lemma, or the axiom of choice, see `http://en.wikipedia.org/wiki/Orthonormal_basis`. But in some infinite dimensional incomplete inner product spaces, there is not an orthonormal basis, which is not in the textbook but can be found in J. Dixmier, *Sur les bases orthonormales dans les espaces prehilbertiens*, Acta Sci. Math. Szeged **15** (1953) 29-30. Hence the incomplete inner product spaces are more troublesome, and we deal with only Hilbert spaces in the remaining part of the section.

We concentrate on Separable Hilbert spaces, where the existence of orthonormal bases is straightforward to prove. Suppose $M = \{m_n\}$ is everywhere dense in H, a Hilbert space. If $\operatorname{span} M$ is a finite dimensional subspace of H, then $H = \overline{\operatorname{span} M} = \operatorname{span} M$ since $\operatorname{span} M$ is closed by Lemma 4, and the result is obvious. If $\operatorname{span} M$ is infinite dimensional, without loss of generality, we assume that all vectors in $M = \{m_n\}_{n=1}^{\infty}$ are linearly independent. Then use the Gram-Schmidt process we construct an orthonormal sequence $\{e_n\}_{n=1}^{\infty}$ such that for any $n \geq 1$, $\operatorname{span}\{e_1, e_2, \dots, e_n\} = \operatorname{span}\{m_1, m_2, \dots, m_n\}$. It is clear that $\operatorname{span}\{e_n\}_{n=1}^{\infty} = \operatorname{span}\{m_n\}_{n=1}^{\infty}$, and then $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis.

The countability of orthonormal basis is equivalent to the separability of the Hilbert space.

Theorem 16. *Let H be a Hilbert space.*

(a) If H is separable, every orthonormal set is countable.

(b) If H has a countable orthonormal basis, then H is separable.

Proof. (a) Suppose H is a separable Hilbert space with a countable dense subset C. Let $M = \{m_\alpha\}_{\alpha \in I}$ be an orthonormal set where the index set I is uncountable. Let $B_\alpha = \{x \in H \mid \|x - m_\alpha\| < 1/2\}$ be an open ball for all $\alpha \in I$. Then these B_α are disjoint to each other. Each B_α contains at least one $c_\alpha \in C$, and these c_α are distinct since B_α are disjoint. So the subset $\{c_\alpha\}_{\alpha \in I}$ of C is uncountable, contradictory to that C is countable.

(b) Suppose $\{e_n\}_{n=1}^\infty$ is an orthonormal basis. Then for any $x \in H$ and $\epsilon > 0$, there are $n_1, \ldots, n_k \in \mathbb{Z}_+$ and $\alpha_1, \ldots, \alpha_k \in K$ such that

$$\|x - (\alpha_1 e_{n_1} + \cdots + \alpha_k e_{n_k})\| < \frac{\epsilon}{2}.$$

Let $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_k$ be either rational numbers if $K = \mathbb{R}$ or rational complex numbers in the form of $p + iq$ where $p, q \in \mathbb{Q}$ if $K = \mathbb{C}$, such that they are close to $\alpha_1, \ldots, \alpha_k$. It suffices to let $|\tilde{\alpha}_i - \alpha_i| < \epsilon/(2\sqrt{n})$. Then we have

$$\|x - (\tilde{\alpha}_1 e_{n_1} + \cdots + \tilde{\alpha}_k e_{n_k})\| < \epsilon.$$

Thus the countable set

$$\{\tilde{\alpha}_1 e_{n_1} + \cdots + \tilde{\alpha}_k e_{n_k} \mid \tilde{\alpha}_1, \ldots, \tilde{\alpha}_k \in \mathbb{Q} \text{ or } \mathbb{Q} + i\mathbb{Q}\}$$

form an everywhere dense set and H is separable.

□

Recall that in a Hilbert space, for any orthonormal sequence $\{e_n\}_{n=1}^\infty$ we have the Bessel inequality that for any $x \in X$

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

If the orthonormal sequence is total, that is, if it is an orthonormal basis, then the Bessel inequality becomes an identity.

Theorem 17 (Parseval relation)**.** *An orthonormal sequence $\{e_n\}_{n=1}^\infty$ is total in a Hilbert space H if and only if for all $x \in H$ the* Parseval relation

$$\sum_{i=1}^\infty |\langle x, e_i \rangle|^2 = \|x\|^2 \tag{29}$$

holds.

Proof. If $\{e_n\}$ is not total, then there is an $x \neq 0$ that is orthogonal to all e_n. For this x, the right-hand side of (29) is positive and the left-hand side of (29) is zero since all terms in the sum are zero. Thus the Parseval relation does not hold.

Otherwise, if $\{e_n\}$ for an orthonormal basis, then for any $x \in X$, we consider

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

By Theorem 14(c), $\{s_n\}$ has a limit $s_\infty = \sum_{i=1}^\infty \langle x, e_i \rangle e_i$. We show that $s_\infty = x$. Let $y = x - s_\infty$. Then for any e_n,

$$\langle y, e_n \rangle = \langle x, e_n \rangle - \langle \lim_{k \to \infty} \sum_{i=1}^k \langle x, e_i \rangle e_i, e_n \rangle = \langle x, e_n \rangle - \langle x, e_n \rangle = 0.$$

By the totality of $\{e_n\}$, we have that $y = 0$ and so

$$x = s_\infty = \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i.$$

Now it is ready to check that

$$\|x\|^2 = \langle \lim_{n\to\infty} s_n, \lim_{n\to\infty} s_n \rangle = \lim_{n\to\infty} \langle s_n, s_n \rangle = \lim_{n\to\infty} \sum_{i=1}^{n} |\langle x, e_n \rangle|^2,$$

and then the Parseval relation (29) holds. $\square$

Note that in the proof of Theorem 17, we derived an important relation that if $\{e_n\}_{n=1}^\infty$ is an orthonormal basis in the Hilbert space H, then for any $x \in H$,

$$x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n. \tag{30}$$

Example 13. Recall that in Example 12 we showed that in the inner product space $C^2[-1,1]$, the Legendre polynomials, after normalization, give an orthonormal sequence

$$\{e_n(t) = \sqrt{\frac{2n+1}{2}} P_n(t)\}_{n=0}^\infty.$$

Regarding $C^2[-1,1]$ as a everywhere dense subspace in $L^2[-1,1]$, these $\{e_n\}$ also form an orthonormal sequence in $L^2[-1,1]$. Actually they form an orthonormal basis in $L^2[-1,1]$. To see it, we take any $x \in L^2[-1,1]$ and $\epsilon > 0$. Since $C^2[-1,1]$ is everywhere dense in $L^2[-1,1]$, there is a continuous function $\tilde{x}(t) \in C^2[-1,1]$ such that $\|x - \tilde{x}\| < \epsilon/2$. By the Weierstrass approximation theorem, there is a polynomial $p(t)$ such that $|p(t) - \tilde{x}(t)| < \epsilon/\sqrt{8}$ for all $t \in [-1,1]$. Then $\|p - \tilde{x}\| < \epsilon/2$, and we have $\|p - x\| < \epsilon$. Since all polynomials are in span$\{e_n\}$, we have that span$\{e_n\}$ is everywhere dense in L^2 and then $\{e_n\}$ is an orthonormal basis. It is obvious that $\{e_n\}$ is not a Hamel basis, since a non-polynomial continous function connot be in span$\{e_n\}$.

At last, we discuss the isomorphism between Hilbert spaces. Let H and $\tilde{H}$ be two Hilbert spaces. A linear operator $T : H \to \tilde{H}$ is an *isomorphism*, and then H and $\tilde{H}$ are *isomorphic*, if and T is bijective and for any $x, y \in H$

$$\langle Tx, Ty \rangle = \langle x, y \rangle. \tag{31}$$

It is clear that the isomorphism T is an isometry. Also if $\{e_n\}$ is an orthonormal basis, then $\{Te_n\}$ is also an orthonormal basis.

Suppose H is a finite dimensional Hilbert space over $\mathbb{R}$ with basis $\{e_1, \dots, e_n\}$. then the linear operator $T : \mathbb{R}^n \to H$ defined as

$$T(\alpha_1, \dots, \alpha_n) = \alpha_1 e_1 + \cdots + \alpha_n e_n$$

is an isomorphism between the Euclidean n-space and H. Thus all n-dimensional Hilbert spaces over $\mathbb{R}$ are isomorphic to the Euclidean n-space. Analogously all n-dimensional Hilbert spaces over $\mathbb{C}$ are isomorphic to the unitary n-space.

The following result for infinite dimensional, but separable, Hilbert spaces is more useful:

Theorem 18. *Let H be a separable and infinite dimensional Hilbert space over K, then H is isomorphic to the ℓ^2-space over K.*

Proof. Without loss of generality, we assume $K = \mathbb{C}$.

Suppose $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis of H. Then define a mapping $T : \ell^2 \to H$ as

$$T(\alpha_1, \alpha_2, \dots) = \sum_{n=1}^{\infty} \alpha_n e_n. \tag{32}$$

First, this mapping is well defined. Since $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$, the series on the right-hand side of (32) converges by Theorem 14(a). It is clear that T is linear. To show that T is surjective, note that for any $x \in H$, by (30), $x = \sum_{n=1}^{\infty} \alpha_n e_n$ where $\alpha_n = \langle x, e_n \rangle$. Then by the Parseval relation, $\sum_{n=1}^{\infty} |\alpha_n|^2 = \|x\|^2 < \infty$, and then $x = T(\alpha_1, \alpha_2, \dots)$ is in the range of T. It is straightforward to show that

$$\|T(\alpha_1, \alpha_2, \dots)\| = \|(\alpha_1, \alpha_2, \dots)\| = \sqrt{|\alpha_1|^2 + |\alpha_2|^2 + \cdots},$$

so T is an isometry and then is injective. At last, by the parallelogram equality, we have that the isometry implies that the identity (31) is satisfied. Thus T is an isomorphism. $\square$

As an example, we have that the Hilbert spaces $L^2[-1, 1]$ and ℓ^2 are isomorphic. An isomorphism is $T : \ell^2 \to L^2[-1, 1]$ such that $T(\alpha_1, \alpha_2, \dots) = \alpha_1 e_0 + \alpha_2 e_1 + \cdots$ where e_n are normalized Legendre polynomial of degree n.

2.3 Bounded linear functionals and operators on a Hilbert space

The bounded linear functionals on a Hilbert space have a simple form:

Theorem 19 (Riesz)**.** *Every bounded linear functional f on a Hilbert space H can be represented by a unique vector $z \in H$ such that*

$$f(x) = \langle x, z \rangle, \tag{33}$$

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

Proof. If f is the zero functional, then $z = 0$ satisfies (33), and it is unique. Below we assume that f is not the zero functional.

The null space $N(f)$ is a closed subspace since f is continuous, and it is not the whole H. Thus there is a $y \in H \setminus N(f)$, such that $f(y) \neq 0$. Use the orthogonal projection onto $N(f)$ we decompose

$$y = y' + z', \quad \text{where } y' \in N(f) \text{ and } z' \in N(f)^{\perp}.$$

Then $z' \neq 0$, and

$$f(z') = f(y) - f(y') = f(y) \neq 0.$$

First we show that any $z'' \in N(f)^{\perp}$ is a scalar multiple of z. Consider the vector

$$w = z'' - \frac{f(z'')}{f(z')} z'.$$

On one hand, $w \in N(f)^{\perp}$. On the other hand,

$$f(w) = f(z'') - \frac{f(z'')}{f(z')} f(z') = 0,$$

so $w \in N(f)$. Thus $w \in N(f) \cap N(f)^{\perp} = \{0\}$ and then $z'' = (f(z'')/f(z'))z'$.

Let

$$z = \frac{\overline{f(z')}}{\|z'\|^2} z'.$$

For any $x \in H$, we decompose it by the orthogonal projection onto $N(f)$ such that

$$x = y'' + z'', \quad \text{where } y'' \in N(f) \text{ and } z'' \in N(f)^{\perp}.$$

Suppose $z'' = \alpha z$. Then

$$f(x) = f(y'') + f(z'') = 0 + f(z'') = \alpha f(z') = \alpha \langle z', z \rangle = \langle z'', z \rangle = \langle z'' + y'', z \rangle = \langle x, z \rangle.$$

Thus we show that for any bounded linear functional f there exists z that satisfies (33).

Suppose $z_1, z_2 \in H$ satisfies that for any $x \in H$

$$\langle x, z_1 \rangle = f(x) = \langle x, z_2 \rangle.$$

Then Take $x = z_1 - z_2$, we find the formula above implies

$$\langle z_1 - z_2, z_1 - z_2 \rangle = 0,$$

and then $z_1 = z_2$. Thus we prove the uniqueness of z.

Since for any $x \in H$,

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

where we use the Schwarz inequality in the last step, we have $\|f\| \leq \|z\|$. On the other hand, take $x = z$, we have

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

so $\|f\| \geq \|z\|$. We conclude that $\|f\| = \|z\|$. □

As an intermediate step in the proof of Theorem 19, we have the following simple but useful lemma:

Lemma 11. *If z_1, z_2 are two vectors in a Hilbert space H (actually inner product space is sufficient here) and $\langle x, z_1 \rangle = \langle x, z_2 \rangle$ for all $x \in H$, then $z_1 = z_2$.*

Now consider bounded linear operators. Let $S : H_1 \to H_2$ be a bounded linear operator from Hilbert space H_1 to Hilbert space H_2. Then the mapping $h : H_1 \times H_2 \to K$ defined as

$$h(x, y) = \langle Sx, y \rangle \tag{34}$$

satisfies

$$h(x_1 + x_2, y) = h(x_1, y) + h(x_2, y), \tag{35}$$
$$h(x, y_1 + y_2) = h(x, y_1) + h(x, y_2), \tag{36}$$
$$h(\alpha x, y) = \alpha h(x, y), \tag{37}$$
$$h(x, \beta y) = \bar{\beta} h(x, y). \tag{38}$$

We call a mapping $h : H_1 \times H_2 \to K$ that satisfies conditions (35)–(35) a *sesquilinear form*. A sesquilinear form h is bounded if there exists $c > 0$ such that

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

The *norm* if a sesquilinear form h is defined as

$$\|h\| = \sup_{\|x\|=1, \|y\|=1} |h(x,y)|.$$

If h is defined by S as in (34), then it is bounded and its norm is $\|S\|$. To see it, we first check that $\|h\| \leq \|S\|$ since

$$|h(x,y)| = \langle Sx, y\rangle \leq \|Sx\|\|y\| \leq \|S\|\|x\|\|y\|.$$

Then take $\{x_n\} \subseteq H_1$ such that $\|x_n\| = 1$ and $\|Sx_n\| \to \|S\|$ as $n \to \infty$, and take $y_n = Sx_n/\|Sx_n\|$, we see that $\sup|h(x_n, y_n)| = \|S\|$ and prove that $\|h\| \geq \|S\|$.

The next theorem shows that the converse is also true.

Theorem 20 (Riesz representation)**.** *Let H_1, H_2 be Hilbert spaces and $h : H_1 \times H_2 \to K$ is a bounded sesquilinear form. Then*

(a) h has a representation

$$h(x,y) = \langle Sx, y\rangle \tag{39}$$

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

(b) h has a representation

$$h(x,y) = \langle x, \tilde{S}y\rangle \tag{40}$$

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

Proof. First we prove part (a). Consider $\bar{h} : H_1 \times H_2 \to K$ such that

$$\bar{h}(x,y) = \overline{h(x,y)}.$$

(If $K = \mathbb{R}$, then $\bar{h}$ is the same as h.) Then for any fixed $x \in H_1$, $\bar{h}(x,y)$ is a linear functional in $y \in H_2$ since

$$\bar{h}(x, \alpha y_1 + \beta y_2) = \overline{\bar{\alpha} h(x, y_1) + \bar{\beta} h(x, y_2)} = \alpha \bar{h}(x, y_1) + \beta \bar{h}(x, y_2).$$

Also $\bar{h}(x,y)$ is a bounded linear functional in y. Then by Theorem 19, for any $x \in H_1$ there is a unique $z_x \in H_2$ such that

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

Below we check that the mapping $S : H_1 \to H_2$ such that $Sx = z_x$ is a linear mapping: For any $x_1, x_2 \in H_1$ and $\alpha, \beta \in K$,

$$\begin{aligned}\bar{h}(\alpha x_1 + \beta x_2, y) = \overline{h(\alpha x_1 + \beta x_2, y)} = \overline{\alpha h(x_1, y) + \beta h(x_2, y)} = \bar{\alpha}\bar{h}(x_1, y) + \bar{\beta}\bar{h}(x_2, y) \\ = \bar{\alpha}\langle y, z_{x_1}\rangle + \bar{\beta}\langle y, z_{x_2}\rangle = \langle y, \alpha z_{x_1} + \beta z_{x_2}\rangle,\end{aligned}$$

so

$$S(\alpha x_1 + \beta x_2) = z_{\alpha x_1 + \beta x_2} = \alpha z_{x_1} + \beta z_{x_2} = \alpha S(x_1) + \beta S(x_2).$$

Next we show that S is bounded and $\|S\| = \|h\|$. To see it, we compute that for any $x \in H_1$

$$\|Sx\| = \|z_x\| = \|\bar{h}(x, \cdot)\| = \sup_{\|y\|=1} |\bar{h}(x, y)| = \sup_{\|y\|=1} |h(x, y)|.$$

So

$$\|S\| = \sup_{\|x\|=1} \|Sx\| = \sup_{\|x\|=1} \sup_{\|y\|=1} |h(x, y)| = \|h\|.$$

At last, we conclude that the bounded linear operator S defined above satisfy (39). S is unique, since the bounded linear operator have to satisfy $Sx = z_x$.

Next we prove part (b) in a parallel way. For any fixed y, $h(x, y)$ is a bounded linear functional in $x \in H_1$. Then by Theorem 19, for any $y \in H_2$, there is a unique $\tilde{z}_y \in H_1$ such that

$$h(x, y) = \langle x, \tilde{z}_y \rangle.$$

We define the mapping $\tilde{S} : H_2 \to H_1$ as $Sy = \tilde{z}_y$. The mapping S is a bounded linear operator with $\|\tilde{S}\| = \|h\|$, but we omit the proof since it is analogous to the proof of the properties of S. Also $\tilde{S}$ is unique, proved in the same way as the uniqueness of S. □

The two linear operators S and $\tilde{S}$ are related. $\tilde{S}$ is the Hilbert-adjoint of S, and vice versa.

Definition 11. Let $T : H_1 \to H_2$ be a bounded linear operator between Hilbert spaces H_1 and H_2. Then the *Hilbert-adjoint operator* T^* of T is the operator from H_2 to H_1 such that

$$\langle Tx, y \rangle = \langle x, T^* y \rangle \quad \text{for all } x \in H_1,\ y \in H_2.$$

Theorem 21. *The Hilbert-adjoint operator T^* of T exists, is unique, and is a bounded linear operator with norm $\|T^*\| = \|T\|$.*

Proof. The mapping $h : H_1 \times H_2 \to K$ defined by $h(x, y) = \langle Tx, y \rangle$ is a bounded sesquilinear form, and $\|h\| = \|T\|$. By part ((b)) of Theorem 20, there is a unique bounded linear operator $\tilde{S} : H_2 \to H_1$ such that $h(x, y) = \langle x, \tilde{S} y \rangle$ and this is the Hilbert-adjoint operator T^* of T by definition. Also we have $\|T^*\| = \|h\| = \|T\|$ by Theorem 20. □

Before presenting more properties of Hilbert adjoint operators, we state a technical lemma.

Lemma 12. *Let X and Y be inner product spaces and $Q : X \to Y$ be a bounded linear operator. Then*

(a) $Q = 0$ if and only if $\langle Qx, y \rangle = 0$ for all $x \in X$ and $y \in Y$.

(b) If $Q : X \to X$ where X is complex, (that is, the scalar field is $\mathbb{C}$), and $\langle Qx, x \rangle = 0$ for all $x \in X$, then $Q = 0$.

The proof of part (a) is an easy exercise, and the proof of part (b), though not easy, is an tutorial problem.

Theorem 22. *Let H_1 and H_2 be Hilbert space, $S : H_1 \to H_2$ and $T : H_1 \to H_2$ be two bounded linear operators and α is a scalar. Then we have*

$$\langle T^*y, x\rangle = \langle y, Tx\rangle \quad \textit{for all } x \in H_1 \textit{ and } y \in H_2, \tag{41}$$
$$(S+T)^* = S^* + T^*, \tag{42}$$
$$(\alpha T)^* = \bar{\alpha} T^*, \tag{43}$$
$$(T^*)^* = T, \tag{44}$$
$$\|T^*T\| = \|TT^*\| = \|T\|^2, \tag{45}$$
$$T^*T = 0 \Leftrightarrow T = 0, \tag{46}$$
$$(ST)^* = T^*S^*, \quad \textit{if } H_1 = H_2. \tag{47}$$

Proof. Properties (41), (42) and (43) are straightforward to check. For example, (41) is from

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

and (42) and (43) are left as exercises.

(41) is equivalent to (44).

To prove (45), we use

$$\|T^*T\| = \sup_{\substack{\|x_1\|=1,\\ \|x_2\|=1}} \langle T^*Tx_1, x_2\rangle = \sup_{\substack{\|x_1\|=1,\\ \|x_2\|=1}} \langle Tx_1, Tx_2\rangle = \sup_{\|x\|=1} \|Tx\|^2 = \|T\|^2,$$

and

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

(45) implies (46).

(47) is because

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

□

Definition 12. A bounded linear operator $T : H \to H$ on a Hilbert space H is *self-adjoint* (or *Hermitian*) if $T^* = T$, *unitary* if $T^* = T^{-1}$ and *normal* if $T^*T = TT^*$.

It is clear that if an operator is self-adjoint or unitary, then the operator is normal.

Example 14. In the unitary n space, we express the vectors by collumn vectors, and denote the vectors $e_k = (\underbrace{0, \dots, 0}_{k-1}, 1, 0, \dots, 0)^T$ $(k = 1, \dots, n)$. An bounded linear operator $T : \mathbb{C}^n \to \mathbb{C}^n$ is defined by its values on $e_1, \dots, e_n$. Suppose

$$\begin{aligned} Te_1 &= a_{11}e_1 + \dots + a_{n1}e_n = (a_{11}, \dots, a_{n1})^T, \\ Te_2 &= a_{12}e_1 + \dots + a_{n2}e_n = (a_{12}, \dots, a_{n2})^T, \\ &\dots, \\ Te_n &= a_{1n}e_1 + \ldots + a_{nn}e_n = (a_{1n}, \dots, a_{nn}). \end{aligned}$$

Then

$$T\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix} = \alpha_1 Te_1 + \dots + \alpha_n Te_n = \alpha_1 \begin{pmatrix}a_{11}\\ \vdots\\ \alpha_{n1}\end{pmatrix} + \dots + \begin{pmatrix}a_{1n}\\ \vdots\\ a_{nn}\end{pmatrix} = \begin{pmatrix}a_{11} & \dots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \dots & \dots a_{nn}\end{pmatrix}\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}.$$

So the bounded linear operators have 1-to-1 correspondence with $n \times n$ complex matrices. Below we denote the matrix representing the operator T as

$$A = \begin{pmatrix} \langle Te_1, e_1 \rangle & \langle Te_2, e_1 \rangle & \dots & \langle Te_n, t_1 \rangle \\ \vdots & & \vdots & \\ \langle Te_1, e_n \rangle & \langle Te_2, e_n \rangle & \dots & \langle Te_n, t_n \rangle \end{pmatrix}. \tag{48}$$

Let $x = (\xi_1, \dots, \xi_n)^T$ and $y = (\eta_1, \dots, \eta_n)^T$ in $\mathbb{C}^n$. Then

$$\langle x, y \rangle = x^T \bar{y},$$

and

$$\langle Tx, y \rangle = (Ax)^T \bar{y} = x^T A^T \bar{y}.$$

Suppose T^*, the Hilbert adjoint operator of T, is represented by the $n \times n$ complex matrix B. Then

$$\langle x, T^* y \rangle = x^T \overline{By} = x^T \bar{B} \bar{y}.$$

By the definition, $\langle Tx, y \rangle = \langle x, T^* y \rangle$, so we obtain that $A^T = \bar{B}$, or equivalently, $B = \bar{A}^T$. Hence we have that T is self-adjoint if and only if A is Hermitian ($\bar{A}^T = A$); T is unitary if and only if A is unitary ($\bar{A}^T = A^{-1}$) and T is normal if and only if A is normal ($A\bar{A}^T = \bar{A}^T A$).

Similarly, if we consider the (real) Euclidean n-space, the bounded linear operators have 1-to-1 correspondence with $n \times n$ real matrices, and the operator T is represented by also (48). We have that T^* is represented by A^T, and hence that T is self-adjoint if and only if A is symmetric ($A^T = A$); T is unitary if and only if A is orthogonal ($A^T = A^{-1}$).

Note that for any finite dimensional Hilbert space H with basis $\{e_1, \dots, d_n\}$, the linear bounded operator has 1-to-1 correspondence with $n \times n$ matrices given by (48), and the above mentioned-results hold.

Now we go back to infinite dimensional Hilbert spaces, and state an important and simple criteron for self-ajjointness in complex Hilbert spaces (that does not apply to real Hilbert spaces).

Theorem 23. *Let H be a complex Hilbert space and $T : H \to H$ be a bounded linear operator. T is self-adjoint if and only if $\langle Tx, x \rangle$ is real for all $x \in H$.*

Proof. If T is self-adjoint, then

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

On the other hand, as a general property of inner product,

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

Thus $\langle x, Tx \rangle$ is equal to its complex conjugate and is real, so is $\langle Tx, x \rangle$.

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

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

Then for any x, $\langle (T - T^*)x, x \rangle = 0$. By the result of Lemma 12(b), we have that $T - T^* = 0$ and T is self-adjoint. $\square$

Now we study some properties of self-adjoint operators and unitary operators.

Theorem 24. *The product of two bounded self joint operators $S : H \to H$ and $T : H \to H$ on a Hilbert space H is self-adjoint if and only if S and T commute, that is, $ST = TS$.*

Proof. Since S and T are self-adjoint,

$$(ST)^* = T^* S^* = TS.$$

So ST is self-adjoint if and only if $ST = TS$. □

Theorem 25. *Let $\{T_n : H \to H\}$ be a sequence of bounded self-adjoint linear operators and $\{T_n\}$ converge to $T : H \to H$ in operator norm, that is, $\|T_n - T\| \to 0$ as $n \to \infty$. Then T is self-adjoint.*

Proof. For any $x, y \in H$,

$$\langle x, Ty\rangle = \lim_{n\to\infty} \langle x, T_n y\rangle = \lim_{n\to\infty} \langle T_n x, y\rangle = \langle Tx, y\rangle.$$

So T is self-adjoint. □

Theorem 26. *Let H be a Hilbert space and U, V be two unitary operators on H. Then*

(a) U is isometric, that is, $\|Ux\| = \|x\|$.

(b) $\|U\| = 1$, provided that $H \neq \{0\}$.

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

(d) UV is unitary.

Proof. (a) For any $x \in H$, $\|Ux\| = \langle Ux, Ux\rangle = \langle x, U^*Ux\rangle = \langle x, U^{-1}Ux\rangle = \langle x, x\rangle = \|x\|$.

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

(c) For any $x, y \in H$, $\langle UVx, y\rangle = \langle Vx, U^{-1}y\rangle = \langle x, V^{-1}U^{-1}y\rangle = \langle x, (UV)^{-1}y\rangle$. So $(UV)^* = (UV)^{-1}$ and UV is unitary.

□

Theorem 27. *On a complex Hilbert space H, a bounded linear operator $T : H \to H$ is unitary if and only if T is isometric and surjective.*

Proof. Since T is isometric, so it is injective. Thus T^{-1} exists. Since T is isometric, T^{-1} is also isometric. Note that in a Hilbert space, the norm determines the inner product by the polarization identity, so the isometry T^{-1} satisfies $\langle T^{-1}x, T^{-1}y\rangle = \langle x, y\rangle$ for any $x, y \in H$. We have that for all $x \in H$

$$\langle x, T^*x\rangle = \langle Tx, x\rangle = \langle T^{-1}Tx, T^{-1}x\rangle = \langle x, T^{-1}x\rangle.$$

Thus $\langle x, (T^* - T^{-1})x\rangle = 0$ for all x, and by Lemma 12(b) we have that $T^* - T^{-1} = 0$ and prove that T is unitary. □

3 Fundamental theorems for normed and Banach spaces

3.1 Hahn-Banach theorem

In Theorem 8 in Section 1.4, we proved a special form of Hahn-Banach theorem such that if X is a vector space and Z is a subspace of X such that $X = \{z + \alpha y \mid z \in Z\}$ where y is a fixed vector in X, then a functional on Z can be extended to a function on X, such that the original functional and its extension both satisfy a certain condition. In this section, we prove the Hahn-Banach theorem in its general form, such that Z can be any subspace of X.

We need a rather abstract technical lemma:

Lemma 13 (Zorn)**.** *Let $M \neq \emptyset$ be a partially ordered set. Suppose that every chain $C \subseteq M$ has an upper bound, then M has at least one maximal element.*

This is an axiom that cannot be proved. Below we explain the terms.

Definition 13. A *partially ordered set* is a set that some pair of elements $a, b \in M$ has the relation $\leq$ (while some pairs may not has the relation,) such that

- $a \leq a$ for all $a \in M$,
- if $a \leq b$ and $b \leq a$, then $a = b$,
- If $a \leq b$ and $b \leq c$, then $a \leq c$.

A *chain* is a partially ordered set such that any two elements $a, b \in M$ saitisfies either $a \leq b$ or $b \leq a$. An *upper bound* of a subset W of a partially ordered set M is

$$u \in M \quad \text{such that } x \leq u \text{ for all } x \in W.$$

A *maximal element* of a partially ordered set M is

$$m \in M \quad \text{such that } x \leq m \text{ for all } x \in M.$$

Now we state the Hahn-Banach theorem.

Theorem 28 (Hahn-Banach)**.** *1. Let X be a vector space over $\mathbb{R}$, and p be a sublinear functional on X. Suppose Z is a subspace of X. Furthermore, let f be a linear functional which is defined on Z and satisfies*

$$f(x) \leq p(x) \quad \textit{for all } x \in Z.$$

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

$$\tilde{f}(x) \leq p(x) \quad \textit{for all } x \in X, \tag{49}$$

that is, $\tilde{f}$ is a linear functional on X, satisfies (50) on X and $\tilde{f}(x) = f(x)$ for every $x \in Z$.

2. *Let X be a vector space over $\mathbb{C}$, and p be a mapping from X to $\mathbb{R}$ that satisfies subadditivity and condition* (17)*. Suppose Z is a subspace of X. Furthermore, let f be a linear functional which is defined on Z and satisfies*

$$|f(x)| \leq p(x) \quad \text{for all } x \in Z.$$

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

$$|\tilde{f}(x)| \leq p(x) \quad \text{for all } x \in X. \tag{50}$$

We call $\tilde{f}$ an *extension* of f.

Proof. We only prove part 1 based on part 1 of Theorem 8. The proof of part 2 based on part 2 of Theorem 8 is analogous.

Consider all possible (X_α, f_α) where X_α is a subset of X such that $Z \subseteq X_\alpha$, and f_α is a functional on X_α satisfying $f_\alpha(z) = f(z)$ for all $z \in Z$. These (X_α, f_α) form a partially ordered set, such that $(X_\alpha, f_\alpha) \leq (X_\beta, f_\beta)$ if $X_\alpha \subseteq X_\beta$ and $f_\beta(x) = f_\alpha(x)$ for all $x \in X_\alpha$.

If $\{(X_\alpha, f_\alpha)\}_{\alpha \in I}$ is a set that is a chain with respect to the relation $\leq$, then they have an upper bound (X_I, f_I) defined as

$$X_I = \bigcup_{\alpha \in I} X_\alpha, \quad f_I(x) = f_\alpha(x) \quad \text{for } x \in X_\alpha \subseteq X_I.$$

Then by Zorn's lemma, the set of all possible (X_α, f_α) has a maximal element, say (W, f'). We claim that $W = X$, and this f' is a desired $\tilde{f}$. If $W \neq X$, there is $y \in X \setminus W$, and the subspace $X = \{w + \alpha y \mid w \in W, \ \alpha \in \mathbb{R}\}$ is a subspace of x that is bigger than W. But by part 1 of Theorem 8 shows that there is a functional f'' on Y such that $(W, f') \leq (Y, f'')$, contradictory to that (W, f') is a maximal element. □

Although the general form of the Hahn-Banach theorem is on any vector space equipped with the function p, we note that on a normed space, either real or complex, the conditions on p are satisfied by $r\|\cdot\|$ for all $r \in (0, +\infty)$. So as a special case of Theorem 28, we have:

Theorem 29 (Hahn-Banach on 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 $\|\tilde{f}\| = \|f\|$.*

Proof. Take $p(x) = \|f\|\|x\|$ and apply Theorem 28. □

A useful special case of Theorem 29 is that Z is 1-dimensional, spanned by $\{x_0\}$. Then over either the real or complex field, there is a unique $f \in Z'$ such that $f(x_0) = \|x_0\|$, and we have $\|f\| = 1$. Thus we have

Theorem 30 (Bounded linear functionals)**.** *Let X be a normed space and let $x_0 \neq 0$ be an element of X. Then there exists a bounded linear functional $\tilde{f}$ such that $\|\tilde{f}\| = 1$ and $\tilde{f}(x_0) = \|x_0\|$.*

An important application of the Hahn-Banach theorem is on the norm of an adjoint operator.

Definition 14. Let X and Y be two normed spaces and $T : X \to Y$ be a bounded linear operator. The *adjoint operator* $T^\times$ of T is a mapping $T^\times : Y' \to X'$ such that for any $f \in Y'$,

$$(T^\times f)(x) = f(Tx).$$

Theorem 31. *The adjoint operator $T^\times$ of a bounded linear operator $T : X \to Y$ is also a bounded operator, and $\|T^\times\| = \|T\|$.*

Proof. First we show that $T^\times$ is linear. For any $x \in X$ and $f_1, f_2 \in Y'$,

$$\begin{aligned} T^\times(\alpha_1 f_1 + \alpha_2 f_2)(x) &= (\alpha_1 f_1 + \alpha_2 f_2)(Tx) \\ &= \alpha_1 f_1(Tx) + \alpha_2 f_2(Tx) \\ &= \alpha_1 (T^\times f_1)(x) + \alpha_2 (T^\times f_2)(x). \end{aligned}$$

So $T^\times(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 T^\times f_1 + \alpha_2 T^\times f_2$.

Next we show that $T^\times$ is bounded. For any $x \in X$ with $\|x\| = 1$,

$$|(T^\times f)(x)| = |f(Tx)| \le \|f\|\|T\|\|x\| = \|f\|\|T\|.$$

So $\|T^\times f\| \le \|T\|\|f\|$ and then $T^\times$ is bounded and $\|T^\times\| \le \|T\|$.

Finally we prove that $\|T^\times\| \ge \|T\|$ and finish the proof of the theorem. Without loss of generality, we assume that $T \ne 0$. It suffices to show that for any $\epsilon > 0$, there is an $f \in Y'$ such that $\|f\| = 1$ and

$$\|T^\times f\| \ge (1 - \epsilon)\|T\|. \tag{51}$$

Furthermore, a sufficient condition for (51) is that their is $x \in X$ such that $\|x\| = 1$ and

$$|(T^\times f)(x)\| = |f(Tx)| \ge (1 - \epsilon)\|T\|. \tag{52}$$

By the definition of operator norm, there is an $x_\epsilon \in X$ such that $\|x_\epsilon\| = 1$ and $\|Tx_\epsilon\| \ge (1 - \epsilon)\|T\|$. Then by Theorem 30, their is a linear functional $f_\epsilon \in Y'$ such that

$$\|f_\epsilon\| = 1 \quad \text{and} \quad |f_\epsilon(Tx_\epsilon)| = \|Tx_\epsilon\|.$$

Taking $f = f_\epsilon$ and $x = x_\epsilon$ in (52), we see that (52) is satisfied, and so $f = f_\epsilon$ also make (51) satisfied. Thus we prove $\|T^\infty\| \ge \|T\|$. □

Example 15 (Example of adjoint operator). Let H_1 and H_2 be Hilbert spaces and $T : H_1 \to H_2$ be a bounded linear operator. There are mappings $A_1 : H_1' \to H_1$ and $A_2 : H_2' \to H_2$ such that any $f \in H_*'$ can be represented as $f(x) = \langle x, A_*(f)\rangle$ where $* = 1, 2$. Then for any $x \in H_1$ and $f \in H_2'$, the adjoint operator $T^\times$ satisfies

$$(T^\times f)(x) = f(Tx) = \langle Tx, A_2 f\rangle = \langle x, T^* A_2 f\rangle.$$

So the adjoint operator is related to the Hilbert joint operator by $A_1 T^\times = T^* A_2$, or $T^\times = A_1^{-1} T^* A_2$.

3.2 Uniform boundedness theorem

We consider properties of Banach space in this subsection and later.

First we recall the Baire's theorem

Theorem 32 (Baire). *Any complete metric space cannot be expressed as the union of countably many nowhere dense subsets, where a subset is nowhere dense if its closure does not contain any open set.*

The proof was given in Mathematical Analysis III. A proof can be found in the textbook. By Baire's theorem, we have the following important theorem for Banach spaces.

Theorem 33 (Uniform boundedness, or Banach-Steinhaus). *Let $\{T_n : X \to Y\}_{n=1}^{\infty}$ be a sequence of bounded linear operators from a Banach space X to a normed space Y such that $\{\|T_n x\|\}_{n=1}^{\infty}$ is bounded for every $x \in X$, that is, for any $x \in X$, $\|T_n x\| \leq c_x$ for all T_n where c_x is a real number depending on x. Then the sequence of the norms $\|T_n\|$ is bouded, that is, there is a real number c such that $\|T_n\| < c$.*

Proof. Define for all $k \in \mathbb{Z}_+$

$$A_k = \{x \in X \mid \|T_n x\| \leq k \text{ for all } n\} = \bigcap_{n=1}^{\infty} \{x \in X \mid \|T_n\| \leq k\}.$$

Then A_k is closed for all k, and $\bigcup_{k=1}^{\infty} X$. By Baire's theorem, there is at least one $\overline{A_k} = A_k$ contains an open set. Without loss of generality, we assume that A_K contains $B_{x_0}(r) = \{x \in X \mid \|x - x_0\| < r\}$. Then for any x such that $\|x\| < r$ and for any N,

$$\|T_n x\| \leq \|T_n x_0\| + \|T_n(x - x_0)\| \leq 2K.$$

So $\|T_n\| \leq 2K/r$ for all n. $\square$

With the Hahn-Banach theorem and uniform boundedness theorem at our disposal, we can discuss the concept of weak convergence of a sequence, that is different from the usual strong convergence if the normed space is infinite dimensional.

Definition 15. A sequence $\{x_n\}_{n=1}^{\infty}$ in a normed space X is said to be *weakly convergent* if there is an $x \in X$ (the *weak limit*) such that for every $f \in X'$,

$$\lim_{n \to \infty} f(x_n) = f(x).$$

We write $x_n \rightharpoonup x$.

In contrast with the weak convergence, we say a seqeunce $\{x_n\}_{n=1}^{\infty}$ is *strongly convergent*, to the limit x, written as $x_n \to x$, if $\|x_n - x\| \to 0$.

As suggested by their name, strong convergence is a stronger condition than weak convergence.

Lemma 14. *Let $\{x_n\}$ be a sequence in a normed space X. Then*

(a) If $x_n \to x$, then $x_n \rightharpoonup x$.

(b) *If X is finite dimensional and $x_n \rightharpoonup x$, then $x_n \to x$.*

Proof. Part (a) is an exercise.

To prove part (b), we note that on any normed space $(X, \|\cdot\|)$ of dimension k, we can pick a basis $e_1, \ldots, e_k$ and defined another norm $\|\cdot\|$ such that $\|\alpha_1 e_1 + \cdots + \alpha_k e_k = (|\alpha_1|^2 + \cdots + |\alpha_k|^2)^{1/2}$. We have that $\|\cdot\|$ $\|\cdot\|_2$ are equivalent norms, and a linear functional f on X is boudned with respect to $\|\cdot\|$ if and only if it is bounded with respect to $\|\cdot\|_2$. Also note that $(X, \|\cdot\|_2)$ is isomorphic to the Euclidean n-space or the unitary n-space.

If a sequence $\{x_n = \alpha_1^{(n)} e_1 + \cdots + \alpha_k^{(n)} e_k\}$ is a weak convergence sequence with limit $x = (\alpha_1, \ldots, \alpha_k)$ in the Euclidean n-space or unitary n-space, then

$$f_j(x_n) \to f_j(x), \quad \text{where} \quad f_j(z) = \langle z, e_j \rangle, \quad j = 1, \ldots, k.$$

So $x_n \to x$. □

If X is inifinite dimensional, then converse of Lemma 14(b) is generally not true.

Example 16. In ℓ^2 space, the sequence $\{x_n = (\underbrace{0, \ldots, 0}_{n}, 1, 0, \ldots)\}$ converge weakly to $x = (0, 0, \ldots)$. To see it, represent a linear functional f by the inner product to a vector z_f, and we have for any $z_f = (\zeta_1, \zeta_2, \ldots) \in \ell^2$ we have

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} \langle x_n, z_f \rangle = \lim_{n\to\infty} \bar{\zeta}_n = 0.$$

But $\{x_n\}$ does not converge to x, since for any n, $\|x_n - x\| = 1$ for all x_n, so the strong limit of $\{x_n\}$ is not x.

Lemma 15. *Let $\{x_n\}$ be a weakly convergent sequence with weak limit x in a normed space X. Then*

(a) *The weak limit x is unique.*

(b) *Every subsequence of x_n converges to x weakly.*

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

Proof. (a) Let $x \neq y$ be two vectors in X. By Theorem 30, there is a linear functional f such that $\|f\| = 1$ and $f(x) - f(y) = f(x - y) = \|x - y\| \neq 0$. Then the limit identities $f(x_n) \to f(x)$ and $f(x_n) \to f(y)$ cannot both hold, and so x and y cannot be both weak limits of $\{x_n\}$.

(b) Obvious be definition.

(c) Recall that X' is a Banach space. There is a canonical mapping $C : X \to X''$ from X to the second dual (the dual space of X') such that $(Cx)(f) = f(x)$. Below in Lemma 16 we show that $\|Cx\| = \|x\|$. (If X is finite dimensional, then it is already shown in Section 1.4.)

Since for any $f \in X'$, $\{(Cx_n)(f)\} = \{f(x_n)\}$ converge to $f(x)$, so they are bounded. By the uniform bounded theorem, we have that $\{\|Cx_n\|\}$ are bounded, and so are $\{\|x_n\|\}$.

□

Lemma 16. *For any normed space X, the canonical mapping $C : X \to X''$ satisfies $\|Cx\| = \|x\|$.*

Proof. It is already shown in (13) that $\|Cx\| \le \|x\|$. Thus we need only to show that $\|Cx\| \ge \|x\|$. By Theorem 30, there is $f \in X'$ such that $\|f\| = 1$ and $f(x) = \|x\|$. Then $|(Cx)(f)| = \|x\|$, and we find that $\|Cx\| = \sup_{\|f\|=1}|(Cx)(f)| \ge \|x\|$. □

Recall that a set M is total in a normed space X if $\overline{\operatorname{span} M} = X$. A characterization of weak convergence by a total set is:

Lemma 17. *In a normed space X, we have $x_n \rightharpoonup x$ if and only if the sequence $\{\|x_n\|\}$ is bounded and for any $f \in M$, a total set of X, $f(x_n) \to f(x)$.*

Proof. The "only if" part is already done. To prove the "if" part, we need to show that for any $g \in X'$, $g(x_n) \to g(x)$, or equivalently, for any $\epsilon > 0$, there is N such that $|g(x_n) - g(x)| < \epsilon$ for all $n \ge N$.

Denote $K = \max(\|x\|, \sup\|x_n\|)$. For any $\epsilon > 0$ there is $f \in \operatorname{span} M$ such that $\|f - g\| < \epsilon/(3K)$. Then

$$|g(x_n) - g(x)| \le |g(x_n) - f(x_n)| + |f(x_n) - f(x)| + |f(x) - g(x)| \le \frac{2\epsilon}{3} + |f(x_n) - f(x)|.$$

We need only to show that there is N such that for all $n \ge N$, $|f(x_n) - f(x)| < \epsilon/3$. This is staightforward from the assumption. □

If we consider the convergence in the space $B(X, Y)$, the bounded linear operators between two normed spaces X and Y, we have more refined definitions of convergence.

Definition 16. Let X and Y be normed spaces. A sequence of operators $\{T_n\}_{n=1}^{\infty} \subseteq B(X, Y)$ is said to be

(a) *uniformly operator convergent* to $T \in B(X, Y)$ if $\|T_n - T\| \to 0$.

(b) *strongly operator convergent* to $T \in B(X, Y)$ if for any $x \in X$, $\|T_n x - Tx\| \to 0$.

(c) *weakly operator convergent* to $T \in B(X, Y)$ if for any $x \in X$ and $f \in Y'$, $|f(T_n x) - f(Tx)| \to 0$.

T is called the *uniform*, *strong* and *weak operator limit* of $\{T_n\}$ respectively.

Note that from the definition, the uniform operator limit T has to be a bounded linear operator. The strong and weak operator limits, however, may not be bounded by definition, although it is not hard to show that they are linear operators. It is easy to check that if T is a uniform operator limit of $\{T_n\}$, then it is also a strong operator limit of $\{T_n\}$, and if T is a strong operator limit of $\{T_n\}$, it is also a weak operator limit of $\{T_n\}$. The converse are generally not true, as shown in the following example.

Example 17. In the ℓ^2-space, define the sequence of operators $\{T_n : \ell^2 \to \ell^2\}$ by

$$T_n(\xi_1, \xi_2, \dots) = (\underbrace{0, \dots, 0}_{n}, \xi_{n+1}, \xi_{n+2}, \dots).$$

They are strongly operator convergent to the zero operator, since for any $x = (\xi_1, \xi_2, \dots) \in \ell^2$,

$$\lim_{n\to\infty} \|T_n x - 0\| = \lim_{n\to\infty} \left(\sum_{i=n+1}^{\infty} |\xi_i|^2 \right)^{\frac{1}{2}} = 0.$$

But since $\|T_n\| = 1$ for all n, they are not uniformly opeator convergent to 0. (Actually they are not uniformly operator convergent to any thing.)

Consider another sequence $\{S_n : \ell^2 \to \ell^2\}$ defined by

$$S_n(\xi_1, \xi_2, \dots) = (\underbrace{0, \dots, 0}_{n}, \xi_1, \xi_2, \dots).$$

They are weakly operator convergent to the zero operator, since for any $x = (\xi_1, \xi_2, \dots) \in \ell^2$ and any bounded linear functional $f : \ell^2 \to K$, such that $f(x) = \langle x, z_f \rangle$ where $z_f = (\alpha_1, \alpha_2, \dots)$,

$$\lim_{n\to\infty} |f(S_n x) - 0| = \lim_{n\to\infty} |\alpha_{n+1}\xi_1 + \alpha_{n+2}\xi_2 + \dots| \le \lim_{n\to\infty} \left(\sum_{i=n+1}^{\infty} |\alpha_i|^2 \right)^{\frac{1}{2}} \|x\| = 0.$$

But since for all $x \neq 0$, $\|S_n x\| = \|x\| \neq 0$, $\{S_n\}$ are not strongly operator convergent to 0. (Actually they are not strongly convergent to any thing.)

The following lemma shows that the limit of a strongly operator convergent sequence is a bounded linear operator, if the domain space is a Banach space.

Lemma 18. *Let X be a Banach space and Y be a normed space. If $\{T_n\} \subseteq B(X, Y)$ is strongly operator convergent with limit T, then $T \in B(X, Y)$.*

Proof. The linearity of T follows from the linearity of T_n. Then we need to show that there exists $c > 0$ such that $\|Tx\| \le c\|x\|$ for all $x \in X$. Since X is a Banach space and at any $x \in X$, $\{T_n x\}$ is bounded since $\|T_n x - Tx\| \to 0$. Using the uniform boundedness theorem, we prove the existence of this c. □

The following theorem shows that the strong operator convergence is similar to the weak convergence.

Theorem 34. *Let X and Y be Banach spaces and M be a total set of X. A sequence $\{T_n\} \subseteq B(X, Y)$ is strongly operator convergent if and only if the sequence $\{\|T_n\|\}$ is bounded and for every $x \in M$, the sequence $\{T_n x\}$ is a Cauchy sequence.*

Proof. First we prove the "only if" part and assume that $\{T_n\} \subseteq B(X, Y)$ is strongly operator convergent. Since in the proof of Lemma 18, we showed that the under the condition that X is a Banach space, there exists $c > 0$ such that $\|Tx\| \le c\|x\|$ for all $x \in X$, the "only if" part is straightforward.

Next we prove the "if" part. The proof is parallel to that of Lemma 17. It suffices to show that for any $z \in X$, $\{T_n z\}$ is a Cauchy sequence. Hence this sequence has a limit since Y is a Banach space. Define the operator $T : X \to Y$ as $Tz = \lim_{n\to\infty} T_n z$, and it is clear that T is a bounded linear operator and it is the strong operator limit of $\{T_n\}$.

To show that $\{T_n z\}$ is Cauchy, we need to find for any $\epsilon > 0$ an N such that for all $m, n > N$, $\|T_m z - T_n z\| < \epsilon$. Suppose $K = \sup\|T_n\|$. Choose $x \in \operatorname{span} M$ such that $\|x - z\| < \epsilon/(3K)$. Then

$$\|T_m z - T_n z\| \leq \|T_m z - T_m x\| + \|T_m x - T_n x\| + \|T_n x - T_n z\| \leq \frac{2\epsilon}{3} + \|T_m x - T_n x\|.$$

By the assumption, there is an N such that for all $m, n > N$, $\|T_m x - T_n x\| < \epsilon/3$. This N satisfies our requirement. □

Since bounded linear functionals are special bounded linear operators, we can define the uniform, strong and weak operator convergence on them. But in this special case, strong and weak operator convergence are equivalent. To see it, we note that the only bounded linear operators on K, the 1-dimensional normed space, are the scalar multiplications $x \to \alpha x$. So the weak operator convergence of $\{f_n\} \subseteq X'$ to an linear operator $f : X \to K$ means that for any $\alpha \in K$ and any $x \in X$

$$\lim_{n\to\infty} |\alpha f_n(x) - \alpha f(x)| = 0.$$

This implies that $|f_n(x) - f(x)| \to 0$ and then $\{f_n\}$ are strongly operator convergent to f. Thus for bounded linear functionals, we change the terminologies "uniform operator convergence" into "strong convergence" and "strong/weak operator convergence" into "weak* convergence".

Definition 17. Let $\{f_n\} \subseteq X'$ be a sequence of bounded linear functionals on a normed space X. Then the *strong convergence* of $\{f_n\}$ means that there is an $f \in X'$ such that $\|f_n - f\| \to 0$, written as

$$f_n \to f;$$

the weak* convergence of $\{f_n\}$ means that there is an $f \in X'$ such that $f_n(x) \to f(x)$ for all $x \in X$, written as

$$f_n \overset{w*}{\to} f.$$

Since the weak* convergence is a special case of the strong operator convergence, Theorem 34 has a corollary:

Corollary 3. *Let X be a Banach space and M be a total set of X. A sequence $\{f_n\} \subseteq X'$ is weak* convergent with the limit being a bounded linear functional, if and only if the seqeunce $\{\|f_n\|\}$ is bounded and the sequence $\{f_n(x)\}$ is Cauchy for all $x \in M$.*

In the end of this subsection, we give an application of weak* convergence in numerical integration. To evaluate the integral

$$f(x) = \int_a^b x(t)dt$$

numerically for continuous function $x(t)$ on $[a, b]$, softwares usually first decide a number n, and then choose $n+1$ "nodes" $a \leq t_1^{(n)} < t_2^{(n)} < \cdots < t_n^{(n)} \leq b$ and $n+1$ real "coefficients" $\alpha_0^{(n)}, \alpha_1^{(n)}, \ldots, \alpha_n^{(n)}$ and compute

$$f_n(x) = \sum_{k=0}^{n} \alpha_k^{(n)} x(t_k^{(n)}), \tag{53}$$

in the hope that if n is large enough, then $f(x)$ is close to $f(x)$. The question is: for a certain way of choosing $t_k^{(n)}$ and $\alpha_k^{(n)}$ for all n, to tell if

$$f_n(x) \to f(x) \quad \text{for all } x(t) \text{ continuous on } [a,b]. \tag{54}$$

It is difficult to check (54), but the following condition is easy to check:

$$f_n(x) \to f(x) \quad \text{for all polynomials } x(t). \tag{55}$$

We have the following theorem that shows that the conditions (54) and (55) are closely related.

Theorem 35. *Let the numerical integrations f_n on $[a,b]$ are defined by* (53) *where $t_k^{(n)}$ and $\alpha_k^{(n)}$ are chosen for any $n \in \mathbb{N}$. Then the convergence* (54) *holds if and only if the convergence* (55) *holds and there exists $c > 0$ such that*

$$\sum_{k=0}^{\infty} |\alpha_k^{(n)}| < c \quad \textit{for all } n. \tag{56}$$

Proof. Consider the Banach space $C[a,b]$ that consists of continuous functions on $[a,b]$ with the norm $\|x(t)\| = \max_{t\in[a,b]}|x(t)|$. It is clear that f_n and f are bounded linear functionals on $C[a,b]$, and the convergence (54) is equivalent to that $f_n \overset{w*}{\to} f$. Note that the set of polynomials is total in $C[a,b]$ by the Weierstrass approximation theorem. Thus by Corollary 3, the weak* convergence is equivalent to the convergence (55) and the boundedness of $\{\|f_n\|\}$. It is not hard to check that

$$\|f_n\| = \sum_{k=0}^{\infty} |\alpha_k^{(n)}|.$$

So (55) and (56) are equivalent to (54). $\square$

3.3 Open mapping theorem

Recall a linear operator between two normed spaces is bounded if and only if it is continuous, that is, the preimage of an open set is open. Given two metric spaces X and Y, we say a mapping $f : X \to Y$ is an *open mapping* if for any open set $U \subseteq X$, $f(U)$ is an open set in Y. Not all continuous mappings are open. If a bijective continuous mapping is open, then its inverse is also a continuous mapping. Then if a bounded linear operator is bijective and open, then its inverse is also a bounded linear operator.

Theorem 36. *Let X and Y be two Banach spaces. A surjective bounded linear operator $T : X \to Y$ is an open mapping. Hence if T is bijective, T^{-1} is continuous and thus bounded.*

To prove the theorem, we use the following lemma.

Lemma 19. *Let X and Y be two Banach spaces. A surjective bounded linear operator $T : X \to Y$ has the property that $T(B(0,1))$ contains an open ball centered at $0 \in Y$.*

Proof. We prove the lemma in three steps. First, we show that $\overline{T(B(0,1))}$ contains an open ball. We use the notation that for any set A in a normed space, $\alpha A = \{\alpha x \mid x \in A\}$. It is obvious that for any $\alpha \in \mathbb{R}_+$,

$$T(B(0,\alpha)) = \alpha T(B(0,1)).$$

Then it is not hard to check that

$$\overline{T(B(0,\alpha))} = \alpha \overline{T(B(0,1))}. \tag{57}$$

Since T is surjective, we have that

$$Y = \bigcup_{k=1}^{\infty} T(B(0,k)) = \bigcup_{k=1}^{\infty} \overline{T(B(0,k))}.$$

By Baire's theorem, there is a $\overline{T(B(0,k))}$ that contains an open ball $B(y,r) \subseteq Y$. By (57), we have that

$$\overline{T(B(0,1))} \supseteq \frac{1}{k} B(y,r) = B(\frac{1}{k}y, \frac{r}{k}) \subseteq Y.$$

It is easy to see that if $\overline{T(B(0,1))}$ contains z, then it contains $-z$. So

$$\overline{T(B(0,1))} \supseteq B(-\frac{1}{k}y, \frac{r}{k}).$$

Next we show that $\overline{T(B(0,1))}$ contains an open ball centered at 0. For any two sets A and B in a normed space, we denote $A + B = \{x + y \mid x \in A,\ y \in B\}$. Note that

$$B(0,2) \supseteq B(0,1) + B(0,1).$$

Then we have

$$T(B(0,2)) \supseteq T(B(0,1)) + T(B(0,1))$$

and

$$\overline{T(B(0,2))} \supseteq \overline{T(B(0,1))} + \overline{T(B(0,1))} \supseteq B(\frac{1}{k}y, \frac{r}{k}) + B(-\frac{1}{k}y, \frac{r}{k}) \supseteq B(0, \frac{r}{k}).$$

By (57), we have

$$\overline{T(B(0,1))} \supseteq B(0, \frac{r}{2k}),$$

or more generally, for any $\alpha \in \mathbb{R}_+$,

$$\overline{T(B(0,\alpha))} \supseteq B(0, \frac{\alpha r}{2k}).$$

At last we show that $T(B(0,1))$ contains an open ball centered at 0. Take any $y \in B(0, r/(4k)) \subseteq Y$. Since $y \in B(0, r/(4k)) \subseteq \overline{T(B(0,1/2))}$, there is $x_1 \in B(0,1/2)$ such that $\|y - Tx_1\| < r/(8k)$. We denote $y_1 = Tx_1$. Inductively, we construct $x_2, x_3, \ldots$ and $y_2 = Tx_2, y_3 = Tx_3, \ldots$. Suppose we have $y_1, \ldots, y_i$ such that

$$\|y - (y_1 + \cdots + y_i)\| < \frac{r}{2^{i+2}k},$$

or equivalently,

$$y - (y_1 + \ldots + y_i) \in B(0, \frac{r}{2^{i+2}k}) \subseteq \overline{T(B(0, \frac{1}{2^{i+1}}))} \subseteq Y.$$

Then we choose an $x_{i+1} \in B(0, \frac{1}{2^{i+1}})$ such that

$$\|(y - (y_1 + \cdots + y_i)) - y_{i+1}\| < \frac{r}{2^{i+3}k},$$

and denote $y_{i+1} = Tx_{i+1}$.

From the construction of x_i and y_i, we have that

$$y = \lim_{n\to\infty} \sum_{i=1}^{n} y_i = T\left(\sum_{i=1}^{n} x_i\right).$$

Since by the construction, $\|x_i\| < 2^{-i-1}$ for all i, the sum $\sum_{i=1}^{n} x_i$ has a limit, say x, and $x \in B(0,1)$. We obtain that $y = T(x) \in T(B(0,1))$. By the arbitrariness of y, we have that $T(B(0,1))$ contains $B(0, r/(4k))$, and prove the lemma. □

Proof of Theorem 36. Let $U \subseteq X$ be an open set and $x \in U$ be an arbitrary vector. Then there is an open ball $B(x,r)$ such that $x \in B(x,r) \subseteq U$. By Lemma 19, there is an $B(0,r') \subseteq Y$ such that $T(B(0,1)) \supseteq B(0,r')$. Hence $T(B(x,r)) \supseteq B(Tx, rr')$. So

$$Tx \in B(Tx, rr') \subseteq T(U),$$

and we conclude that $T(U)$ is open and T is an open mapping. □

3.4 Closed graph theorem

Caveat The statement of the closed graph theorem is different from the statement in the textbook, because we use a simplified definition of linear operator. (The domain $D(T)$ is always the whole space.)

In this subsection we consider closed linear operators. Given two normed spaces X and Y, we define the *product space* $X \times Y$ as the normed space with point set $\{(x,y) \mid x \in X,\ y \in Y\}$ the natural vector space structure, and the norm $\|(x,y)\| = \|x\| + \|y\|$.

Definition 18. Let X and Y be normed spaces and $T : X \to Y$ be a linear operator. T is called a *closed linear operator* if its graph

$$G(T) = \{(x,y) \mid x \in X,\ y = Tx\}$$

is closed in the normed space $X \times Y$.

Example 18. Consider the subspace $C^1[a,b] = \{x(t) \in C[a,b] \mid x'(t) \text{ exists and } x'(t) \in C[a,b]\}$, where $C[a,b]$ is the normed space consisting of continuous functions on $[a,b]$ with the norm $\|x(t)\| = \max_{t\in[a,b]}|x(t)|$. Then linear operator $T : C^1[a,b] \to C[a,b]$ such that $Tx(t) = x'(t)$ is unbounded. ($Tx^n = nx^{n-1}$.) But T is a closed operator. Let $\{(x_n(t), x_n'(t))\} \subseteq G(T)$ be any convergent sequence with limit $(x(t), y(t))$, that is, $x(t) \in C^1[a,b]$ and $y(t) \in C[a,b]$ and

$$\lim_{n\to\infty} \|x_n - x\| + \|x_n' - y\| = 0,$$

which is equivalent to that

$$\lim_{n\to\infty}\|x_n - x\| = 0, \quad \text{and} \quad \lim_{n\to\infty}\|x_n' - y\| = 0.$$

In more concrete language, it is equivalent to say that

$$\begin{aligned}\{x_n(t)\} \text{ is a seqeunce of functions uniformly converge to } x(t) \text{ on } [a,b],\\ \text{they are differentiable with continuous derivatives, and}\\ \{x_n'(t)\} \text{ converge uniformly to } y(t) \text{ on } [a,b]. \quad (58)\end{aligned}$$

We claim that $(x(t), y(t)) \in G(T)$, that is, $x'(t) = y(t)$. Hence T is a closed operator. In more concrete language, the more concrete version of the claim is that if $\{x_n\}$ satisfies the condition (58), then $x'(t) = y(t)$. This is proved in Mathematical Analysis II.

From the example, we see

Lemma 20. *The closedness of operator $T : X \to Y$ is equivalent to that if $\{x_n\} \to x$ and $\{Tx\} \to y$, then $Tx = y$.*

Theorem 37 (closed graph)**.** *Let X and Y be Banach spaces and $T : X \to Y$ be a closed linear operator. Then T is a bounded linear operator.*

Proof. First, since X and Y are both Banach spaces, the product space $X \times Y$ is also a Banach space. The proof is left as an exercise. Since $G(T)$ is a closed set and obviously a subspace of $X \times Y$, it is also a Banach space as a subspace of $X \times Y$. The mapping $\tilde{T} : X \to G(T)$ such that $\tilde{T}(x) = (x, Tx)$ is a bijective linear operator. If we show that $\tilde{T}$ is bounded, then T is bounded, since

$$\|\tilde{T}\| = \sup_{\|x\|=1}(\|x\| + \|Tx\|) = 1 + \sup_{\|x\|=1}\|Tx\| = 1 + \|T\|.$$

Although it is not clear that $\|\tilde{T}x\| < c\|x\|$ for some c, it is clear that

$$\|\tilde{T}\| \geq \|x\|. \tag{59}$$

Since $\tilde{T}$ is bijective, it has an inverse mapping $\tilde{T}^{-1} : G(T) \to X$. It is a linear operator, and it is bounded since by (59) $\|\tilde{T}^{-1}\| \leq 1$. By the open mapping theorem, we have that $\tilde{T} = (\tilde{T}^{-1})^{-1}$ is a bounded. Thus we prove the theorem. □

Corollary 4 (Closed graph theorem as stated in the textbook)**.** *Let X and Y be Banach spaces and $U \subseteq X$ be a subspace. Let $T : U \to Y$ be a linear operator such that the graph*

$$G(T) = \{(x, y) \mid x \in U,\ y = Tx\}$$

is a closed set in $X \times Y$. Then if U is a closed set of X, the operator T is bounded.

Proof. If the subspace U is closed in the Banach space X, then U is a Banach space itself. We have

$$G(T) \subseteq U \times Y \supseteq X \times Y$$

where $U \times Y$ is a subspace of $X \times Y$. Since $G(T)$ is closed in $X \times Y$, it is closed in $U \times Y$. Applying Theorem 37 we prove the result.

At last we finish the discussion on closed linear operator by a simple lemma.

Lemma 21. *Let $T : X \to Y$ be a bounded linear operator. Then T is a closed linear operator.*

Proof. If $\{x_n\} \subseteq X$ converges to x and $\{Tx_n\}$ converges to y, then the continuity of T implies that $Tx = y$. Using Lemma 20 we have that T is closed. □

□

4 Basic spectral theory on normed spaces

First we review the concepts of eigenvalue and eigenvector that are well known in linear algebra.

Definition 19. Let X be a normed space and $T : X \to X$ be a linear operator. An *eigenvalue* of T is a scalar $\lambda \in K$ such that there exists a nonzero $x \in X$ such that

$$Tx = \lambda x.$$

This x is called an *eigenvector* of T corresponding to the eigenvalue λ. The eigenvectors corresponding to the eigenvalue λ with the zero vector form a subspace of X, which is called the *eigenspace* of T corresponding to λ.

Example 19. The linear operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 5a+4b \\ a+2b \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}$$

has two eigenvalues, 1 and 6. (It is not a coincidence that they are also the eigenvalues of the matrix $\left(\begin{smallmatrix} 5 & 4 \\ 1 & 2 \end{smallmatrix}\right)$,) and

$$\begin{pmatrix} 5 & 4 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 5 & 4 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 4 \\ 1 \end{pmatrix} = 6\begin{pmatrix} 4 \\ 1 \end{pmatrix}.$$

So $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 4 \\ 1 \end{smallmatrix}\right)$ are eigenvectors of T, corresponding to the eigenvalues 1 and 6 respectively. It is easy to see that $\mathrm{span}\{\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)\}$ and $\mathrm{span}\{\left(\begin{smallmatrix} 4 \\ 1 \end{smallmatrix}\right)\}$ are the corresponding eigenspaces.

Example 20. The linear operator $S : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$S\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}$$

has no real eigenvalue. But if we define S by the same formula above as a complex linear operator from $\mathbb{C}^2$ to $\mathbb{C}^2$, then it has two eigenvalues i and $-i$.

From Example 20 we see that the discussion of eigenvalues is more convenient if we allow complex numbers. Later in this section, we assume the normed spaces are complex unless otherwise stated.

Let $X \neq \{0\}$ be a complex normed space and $T : X \to X$ be a linear operator. If λ is not an eigenvalue of T, then the null space $N(T - \lambda I)$ is $\{0\}$, and then $T - \lambda I$ is an injective linear operator from X to X, or a bijective operator from X to $R(T - \lambda I)$, the range. Thus we can define the inverse operator

$$R_\lambda(T) = (T - \lambda I)^{-1} : R(T - \lambda I) \to X,$$

which is called the *resolvent* of T.

Definition 20. Let X be a complex normed space and $T : X \to X$ be a linear operator. A *regular value* λ of T is a complex number such that

R1 $R_\lambda(T)$ exists, that is, $N(T - \lambda I) = \{0\}$, or equivalently, λ is not an eigenvalue of T.

R2 The operator $R_\lambda(T) : R(T - \lambda I) \to X$ is a bounded linear operator.

R3 The domain of $R_\lambda(T)$, or equivalently the range of $T - \lambda I$, is everywhere dense in X.

The *resolvent set* $\rho(T)$ of T is the set of all regular values λ of T. Its complement in the complex plane, $\sigma(T) = \mathbb{C} \setminus \rho(T)$, is called the *spectrum* of T, and a $\lambda \in \sigma(T)$ is called a *spectral value* of T.

Furthermore, we divide the spectrum into three disjoint sets.

Definition 21.

- *Point spectrum* or *discrete spectrum* $\sigma_p(T)$: The set of eigenvalues.
- *Continuous spectrum* $\sigma_c(T)$: The set of λ such that $R_\lambda(T)$ exists and satisfies R3 but not R2, that is, $R_\lambda(T)$ is unbounded.
- *Residual spectrum* $\sigma_r(T)$: The set of λ such that $R_\lambda(T)$ exists but does not satisfy R3, that is, the range of $T - \lambda I$ is not everywhere dense in X.

The operator T in Example 19 has $\sigma(T) = \sigma_p(T) = \{1, 6\}$. Actually for any linear operator T from a finite dimensional normed space to itself, $\sigma(T) = \sigma_p(T)$.

Example 21. In the normed space $C[0,1]$ that consists of (complex-valued) continuous functions on $[0,1]$ with the maximal norm $\|x(t)\| = \max_{t\in[0,1]}|x(t)|$, the bounded linear operator T defined as

$$Tx(t) = tx(t) \tag{60}$$

has no eigenvalue. If λ is an eigenvalue of T, then

$$tx(t) = \lambda x(t), \quad \text{or equivalently,} \quad (t - \lambda)x(t) = 0$$

is satisfied by a continuous function $x(t) \neq 0$, but it is not possible. All $\lambda \in \mathbb{C} \setminus [0,1]$ is a regular point, since the range of $T - \lambda I$ is the whole space $C[0,1]$ and the resolvent is a bounded linear operator $R_\lambda(T)x(t) = (t-\lambda)^{-1}x(t)$. But for any $\lambda \in [0,1]$,

$$\begin{aligned} R(T - \lambda I) &= \{(t-\lambda)x(t) \mid x(t) \text{ is continous on } [0,1]\} \\ &\subseteq \{x(t) \mid x(t) \text{ is continous on } [0,1] \text{ and } x(\lambda) = 0\}, \end{aligned}$$

so the domain of $R_\lambda(T)$ is not everywhere dense in X. Thus $\sigma(T) = \sigma_r(T) = [0,1]$.

Example 22. The operator T is defined by (60), but the normed space is changed into $C^2[0,1]$, where the vectors in $C^2[0,1]$ are still continuous functions on $[0,1]$, but the norm is given by

$$\|x(t)\| = \left(\int_0^1 |x(t)|^2 dt\right)^{\frac{1}{2}}.$$

Here we also have that T has no eigenvalue, and all $\lambda \in \mathbb{C} \setminus [0,1]$ are regular points. But in the normed space $C^2[0,1]$, for all $\lambda \in [0,1]$, $R(T - \lambda I)$ is everywhere dense, while it is not hard to see that $R_\lambda(T)$ is unbounded. Thus $\sigma(T) = \sigma_c(T) = [0,1]$.

4.1 Bounded linear operators

We derive some basic properties for bounded linear operators on a complex Banach space. In this subsection, we assume that X is a complex Banach space and $T : X \to X$ is a bounded linear operator.

In the definition of resolvent, the domain of $R_\lambda(T)$ brings some technical difficulty. But under some natural condition, this difficulty does not occur.

Lemma 22. *For any $\lambda \in \rho(T)$, $R_\lambda(T)$ is defined on the whold space X and is bounded.*

Proof. We need to show that the domain of $R_\lambda(T)$, that is, the range of $T - \lambda I$, is X. Since by R2 it is already dense in X, we need only to show that $R(T - \lambda I)$ is closed. Let $x \in X$ be the limit of $\{x_n\} \subseteq R(T - \lambda I)$ and $x_n = Ty_n - \lambda y_n$, or equivalently, $y_n = R_\lambda(T)x_n$. By the boundedness of $R_\lambda(T)$, $\{y_n\}$ is a Cauchy sequence, and then it has a limit $y \in X$. By the boundedness of T, we have that $Ty - \lambda y = x$. □

The next theorem shows that we can think about an operator as a scalar, if the convergence is not a problem.

Theorem 38. *Let $T \in B(X, X)$ and $\|T\| < 1$, then $(I - T)^{-1}$ exists as a bounded linear operator on the whole space X and*

$$(I - T)^{-1} = \sum_{j=0}^{\infty} T^j = I + T + T^2 + \cdots .$$

Proof. First, we denote $S_n = I + T + T^2 + \cdots + T^{n-1}$, and have that $S_n \in B(X, X)$, since

$$\|S_n\| \leq \|I\| + \|T\| + \cdots + \|T^{n-1}\| \leq 1 + \|T\| + \cdots + \|T\|^{n-1} < (1 - \|T\|)^{-1}.$$

Similarly, we have that if $n \geq m$,

$$\|S_n - s_m\| \leq \|T\|^m + \|T\|^{m+1} + \cdots + \|T\|^{n-1} < \|T\|^m (1 - \|T\|)^{-1}.$$

By Theorem 7, $B(X, X)$ is a Banach space, so (S_n) has a limit, denoted as $S \in B(X, X)$, which is expressed $I + T + T^2 + \cdots$.

Then we have that for any $x \in X$,

$$\begin{aligned} S(I - T)x &= \lim_{n\to\infty} S_n(I - T)x = \lim_{n\to\infty} (I - T^n)x = x, \\ (I - T)Sx &= (I - T) \lim_{n\to\infty} S_n x = \lim_{n\to\infty} (I - T^n)x = x. \end{aligned}$$

Hence S is an invert of T. The uniqueness of the inverse is clear. □

The theorem above implies an important property of the spectrum of a bounded linear operator.

Theorem 39. *The resolvent set $\rho(T)$ is open; hence the spectrum $\sigma(T)$ is closed.*

Proof. Let $\lambda \in \rho(T) \subseteq \mathbb{C}$. Then $R_\lambda = (T - \lambda I)^{-1}$ is a bounded linear operator from X to X, by Lemma 22. Let $\zeta \in \mathbb{C}$, we have

$$T - \zeta I = (T - \lambda I) - (\zeta - \lambda)I = (T - \lambda I)(I - (\zeta - \lambda)R_\lambda).$$

We note that if $|\zeta - \lambda| < \|R_\lambda\|^{-1}$, then by Theorem 38, there exists $S_\zeta = (I - (\zeta - \lambda)R_\lambda)^{-1}$. We conclude that $(S_\zeta)^{-1}R_\lambda$ is the inverse of $T - \zeta I$ and $\zeta \in \rho(T)$, if $|\zeta - \lambda| < \|R_\lambda\|^{-1}$. Hence λ is an interior point of $\rho(T)$, and $\rho(T)$ is open since λ is arbitrary. □

It is not hard to see that if λ is an eigenvalue of T, then $|\lambda| \leq \|T\|$. The proof is an exercise. The continous spectral points and residual spectral points, although not as simple as eigenvalues, have the same property.

Definition 22. The *spectrum radius* $r_\sigma(T)$ is the radius $\sup_{\lambda \in \sigma(T)} |\lambda|$.

Theorem 40. *$r_\sigma(T) \leq \|T\|$.*

Proof. It suffices to show that if $\lambda \in \mathbb{C}$ and $|\lambda| > \|T\|$, then $T - \lambda I$ is invertible. This is equivalent to that $I - \lambda^{-1}T$ is invertible. It follows from Theorem 38 with T replaced by $\lambda^{-1}T$. □

Actually, we can show that $r_\sigma(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$, but we do not have enough time in this semester. (Think: Why does the right-hand side limit exist?)

4.2 Compact linear operators

In the remaining part of this short section we focus on a special type of linear operators and their spectral properties.

Definition 23. Let X and Y be (complex) normed spaces. An operator $T : X \to Y$ is called a *compact linear operator* if T is linear and for any bounded subset, if T is linear and for any bounded set $M \subseteq X$, $\overline{TM}$ is compact.

An equivalent description of compact linear operator is that T is linear and for any bounded sequence $\{x_n\} \subseteq X$, the image $\{Tx_n\}$ has a convergent subsequence. (The proof of the equivalence is an exercise.)

It is not hard to see that

Lemma 23. *A compact linear operator is a bounded linear operator.*

Although this result is important, we leave the proof as an exercise.

Example 23. Let $K(x, y)$ be a continuous function on $[a, b] \times [a, b]$. Then the operator $T : C[a, b] \to C[a, b]$ given by

$$Tx(t) = \int_a^b K(t, s)x(s)ds$$

is a compact linear operator. To see that it is compact, we take any sequence $\{x_n(t)\} \subseteq C[a, b]$ bounded in the maximal norm on $C[a, b]$. It is easy to check that $\{Tx_n(t)\}$ is also a bounded sequence, or in more concrete term, the sequence of functions are uniform bounded. After some work, we can also show that $\{Tx_n\}$ are equicontinuous. By the Alzelá-Ascoli theorem, there is a subsequence of $\{Tx_n(t)\}$ that converges uniformly, and thus is a convergent subsequence in $C[a, b]$. Hence we prove the compactness of T.

Example 24. The operator $T : \ell^2 \to \ell^2$ defined by

$$T(\alpha_1, \alpha_2, \dots) = (\alpha_1, \frac{\alpha_2}{2}, \frac{\alpha_3}{3}, \dots)$$

is a compact operator. We leave it as an exercise. (A direct proof is possible but tricky. You are encouraged to read the textbook and use the powerful tools there to solve the problem.)

Below we state and prove spectral properties of compact linear operators. We Let X be a (complex) normed space and $T : X \to X$ be a compact linear operator unless otherwise stated.

Theorem 41. *The set of eigenvalues of T is countable, and the only possible limit point is $\lambda = 0$.*

Proof. We only need to show that for any $\epsilon > 0$, there are at most finitely many eigenvalues of T with absolute values greater than ϵ. Suppose to the contray, there are a seqeunce of distinct eigenvalues $\{\lambda_n\}_{n=1}^{\infty} \subseteq \sigma_p(T)$ such that $|\lambda_1| > \epsilon$. Then there a sequence $\{x_n\} \subseteq X$ such that x_n is an eigenvector corresponding to λ_n. A simple fact (to be proved later in Lemma 24) is that the eivenvectors $\{x_n\}$ are linearly independent. Let $X_n = \operatorname{span}\{x_1, \dots, x_n\}$ be nested subspaces. Below we construct a seqeunce $\{z_n\}$ such that $z_n \in X_n \setminus X_{n-1}$ and $\|z_n\| = 1$ such that no subsequence of $\{z_n\}$ converges. Thus there cannot be infinitely many eigenvalues of T whose absolute values are greater than ϵ and we prove the result.

The construction of $\{z_n\}$ is as follows. First, let $z_1 = \|x_1\|^{-1}x_1$. Then we pick z_n by as a vector in X_n such that $\|z_n\| = 1$ and $\|z_n - x\| > 1/2$ for all $x \in X_{n-1}$. By using Riesz's lemma and the fact that the finite dimensional subspace X_{n-1} is a closed subspace of X_n, we have that such z_n exists. We show that for any $m \neq n$, $\|Tz_m - Tz_n\| \geq \epsilon/2$, so no subsequence of $\{Tz_n\}$ is Cauchy, let alone convergent. The reason is simple: Assume that $n > m$ and $z_n = \alpha_n x_n + \alpha_{n-1}x_{n-1} + \cdots + \alpha_1 x_1$ where $\alpha_n \neq 0$,

$$\begin{aligned}
\|Tz_n - Tz_m\| &= \|(\lambda_n\alpha_n x_n + \cdots + \lambda_1\alpha_1 x_1) - Tz_m\| \\
&= \|\lambda_n z_n + (\underbrace{(\lambda_{n-1} - \lambda_n)\alpha_{n-1}x_{n-1} + \cdots + (\lambda_1 - \lambda_n)\alpha_1 x_1}_{\in X_{n-1}}) - \underbrace{Tz_m}_{\in X_{n-1}}\| \\
&= |\lambda_n|\|z_n - \underbrace{((1 - \frac{\lambda_{n-1}}{\lambda_n})\alpha_{n-1}x_{n-1} + \cdots + (1 - \frac{\lambda_1}{\lambda_n})\alpha_1 x_1 + \frac{1}{\lambda_n}Tz_m)}_{\in X_{n-1}}\| \\
&> \frac{1}{2}|\lambda_n| > \frac{\epsilon}{2}.
\end{aligned}$$

□

Now we prove a technical lemma:

Lemma 24. *Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be distinct eigenvalues of T, and $x_1, x_2, \dots, x_n$ are eigenvectors corresponding to these eigenvalues respectively, then $x_1, \dots, x_n$ are linearly independent.*

Proof. Suppose to the contrary, $x_1, \dots, x_n$ are linearly dependent. Then there is a smallest k such that $x_1, \dots, x_k$ are linearly dependent, that is,

$$\alpha_1 x_1 + \cdots + \alpha_k x_k = 0,$$

where $\alpha_1, \dots, \alpha_{n-1}$ are not all 0. Then

$$(T - \lambda_k I)(\alpha_1 x_1 + \cdots + \alpha_k x_k) = (\lambda_1 - \lambda_k)\alpha_1 x_1 + \cdots + (\lambda_{k-1} - \lambda_k)\alpha_{k-1} x_{k-1} = 0,$$

so $x_1, \dots, x_{k-1}$ are linearly depencent, contradictory to that k is the least number to make $x_1, \dots, x_k$ linearly dependent. $\square$

Below we denote $T_\lambda = T - \lambda I$. If $\lambda \notin \sigma_p(T)$ then $N(T_\lambda) = \{0\}$, otherwise $N(T_\lambda)$ is the eigenspace corresponding to λ.

Theorem 42. *For every $\lambda \neq 0$, the null space $N(T_\lambda)$ is finite dimensional.*

Proof. If $N(T_\lambda)$ is not finite dimensional, then its unit closed ball $B_\lambda = \{x \in N(T_\lambda) \mid \|x\| \leq 1\}$ is closed but noncompact as a set in the subspace $N(T_\lambda)$, as proved by Riesz's lemma in Section 1.5. Then we have that B_λ is noncompact in X, and using the fact that $N(T_\lambda)$ is a closed subspace of X, we also have that it is a closed set in X. Since

$$\overline{T(B_\lambda)} = \overline{\lambda B_\lambda} = \lambda B_\lambda = \{x \in N(T_\lambda) \mid \|x\| \leq \lambda\},$$

that is a noncompact set, we conclude that T is not a compact linear operator and derive a contradiction. $\square$

Theorem 43. *For any $\lambda \neq 0$, the range $R(T_\lambda)$ is a closed subspace.*

Proof. We show the closedness by contradiction. Suppose $\{T_\lambda x_n\}$ is a convergent sequence in $R(T_\lambda)$ whose limit is $y \neq 0$, and assume that $y \notin R(T_\lambda)$. Without loss of generality, we assume that $\{x_n\}$ are linearly independent. Below we make a technical assumption that the distance between each x_n and the subspace $N(T_\lambda)$ is no less than than $\|x_n\|/2$, that is, $\inf_{z \in N(T_\lambda)} \|x_n - z\| \geq \|x_n\|/2$. If $N(T_\lambda) = \{0\}$, this condition is trivially true. Otherwise, we have that $N(T_\lambda)$ is finite dimensional and thus closed, so by Riesz's lemma, we have a $z_n \in N(T_\lambda)$ such that $\inf_{z \in N(T_\lambda)} \|(x_n - z_n) - z\| \geq \|x_n - z_n\|/2$, and then we rewrite $x_n - z_n$ as x_n, since $T_\lambda(x_n - z_n) = T_\lambda(x_n)$.

If $\{\|x_n\|\}$ is bounded, then we have a contradiction to the assumption. To see it, we have by the compactness of T that a subsequence $\{x_{n_k}\}$ satisfies that $\{Tx_{n_k}\}$ converges to $z \in X$. Then $\{\lambda x_{n_k}\}$ converges to the limit of $\{Tx_{n_k} - T_\lambda x_{n_k}\}$ that is $z - y$. By the continuity of T_λ,

$$T_\lambda(\frac{1}{\lambda}(z - y)) = \lim_{n \to \infty} T_\lambda(x_{n_k}) = y,$$

and $y \in R(T_\lambda)$.

Thus $\{x_n\}$ cannot be bounded, and without loss of generality we assume $\|x_n\| \to \infty$. Write $y_n = \|x_n\|^{-1} x_n$, we have that $\|y_n\| = 1$ for all y_n and $\{T_\lambda y_n\}$ converges to 0. Use the compactness of T again, we have that a subsequence y_{n_k} satisfies that $\{Ty_{n_k}\}$ converges to a limit $w \in X$. Then $\{\lambda y_{n_k}\}$ converges to the limit of $\{Ty_{n_k} - T_\lambda y_{n_k}\}$ that is also w, or equivalently, $\{y_{n_k}\} \to \lambda^{-1} w$. By the continuity of T_λ, we have that $T_\lambda \lambda^{-1} w) = 0$, that is, $\lambda^{-1} w \in N(T_\lambda)$. But we assumed that x_{n_k} satisfies $\inf_{z \in N(T_\lambda)} \|x_{n_k} - z\| \geq \|x_{n_k}\|/2$, so $\inf_{z \in N(T_\lambda)} \|y_{n_k} - z\| \geq \|y_{n_k}\|/2 = 1/2$, and then y_{n_k} cannot converge to a point in $N(T_\lambda)$, and we derive a contradiction. Thus $R(T_\lambda)$ is closed. $\square$

Theorem 44. *Suppose X is a Banach space. If $\lambda \neq 0$ is not an eigenvalue, then λ is a regular point.*

Proof. The condition that λ is not an eigenvalue means that $N(T_\lambda) = \{0\}$ and then T_λ is injective. Below we show that T_λ is also surjective. Denote $X_n = R(T_\lambda^n)$. Suppose to the contrary, T_λ is not surjective, then X_1 is a proper subset of X. Since $X_2 = T_\lambda(X_1)$ and T_λ is injective, we have that X_2 is a proper subset of $T(X) = X_1$. Similarly, all the inclusions

$$X_1 \supset X_2 \supset X_3 \supset X_4 \supset \cdots$$

are proper.

We know that X_1 is a closed subspace. Actually all X_n are closed subspaces. To see it, by binomial expansion we write

$$T_\lambda^n = (T - \lambda I)^n = \underbrace{T^n - n\lambda T^{n-1} + \cdots + (-1)^{n-1}\lambda^{n-1}}_{\text{compact linear operator}} + (-1)^n \lambda^n I,$$

and then use Theorem 43. Here we use two properties of compact linear operators whose proofs are left as exercise: (1) The power of a compact linear operator is compact, and (2) the sum of two compact linear operators is compact.

Hence by Riesz's lemma, we can construct a sequence $\{z_n\}_{n=1}^\infty$ such that $z_n \in X_n \setminus X_{n+1}$, $\|z_n\| = 1$ and $\|z_n - x\| > 1/2$ for all $x \in X_{n+1}$. The existence of such a sequence is due to the same reason of the existence of $\{z_n\}$ in the proof of Theorem 41. Then we have that for all $n < m$,

$$\begin{aligned}
\|Tz_n - Tz_m\| &= \|\lambda z_n + \underbrace{(T - \lambda I)z_n}_{\in X_{n+1}} - \underbrace{\lambda z_m}_{\in X_{n+1}} - \underbrace{(T - \lambda I)z_m}_{\in X_{n+1}}\| \\
&= |\lambda| \|z_n - \underbrace{(z_m + \frac{1}{\lambda}(T - \lambda I)z_m - \frac{1}{\lambda}(T - \lambda I)z_n)}_{\in X_{n+1}}\| \\
&> \frac{1}{2}|\lambda|.
\end{aligned}$$

So no subsequence of $\{z_n\}$ converges, contradictory to that T is compact. Thus we prove that T_λ is surjective.

By the open mapping theorem, we have that $R_\lambda(T) = T_\lambda^{-1}$ is a bounded linear operator, and so λ is a regular point. $\square$

The last theorem shows that the for a compact linear operator T, $\sigma(T) = \sigma_p(T)$, and there are neither continuous spectrum nor residual spectrum.

5 Spectral properties of bounded self-adjoint linear operators

In this section, we assume that H is a complex Hilbert space.

Let $T : H \to H$ be a bounded linear operator. The Hilbert adjoint of T is an operator $T^* : H \to H$ that satisfies

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

for any $x, y \in H$. We say that T is *self-adjoint* if $T = T^*$. Below we assume $T : H \to H$ is a bounded self-adjoint linear operator.

Self-adjoint operators are very important in quantum mechanics, where all physically meaningful quantities are expressed by an operator on a complex Hilbert space. Actually most meaningful operators in quantum mechanics are *unbounded*. But it is out of the scope of this module.

An interesting observation is that in quantum mechanics, one talks about the point/discrete spectrum and the continuous spectrum, but never the residual spectrum. We can prove that the residual spectrum is empty for a bounded self-adjoint operator.

Theorem 45. *1. All the eigenvalues of T are real.*

2. Eigenvectors corresponding to different eigenvalues are orthogonal.

Proof. For part 1, we note that if $\lambda \in \sigma_p(T)$, thenthere exists an eigenvector $x \in X$ with $\|x\| = 1$ corresponding to λ. We have

$$\lambda = \langle \lambda x, x\rangle = \langle Tx, x\rangle = \langle x, Tx\rangle = \langle x, \lambda x\rangle = \bar{\lambda}.$$

We conclude that $\lambda \in \mathbb{R}$.

For part 2, we assume that x, y are eigenvectors associated to λ and ξ respectively, where $\lambda \neq \xi$ are real numbers. Then

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

We conclude that $\langle x, y\rangle = 0$, that is, x, y are orthogonal. □

Theorem 46. *Complex number $\lambda \in \rho(T)$ if and only if there exists $c > 0$ such that $(T_\lambda = T - \lambda I)$*

$$\|T_\lambda x\| \geq c\|x\|.$$

Proof. For the "only if" part, we note that if $\lambda \in \rho(T)$, then $R_\lambda = T_\lambda^{-1}$ is well defined and bounded, so if $T_\lambda x = y$, we have $R_\lambda y = x$ and $\|x\| \leq c^{-1}\|y\|$ for some $c > 0$. This proves the result.

For the "if" part, we note that if the inequality $\|T_\lambda x\| \geq c\|x\|$ holds, then T_λ is injective, and $R_\lambda : R(T_\lambda) \to H$ is bounded. Hence λ cannot be in $\sigma_p(T)$ or $\sigma_c(T)$. We only need to show that λ is not in $\sigma_r(T)$, that is, $\overline{R(T_\lambda)}$ is not a proper subspace of H.

Suppose not, then there exists $y \in H$ such that $\|y\| > 0$ and $y \perp \overline{R(T_\lambda)}$. We have that for all $x \in H$,

$$\langle x, (T - \bar{\lambda} I)y\rangle = \langle (T - \lambda I)x, y\rangle = 0,$$

so y is an eigenvector corresponding to $\bar{\lambda}$. So $\lambda \in \mathbb{R}$, and $\lambda = \bar{\lambda} \in \sigma_p(T)$, contradictory to the assumption. □

Theorem 47. *Let $T : H \to H$ be a bounded self-adjoint linear operator. Then the spectrum $\sigma(T)$ is real.*

Proof. It suffices to show that for any $\lambda \in \mathbb{C}\backslash\mathbb{R}$, there exists $c > 0$ such that $\|T_\lambda x\| \geq c\|x\|$ for all $x \in H$.

We note that for any $x \in H$, $\langle Tx, x\rangle = \langle, x, Tx\rangle \in \mathbb{R}$. The identity

$$\langle (T - \lambda I)x, (T - \lambda I)x\rangle = \langle Tx, Tx\rangle + |\lambda|^2\langle x, x\rangle - \lambda\langle x, Tx\rangle - \bar{\lambda}\langle Tx, x\rangle$$

and the inequality

$$|\langle Tx, x\rangle| = |\langle x, Tx\rangle| \leq \|Tx\|\|x\|$$

imply that

$$\begin{aligned}\langle (T - \lambda I)x, (T - \lambda I)x\rangle &\geq \|Tx\|^2 + (\Re\lambda)^2\|x\|^2 - 2\Re\lambda\|Tx\|\|x\| + (|\lambda|^2 - (\Re\lambda)^2)\|x\| \\ &= (\|Tx\| - \Re\lambda\|x\|)^2 + (|\lambda|^2 - (\Re\lambda)^2)\|x\|.\end{aligned}$$

Hence we can take $c = (|\lambda|^2 - (\Re\lambda)^2)$. So $\lambda \in \mathbb{C}$ are in $\rho(T)$. □

Theorem 48. *The residual spectrum $\sigma_r(T) = \emptyset$.*

Proof. Suppose $\lambda \in \sigma_r(T)$. By last theorem, $\lambda \in \mathbb{R}$. Hence $T_\lambda = T - \lambda I$ is also a bounde self-adjoint linear operator and $0 \in \sigma_r(T_\lambda)$. Hence we only need to show that $0 \notin \sigma_r(T)$.

Suppose not, then $R(T)$ is not everywhere dense in H, and then by Lemma 10 there exists nonzero $y \in H$ that is orthogonal to $R(T)$. We have that for all $x \in H$

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

so by Lemma 12(a) we have $Ty = 0$, that is, $0 \in \sigma_p(T)$, contradictory to that $0 \notin \sigma_r(T)$. □

Theorem 46 implies more spectral properties of a bounded self-adjoint operator.

Theorem 49. *Let $m = \inf_{x\in H,\ \|x\|=1}\langle Tx, x\rangle$ and $M = \sup_{x\in H,\ \|x\|=1}\langle Tx, x\rangle$.*

1. *$\sigma(T)$ is contained in the closed interval $[m, N]$.*
2. *m and M are in $\sigma(T)$.*

Before the proof, we introduce the concept of *positive* self-adjoint operators: T is positive means that $\langle Tx, x\rangle \geq 0$ for all $x \in H$. (Note that for all self-adjoint T, $\langle Tx, x\rangle$ is a real number (exercise)).

Proof. To prove part 1, we only need to show that for $\lambda \in \mathbb{R}$ such that either $\lambda > M$ or $\lambda < m$, there is $c > 0$ such that $\|T_\lambda x\| \geq c\|x\|$ for all $x \in H$. Note that if T is bounded and self-adjoint, then $-T$ is too, and m, M for $-T$ are exactly M, m for T, and $\lambda \in \sigma(T)$ if and only if $-\lambda \in \sigma(-T)$. So we only need to show that all $\lambda < m$ are regular. Furthermore, $T_m = T - mI$ is a positive self-adjoint operator and $\lambda \in \sigma(T)$ if and only if $\lambda - m \in \sigma(T_m)$. Hence we only need to prove that if $\lambda < 0$ and T is positive, then $\lambda \in \sigma(T)$.

Suppose $\lambda < 0$ and T is positive, then for any $x \in H$,

$$\|T_\lambda x\|^2 = \langle T_\lambda x, T_\lambda(x)\rangle = \|Tx\|^2 + \lambda^2\|x\|^2 - 2\lambda\langle Tx, x\rangle \geq \lambda^2\|x\|,$$

so $\|T_\lambda x\| \geq |\lambda|\|x\|$, and so λ is a regular point.

To prove part 2, we use the argument above and reduce it to the proof that if T is positive and

$$\inf_{x\in H,\ \|x\|=1} \langle Tx, x\rangle = 0,$$

then 0 is a spectral point of T.

If not, then there exists $c > 0$ such that $\|Tx\| \geq c\|x\|$ for all $x \in H$. Suppose $\|T\| = C$. We can find $x \in H$ such that $\|x\| = 1$ and $\langle Tx, x\rangle < c^4/C^3$. Then letting $a = C^3/c^2$, we have

$$\langle T(Tx - ax), Tx - ax\rangle = \langle T^2x, Tx\rangle - 2a\|Tx\|^2 + a^2\langle Tx, x\rangle < C^3 - 2ac^2 + a^2(c^4/C^3) = 0,$$

contradictory to the assumption that T is positive. Hence 0 is a spectral point. $\square$