II Linear Analysis

Ishan Nath, Michaelmas 2023

Based on Lectures by Prof. Imre Leader

January 30, 2024

Contents

1   Normed Spaces and Linear Operators

A *normed space* is a pair $(X, \|\cdot\|)$, where X is a real or complex vector space, and $\|\cdot\|$ is a *norm* on X, meaning a function from $X \to \mathbb{R}^+$ satisfying:

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

- $\|\lambda x\| = |\lambda| \|x\|$ for all scalars λ and $x \in X$,

- $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

The last condition is known as the triangle inequality (or $\triangle$ inequality).

Example 1.1.   (Normed spaces)

1. ℓ_2^n: $X = \mathbb{R}^n$, $\|x\| = (\sum |x_i|^2)^{1/2}$. Here $X = \mathbb{C}^n$ is also possible.

   $\triangle$ inequality follows from Cauchy-Schwarz.

2. ℓ_1^n: $X = \mathbb{R}^n$, $\|x\| = \sum |x_i|$. This is called the 1-norm, written $\|x\|_1$.

   $\triangle$ inequality follows as $|\cdot|$ satisfies the $\triangle$ inequality.

3. ℓ_∞^n: $X = \mathbb{R}^n$, $\|x\| = \max |x_i|$. This is called the ∞-norm, written $\|x\|_\infty$.

A norm on X gives rise to a metric on X: $d(x, y) = \|x - y\|$. The conditions for being a metric can be checked.

This metric is 'translation-invariant': $d(x + z, y + z) = d(x, y)$.

As a consequence of having a metric, we can talk about open sets, closed sets, convergent subsequences, etc. Under this metric, $+$ is continuous: if $x_n \to x$, $y_n \to y$ then $x_n + y_n \to x + y$.

Moreover $\cdot$ is continuous: if $x_n \to x$, $\lambda_n \to \lambda$, then $\lambda_n x_n \to \lambda x$. And thankfully $\|\cdot\|$ is continuous: if $d(x, y) < \varepsilon$, then $\|x - y\| < \varepsilon$ so $|\|x\| - \|y\|| < \varepsilon$.

In order to visualise normed spaces it is often helpful to consider the unit ball $B(X)$ of the space. This is the set

$$B = \{x \in X \mid \|x\| \leq 1\}.$$

The balls for ℓ_2^2, ℓ_1^2 and ℓ_∞^2 look like a circle, a diamond and a square respectively. Maybe I will include some pictures; probably not.

B uniquely defines the norm: we can let

$$\|x\| = \inf\{t > 0 \mid x \in tB\}.$$

Moreover B is always convex: if $x, y \in B$ and $0 < \lambda < 1$, then $\lambda x + (1 - \lambda)y \in B$. This is an easy consequence of the $\triangle$ inequality.

An interesting fact is that any set $B \subset \mathbb{R}^n$ which is closed, bounded, convex, symmetric, and a neighbourhood of 0 defines a norm in the same way as above: $\|x\| = \inf\{t > 0 \mid x \in tB\}$. Moreover under this construction, B is precisely the unit ball.

A *Banach space* is a complete normed space. The three examples we saw earlier are all Banach spaces, since $x^{(n)} \to x \iff x_i^{(n)} \to x_i$.

We will look at some more examples of normed spaces. In the following let S be the vector space of all scalar sequences, where operations are done pointwise.

Example 1.2.   (Normed spaces contd.)

4. For $1 \le p < \infty$, ℓ_p^n: $X = \mathbb{R}^n$, $\|x\|_p = \left(\sum |x_i|^p\right)^{1/p}$.

   $\triangle$ inequality follows from Minkowski's inequality.

5. ℓ_1: $X = \{x \in S \mid \sum |x_n| < \infty\}$, $\|x\|_1 = \sum |x_n|$.

   $\triangle$ inequality is obvious.

6. For $1 \le p < \infty$, ℓ_p: $X = \{x \in S \mid \sum |x_n|^p < \infty\}$, $\|x\|_p = \left(\sum |x_n|^p\right)^{1/p}$.

   $\triangle$ inequality follows from Minkowski's.

7. ℓ_∞: $X = \{x \in S \mid x \text{ bounded}\}$, $\|x\|_\infty = \sup |x_n|$.

   $\triangle$ inequality is easy.

8. C_0: $X = \{x \in S \mid x \to 0\}$, $\|x\|_\infty$ as before.

We will later see that ℓ_p and C_0 are Banach spaces.

1.1   Inequalities

Say $f : \mathbb{R} \to \mathbb{R}$ is *convex* if

$$f(\lambda x + (1 - \lambda)y) \le \lambda f(x) + (1 - \lambda)f(y),$$

for all $x, y \in \mathbb{R}$, $0 < \lambda < 1$. f is *concave* if $-f$ is convex.

Lemma 1.1. *For $1 \le p < \infty$, the function $x \mapsto x^p$ is convex over $\mathbb{R}^+$.*

Proof: We need to prove that $(\lambda x + (1-\lambda)y)^p \leq \lambda x^p + (1-\lambda)y^p$. Without loss of generality, let $x \leq y$. If $x = y$, we are done and both sides are equal, so let $x < y$.

Then the derivative of both sides with regards to x are

$$\frac{\mathrm{d}}{\mathrm{d}x}(\lambda x + (1-\lambda)y)^p = \lambda p(\lambda x + (1-\lambda)y)^{p-1}, \qquad \frac{\mathrm{d}}{\mathrm{d}x}(\lambda x^p + (1-\lambda)y^p) = \lambda p x^{p-1},$$

hence the derivative of the left hand side is greater than the derivative of the right hand side, for all $x \leq y$. This finishes the proof.

Alternatively, note that $x \mapsto x^p$ has a derivative that is increasing.

Theorem 1.1 (Minkowski's inequality). *Let $1 \leq p < \infty$, and $x, y \in \ell_p$.*

Then $x + y \in \ell_p$, with $\|x + y\|_p \leq \|x\|_p + \|y\|_p$.

Proof: First we show that if $\|x\|_p, \|y\|_p = 1$, then $\|\lambda x + (1-\lambda)y\|_p \leq 1$, for all $0 < \lambda < 1$.

We have that $|\lambda x_n + (1-\lambda)y_n|^p \leq \lambda |x_n|^p + (1-\lambda)|y_n|^p$, so by summing

$$\sum_{n=1}^{k} |\lambda x_n + (1-\lambda)y_n|^p \leq \lambda \sum_{n=1}^{k} |x_n|^p + (1-\lambda) \sum_{n=1}^{k} |y_n|^p.$$

Letting $k \to \infty$, we get

$$\sum_{n=1}^{k} |\lambda x_n + (1-\lambda)y_n|^p \leq \lambda \cdot 1 + (1-\lambda) \cdot 1 = 1.$$

For general x, y non-zero, apply the above to $\frac{x}{\|x\|_p}$ and $\frac{y}{\|y\|_p}$, to get that

$$\frac{\|x\|_p}{\|x\|_p + \|y\|_p} \frac{x}{\|x\|_p} + \frac{\|y\|_p}{\|x\|_p + \|y\|_p} \frac{y}{\|y\|_p}$$

has norm at most 1, or $\|x + y\|_p \leq \|x\|_p + \|y\|_p$.

Note that if $x \in \ell_1$ and $y \in \ell_\infty$, then $x_n y_n \in \ell_1$, with $\|(x_n y_n)\|_1 \leq \|x\|_1 \|y\|_\infty$.

For $1 < p < \infty$, the *conjugate index* is the $1 < q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Note that 2 is self-conjugate.

Our aim is to prove that if p, q are conjugate, then $|x \cdot y| \leq \|x\|_p \|y\|_q$. This is a generalization of Cauchy-Schwarz.

Lemma 1.2. *Let $1 < p, q < \infty$ be conjugate. Then*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q},$$

for all $a, b \geq 0$.

> **Proof:**   We may assume that $ab > 0$. Let $x = a^p, y = b^q$. Then we need to prove that
> $$x^{1/p} y^{1/q} \leq \frac{x}{p} + \frac{y}{q},$$
> or equivalently, taking the log of both sides,
> $$\frac{1}{p} \log x + \frac{1}{q} \log y \leq \log\left(\frac{x}{p} + \frac{y}{q}\right).$$
> This is true by the concavity of log.

Theorem 1.2 (Hölder's inequality). *Let $x \in \ell_p$, $y \in \ell_q$, where $1 < p, q < \infty$ are conjugate.*

Then $(x_n y_n) \in \ell_1$, and
$$\|(x_n y_n)\|_1 \leq \|x\|_p \|y\|_q.$$

> **Proof:**   Without loss of generality, assume $x, y \neq 0$. By scaling, assume $\|x\|_p = \|y\|_q = 1$. Then
> $$|x_n y_n| \leq \frac{|x_n|^p}{p} + \frac{|y_n|^q}{q}.$$
> By summing,
> $$\sum_{n=1}^{\infty} |x_n y_n| \leq \frac{1}{p} + \frac{1}{q} = 1.$$

1.2   Complete, Dense and Separable

We can prove that each ℓ_p is complete, which we will do later.

For now we will give some more examples of normed vector spaces.

Example 1.3.

Here you are.

1. $C[0,1]$: $X = \{f : [0,1] \to \mathbb{R} \mid f \text{ continuous}\}$, $\|f\|_\infty = \sup_x |f(x)|$.

   This is complete, by the general principle of uniform convergence, plus the fact that a uniform limit of continuous functions is continuous.

2. More generally we can take $C(K)$, where K is compact and Hausdorff, with $\|f\|_\infty = \sup_x \in K |f(x)|$. We can also consider complex valued functions.

3. We can take $C[0,1]$ with $\|f\|_1 = \int_0^1 |f|$.

   This is incomplete, as we can approximate $H(0.5)$.

   Similarly, we can take $\|f\|_p = (\int_0^1 |f|^p)^{1/p}$ for any $1 \le p < \infty$. These are again incomplete. However, if we take $L_p = \{f : [0,1] \to \mathbb{R} \mid f \text{ Lebesgue integrable}, \|f\|_p < \infty\}$, then this is complete.

4. We can take $C^1[0,1] = \{f \in C[0,1] \mid f \text{ continuously differentiable}\}$, with $\|f\|_\infty$.

   This is incomplete, by taking approximations of $|x - 1/2|$. However, it is complete with the norm $\|f\| = \|f\|_\infty + \|f'\|_\infty$.

   In general, we often need the norm to take into account the definition of the vector space to have completeness.

5. Let $\triangle = \{z \in \mathbb{C} \mid |z| \le 1\}$, the closed unit disc. Write $A(\triangle)$ for $\{f \in C(\triangle) \mid f \text{ analytic in } \triangle\}$, with the sup norm $\|f\|_\infty$. Then this is complete, as the limit of analytic functions is analytic.

In ℓ_p, we write e_n for the element with a 1 in the n'th coordinate, and 0's in all other coordinates. Then ℓ_p is not the linear span of the e_n: indeed, $\langle e_n \rangle$ is F, the space of all finite sequences

$$F = \{x \in S \mid x_n = 0 \text{ for all } n \ge n_0 \text{ for some } n_0\}.$$

Note F does 'get close': for $1 \le p < \infty$ the closed linear span of the e_n is ℓ_p: given $x \in \ell_p$, we have $\sum_{n=1}^k x_n e_n \to x$ as $k \to \infty$, because $x - \sum_{i=1}^k x_n e_n$ has norm

$$\left(\sum_{n=k+1}^\infty |x_n|^p \right)^{1/p},$$

which goes to 0 because $\sum_{i=1}^\infty |x_n|^p$ converges.

However, this is not true in ℓ_∞. Take $x = (1, 1, 1, \ldots)$. Then for all $y \in F$, $\|x - y\|_\infty \geq 1$. So F is not dense in ℓ_∞.

We call a subset A of a topological space X is *dense* if $\overline{A} = X$. We say a topological space X is *separable* if it has a countable dense subset. For example, $\mathbb{Q}$ is dense in $\mathbb{R}$, so $\mathbb{R}$ is separable.

The above remarks show that ℓ_p for $1 \leq p < \infty$ is separable. However, ℓ_∞ is not separable (fun exercise!)

Note that subspaces do not need to be closed. For example take $F \subset \ell_1$. Then $F \neq \ell_1$, but $\overline{F} = \ell_1$.

Proposition 1.1. *Let X be a normed space, and Y a subspace of X.*

(i) *If Y is complete, then Y is closed.*

(ii) *If X is complete and Y is closed, then Y is complete.*

> **Proof:**   Let's prove these one by one.
>
> (i) Given $x \in \overline{Y}$, we have $y_1, y_2, \ldots$ in Y with $y_n \to x$. So (y_n) is Cauchy. Then as Y complete, $y_n \to y$ for some $y \in Y$. Then $x = y$, so $x \in Y$.
>
> (ii) Given a sequence (y_n) Cauchy in Y, it is also Cauchy in X. As X complete, $y_n \to x$ for some $x \in X$. But as Y is closed, $x \in Y$.

2   Linear Operators

Theorem 2.1. *Let X, Y be normed spaces, and $T : X \to Y$ linear. Then the following are equivalent:*

  (i) *T is continuous.*

  (ii) *T is continuous at 0.*

  (iii) *There exists $K > 0$ such that $\|Tx\| \le K\|x\|$ for all $x \in X$ (i.e. T bounded).*

> **Proof:**   First note that (i) immediately implies (ii).
>
> To show (ii) $\implies$ (iii), note $B(Y)$ is a neighbourhood of $0 \in Y$, so there exists $\delta > 0$ such that
>
> $$\|x\| \le \delta \implies \|Tx\| \le 1,$$
>
> for all $x \in X$. Hence $\|Tx\| \le \delta^{-1}\|x\|$ for all $x \in X$.
>
> Finally, for (iii) $\implies$ (i), take $\varepsilon > 0$. Then for all $x, y \in X$,
>
> $$\|x - y\| \le \frac{\varepsilon}{K} \implies \|Tx - Ty\| \le \varepsilon.$$
>
> So in fact, T is uniformly continuous.

We write $L(X, Y)$ for $\{T : X \to Y \mid T \text{ linear and continuous}\}$. For $T \in L(X, Y)$, the *norm* or *operator norm* of T is

$$\|T\| = \|T\|_{\mathrm{op}} = \sup\{\|Tx\| \mid \|x\| \le 1\}.$$

Thus, $\|Tx\| \le \|T\|\|x\|$.

Given these definitions, it makes sense to show that $L(X, Y)$ is a normed space with norm $\|\cdot\|_{\mathrm{op}}$. Indeed, note that $L(X, Y)$ is a vector space. Given $S, T \in L(X, Y)$,

$$\|(S + T)x\| \le \|Sx\| + \|Tx\| \le \|S\|\|x\| + \|T\|\|x\| = (\|S\| + \|T\|)\|x\|.$$

So $S + T \in L(X, Y)$ with $\|S + T\| \le \|S\| + \|T\|$. This show $\|\cdot\|_{\mathrm{op}}$ is a norm on $L(X, Y)$.

Proposition 2.1. *Let X, Y, Z be normed spaces, $S \in L(X, Y)$, $T \in L(Y, Z)$. Then $T \circ S \in L(X, Z)$, and moreover $\|T \circ S\| \le \|S\|\|T\|$.*

Proof: Really easy. For all $x \in X$,

$$\|(T \circ S)(x)\| = \|T(S(x))\| \leq \|T\| \|S(x)\| \leq \|T\| \|S\| \|x\|.$$

Example 2.1.

Start with some boring examples.

1. Take $T : \ell_2^n \to \ell_2^n$ given by $T((x_1, \ldots, x_n)) = (x_1, \ldots, x_r, 0, \ldots, 0)$ for some fixed $1 \leq r \leq n$. Then T is continuous, with $\|T\| \leq 1$. In fact, $\|T\| = 1$ as $T(e_1) = e_1$.

2. Take the right shift T on ℓ_p:

$$T(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots).$$

   Then $\|T\| = 1$, and indeed T is an *isometry*: $\|Tx\| = \|x\|$ for all x.

   Note that T is injective, but not surjective.

3. We also have the friend, the left shift on ℓ_p:

$$T(x_1, x_2, \ldots) = (x_2, x_3, \ldots).$$

   This is continuous as the norm decreases, and in fact $\|T\| = 1$ as $T(e_2) = (e_1)$. Also, T is surjective, but not injective.

4. Fix p, q with $1 \leq p, q \leq \infty$ which are conjugate: $\frac{1}{p} + \frac{1}{q} = 1$. For any $x \in \ell_p$, $y \in \ell_q$, write $x \cdot y$ for $\sum x_n y_n$. This converges absolutely, with $\sum |x_n y_n| \leq \|x\|_p \|y\|_q$ by Hölder.

   For fixed $y \in \ell_q$, define $\phi_y : \ell_p \to \mathbb{R}$ given by $x \mapsto x \cdot y$.

   Then ϕ_y is linear. Also $\|\phi_y(x)\| = \|x \cdot y\| \leq \|x\|_p \|y\|_q$, so ϕ_y is continuous. Hence $\phi_y \in L(\ell_p, \mathbb{R})$ with $\|\phi_y\| \leq \|y\|_q$.

5. Let us cook up a discontinuous linear map. Define $T : F \to \mathbb{R}$ by

$$T\left(\sum_{n=1}^k x_n e_n \right) = \sum_{n=1}^k n x_n.$$

   Then T is not bounded, as $|T(e_n)| = n$.

6. Define $T : \ell_1 \to \ell_2$ by $T(x) = x$. Note that $x \in \ell_1 \implies x \in \ell_2$, as

$$\sum |x_n| \leq 1 \implies \sum |x_n|^2 \leq 1.$$

> We have $T \in L(\ell_1, \ell_2)$, with $\|T\| \le 1$ (and in fact $T(e_1) = e_1$ so $\|T\| = 1$). This is not surjective, as $(1, 1/2, 1/3, \ldots)$ is not in the image of T. Note that the image T is dense in ℓ_2, as F is contained in the image. Hence the image is not complete!

In general, $L(X, \mathbb{R})$ is called the *dual* of X, written X^*. Elements of X^* are called *functionals*.

For normed spaces X, Y, an *isomorphism* from X to Y is a linear homeomorphism from X to Y. So a bijective linear $T : X \to Y$ is an isomorphism if and only if T and T^{-1} are continuous, or equivalently if there exists $c, d > 0$ such that

$$c\|x\| \le \|Tx\| \le d\|x\|,$$

for all $x \in X$. If T is a bijective linear map with $\|Tx\| = \|x\|$ for all x, we say that T is an *isometric isomorphism*.

The *Banach-Mazur distance* between isomorphic X, Y is

$$d(X, Y) = \inf_{\substack{T \in L(X,Y) \\ T \text{ isomorphism}}} \|T\|\|T^{-1}\|.$$

Norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a vector space V are *equivalent* if their metrics induce the same topology. That is, if the identity from $(V, \|\cdot\|_1)$ to $(V, \|\cdot\|_2)$ is continuous in both ways, i.e. there exists $c, d > 0$ with

$$c\|x\|_1 \le \|x\|_2 \le d\|x\|_1,$$

for all $x \in V$. We can also think of this as scalings of the unit balls being contained in each other.

Note that if $\|\cdot\|_1$ is equivalent to $\|\cdot\|_2$, and one is complete, then so is the other. This is false in general metric spaces, e.g. $(0, 1) \sim \mathbb{R}$.

Remark.

1. $C[0, 1]$ is complete with $\|\cdot\|_\infty$, and incomplete with $\|\cdot\|_1$, hence $\|\cdot\|_1$ is not equivalent to $\|\cdot\|_\infty$.

2. On F, $\|\cdot\|_1$ and $\|\cdot\|_2$ are inequivalent. Take $x = (1, \ldots, 1, 0, \ldots)$, with n 1's. Then $\|x\|_1 = n$, but $\|x\|_2 = \sqrt{n}$.

3. The identity $T : (C[0, 1] : \|\cdot\|_\infty) \to (C[0, 1], \|\cdot\|_1)$ is a linear bijection with T continuous, but T^{-1} not continuous.

Note that if $T_1, T_2, \ldots \in L(X, Y)$ and $T_n \to T$ in operator norm, then $T_n x \to Tx$ for all $x \in X$. Indeed,

$$\|Tx - T_n x\| \leq \|T - T_n\| \|x\|.$$

The converse is false, e.g. take $T_1, T_2, \ldots \in \ell_1^*$ given by $T_n(x) = x_n$ for all $x \in X$. Then $T_n x \to 0$ for all $x \in X$, but T_n does not tend to 0 in operator norm, as $\|T_n\| = 1$ for all n.

Theorem 2.2. *Let X, Y be normed spaces with Y complete. Then $L(X, Y)$ complete. In particular, X^* is complete.*

Proof: Let (T_n) be Cauchy in $L(X, Y)$. Then, for each $x \in X$, the sequence $(T_n x)$ is Cauchy in Y, as

$$\|T_n x - T_m x\| \leq \|T_n - T_m\| \|x\|.$$

Hence $T_n x \to T(x)$ for some $T(x) \in Y$. We now need to show that $T \in L(X, Y)$ and $T_n \to T$ in $\| \cdot \|_{\mathrm{op}}$.

First we show T is linear. Note that as T_n linear,

$$T_n(\lambda x + \mu x') = \lambda T_n(x) + \mu T_n(x'),$$

and taking $n \to \infty$ we get

$$T(\lambda x + \mu x') = \lambda T(x) + \mu T(x').$$

To show T is bounded, first note that given $\varepsilon > 0$, there exists N such that $\|T_m - T_n\| < \varepsilon$ for all $m, n \geq N$. Hence

$$\|T_m x - T_n x\| \leq \varepsilon \|x\|,$$

for all $x \in X$. Letting $n \to \infty$, we get

$$\|T_m x - Tx\| \leq \varepsilon \|x\|. \tag{$*$}$$

So $\|Tx\| \leq (\|T_m\| + \varepsilon) \|x\|$, hence T is bounded.

To show that $T_n \to T$, note that $(*)$ says that $\|(T_m - T)x\| \leq \varepsilon \|x\|$ for all $m \geq N$. So $\|T_n - T\| \leq \varepsilon$ for all large enough n.

2.1   Dual of ℓ_p

Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. For each $y \in \ell_q$, we have $\phi_y \in \ell_p^*$, where

$$\phi_y(x) = \sum |x_n y_n|.$$

By Hölder, we have $\sum |x_n y_n| \leq \|x\|_p \|y\|_q$, so $\|\phi_y\| \leq \|y\|_q$.

Let's determine $\|\phi_y\|$. Let $x_n = |y_n|^{q/p} \operatorname{sgn}(y_n)$. Then $x \in \ell_p$, since

$$\sum |x_n|^p = \sum |y_n|^q \implies \|x\|_p = \left(\sum |y_n|^q\right)^{1/p} = \|y\|_q^{q/p}.$$

Then we get

$$\phi_y(x) = \sum x_n y_n = \sum |y_n|^{q/p+1} = \sum |y_n|^q = \|y\|_q^q.$$

Hence, for $y \neq 0$,

$$\|\phi_y\| \geq \frac{\|y\|_q^q}{\|y\|_q^{q/p}} = \|y\|_q.$$

Hence, $\|\phi_y\| = \|y\|_q$.

Theorem 2.3. *Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then the map $q : \ell_q \to \ell_p^*$ by $y \mapsto \phi_y$ is an isometric isomorphism.*

Proof: First, θ is linear, as $\phi_{\lambda y + \mu y'} = \lambda \phi_y + \mu \phi_{y'}$.

Then ϕ is an isometry, as $\|\phi_y\| = \|y\|_q$, hence θ is injective.

Finally, we have to show that θ is surjective. Given $T \in \ell_p^*$, define $y_n = T e_n$ for each n. We will show $y \in \ell_q$ and $T = \phi_y$. Let

$$x_n = \begin{cases} |y_n|^{q/p} \operatorname{sgn}(y_n) & n \leq N, \\ 0 & n > N, \end{cases}$$

where N is fixed. Then $x \in \ell_p$, with

$$\|x\|_p = \left(\sum_{n=1}^{N} |y_n|^q\right)^{1/p}.$$

Now we get

$$Tx = \sum_{n=1}^{N} x_n y_n = \sum_{n=1}^{N} |y_n|^{q/p+1} = \sum_{n=1}^{N} |y_n|^q.$$

So, we get

$$\sum_{n=1}^{n} |y_n|^q \le \|T\| \left(\sum_{n=1}^{N} |y_n|^q \right)^{1/p},$$

$$\implies \left(\sum_{n=1}^{N} |y_n|^q \right)^{1-1/p} \le \|T\| \implies \sum_{n=1}^{N} |y_n|^q \le \|T\|^q.$$

Hence $y \in \ell_q$.

Now to show $T = \phi_y$, note that $T = \phi_y$ on $e_1, e_2, \ldots$. Hence $T = \phi_y$ on the closed linear span of $e_1, e_2, \ldots$ as T, ϕ_y are linear and continuous.

But the closed linear span is simply ℓ_p, so $T = \phi_y$ on ℓ_p.

The same proof shows that $\ell_1^* = \ell_\infty$ and $C_0^* = \ell_1$. But it does not show that $\ell_\infty^* = \ell_1$, as the closed linear span of (e_i) is not ℓ_∞.

Interestingly, we cannot have an infinite dimensional X with $X^* = \{0\}$, by the Hahn-Banach theorem.

We now know that ℓ_p is complete, as $\ell_p = \ell_p^*$, and we just proved that the dual of a space is always complete.

3　Finite-Dimensional Normed Spaces

Theorem 3.1. *Let V be a finite dimensional vector space. Then any two norms on V are equivalent.*

Proof: Without loss of generality, $V = \mathbb{R}^n$. We will show that any norm $\|\cdot\|$ is equivalent to the usual Euclidean norm $\|\cdot\|_2$. For $x \in \mathbb{R}^n$,

$$\|x\| = \left\|\sum_{i=1}^{n} x_i e_i\right\| \leq \sum_{i=1}^{n} |x_i| \|e_i\| \leq n\|x\|_2 \max\{\|e_1\|, \ldots, \|e_n\|\}.$$

Hence $\|x\| \leq C\|x\|_2$ for some constant $C > 0$.

Now on $S = \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$, then the function $f(x) = \|x\|$ is continuous, by the above. But S is compact, so f attains its lower bound.

Hence there exists $\delta > 0$ such that $\|x\|_2 = 1 \implies \|x\| \geq \delta$.

From this fun fact we will get a lot of corollaries.

Corollary 3.1. *Let T be a linear map from X to Y with $\dim X < \infty$. Then T is continuous.*

Proof: Define a new norm on X by $\|x\|' = \|x\|_X + \|Tx\|_Y$. Then there exists $C > 0$ such that $\|x\|' \leq C\|x\|_X$ for all $x \in X$, hence

$$\|Tx\|_Y \leq (C-1)\|x\|_X.$$

Corollary 3.2. *Let X, Y be normed spaces with $\dim X = \dim Y < \infty$. Then X, Y are isomorphic.*

Proof: Let T be a linear bijection from X to Y. Then T, T^{-1} are continuous by the previous corollary.

Corollary 3.3.

(i) *Every finite dimensional normed space is complete.*

(ii) *If X is normed, Y a finite dimensional subspace of X, then Y is closed.*

Proof: Chill.

(i) This is true in $\|\cdot\|_2$, so true in any norm.

(ii) In any metric space, complete implies closed, and Y is complete.

Corollary 3.4. *Let X be a finite dimensional normed space. Then $B(X)$ is compact.*

Proof: We have $B(X)$ is bounded in $\|\cdot\|_2$, as it is bounded in $\|\cdot\|$, and $B(X)$ is closed in $\|\cdot\|_2$, as it is closed in $\|\cdot\|$.

Hence $B(X)$ is compact in $\|\cdot\|_2$, by Heine-Borel. So $B(X)$ is compact in $\|\cdot\|$, as the two norms induce the same topology.

Note that this is false in general. In ℓ_p, we have $\|e_n\| = 1$, but $\|e_n - e_m\| \geq 1$, so $B(\ell_p)$ is not sequentially compact. In fact, we will show this is true for all infinite dimensional spaces.

For a metric space X, $x \in X$ and closed $Y \subset X$, the *distance* from x to Y is

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

If $d(x, Y) = 0$, then $x \in Y$, since if $y_1, y_2, \ldots \in Y$ with $d(x, y_n) \to 0$, then $x \in \overline{Y} = Y$. Note that in general, this distance need not be attained. For example in ℓ_1, take $x = 0$ and $Y = \{(1 + \frac{1}{n})e_n \mid n \in \mathbb{N}\}$.

Proposition 3.1 (Riesz's Lemma). *Let X be a normed space and Y a proper subspace. Then*

(i) *For all $\varepsilon > 0$ there exists $x \in X$, $\|x\| = 1$ with $d(x, Y) \geq 1 - \varepsilon$.*

(ii) *If $\dim X < \infty$, then there exists $x \in X$, $\|x\| = 1$ with $d(x, Y) = 1$.*

Proof:

(i) Take $x \in X$ with $\|x\| = 1$, $x \notin Y$. Pick $y \in Y$ with $\|x - y\| \leq (1 + \varepsilon)d(x, Y)$, and let

$$z = \frac{x - y}{\|x - y\|}.$$

Then we will show $d(z, Y) \geq 1 - \varepsilon$. Indeed, for any $y' \in Y$,

$$z - y' = \frac{1}{\|x - y\|}(x - y - \|x - y\|y'),$$

where the latter element is in Y. Hence

$$\|z - y'\| \geq \frac{1}{\|x - y\|} d(x, Y) \geq \frac{1}{1 + \varepsilon}.$$

(ii) On $\{x \mid \|x\| = 1\}$, which is compact, the function $x \mapsto d(x, Y)$ is continuous, and so attains its upper bound.

Corollary 3.5. *Let X be an infinite dimensional normed space. Then $B(X)$ is not compact.*

Proof: We will find $x_1, x_2, \ldots \in X$ with $\|x_n\| = 1$ and $\|x_n - x_m\| \geq 1$ for all $n \neq m$.

Having chosen $x_1, \ldots, x_n$, put $Y = \langle x_1, \ldots, x_n \rangle$ and $X' = \langle x_1, \ldots, x_n, x \rangle$ for any $x \notin Y$.

By Riesz's lemma, there exists $x_{n+1} \in X'$ with $\|x_{n+1}\| = 1$ and $d(x_{n+1}, Y) \geq 1$. Then we are done by induction.

Now for some Analysis and Topology revision.

Let X be a metric space. Then X is compact $\iff$ X is sequentially compact $\iff$ X is complete and totally bounded.

Note that for any $A \subset X$, then A is totally bounded if and only if $\overline{A}$ is totally bounded. Hence, in a complete metric space X, a subset $A \subset X$ has

$$A \text{ compact} \iff A \text{ totally bounded}.$$

3.1   Compact Operators

For X, Y normed and $T : X \to Y$ linear, T is *compact* if $\overline{T(B_X)}$ is compact. In other words, T is small.

Remark.

1. Suppose T has finite rank, i.e. $T(X)$ is finite dimensional. Then $T(B_X)$ is bounded, and $\overline{T(B_X)} \subset T(X)$. Therefore $\overline{T(B_X)}$ is a closed and bounded subset of $T(X)$, so $\overline{T(B_X)}$ is compact. Hence every finite rank operator is always compact.

2. T compact $\implies T(B_X)$ bounded $\implies T$ continuous.

3. For X infinite dimensional, $\mathrm{id}_X : X \to X$ is not a compact operator, as B_X is not compact.

4. $T : X \to Y$ is compact $\iff$ every bounded sequence (x_n) in X has a subsequence (x_{n_i}) such that (Tx_{n_i}) is convergent in Y, as compact is equivalent to sequentially compact.

5. If Y is Banach, then $T : X \to Y$ compact $\iff$ $T(B_x)$ is totally bounded.

Proposition 3.2. *Let X, Y be normed with Y complete. Then the compact operators from X to Y form a closed subspace of $L(X,Y)$.*

Proof: Clearly scaling a compact operator gives a compact operator. Hence we want to show that if S, T are compact, then $S+T$ compact, and if $T_n \to T$ for each T_n compact, then T is compact.

(i) Given (x_n) bounded in X, we have a subsequence (x_{n_i}) such that $S(x_{n_i}) \to y$, for example. Moreover, (x_{n_i}) has a subsequence (x_{m_i}) such that $T(x_{m_i}) \to z$, say.

Then $(S + T)(x_{m_i}) = S(x_{m_i}) + T(x_{m_i}) \to y + z$.

(ii) Given $\varepsilon > 0$, we can choose n with $\|T_n - T\| < \varepsilon$. Now we have that

$$T_n(B_x) \subset \bigcup_{i=1}^{k} B(T_n x_i, \varepsilon)$$

for some $x_1, \ldots, x_k \in B_x$, as $T_n(B_x)$ is totally bounded. Hence

$$T(B_x) \subset \bigcup_{i=1}^{k} B(T_n x_i, 2\varepsilon) \subset \bigcup_{i=1}^{k} B(T x_i, 3\varepsilon).$$

Hence $T(B_x)$ is totally bounded, and so compact.

In particular, any limit of finite-rank operators is compact.

Example 3.1.

Let $X = \ell_p$, and let p_n be the projection onto the n'th coordinate, so

$$p_n(x) = (0, \ldots, 0, x_n, 0, \ldots).$$

Then p_n is of finite rank and $\|p_n\| = 1$. Let

$$p = \sum_{n=1}^{\infty} \frac{1}{n^2} p_n.$$

This is convergent as the sum is Cauchy. So

$$p(x) = \left(x_1, \frac{x_2}{4}, \frac{x_3}{9}, \frac{x_4}{16}, \dots \right).$$

Hence p is a compact operator, not of finite rank.

Proposition 3.3. *Let X, Y, Z be normed spaces, $S \in L(X < Y)$ and $T \in L(Y, Z)$.*

(i) *If S is compact, then $T \circ S$ is compact.*

(ii) *If T is compact, then $T \circ S$ is compact.*

Proof:

(i) Given (x_n) bounded in X, we have a subsequence (x_{n_i}) with $Sx_{n_i} \to y$ for some $y \in Y$. So $TSx_{n_i} \to Ty$ as T is continuous.

(ii) Given (x_n) bounded in X, then (Sx_n) is bounded in Y, as S is continuous. Hence there exists a subsequence (Sx_{n_i}) such that $TSx_{n_i} \to z$, for some $z \in Z$.

3.2   Open Mapping Lemma

We know that if $T \in L(X, Y)$ is surjective, then X Banach does not necessarily mean Y is Banach. The normal example is $T : (l_1, \| \cdot \|_1) \to (l_2, \| \cdot \|_2)$ by $x \mapsto x$.

Then the right hand space is dense in $\ell_2 \neq \ell_1$, hence the image is not complete. But we will now see that in some cases we can make this claim.

Theorem 3.2 (Open Mapping Lemma)**.** *Let $T \in L(X, Y)$, where X, Y are normed with X complete. Suppose that $B_Y \subset \overline{T(B_X)}$. Then,*

(i) $B_Y \subset T(2B_X)$.

(ii) Y *is complete.*

Proof:

(i) Given $y \in B_Y$, we seek $x \in X$ with $\|x\| \leq 2$ and $Tx = y$. As $T(B_X)$ is dense in B_Y, there exists $x_1 \in X$ with $\|x_1\| \leq 1$ and $\|y - Tx_1\| \leq \frac{1}{2}$.

Moreover, as $T(\frac{1}{2}B_X)$ is dense in $\frac{1}{2}B_Y$, there exists $x_2 \in X$ with $\|x_2\| \leq \frac{1}{2}$ and $\|y - Tx_1 - Tx_2\| \leq \frac{1}{4}$.

Continuing, we obtain $x_1, x_2, \ldots \in X$ with $\|x_n\| \leq \frac{1}{2^{n-1}}$ and $\|y - Tx_1 - \cdots - Tx_n\| \leq \frac{1}{2^n}$. Let $x = \sum x_n$, which is convergent as it is Cauchy.

Then $\|x\| \leq \sum \|x_n\| \leq 2$, and $\|y - Tx\| = 0$, i.e. $y = Tx$.

(ii) Given (y_n) Cauchy in Y, pass to a subsequence such that $\|y_n - y_{n-1}\| \leq \frac{1}{2^n}$ for all n.

Pick $x_1 \in X$, $\|x_1\| \leq 2\|y_1\|$ and $Tx_1 = y_1$. Similarly pick $x_2, x_3, \ldots \in X$ with $Tx_n = y_n - y_{n-1}$ and $\|x_n\| \leq \frac{1}{2^{n-1}}$.

Thus $T(x_1) = y_1$, $T(x_1 + x_2) = y_2$, and so on. Let $x = \sum x_n$, which converges as it is Cauchy. Hence

$$\sum_{i=1}^{k} x_i \to x \implies y_k \to Tx.$$

For X, Y topological spaces, $f : X \to Y$ is open if $f(A)$ is open in Y for every open A in X. Continuous functions need not be open: for example, take $x \mapsto x^2$. Then the image of $(-1, 1)$ is $[0, 1)$.

For X, Y normed, and $T \in L(X, Y)$, we have

$$T \text{ open} \iff T(\text{neighbourhood of } 0) \text{ is a neighbourhood of } 0$$
$$\iff T(B_x) \supset \frac{1}{k}B_y \text{ for some } k > 0$$
$$\iff \text{for all } y \in Y \text{ there is } x \in X \text{ with } \|x\| < K\|y\|, Tx = y.$$

So the open mapping lemma says:

If $T(B_x)$ dense in B_Y, then T is an open mapping.

3.3   Creating New Spaces

Given normed spaces X, Y, we can make $X \oplus Y$ into a normed space by setting

$$\|(x, y)\| = \|x\| + \|y\|,$$

called $X \oplus_1 Y$. Similarly, for any $1 \le p < \infty$ we can define $X \oplus_p Y$. These are all equivalent, and induce the usual product topology on $X \oplus Y$. This is the *product* of X and Y.

For X, Y complete, $X \oplus Y$ is complete. Moreover, X, Y are closed in $X \oplus Y$.

We can also define the *quotient*. Let X be normed and N a closed subspace of X. Then we can consider the quotient vector space X/N, with the projection map

$$\pi : X \to X/N$$
$$x \mapsto x + N.$$

Proposition 3.4. *Let X be normed and N a closed subspace of X, with $\pi : X \to X/N$ the projection map.*

Then $\|z\| = \inf\{\|x\| \mid \pi x = z\}$ is a norm on X/N. Moreover,

(i) *π is continuous.*

(ii) *If X is Banach, then X/N is Banach.*

> **Proof:**   To check this is a norm, we need to check that $\|z\| = 0 \implies z = 0$. Suppose we have $\pi(x)$ with $\|\pi(x)\| = 0$. Hence there exists $x_1, x_2, \ldots$ with $\|x_n\| \to 0$ and $\pi(x_n) = \pi(x)$.
>
> Then $x - x_n \in N$ for all n. So as $x_n \to 0 \in N$, and N is closed, $x \in N$. So $\pi(x) = 0$.
>
> As $\|\pi(x)\| \le \|x\|$, π is continuous.
>
> Now to show that X/N is Banach if X is Banach, wuppose we have $z \in X/N$, with $\|z\| < 1$. Then there exists $x \in X$ with $\|x\| < 1$, $\pi(x) = z$. Therefore, $\pi(B_X)$ is dense in $B_{X/N}$. Hence X/N is complete by open mapping lemma.

If we have $T \in L(X, Y)$ and $T|_N = 0$, then T induces a linear $\tilde{T} : X/N \to Y$. In fact, $\tilde{T}$ is continuous.

But, we do not need to have $X/\operatorname{Ker} T$ isomorphic to $\operatorname{Im} T$. Take the identity $\ell_1 \to \ell_2$. This has $X/\ker T$ being ℓ_1, but $\operatorname{Im} T$ is not complete.

4   Baire Category Theorem and Applications

The moral of this section is that complete spaces are very nice.

We begin by looking at dense open sets. Some trivial examples in $\mathbb{R}$ are $\mathbb{R} \setminus \{0\}$ and $\mathbb{R} \setminus \{\frac{1}{1}, \frac{1}{2}, \ldots\} \cup \{0\}$.

Theorem 4.1. *Let X be a (non-empty) complete metric space, and let $O_1, O_2, \ldots$ be dense open sets. Then,*

$$\bigcap_{n=1}^{\infty} O_n \neq \emptyset.$$

Proof: Let $B(x, \varepsilon)$ be the open ball with radius ε centred at x, and $\overline{B}(x, \varepsilon)$ the closed ball.

Since O_1 isn't empty, there is a closed ball $\overline{B}(x_1, \varepsilon + 1) \subset O_1$ for some $x_1 \in X$ and $\varepsilon_1 < 1$. Now as O_2 is dense, it meets $B(x_1, \varepsilon_1)$, so there exists

$$\overline{B}(x_2, \varepsilon_2) \subset \overline{B_1}(x_1, \varepsilon_1) \cap O_2,$$

with $x_2 \in O_2, \varepsilon_2 < \frac{1}{2}$. Continuing in this manner, we obtain

$$\overline{B}(x_1, \varepsilon_1) \supset \cdots \supset \overline{B}(x_n, \varepsilon_n) \supset \cdots$$

with $\varepsilon_n \to 0$ and $\overline{B}(x_n, \varepsilon_n) \subset O_n$. Note that (x_n) is Cauchy, as if $n, m > N$ then $x_m, x_n \in \overline{B}(x_N, \varepsilon_N)$ so $d(x_n, x_m) \leq 2\varepsilon_N$. Hence say $x_n \to x$.

Then for all n, $x \in \overline{B}(x_n, \varepsilon_n)$, as $x_m \in \overline{B}(x_n, \varepsilon_n)$ for all $m \geq n$, and $\overline{B}(x_m, \varepsilon_m)$ is closed. In particular, $x \in O_n$ for all n.

Remark. The same proof shows that $\bigcap O_n$ is dense.

Note that we need X is complete: otherwise enumerate $\mathbb{Q} = \{q_1, q_2, \ldots\}$ and let $O_n = \mathbb{Q} \setminus \{q_n\}$.

For an open, dense set O, we have that O^c is closed and contains no open ball, i.e. it has an empty interior. Therefore, we get the following dual of our theorem:

Theorem 4.2. *Let $X \neq \emptyset$ be a complete metric space, and $F_1, F_2, \ldots$ be closed in X. Then if $\bigcup F_i = X$, then some F_n has an interior.*

This is highly non-obvious, even for $X = [0, 1]$.

A set $A \subset X$ is *nowhere dense* if it is not dense on any open ball, i.e. $\overline{A}$ has an empty interior. Rewriting the above theorem, we get the following.

Theorem 4.3. *Let $X \neq \emptyset$ be a complete metric space, and $F_1, F_2, \ldots$ closed and nowhere-dense. Then $\bigcup F_i \neq X$.*

Say A is *meagre* if A is a countable union of nowhere dense sets.

Hence the above theorem says, if X is complete, then X is not meagre in X. We can think of meagre sets as small sets.

Remark.

1. Meagre sets can be dense, for example $\mathbb{Q}$ in $\mathbb{R}$.

2. Nowhere dense sets can be uncountable, for example in $\mathbb{R}$ we have the Cantor set
$$C = \left\{ \sum_{n=1}^{n} a_n 3^{-n} \mid a_n \in \{0, 2\} \right\}.$$

3. Meagre is a bit like measure 0, but they are not the same.

4. Sometimes, being meagre is called being 'of first category', and being non-meagre is called being 'of second category'.

Let $f_1, f_2, \ldots \in C[0, 1]$. Suppose the f_n are pointwise bounded. Must they be uniformly bounded? The answer is no, by taking suitable spikes.

However, we can prove the following:

Theorem 4.4 (Osgood's Theorem). *Let $f_1, f_2, \ldots \in C[0, 1]$ be pointwise bounded. Then there exists an interval (a, b) on which they are uniformly bounded.*

> **Proof:**   For $n = 1, 2, \ldots$ let $t_n = \{x \in [0, 1] \mid |f_i(x)| \leq n \forall i\}$.
>
> Then $\bigcup t_n = [0, 1]$, as the f_i are pointwise bounded, and also each t_n is closed, as it is an intersection of closed sets. Hence by Baire, we have $t_n \supset (a, b)$.

This is a typical 'anti-pathology' application of Baire.

Theorem 4.5 (Principle of Uniform Boundedness). *Let X be Banach, Y be a normed space and let $A \subset L(X, Y)$ be pointwise bounded, i.e. for all $x \in X$, there is k such that $\|T(x)\| \leq k$ for all $T \in A$.*

Then A is bounded, i.e. there exists k such that $\|T\| \leq k$ for all $T \in A$.

> **Proof:**   For $n = 1, 2, \ldots$ let $t_n = \{x \in X \mid \|T(x)\| \leq n \forall T \in A)$. Then $\bigcup t_n = X$, and each t_n is closed.

So by Baire, we have $t_n \supset B(x, \varepsilon)$ for some n and some $x \in X, \varepsilon > 0$.

Then for all $y \in X$ with $\|y\| < \varepsilon$, we have $y = (x + y) - x$, with $x + y, x \in B(x, \varepsilon)$.

Hence $\|Ty\| \leq 2n$ for all $T \in A$, and hence $\|T\| \leq \frac{2n}{\varepsilon}$ for all $T \in A$.

Some motivation: we know that if $T_n \to T$ pointwise, then $T_n \nrightarrow T$ in norm: take $T_n : \ell_1 \to \mathbb{R}$ be given by 'taking the n'th component'. Also, we know that if $f_n \to f$ pointwise with f_n continuous, we don't necessarily have f continuous. But with completeness, we can say the following.

Corollary 4.1 (Banach-Steinhaus Theorem)**.** *Let X be Banach, and Y complete. Suppose $T_1, T_2, \ldots \in L(X, Y)$ have $(T_n x)$ convergent for all x, and say $T_n x \to Tx$. Then T is a continuous linear map.*

Proof: First, it is easy to show T is linear, as Y is complete so

$$T(\lambda x + \mu y) = \lim_{n \to \infty} T_n(\lambda x + \mu y) = \lim_{n \to \infty} \lambda T_n(x) + \mu T_n(y) = \lambda T(x) + \mu T(y).$$

Now to prove continuity, we notice that $(T_n x)$ is convergent, hence bounded for all x. Therefore, by the principle of uniform boundedness, the T_n are uniformly bounded, say by k. So $\|T_n x\| \leq \|T_n\| \|x\| \leq k\|x\|$.

Letting $n \to \infty$, we get $\|Tx\| \leq k\|x\|$, so T is continuous.

Recall the open mapping lemma, that if $T(B_X) \supset B_Y$ then T is an open mapping. Let's prove a slightly stronger result.

Theorem 4.6 (Open Mapping Theorem)**.** *Let X, Y be Banach and $T \in L(X, Y)$ be surjective. Then T is an open mapping.*

Proof: For $n = 1, 2, \ldots$, let $E_n = \overline{T(nB_X)}$. Then each E_n is closed, and $\bigcup E_n = Y$ as T is surjective.

So some $E_n \supset B(y, \varepsilon)$ for some $y \in Y, \varepsilon > 0$. By scaling, we can without loss of generality assume $n = 1$, so $\overline{T(B_X)} \supset B(y, \delta)$.

It follows that $\overline{T(B_X)} \supset B(0, \delta)$, from convexity. Indeed, given $z \in B(0, \delta)$ and $\varepsilon > 0$, $y + z, y_z \in B(y, \delta)$. So there exists $x, x' \in X$ with $\|x\|, \|x'\| \leq 1$ with $\|Tx - (y + z)\| < \varepsilon$, $\|Tx' - (y - z)\| < \varepsilon$.

Then $T(\frac{x-x'}{2})$ satisfies $\|T(\frac{x-x'}{2}) - z\| < \varepsilon$. Hence T is an open mapping, by open mapping lemma applied to $\frac{T}{\delta}$.

Corollary 4.2 (Inversion Theorem). *Let X, Y be Banach, and $T \in L(X, Y)$ bijective. Then T is an isomorphism.*

Proof: By open mapping theorem, we know T is an open mapping. That is, for every $y \in Y$, we have $\|T^{-1}y\| \leq k\|y\|$ for some fixed $k > 0$.

Hence T^{-1} is continuous.

Remark. If X, Y are Banach and $T \in L(X, Y)$ is surjective, then we do have

$$X/\ker T \cong Y,$$

because the induced mapping is a continuous linear bijection, hence by inversion theorem it is an isomorphism.

Corollary 4.3 (Comparison Theorem). *Let $\| \cdot \|_1$ and $\| \cdot \|_2$ be complete norms on vector space V. Suppose there exists $c > 0$ such that*

$$\|x\|_2 \leq c\|x\|_1,$$

for all $x \in V$. Then $\| \cdot \|_1$ and $\| \cdot \|_2$ are equivalent.

Proof: Consider the identity $\mathrm{id} : (V, \| \cdot \|_1) \to (V, \| \cdot \|_2)$. Then id is a linear bijection and continuous, so T^{-1} is also continuous.

Hence there exists $d > 0$ such that $\|x\|_1 \leq d\|x\|_2$, for all $x \in V$.

In essence, comparable norms are equivalent.

Remark. This is one way to prove the integral norm on $C[0, 1]$ is incomplete.

Let X, Y be metric spaces, and $f : X \to Y$. The *graph* of f is

$$\mathrm{Graph}(f) = \{(x, y) \in X \times Y \mid y = f(x)\}.$$

If f is continuous, then certainly $\mathrm{Graph}(f)$ is closed. Indeed, suppose $(x_n, f(x_n)) \to (x, y)$ for some $x, y \in X \times Y$.

Then $x_n \to x$ and $f(x_n) \to y$. But f is continuous, so $f(x_n) \to f(x)$, and thus $y = f(x)$, i.e. $(x, y) \in \mathrm{Graph}(f)$.

The converse is not true, by taking $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} \frac{1}{x} & x \neq 0, \\ 0 & x = 0. \end{cases}$$

But, in Banach spaces this is true.

Theorem 4.7 (Closed Graph Theorem). *Let X, Y be Banach and $T : X \to Y$ be linear. Then T is continuous if and only if T has closed graph.*

Proof: The forwards direction is already done above.

For the other direction, give $X \times Y$ the 1-norm, so $\|(x, y)\| = \|x\|_X + \|y\|_Y$. Then $X \times Y$ is Banach, and $\text{Graph}(T)$ is closed, so $\text{Graph}(T)$ is Banach.

Now we have an elementary embedding $S : X \to \text{Graph}(T)$ by $x \mapsto (x, Tx)$, and we want to prove that S is continuous.

Now S is linear, bijective and S^{-1} is continuous. Then by the inversion theorem, S is continuous, and hence T is continuous.

Why is the closed graph theorem interesting? It says that for linear $T : X \to Y$,

$$\begin{aligned} T \text{ has closed graph} \iff & (x_n, Tx_n) \to (x, y) \implies y = Tx \\ \iff & x_n \to x, Tx_n \to y \implies y = Tx \\ \iff & x_n \to 0, Tx_n \to y \implies y = 0. \end{aligned}$$

So usually, to show that if T is continuous, we show that $x_n \to 0 \implies Tx_n \to 0$.

However, this theorem says that it is enough to show $x_n \to 0$, $Tx_n \to y$ then $y = 0$.

Example 4.1.

Suppose X is Banach, and Y, Z closed subspaces with $X = Y \oplus Z$ as vector spaces. Then the projections from X to Y and Z are continuous.

Using closed graph theorem, it is enough to show the following: if $y_n + z_n \to 0$ and $y_n \to y$, then $y = 0$.

Since $y_n \to y$, we must have $z_n \to -y$. But $y \in Y$ as Y is closed, and $-y$ belongs to Z as Z is closed. So $y = 0$.

Let's look at a typical counterexample application of Baire: finding a continuous nowhere-differentiable function.

Example 4.2.

Suppose $f \in C[0,1]$ is differentiable at x. Then $\frac{f(x+t)-f(x)}{t}$ approaches a limit as $t \to 0$. In particular, it is bounded over all $t \neq 0$, $x + t \in [0,1]$. Say that $|\frac{f(x+t)-f(x)}{t}| \leq n$ for all $t \neq 0$, $x + t \in [0,1]$. Then we can define

$$E_n = \left\{ f \in C[0,1] \,\middle|\, \exists x \in [0,1], \left| \frac{f(x+t)-f(x)}{t} \right| \leq n \; \forall t, t \neq 0, x+t \in [0,1] \right\}.$$

Now f is differentiable at some x implies $f \in E_n$ for some n. We will now show that each E_n is closed, and nowhere dense. Then we are done, as $\bigcup E_n \neq C[0,1]$ by Baire. In fact, the differentiable functions are meagre.

To show that E_n is closed, suppose $f_1, f_2, \ldots \in E_n$, with $f_i \to f$ uniformly. Then for each i, we have $x_i \in [0,1]$ with

$$\left| \frac{f_i(x_i + t) - f_i(x_i)}{t} \right| \leq n,$$

for all sensible t. Passing to a subsequence, we can let $x_i \to x$ for some $x \in [0,1]$. Now, for each t, we have $f_i(x_i + t) - f_i(x_i) \to f(x + t) - f(x)$. Hence we get

$$\left| \frac{f(x+t)-f(x)}{t} \right| \leq n,$$

for all $t \neq 0$ with $x + t \in (0,1)$. So $f \in E_n$.

Now we will show that E_n is nowhere dense. In fact, for $f \in C[0,1]$ we will show there exists g with $\|g\| < 4\varepsilon$ with $f + g \notin E_n$. As f is uniformly continuous, there exists δ such that $|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$.

We will now look at $f + g$, where g is spiky with spikes of height 4ε and width w, where $w < \delta$ and $w < \varepsilon/n$. So for all x, there exists y such that $|x - y| < \delta$, with $|g(x) - g(y)| \geq 2\varepsilon$, hence

$$|(f+g)(x) - (f+g)(y)| > \varepsilon \implies \left| \frac{(f+g)(y) - (f+g)(x)}{y - x} \right| \geq n.$$

Or directly, we can add a lot of small spiky functions to construct an explicit example.

5   Spaces of Continuous Functions

In this section we will look at studying $C(K)$, where K is a compact, Hausdorff space. These are not all given by a metric.

We will look at the following:

1. The existence of continuous and non-constant functions.

2. The subsets of $C(K)$ that are compact.

3. The subspaces of $C(K)$ are dense.

$C[0,1]$ are good examples to thing of for the last two subsections.

5.1   Existence of Continuous Functions

If X is metric, then we have infinitely many continuous functions to $\mathbb{R}$: given by $x \mapsto d(x, x_0)$ for x_0 fixed. Or perhaps we can take $x \mapsto d(x, A)$ where $A \subset X$ is closed.

What if more generally our space is compact and Hausdorff? Let us say that a Hausdorff space X is *normal* if any two disjoint closed sets can be separated by open sets, that is, for all A, B closed with $A \cap B = \emptyset$, then there exists open $U \supset A$ and $V \supset B$ with $U \cap V = \emptyset$.

For example, any metric space is normal: let

$$U = \{x \mid d(x, A) < d(x, B)\}, \qquad V = \{x \mid d(x, B) < d(x, A)\}.$$

Proposition 5.1. *Let K be compact and Hausdorff. Then K is normal.*

> **Proof:**   Given disjoint, closed $A, B \subset X$, for all $a \in A$ and $b \in B$ there exist open U_{ab} and V_{ab} disjoint with $a \in U_{ab}$, $b \in V_{ab}$.
>
> For fixed $a \in A$, the sets V_{ab} cover B. So there exists a finite subcover (as B is compact), say $V_a = V_{ab_1} \cup \cdots \cup V_{ab_n}$.
>
> Let $U_a = U_{ab_1} \cap \cdots \cap U_{ab_n}$, and V_a as above. Then U_a, V_a are open and disjoint, with $a \in U_a$ and $B \subset V_a$.
>
> Now, the U_a cover A, so there exists a finite subcover $U = U_{a_1} \cup \cdots \cup U_{a_k}$. Let $V = V_{a_1} \cap \cdots \cap V_{a_k}$, then U, V are open, disjoint and $A \subset U$, $B \subset V$.

Note that the space X is normal if and only if, for any closed A and open $V \supset A$, there exists an open set U with $A \subset U$, $\overline{U} \subset V$.

We'd like to construct a continuous function as follows: pick A, B disjoint and closed sets in X. Then we want to find a continuous function $f : X \to \mathbb{R}$ such that $f|_A = 0$ and $f|_B = 1$.

This is a stronger separation property than normality. Given such an f, let

$$U = \{x \mid f(x) < 1/2\}, \qquad V = \{x \mid f(x) > 1/2\}.$$

This is certainly true if X is metric: let

$$f(x) = \frac{d(x, A)}{d(x, A) + d(x, B)}.$$

Theorem 5.1 (Urysohn's Lemma)**.** *Let X be normal and A, B disjoint, closed subsets of X. Then there exists a continuous $f : X \to \mathbb{R}$ such that $f|_A = 0$, $f|_B = 1$.*

> **Proof:** Let $V = B^c$. We know that A is closed, and $A \subset V$ is open. So there exists an open $O_{1/2}$ with $A \subset O_{1/2}$ and $\overline{O}_{1/2} \subset V$.
>
> Now as $\overline{O}_{1/2} \subset V$ is closed, there exists open $O_{3/4}$ with $\overline{O}_{1/2} \subset O_{3/4}$ and $\overline{O}_{3/4} \subset V$.
>
> Moreover, since $A \subset O_{1/2}$, there exists open $O_{1/4}$ with $A \subset O_{1/4}$ and $\overline{O}_{1/4} \subset O_{1/2}$.
>
> We can keep going like this: we obtain open sets O_q for each $q \in (0, 1)$ dyadic. Then $\overline{O}_q \subset r$ for all $q < r$, and $A \subset O_q$, $\overline{O}_q \subset V$ for all q. Also let $O_1 = X$.
>
> Define $f : X \to \mathbb{R}$ by
> $$f(x) = \inf\{q \mid x \in O_q\}.$$
>
> Then $f = 0$ on A and $f = 1$ on B. To show f is continuous, it is enough to show that $\{x \mid f(x) < a\}$ and $\{x \mid f(x) > a\}$ are open for all a.
>
> This is because the inverse image of any (a, b) is open, as it is the intersection of two open sets. Hence the inverse image of any open set is open, as any open set is the union of intervals.
>
> First, let us look at $\{x \mid f(x) < a\}$. Note that
>
> $$f(x) < a \iff x \in O_q \text{ for some } q < a$$
> $$\iff x \in \bigcup_{q < a} O_q,$$

which is open as it is a union of open sets. Now for the other part, $\{x \mid f(x) > a\}$. Note that

$$
\begin{aligned}
f(x) > a &\iff \exists q > a \text{ with } x \notin O_q \\
&\iff \exists r > a \text{ with } x \notin \overline{O}_r \\
&\iff x \in \bigcup_{r>a} \overline{O}_r^{\,c},
\end{aligned}
$$

which is again the union of open sets.

Corollary 5.1. *Let K be an infinite compact Hausdorff space. Then $C(K)$ is infinite dimensional.*

Proof: Take distinct $x_1, x_2, \ldots \in K$. For each n, we have $(x_1, \ldots, x_n)$ and x_{n+1} are disjoint closed sets. So there exists $f_n \in C(K)$ with $f_n(x_i) = 0$ for all $1 \le i \le n$, and $f_n(x_{n+1}) = 1$.

Then $f_1, f_2, \ldots$ are linearly independent.

Theorem 5.2 (Tietze Extension Theorem)**.** *Let X be normal, $Y \subset X$ closed and $f : Y \to \mathbb{R}$ a bounded continuous function.*

Then f extends to a continuous $g : X \to \mathbb{R}$ (i.e. $g|_Y = f$), with $\|g\|_\infty = \|f\|_\infty$.

Remark.

1. The condition $\|g\|_\infty = \|f\|_\infty$ is trivial: if g extends f (with $\|f\|_\infty = 1$), then so does g', where

$$
g'(x) = \begin{cases} 1 & g(x) \ge 1, \\ -1 & g(x) \le -1, \\ g(x) & \text{otherwise.} \end{cases}
$$

2. Tietze extends Urysohn, as Urysohn is the case $Y = A \cup B$ with $f : Y \to \mathbb{R}$ being 0 on A, and 1 on B.

3. We do need Y closed, as otherwise take $X = [0,1]$, $Y = (0,1]$ and $f(x) = \sin(1/x)$.

4. Tietze is not obvious when X is metric.

Remark.

1. The proof shows that S is totally bounded $\iff$ S bounded, and equiv.???

2. This has the same proof for $C_{\mathbb{C}}(K)$.

Ascoli-Arzelá is often useful in showing that operators to $C(K)$ are compact.

Example 5.1.

We will consider the integral operator.

Let $g \in C([0,1]^2)$ be fixed. Then given $f \in C[0,1]$, we can define

$$T(f)(x) = \int_0^1 g(x,t)f(t)\,\mathrm{d}t.$$

Now $T(f)$ is continuous. Indeed, g is uniformly continuous. So given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|x - x'| < \delta \implies |g(x,t) - g(x',t)| < \varepsilon,$$

for all t. Hence,

$$|T(f)(x) - T(f)(x')| \leq \int_0^1 |g(x,t) - g(x',t)||f(t)\,\mathrm{d}t \leq \varepsilon\|f\|, \qquad (*)$$

Therefore the mapping $T : C[0,1] \to C[0,1]$ is linear, and is continuous. Moreover,

$$|T(f)(x)| \leq \|g\|\|f\|,$$

by the definition, so $\|T(f)\| \leq \|g\|\|f\|$, hence T is continuous, and $\|T\| \leq \|g\|$.

We call T the *integral operator with kernel g*.

We claim that T is a compact operator.

Indeed, we need the image of the unit ball to be equicontinuous (as it is bounded, so we immediately get that it is totally bounded).

But by $(*)$, given f with $\|f\| < 1$ and $\varepsilon > 0$, we get that

$$|x - x'| < \delta \implies |T(f)(x) - T(f)(x')| \leq \varepsilon,$$

hence T is equicontinuous at x.

In fact, we can even show that $T : (C[0,1], \|\cdot\|_1) \to (C[0,1], \|\cdot\|_\infty)$ is compact.

5.2   Density

A *subalgebra* A of $C(K)$ is a subspace such that if $f, g \in A$, then $f \cdot g \in A$.

> **Example 5.2.**
>
> 1. In $C[0,1]$, $\{f \mid f \text{ is a poly}\}$ is a subalgebra. We will later see that this is dense.
>
> 2. In $C(K)$, $\{f \mid f(x) = 0 \text{ for some fixed } x\}$ is a subalgebra which is not dense.
>
> 3. In $C(K)$, $\{f \mid f(x) = f(y) \text{ for some fixed } x, y \in K\}$ is a subalgebra which is again not dense.

We aim to show that under some very mild conditions, subalgebra are dense.

For $f, g \in C(K)$, we write $f \vee g$, which is "f max g" or "f join g", for the function

$$f \vee g : K \to \mathbb{R},$$
$$x \mapsto \max(f(x), g(x)).$$

We also define $f \wedge g$, which is "f min g" or "f meet g" for the function

$$f \wedge g : K \to \mathbb{R},$$
$$x \mapsto \min(f(x), g(x)).$$

A *sublattice* of $C(K)$ is a subset $A \subset C(K)$, which is closed under the lattice operations:

$$f, g \in A \implies f \vee g, f \wedge g \in A.$$

> **Example 5.3.**
>
> In $C[0,1]$,
>
> 1. $\{f \mid f(x) \geq 0 \text{ for all } x\}$ is a sublattice.
>
> 2. $\{f \mid \|f\| \leq 1\}$ is a sublattice.
>
> 3. $\{f \mid f(x) = 0 \text{ for fixed } x\}$ is a sublattice.
>
> 4. $\{f \mid f \text{ is a polynomial}\}$ is a subalgebra, but not a sublattice.

Lemma 5.1. *Let K be compact and Hausdorff, and A a sublattice of $C(K)$ with $f \in C(K)$. Suppose A can approximate f at any two points, so for all $x, y \in K$ there exists $g \in A$ such that*

$$|g(x) - f(x)|, |g(y) - f(y)| < \varepsilon,$$

for any $\varepsilon > 0$.

Then A approximates f uniformly.

Proof: Given $x, y \in K$ and $\varepsilon > 0$, there exists $g_{xy} \in A$ such that

$$|g_{xy}(x) - f(x)|, |g_{xy}(y) - f(y)| < \varepsilon.$$

Write U_{xy} for
$$\{z \in K \mid |g_{xy}(z) - f(z)| < \varepsilon\}.$$

This is an open set containing x and y.

For x fixed, the sets U_{xy} cover K, so as K is compact,

$$K = U_{xy_1} \cup \cdots \cup U_{xy_n},$$

for some $y_1, \ldots, y_n$. Then let

$$g_x = g_{xy_1} \wedge \cdots \wedge g_{xy_n}.$$

Then $|g_x(x) - f(x)| < \varepsilon$, and for all $y \in K$ we have $g_x(y) < f(y) + \varepsilon$.

Now define U_x as
$$\{z \mid |g_x(z) - f(z)| < \varepsilon\},$$

which is an open set containing x. Again, as K is compact, we can write

$$K = U_{x_1} \cup \cdots \cup U_{x_m},$$

for some $x_1, \ldots, x_m \in K$. Putting

$$g = g_{x_1} \vee \cdots \vee g_{x_n},$$

we have for all $y \in K$, $g(y) < f(y) + \varepsilon$ as $g_x(y) < f(y) + \varepsilon$, and also $g(y) > f(y) - \varepsilon$ as $y \in U_{x_i}$ for some i.

Hence g approximates f.

Lemma 5.2. *The function $|x|$ on $[-1, 1]$ can be uniformly approximated by polynomials.*

Proof: Given $\varepsilon > 0$, it is sufficient to uniformly approximate the function

$(x^2 + \varepsilon)^{1/2}$, as

$$\big||x| - (x^2 + \varepsilon)^{1/2}\big| = \frac{\varepsilon}{|x| + (x^2 + \varepsilon)^{1/2}} < \frac{\varepsilon}{\varepsilon^{1/2}} = \varepsilon^{1/2}.$$

Now, the function $(x^2 + \varepsilon)^{1/2}$ has a Taylor series about $x = 1/2$, which uniformly converges on $[0, 1]$. This is because $(z + \varepsilon)^{1/2}$ is analytic in $\Re z > -\varepsilon$.

Hence there exists N such that

$$\left|(x + \varepsilon)^{1/2} - \sum_{i=1}^{N} a_i \left(x - \frac{1}{2}\right)^n\right| < \varepsilon,$$

for all $x \in [0, 1]$. Hence

$$\left|(x^2 + \varepsilon)^{1/2} - \sum_{i=1}^{N} a_i \left(x^2 - \frac{1}{2}\right)^n\right| < \varepsilon,$$

for all $x \in [-1, 1]$.

Note that we can assume this polynomial approximation has no constant term. Suppose $\big||x| - P(x)\big| < \varepsilon$. Then $|P(0)| < \varepsilon$, so putting $Q(x) = P(x) - P(0)$, we get

$$\big||x| - Q(x)\big| < 2\varepsilon.$$

Theorem 5.3. *Let K be compact, Hausdorff and $A \subset C(K)$ a closed subalgebra. Then A is a sublattice.*

Proof: It is sufficient to show that if $f \in A$, then $|f| \in A$. Indeed, note

$$(f \vee g)(x) = \frac{1}{2}(f(x) + g(x) + |f(x) - g(x)|),$$

$$(f \wedge g)(x) = \frac{1}{2}(f(x) + g(x) - |f(x) - g(x)|).$$

Given $f \in A$ with $\|f\|_\infty \leq 1$ and $\varepsilon > 0$, there exists a polynomial P with $P(0) = 0$, and

$$\big||t| - P(t)\big| < \varepsilon,$$

for all $t \in [-1, 1]$. So consider the function $P(f)$. Then certainly $P(f) \in A$ as P has no constant term. Also, for all $x \in X$, we have $f(x) \in [-1, 1]$, so

$$\big||f(x)| - P(f)(x)\big| < \varepsilon,$$

hence $\|\,|f| - P(f)\,\|_\infty \leq \varepsilon$. As A is closed, we get $|f| \in A$.

Theorem 5.4 (Stone-Weierstrass Theorem). *Let K be compact, Hausdorff and $A \subset C(K)$ a subalgebra. Suppose that:*

(i) *A contains the constants.*

(ii) *A separates the points of K, i.e. for all $x, y \in K$ distinct, there exists $f \in A$ such that $f(x) \neq f(y)$.*

Then A is dense in $C(K)$.

Proof: Consider $\overline{A}$. Then this is a closed subalgebra, hence $\overline{A}$ is a sublattice.

Also for any distinct $x, y \in K$ and $s, t \in \mathbb{R}$, there exists $g \in A$ such that $g(x) = s$, $g(y) = t$, by the given conditions.

In particular, we can approximate (exactly) any function $f \in C(K)$ at x, y. Hence for all $f \in C(K)$, we can approximate f within ε on K by some $g \in \overline{A}$.

But this means that $f \in \overline{(\overline{A})} = \overline{A}$, so $\overline{A} = K$ and A is dense.

Remark. We could replace the given conditions by: for all distinct $x, y \in K$, there exists $f \in A$ such that $f(x), f(y) \neq 0$ and $f(x) \neq f(y)$. Then we can look at f, f^n.

From Stone-Weierstrass, we get the following:

1. Polynomials are dense in $C[0, 1]$.

2. For compact $K \subset \mathbb{R}^n$, the polynomials in n variables are dense.

3. Hence for all $f \in C([0, 1]^2)$, we have

$$\int_0^1 \int_0^1 f(x, y) \, \mathrm{d}x \, \mathrm{d}y = \int_0^1 \int_0^1 f(x, y) \, \mathrm{d}y \, \mathrm{d}x.$$

   This is because these are linear functionals, which agree on a dense set (the polynomials), hence must be equal on the entire set.

4. For K, L compact, functions of the form

$$(x, y) \mapsto \sum_{i=1}^n f_i(x) g_i(y),$$

   where $f_i \in C(K), g_i \in C(L)$, are dense in $C(K \times L)$.

5. Now for measure theory: as the polynomials are dense in $(C[0,1], \| \cdot \|_\infty)$, they are dense in $(C[0,1], \| \cdot \|_p)$. Therefore as the continuous functions are dense in $L_p[0,1]$, we get that the polynomials are dense in $L_p[0,1]$.

Note that Stone-Weierstrass is false in $C_\mathbb{C}(K)$. For example, let $\Delta = \{z \mid \|z\| \leq 1\} \subset \mathbb{C}$, and
$$A(\Delta) = \{f \in C_\mathbb{C}(K) \mid f \text{ analytic on } \Delta\}.$$

Then $A(\Delta)$ contains constants, separates points, is a subalgebra and is closed. However it is not equal to $C_\mathbb{C}(\Delta)$ as there are some non-analytic functions.

Theorem 5.5 (Complex Stone-Weierstrass)**.** *Let K be compact Hausdorff, and $A \subset C_\mathbb{C}(K)$ a subalgebra. Suppose that:*

(i) *A contains the constants.*

(ii) *A separates the points of K.*

(iii) *If $f \in A$, then $\bar{f} \in A$.*

Then A is dense in $C_\mathbb{C}(K)$.

> **Proof:**   Let $B = \{f \in A \mid f(x) \in \mathbb{R} \text{ for all } x \in K\}$. Note that $f \in A \implies \Re f, \Im f \in A$, as
> $$\Re(z) = \frac{z + \bar{z}}{2}, \qquad \Im(z) = \frac{z - \bar{z}}{2}.$$
> Hence B is a subalgebra of $C(K)$ that contains the constants and is separates K. So $B \subset C(K)$ has $\overline{B} = C(K)$ by Stone-Weierstrass.
>
> Hence for all $f \in C_\mathbb{C}(K)$, we have $\Re f, \Im f \in \overline{B}$, so certainly in $\bar{A}$. Thus
> $$f = \Re f + i \Im f \in \overline{A}.$$

> **Example 5.4.**
>
> Let $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$. The *trigonometric polynomials* are
> $$\sum_{n=-N}^{N} a_n e^{-in\theta},$$
> for $a_n \in \mathbb{C}$ and $N \in \mathbb{N}$. These are dense in $C_\mathbb{C}(T)$.
>
> Equivalently, on the space $\{f \in C_\mathbb{C}[0, 2\pi]\}$ with $f(0) = f(2\pi)$, every f may

be uniformly approximated by functions of the form

$$\theta \mapsto \sum_{n=-N}^{N} a_n e^{in\theta}.$$

Fourier series lol.

6   Hilbert Spaces

Let X be a real or complex vector space. An *inner product* on X is a function $(\cdot, \cdot)$ from X^2 to the scalars $\mathbb{R}$ (or $\mathbb{C}$) such that:

(i) For all $x, y, z \in X$ and scalars λ, μ, $(\lambda x + \mu y, z) = \lambda(x, z) + \mu(y, z)$.

(ii) For all $x, y \in X$, $(x, y) = \overline{(y, x)}$.

(iii) For all $x \in X$, $(x, x) \geq 0$ with equality if and only if $x = 0$.

Example 6.1.

1. The usual Euclidean spaces: $\mathbb{R}^n$ or $\mathbb{C}^n$ with

$$(x, y) = \sum_{i=1}^{n} x_i \overline{y_i}, \qquad \|x\| = \left(\sum |x_i|^2 \right)^{1/2}.$$

2. On ℓ_2, we can define a similar thing:

$$(x, y) = \sum_{i=1}^{\infty} x_i \overline{y_i}, \qquad \|x\| = \left(\sum |x_i|^2 \right)^{1/2}.$$

3. On $C[0, 1]$, we have

$$(f, g) = \int_0^1 f(x)\overline{g(x)} \, \mathrm{d}x, \qquad \|f\| = \left(\int_0^1 |f(x)|^2 \, \mathrm{d}x \right)^{1/2}.$$

An *inner product space* is a vector space X with an inner product $(\cdot, \cdot)$.

Proposition 6.1. *Let X be an inner product space.*

(i) *Cauchy-Schwarz: for all $x, y \in X$, $\|(x, y)\| \leq \|x\|\|y\|$.*

(ii) *Triangle inequality: for all $x, y \in X$, $\|x + y\| \leq \|x\| + \|y\|$.*

Proof:

(i) Without loss of generality, let (x, y) be real (by replacing x with $xe^{i\theta}$ for suitable θ). Then, for all $t \in \mathbb{R}$,

$$\|x + ty\|^2 = (x + ty, x + ty) = \|x\|^2 + t^2\|y\|^2 + 2t(x, y).$$

For $y \neq 0$, let
$$t = \frac{-(x,y)}{\|y\|^2},$$
then we get
$$\|x\|^2 + \frac{(x,y)^2}{\|y\|^2} - 2\frac{(x,y)^2}{\|y\|^2} \geq 0,$$
which rearranges to
$$(x,y)^2 \leq \|x\|^2 \|y\|^2.$$

(ii) $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\Re(x,y) \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2.$

Hence we can talk about norms, continuity, convergence and everything we did in normed vector spaces.

Note that from Cauchy-Schwarz, the inner product is continuous, that is if $x_n \to x$ and $y_n \to y$, then $(x_n, y_n) \to (x,y)$. Indeed,

$$(x,y) - (x_n, y_n) = (x, y - y_n) + (x - x_n, y_n) \to 0.$$

A *Hilbert space* is a complete inner product space.

For example, $\mathbb{R}^n$ and $\mathbb{C}^n$ are Hilbert spaces, but $C[0,1]$ is not.

Proposition 6.2 (Polarization identity)**.** *Let X be an inner product space, and $x, y \in X$. Then:*

- *If X is real, $(x,y) = \frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2)$.*
- *If X is complex, $(x,y) = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2)$.*

Proof: Basically just expand everything.

Therefore, the norm determines the inner product space. Hence we can talk about a normed space "being an inner product space", meaning that there exists an inner product that gives the norm.

Proposition 6.3 (Parallelogram Law)**.** *Let X be an inner product space, and $x, y \in X$. Then*
$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Proof: Expand again.

It turns out that the parallelogram law characterizes inner product spaces: if a normed space X satisfies the parallelogram law, then we can define (x, y) by the polarization identity, and check that $(\cdot, \cdot)$ is an inner product, giving the original norm. Hence a normed space X is an inner product space if and only if each two-dimensional subspace of X is an inner product space.

For $x, y \in X$, say x, y are *orthogonal*, written $x \perp y$, if $(x, y) = 0$.

Proposition 6.4 (Pythagoras' Theorem)**.** *Let X be an inner product space, $x, y \in X$. Then*

$$x \perp y \implies \|x + y\|^2 = \|x\|^2 + \|y\|^2$$

Recall that, in a normed space, the distance from a point x to a closed set S need not be attained, even in ℓ_2. However in inner product spaces, we see that this is the case.

Theorem 6.1 (Closed Point Theorem)**.** *Let H be a Hilbert space, $x \in H$, and S a closed subspace of H.*

Then the distance from x to S is attained, that is, there is $y \in S$ such that $\|x - y\| = d(x, S)$. Moreover, y is unique.

> **Proof:**   Choose $y_1, y_2, \ldots$ in S with $\|x - y_n\| \to d(x, S)$. Then we want to show that (y_n) is Cauchy.
>
> Indeed, we have
>
> $$2\|x - y_n\|^2 + 2\|x - y_m\|^2 = \|y_n - y_m\|^2 + \|2x - y_n - y_m\|^2,$$
>
> hence
>
> $$\begin{aligned} \|y_n - y_m\|^2 &= 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\left\|x - \frac{y_n + y_m}{2}\right\| \\ &\leq 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - 4d(x, s) \\ &\to 0, \end{aligned}$$
>
> as $\min(m, n) \to \infty$. Hence $y_n \to y$ for some y, as S is complete, so $\|x - y\| = d(x, S)$.
>
> For uniqueness, suppose $\|x - y\|, \|x - z\| = d(x, S)$. Then, as above,
>
> $$\|y - z\|^2 \leq 2d^2 + 2d^2 - 4d^2 = 0,$$

Note the same proof would have worked for S merely closed and convex.

For an inner product space X and $x \in X$, write $x^\perp$ for $\{y \in H \mid (x, y) = 0\}$. Then $x^\perp$ is a closed subspace of X.

For $S \subset X$, write $S^\perp = \{y \mid (x, y) = 0 \text{ for all } x \in S\}$. Then

$$S^\perp = \bigcap_{x \in X} x^\perp,$$

is also a closed subspace of X. Moreover note trivially that $S \subset T \implies T^\perp \subset S^\perp$.

Theorem 6.2. *Let H be a Hilbert space, and F a closed subspace of H. Then $H = F \oplus F^\perp$, where $F^\perp$ is the* orthogonal complement *of F, and H is the* ortogonoal direct sum *of F and $F^\perp$.*

> **Proof:** We need $F \cap F^\perp = \{0\}$, and $F + F^\perp = H$.
>
> Indeed, if $x \in F, F^\perp$, then $(x, x) = 0$, hence $x = 0$.
>
> Now to show $F, F^\perp$ span. Take $x \in H$. Then we want $y \in F$ such that $x - y \in F^\perp$ Indeed, let $y \in F$ be such that $\|x - y\| = d(x, F)$, then we claim $x - y \in F^\perp$.
>
> Indeed, suppose $x - y \notin F^\perp$, so there exists $z \in F$ with $(x - y, z) \neq 0$. Without loss of generality, assume $(x - y, z)$ is real and positive. Then for $t > 0$,
>
> $$d(x, F)^2 \leq \|x - (y + tz)\|^2 = \|x - y\|^2 - 2t(x - y, z) + t^2 \|z\|^2.$$
>
> As $d(x, F)^2 = \|x - y\|^2$, we have $t^2 \|z\|^2 - 2t(x - y, z) \geq 0$, which is false for small t. Hence $x - y \in F^\perp$.

Note for any $S \subset X$, we have $S \subset (S^\perp)^\perp$, since $x \in S \implies (x, y) = 0$ for all $y \in S^\perp$.

Corollary 6.1. *Let H be a Hilbert space. Then:*

(i) *F is closed subspace implies $(F^\perp)^\perp = F$.*

(ii) *If $S \subset H$, then $(S^\perp)^\perp = \overline{\langle S \rangle}$.*

(iii) *$S \subset H$ is dense $\iff S^\perp = \{0\}$.*

> **Proof:** (i) We know that $F \subset (F^\perp)^\perp$. Conversely, given $x \notin F$, then $x = y + z$ for some $y \in F$ and non-zero $z \in F^\perp$. Hence $(x, z) = (y + z, z) = \|z\|^2 > 0$, hence $x \notin (F^\perp)^\perp$.

(ii) For $S \subset H$,

$$y \in S^\perp \iff (y, x) = 0 \ \forall x \in S$$
$$\iff (y, x) = 0 \ \forall x \in \langle S \rangle$$
$$\iff (y, x) = 0 \ \forall x \in \overline{\langle S \rangle},$$

so $S^\perp = \overline{\langle S \rangle}^\perp$. But then

$$(S^\perp)^\perp = (\overline{\langle S \rangle}^\perp)^\perp = \overline{\langle S \rangle},$$

by (i).

(iii) If S has dense linear span, then $S^\perp = \overline{\langle S \rangle}^\perp = H^\perp = \{0\}$.

If S does not have dense linear span, then $\overline{\langle S \rangle} \neq H$, so $\overline{\langle S \rangle}^\perp \neq \{0\}$.

For a Hilbert space H and fixed $y \in H$, define $\theta_y : H \to \mathbb{C}$ by $x \mapsto (x, y)$. Then θ_y is linear and continuous, as $|(x, y)| \leq \|x\| \|y\|$. Hence $\|\theta_y\| \leq \|y\|$.

In fact, $\|\theta_y\| = \|y\|$, because $\theta_y(y) = \|y\|^2$. We can now generalise $\ell_2^* = \ell_2$:

Theorem 6.3 (Riesz Representation Theorem). *Let H be a Hilbert space, and $f \in H^*$. Then there exists $y \in H$ such that $f = \theta_y$.*

Proof: Without loss of generality, assume $f \neq 0$, and let $E = \ker f$. Then E is a proper closed subspace of H.

Thus $E^\perp \neq \{0\}$, and also $\dim E^\perp < 2$ as for all $x, y \in H$, we have $f(\lambda x + \mu y) = 0$ for some $\lambda, \mu \in \mathbb{C}$ not both zero. Thus, $E^\perp = \langle y \rangle$ for some y with $f(y) \neq 0$.

We may assume that $f(y) = \|y\|^2$ by scaling y. Now for any $x \in H$, write $x = z + \lambda y$ for some $z \in E$ and $\lambda \in \mathbb{C}$. Then

$$f(x) = f(z + \lambda y) = \lambda f(y) = \lambda \|y\|^2,$$

and $\theta_y(x) = (z + \lambda y, y) = \lambda \|y\|^2$.

Corollary 6.2. *For H a Hilbert space, then the map $\theta : H \to H^*$ given by $y \mapsto \theta_y$ is an isometric conjugate-linear isomorphism from H to H^*.*

Proof: We have $\|\theta_y\| = \|y\|$ for all $y \in H$, so θ is isometric, and also surjective by Riesz. Moreover

$$\theta_{\lambda y + \mu z} = \overline{\lambda}\theta_y + \overline{\mu}\theta_z,$$

so θ is conjugate-linear.

For an inner product space X, say $x_1, x_2, \ldots$ in X is an *orthonormal sequence* if $\|x_n\| = 1$ for all n, and $(x_n, x_m) = 0$ for $m \neq n$.

It is an *orthonormal basis* if, in addition, $\overline{\langle x_i \mid i \in \mathbb{N} \rangle} = H$. Note we can use the same terminology for a finite sequence as well.

Example 6.2.

1. In ℓ_2, $e_1, e_2, \ldots$ is an orthonormal basis.

2. In ℓ_2, (e_{2n}) is an orthonormal sequence, but not an orthonormal basis.

3. In $C_{\mathbb{C}}[0, 2\pi]$ with

$$(f, g) = \int_0^{2\pi} f\bar{g},$$

then $(\theta \to e^{in\theta})|_{n=-\infty}^{\infty}$ form an orthonormal basis.

Indeed, they are dense in span in $\{f \in C_{\mathbb{C}}[0, 2\pi] \mid f(0) = f(2\pi)\}, \|\cdot\|_{\infty}$ by Stone-Weierstrass, hence dense in span in $\{f \in C_{\mathbb{C}}[0, 2\pi] \mid f(0) = f(2\pi)\}, \|\cdot\|_2$, and hence dense in span in $C_{\mathbb{C}}[0, 2\pi]$.

Remark.

1. An orthonormal sequence (e_n) is linearly independent, as if $\sum \lambda_i e_i = 0$, then

$$\left(\sum_{i=1}^{N} \lambda_i e_i, e_n \right) = 0,$$

for all $n \leq N$, so $\lambda_i = 0$ for all $i \leq N$.

2. An orthonormal basis is not, in general, an algebraic basis.

3. If inner product space X has an orthonormal basis, then X is separable.

4. In a Hilbert space H, an orthonormal sequence (e_n) is an orthonormal basis

if:

$$\overline{\langle e_1, e_2, \ldots \rangle} = H \iff \overline{\langle e_1, e_2 \ldots \rangle}^\perp = \{0\}$$
$$\iff \forall x \in X \text{ if } (x, e_n) = 0 \text{ for all } n \implies x = 0.$$

Theorem 6.4 (Gram-Schmidt Process). *Let (x_n) be a linearly independent sequence in inner product space X. Then there exists an orthonormal sequence (e_n) such that $\langle e_1, \ldots, e_n \rangle = \langle x_1, \ldots, x_n \rangle$ for all n.*

Proof: We begin by normalising let

$$e_1 = \frac{x_1}{\|x_1\|},$$

so $\langle e_1 \rangle = \langle x_1 \rangle$. Then let

$$e_2' = x_2 - (x_2, x_1)e_1,$$

so $e_2' \perp e_1$. Then set

$$e_2 = \frac{e_2'}{\|e_2'\|}.$$

Note that $e_2' \neq 0$ as the x_i are linearly independent. Then we have $\langle e_1, e_2 \rangle = \langle x_1, x_2 \rangle$. We continue this way. Having chosen $e_1, \ldots, e_{n-1}$, put

$$e_n' = x_n - (x_n, e_1)e_1 - \cdots - (x_n, e_{n-1})e_{n-1},$$

and then set

$$e_n = \frac{e_n'}{\|e_n'\|}.$$

Corollary 6.3. *Let X be a separable inner product space. Then it has an orthonormal basis.*

Proof: Let $x_1, x_2, \ldots$ be dense in X, so in particular $\langle x_1, x_2, \ldots \rangle$ are dense. Without loss of generality, assume the x_n are linearly independent (by removing any x_i dependent on its predecessors).

Then we can apply Gram-Schmidt. We obtain an orthonormal sequence $e_1, e_2, \ldots$ with dense linear span.

Example 6.3.

In $C[-1,1]$ with inner product

$$(f,g) = \int_{-1}^{1} fg,$$

then $1, t, t^2, \ldots$ is dense by Stone-Weierstrass. So by Gram-Schmidt, we obtain an orthonormal basis which starts

$$\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}t, \sqrt{\frac{5}{8}}(3t^2 - 1), \ldots$$

These are the *normalised Legendre polynomials*. The n'th such polynomial is a multiple of

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n}((t^2 - 1)^n).$$

Corollary 6.4. *Let X be an n-dimensional inner product space. Then X is isometrically isomorphic to ℓ_2^n.*

Proof: Let $e_1, \ldots, e_n$ be an orthonormal basis. Define $T : X \to \ell_2^n$ by

$$\sum_{i=1}^{n} \lambda_i e_i \mapsto (\lambda_1, \ldots, \lambda_n).$$

Then T is a linear bijection. Also,

$$\left\| \sum_{i=1}^{n} \lambda_i e_i \right\|^2 = \sum_{i=1}^{n} |\lambda_i|^2 = \|(\lambda_1, \ldots, \lambda_n)\|^2.$$

So $\|T(x)\| = \|x\|$ for all $x \in X$.

Note that any isometric isomorphism also preserves the inner product, by polarisation.

Our goal is to extend the above theorem, to say that every separable Hilbert space is isometrically isometric to ℓ_2.

Proposition 6.5. *Let (e_n) be an orthonormal sequence in Hilbert space H. Then $\sum \lambda_i e_i$ converges $\iff (\lambda_i) \in \ell_2$.*

Proof:

$\implies$ We have

$$\left\|\sum_{i=1}^{N}\lambda_i e_i\right\|^2 \to \left\|\sum_{i=1}^{\infty}\lambda_i e_i\right\|^2 \implies \sum_{i=1}^{N}|\lambda_i|^2 \text{ converges to some limit}$$

$$\implies (\lambda_i) \in \ell_2.$$

$\impliedby$ We have

$$\left\|\sum_{i=N}^{M}\lambda_i e_i\right\|^2 = \sum_{i=N}^{M}|\lambda_n|^2 \to 0$$

as $\min(N, M) \to \infty$. Hence we get $\sum \lambda_i e_i$ is Cauchy, so convergent.

Corollary 6.5 (Riesz-Fischer Theorem). *Let (e_n) be an orthonormal sequence in Hilbert space H. Then for any $c \in \ell_2$, there exists $x \in H$ such that $(x, e_n) = c_n$ for all n.*

Proof: Define the term

$$x = \sum_{i=1}^{\infty}c_i e_i.$$

This converges as $c \in \ell_2$. Then note that

$$\left(\sum_{i=1}^{N}c_i e_i, e_n\right) = c_n$$

for all $n \le N$. Hence $(x, e_n) = c_n$.

For (e_n) an orthonormal sequence in inner product space X, and $x \in X$ the values (x, e_n) are called the *coefficients* or *Fourier coefficients* of x with respect to (e_n).

Our aim is to have an orthonormal basis (e_n) with

$$x = \sum_{i=1}^{\infty}(x, e_i)e_i.$$

There is no such result for general Banach spaces: in $C[0, 1], \|\cdot\|_\infty$, then $1, t, t^2, \ldots$ have dense linear span, but

$$\left|t - \frac{1}{2}\right| \ne \sum_{n=0}^{\infty}c_n t^n$$

for some c_n, as $|t - 1/2|$ is not differentiable.

Theorem 6.5 (Bessel's Inequality). *Let (e_n) be an orthonormal sequence in an inner product space X. Then*

$$((x, e_n))_{n=1}^{\infty} \in \ell_2.$$

Proof: We will show that

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

for all N. Then we are done as $((x, e_n)) \in \ell_2$. Define

$$y = \sum_{i=1}^{N} (x, e_i)e_i,$$

then we want to show that $\|y\|^2 \leq \|x\|^2$. We have

$$(y, e_n) = (x, e_n)$$

for all $n \leq N$, so in particular $x - y \perp e_1, \ldots, e_N$. Hence $x - y \perp \langle e_1, \ldots, e_N \rangle$

Therefore $x - y \perp y$, and $\|x\|^2 = \|y\|^2 + \|x - y\|^2 \geq \|y\|^2$.

Theorem 6.6. *Let H be a Hilbert space, and (e_n) an orthonormal basis for H. Then*

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

Proof: We have that
$$((x, e_n))_{n=1}^{\infty} \in \ell_2,$$

by Bessel's. Hence we have

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

is convergent, to say y.

Whence we have $y - x \in \overline{\langle e_1, e_2, \ldots \rangle} = \{0\}$, so $y = x$.

From this we get that

$$\|x\|^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2, \qquad (x, y) = \sum_{n=1}^{\infty} (x, e_n)\overline{(y, e_n)},$$

for all $x, y \in H$. These two facts are called *Parseval's identity*.

Remark. The above theorem also holds in any inner product space, as we can embed X into a Hilbert space $\overline{X}$, the "completion" of X, in which it is dense.

Example 6.4.

In $C_{\mathbb{C}}[0, 2\pi]$ with inner product

$$(f, g) = \frac{1}{2\pi} \int_0^{2\pi} f\overline{g},$$

we know that the maps $\theta \mapsto e^{in\theta}$ for $n \in \mathbb{Z}$ form an orthonormal basis, hence for all $f \in C_{\mathbb{C}}[0, 2\pi]$ we have

$$f = \sum_{-\infty}^{\infty} c_n e^{in\theta},$$

where $c_n = (f, e^{in\theta})$. This is the *Fourier series* of f, and here convergence is in $\|\cdot\|_2$.

For $f \in C_{\mathbb{C}}[0, 2\pi]$ with $f(0) = f(2\pi)$, write

$$S_k = \sum_{-k}^{k} (f, e^{in\theta}) e^{in\theta}.$$

Then we know $S_k \to f$ in $\|\cdot\|_2$. But it is not true that $S_k \to f$ uniformly, or even pointwise. However Fejer's theorem says that

$$\frac{S_1 + \cdots + S_k}{k} \to f$$

uniformly.

Corollary 6.6. *Let H be a separable infinite dimensional Hilbert space. Then H is isometrically isomorphic to ℓ_2.*

Proof: Pick an orthonormal basis $e_1, e_2, \ldots$ for H.

Define $T : H \to \ell_2$ by
$$Tx = ((x, e_n))_{n=1}^\infty.$$

Then T is linear, and moreover
$$\|x^2\| = \sum_{n=1}^\infty |(x, e_n)|^2.$$

So T is an isometry. Moreover T is surjective by Riesz-Fisher.

6.1   Matrices of Linear Maps

Let H be a Hilbert space. We want to understand the operatos on H, meaning elements of $L(H) = L(H, H)$.

Suppose H has orthonormal basis (e_n). Then the *matrix* of $T \in L(H)$ with respect to (e_n) is
$$A = \begin{pmatrix} Te_1 & Te_2 & \cdots \end{pmatrix}.$$
The array $A = (a_{ij})$ is given by $a_{ij} = (Te_j, e_i)$. Hence

$$Te_j = \sum_{n=1}^\infty a_{ij} e_i.$$

So the matrix of T determines T, since knowing T on an orthonormal basis means we know T on $\overline{\langle e_1, e_2, \ldots \rangle}$.

Example 6.5.

If $T = I$, then the corresponding matrix is
$$\begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

If T is the right shift, then

$$\begin{pmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

If T is the left shift then

$$\begin{pmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

If $T(\sum x_n e_n) = \sum \frac{x_n}{2^n} e_n$, then

$$\begin{pmatrix} 1/2 & 0 & 0 & \cdots \\ 0 & 1/4 & 0 & \cdots \\ 0 & 0 & 1/8 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

Note that not every matrix corresponds to an operator $T \in L(H)$. Even if we insist that each column $c^{(i)}$ is in ℓ_2 with $\|c^{(i)}\| \leq 1$, then this can still fail. Consider

$$\begin{pmatrix} 1 & 1 & 1 & \cdots \\ 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

Indeed, if $T \in L(\ell_2)$ has this matrix, then $T(e_n) = e_1$ for all n. So then $T(e_1 + \cdots + e_n) = ne_1$. But $\|e_1 + \cdots + e_n\| = \sqrt{n}$, and $\|ne_1\| = n$, hence the operator is not bounded.

We can only say that A corresponds to an operator if and only if there exists $K > 0$ such that

$$\left\| \sum_{i=1}^{\infty} \lambda c^{(i)} \right\|_2 \leq K \|\lambda\|_2,$$

for all $\lambda \in \ell_2$.

6.2   Adjoints

Theorem 6.7. *Let H be a Hilbert space, and $T \in L(H)$. Then there exists a unique map $T^* : H \to H$ such that*

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

for all $x, y \in H$. Moreover, $T^* \in L(H)$.

Such a T^* is called the adjoint *of T.*

Remark. If T has a matrix A with respect to orthonormal basis (e_n), then T^* has matrix $\overline{A}^T$. Indeed,

$$(T^*e_j, e_i) = \overline{(e_i, T^*e_j)} = \overline{(Te_i, e_j)} = \overline{a_{ji}}.$$

But we cannot use $\overline{A}^T$ to define T^*, because:

- H might not have an orthonormal basis, for example if H is not separable.

- This might depend on the choice of orthonormal basis.

- It is not clear why $\overline{A}^T$ should correspond to a continuous linear operator.

Example 6.6.

If $H = \ell_2$, and I is the identity, then $I^* = I$, as $(x, y) = (x, y)$.

Looking at matrices, note that $(\lambda I)^* = \overline{\lambda} I$.

Similarly if we look at matrices we can see that $(\text{left shift})^* = \text{right shift}$.

Proof: For fixed $y \in H$, the map $x \mapsto (Tx, y)$ is a continuous linear map, as $|(Tx, y)| \leq \|y\| \|T\| \|x\|$.

Hence there exists unique $z \in H$ such that $(Tx, y) = (x, z)$ for all $x \in H$, by Riesz representation. Then this z is T^*y. Now we need to show that $T^* \in L(H)$.

First, T^* is linear. Note that for all x,

$$(Tx, \lambda y + \mu z) = (x, T^*(\lambda y + \mu z)).$$

But also,

$$(Tx, \lambda y + \mu z) = \overline{\lambda}(Tx, y) + \overline{\mu}(Tx, z) = \overline{\lambda}(x, T^*y) + \overline{\mu}(x, T^*z)$$
$$= (x, \lambda T^*y + \mu T^*z).$$

Since this holds for all x, $T^*(\lambda y + \mu z) = \lambda T^* y + \mu T^* z$.

Now we show T^* is continuous. Note that $|(x, T^* y)| = |(Tx, y)| \leq \|y\| \|T\| \|x\|$, for all x. Therefore, letting $x = T^* y$, we get $\|T^* y\| \leq \|y\| \|T\|$. Therefore, $\|T^*\| \leq \|T\|$.

Proposition 6.6. *Let H be a Hilbert space, and $S, T \in L(H)$. Then:*

(i) $(\lambda S + \mu T)^* = \overline{\lambda} S^* + \overline{\mu} T^*$, *for all $\lambda, \mu \in \mathbb{C}$.*

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

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

(iv) $\|T^*\| = \|T\|$.

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

Proof:

(i) Let's do some algebra and expand everything:

$$((\lambda S + \mu T)x, y) = \lambda(Sx, y) + \mu(Tx, y) = \lambda(x, S^* y) + \mu(x, T^* y)$$
$$= (x, \overline{\lambda} S^* y + \overline{\mu} T^* y),$$

for all $x \in H$. Hence $(\lambda S + \mu T)^* y = (\overline{\lambda} S^* + \overline{\mu} T^*)y$, for any y. Therefore the two maps are equal.

(ii) Again, let's do some manipulations.

$$(STx, y) = (Tx, S^* y) = (x, T^* S^* y),$$

for all $x \in H$. Hence $(ST)^* y = T^* S^* y$, for any y.

(iii) We get that

$$(T^* x, y) = \overline{(y, T^* x)} = \overline{(Ty, x)} = (x, Ty)$$

for all x, so $(T^*)^* y = Ty$.

(iv) We have that $\|T^*\| \leq \|T\|$. But also $\|T\| = \|(T^*)^*\| \leq \|T^*\|$, hence the two norms are equal.

(v) Certainly $\|T^* T\| \leq \|T\|^2$ from (iv). Also, for any x,

$$\|Tx\|^2 = (Tx, Tx) = (x, T^* Tx) \leq \|x\| \|T^* T\| \|x\|,$$

hence $\|Tx\| \leq \|x\| \|T^* T\|^{1/2}$, or $\|T\| \leq \|T^* T\|^{1/2}$.

Say that $T \in L(H)$ is *Hermitian* or *self-adjoint* if $T^* = T$. Equivalently, $(Tx, y) = (x, Ty)$ for all $x, y \in H$.

So, if T has matrix A with respect to orthonormal basis (e_n), then T is hermitian if and only if $\overline{A}^T = A$.

Example 6.7.

1. I is Hermitian.

2. Let $T : \ell_2 \to \ell_2$ given by

$$\sum_{n=1}^{\infty} x_n e_n \mapsto \sum_{n=1}^{\infty} \lambda_n x_n e_n,$$

   where (λ_n) is a bounded sequence. Then T is Hermitian if and only if all λ_n are real.

3. Let F be a closed subspace of H, so $H = F \oplus F^\perp$. The *projection onto F* or *orthogonal projection onto F* is the map $T : H \to H$ given by

$$T(x + y) = x,$$

   when $x = F, y \in F^\perp$. Then certainly $T \in L(H)$. In fact, T is Hermitian, as for all $x, x' \in F$ and $y, y' \in F^\perp$,

$$(T(x + y), x' + y') = (x, x' + y') = (x, x'),$$

   and also

$$(x + y, T(x' + y')) = (x + y, x') = (x, x').$$

Proposition 6.7. *Let H be a complex Hilbert space, and $T \in L(H)$. Then there exists Hermitian $T_1, T_2 \in L(H)$ such that $T = T_1 + iT_2$.*

Proof: We have that $T + T^*$ is Hermitian, and $i(T - T^*)$ is Hermitian. Moreover,
$$T = \frac{1}{2}(T + T^*) + \frac{1}{2}(T - T^*).$$

Remark. T_1, T_2 are unique. Indeed, if also $T = T_1' + iT_2'$, then

$$(T_1 - T_1') + i(T_2 - T_2') = 0, \qquad (T_1 - T_1') - i(T_2 - T_2') = 0,$$

hence $T_1 = T_1'$ and $T_2 = T_2'$.

7   Spectral Theory

Let $T \in L(X)$, where X is a Banach space. Say T is *invertible* if there exists $S \in L(X)$ with
$$ST = TS = I.$$

In particular, we have

$$\begin{aligned}
T \text{ invertible} &\iff T \text{ bijective}, T^{-1} \text{ continuous} \\
&\iff T \text{ bijective} \\
&\iff T \text{ injective}, T \text{ surjective}.
\end{aligned}$$

Recall that we do need both: take the left shift and the right shift. This differs from the finite-dimensional case.

Theorem 7.1. *Let X be a Banach space, and $T \in L(X)$ with $\|T\| < 1$. Then $I - T$ is invertible, with*
$$(I - T)^{-1} = \sum_{n=0}^{\infty} T^n.$$

Note this sum does converge as $\|T\| < 1$, so the sum is Cauchy in $L(X)$.

> **Proof:**   We have
> $$(I - T) \sum_{n=0}^{k} T^n = I - T^{k+1} \implies (I - T) \sum_{n=0}^{\infty} T^n = I,$$
> and moreover
> $$\left(\sum_{n=0}^{k} T^n \right)(I - T) = I - T^{k+1}$$
> as well.

Write $\mathsf{GL}(X)$ for $\{T \in L(X) \mid T \text{ invertible}\}$. This is also written $\mathsf{G}(X)$ or just G.

Theorem 7.2. *Let X be a Banach space, and $G = \mathsf{GL}(X)$. Then:*

  (i) *G is open.*

 (ii) *The map $T \mapsto T^{-1}$ from G to G is continuous.*

(iii) *If $T_1, T_2, \ldots, \in G$ with $T_n \to T$ for some $T \notin G$, then $\|T_n^{-1}\| \to \infty$.*

Proof:

(i) Given $T \in G$, for any $S \in L(X)$, we have $T - S = T(I - T^{-1}S)$. Hence for $\|S\| \leq \frac{1}{\|T^{-1}\|}$, we have

$$\|T^{-1}S\| \leq \|T^{-1}\|\|S\| \leq 1.$$

So $I - T^{-1}S$ is invertible, with inverse

$$\sum_{n=0}^{\infty}(T^{-1}S)^n.$$

Hence $T - S$ is invertible with inverse

$$\sum_{n=0}^{\infty}(T^{-1}S)^nT^{-1}.$$

Therefore there is a ball around T in G.

(ii) Given $T \in G$, then for $\|S\| \leq \frac{1}{\|T^{-1}\|}$, we have

$$(T - S)^{-1} - T^{-1} = \sum_{n=1}^{\infty}(T^{-1}S)^nT^{-1},$$

so we have

$$\|(T - S)^1 - T^{-1}\| \leq \sum_{n=1}^{\infty}\|S\|^n\|T^{-1}\|^{n+1} \to 0$$

as $\|S\| \to 0$.

(iii) Given $\varepsilon > 0$, then for $\|T_n\| < \varepsilon$, we have

$$B\left(T_n, \frac{1}{\|T_n^{-1}\|}\right) \subset G,$$

and also $T \notin G$. Hence T is not in this ball, implying

$$\frac{1}{\|T_n^{-1}\|} \leq \varepsilon \implies \|T_n^{-1}\| \geq \frac{1}{\varepsilon}.$$

From now on, we restrict to *complex* Banach spaces.

For X a complex Banach space, $T \in C(X)$ and $\lambda \in \mathbb{C}$, say λ is an *eigenvalue* of T if there exists $x \neq 0$ with $Tx = \lambda x$, i.e. $T - \lambda I$ is not injective.

Such an x is an *eigenvector* of T for eigenvalue λ.

> **Example 7.1.**
>
> Let $X = \ell_2$, and T be the right-shift.
>
> Then T is not invertible, but 0 is not an eigenvalue of T, as for example $\|Tx\| = \|x\|$ for all x. This is unlike the finite dimensional case.
>
> Moreover, in finite dimensional case, every operator has an eigenvalue. But this T has no eigenvalues because if λ is an eigenvalue with eigenvector x, then $|\lambda| = 1$. Hence
> $$0 = |x_1| = |x_2| = \cdots,$$
> so $x = 0$.

This shows we are not really interested in eigenvalues and eigenvectors in Banach spaces. For X a complex Banach space, and $T \in C(X)$, then the *spectrum* of T is

$$\sigma(T) = \{\lambda \mid T - \lambda I \text{ not invertible}\}.$$

So each eigenvalue of T belongs to $\sigma(T)$, and

$$\sigma(T) = \{\lambda \in \mathbb{C} \mid \lambda \text{ an eigenvalue of } T\}$$

if $\dim X < \infty$. But in general $\sigma(T)$ can be bigger. For example 0 is in the spectrum of the right shift. Now note that

$$\lambda \in \sigma(T) \iff T - \lambda I \text{ not invertible}$$
$$\iff T - \lambda I \text{ not injective or } T - \lambda I \text{ not surjective.}$$

Actually this last condition is not that helpful.

Proposition 7.1. *Let X be a complex Banach space, and $T \in C(X)$. Then $\sigma(T)$ is a closed subset of $\{z \in \mathbb{C} \mid |z| \leq \|T\|\}$ in particular, $\sigma(T)$ is compact.*

> **Proof:** We begin by showing $\sigma(T)$ is closed. Note that $\lambda \in \mathbb{C} \setminus \sigma(T) \iff T - \lambda I \in G$.
>
> But the map $\lambda \mapsto T - \lambda I$ from $\mathbb{C}$ to $L(X)$ is continuous, and G is open. Hence $\mathbb{C} \setminus \sigma(T)$ is open, and so $\sigma(T)$ is closed.

Finally note that $\sigma(T) \subset \{\lambda \mid |\lambda| \leq \|T\|\}$. Indeed, given $\lambda \in \mathbb{C}$ with $|\lambda| > \|T\|$, we have

$$T - \lambda I = -\lambda\left(I - \frac{T}{\lambda}\right),$$

but as $\|\frac{T}{\lambda}\| < 1$, this is invertible with inverse

$$\sum_{n=0}^{\infty} \left(\frac{T}{\lambda}\right)^n \cdot \left(-\frac{1}{\lambda}\right).$$

We often call $\mathbb{C} \setminus \sigma(T)$ the *resolvent set* of T, and the *resolvent function* R_T or R is the function from $\mathbb{C} \setminus \sigma(T) \to L(X)$ given by $\lambda \mapsto (T - \lambda I)^{-1}$.

Note that if λ is in the resolvent set and $|\lambda| > \|T\|$, then

$$R(\lambda) = \sum_{n=0}^{\infty} \left(\frac{T}{\lambda}\right)^n \cdot \left(-\frac{1}{\lambda}\right).$$

Example 7.2.

Let's look at the spectrum of the left-shift S on ℓ_2. For each $|\lambda| < 1$, we have that λ is an eigenvalue, with eigenvector

$$(1, \lambda, \lambda^2, \ldots).$$

So $\sigma(S) \subset \{z \mid |z| \leq 1\}$, and is closed. Hence $\sigma(S) = \{z \mid |z| \leq 1\}$.

Let's look at another way to find spectra.

Example 7.3.

Consider $T : \ell_2 \to \ell_2$ given by

$$\sum_{n=1}^{\infty} x_n e_n \mapsto \sum_{n=1}^{\infty} \frac{x_n}{2^n} e_n.$$

Then 0 is not an eigenvalue, as T is injective. But $\|e_n\| = 1$, and $\|Te_n\| \to 0$, so T cannot be invertible, as T^{-1} could not be continuous. Hence 0 is in the spectra.

For X a complex Banach space and $T \in L(X)$, we say $\lambda \in \mathbb{C}$ is an *approximate eigenvalue* of T if for all $\varepsilon > 0$, there is $x \in X$ with $\|x\| = 1$ and $\|Tx - \lambda x\| < \varepsilon$.

Equivalently, there exists $x_1, x_2, \ldots$ in X with $\|x_n\| = 1$ for all n, and $Tx_n - \lambda x_n \to 0$.

We call $(x_n)_{n=1}^{\infty}$ an *approximate eigenvector* for T with eigenvalue λ.

In the above example, 0 is an approximate eigenvalue, but not an eigenvalue. For the right-shift, $0 \in \sigma(T)$ but 0 is not an approximate eigenvalue as $\|Tx\| = \|x\|$.

Remark.

1. λ is an eigenvalue of $T \implies \lambda$ an approximate eigenvalue of T.

2. λ an approximate eigenvalue of $T \implies \lambda \in \sigma(T)$.

We can write

$$\sigma_P(T) = \{\lambda \mid \lambda \text{ an eigenvalue of } T\},$$
$$\sigma_{AP}(T) = \{\lambda \mid \lambda \text{ an approximate eigenvalue of } T\}.$$

$\sigma_P(T)$ is the *point spectrum* of T, and $\sigma_{AP}(T)$ is the *approximate point spectrum* of T.

We have $\sigma_P(T) \subset \sigma_{AP}(T) \subset \sigma(T)$.

Theorem 7.3. *Let X be a compact Banach space, and $T \in C(X)$. Let $\lambda \in \partial\sigma(T)$. Then λ is an approximate eigenvalue of T.*

> **Proof:** Choose $\lambda_1, \lambda_2, \ldots$ in $\mathbb{C} \setminus \sigma(T)$ with $\lambda_n \to \lambda$.
>
> So $T - \lambda_n \to T - \lambda$. Hence $\|(T - \lambda_n)^{-1}\| \to \infty$.
>
> So there exists $x_1, x_2, \ldots$ in X with $\|x_n\| \to 0$ and $\|(T - \lambda_n)^{-1}x_n\| = 1$. Let $y_n = (T - \lambda_n)^{-1}x_n$. Then $\|y_n\| = 1$ for all n.
>
> But also, $(T - \lambda_n)y_n = x_n \to 0$, whence $(T - \lambda)y_n \to 0$.

Example 7.4.

We find the spectrum of the right-shift T on ℓ_2.

We know that $\sigma(T) \subset \triangle = \{z \in \mathbb{C} \mid |z| \leq 1\}$. We have $\|Tx\| = \|x\|$ for all x, so for $|\lambda| < 1$, we have $\|(T - \lambda)x\| \geq (1 - \lambda)\|x\|$, so λ is not an approximate eigenvalue.

Hence we know $\partial\sigma(T) \subset \pi = \{z \mid |z| = 1\}$.

Hence we must have either $\text{int}(\triangle) \subset \sigma(T)$, or $(\text{int}(\triangle)) \cap \sigma(T) = \emptyset$. But $0 \in \sigma(T)$ as T is not invertible, so $\text{int}(\triangle) \subset \sigma(T)$

Moreover, $\sigma(T) = \triangle$ as $\sigma(T)$ is closed.

Theorem 7.4 (Spectral Mapping Theorem). *Let X be a complex Banach space, and $T \in C(X)$. Let p be a non-constant complex polynomial. Then*

$$\sigma(p(T)) = \{p(\lambda) \mid \lambda \in \sigma(T)\}.$$

Proof: Given $\mu \in \mathbb{C}$, write $p(z) - \mu$ as $c(z - \lambda_1)\cdots(z - \lambda_n)$ for some $c, \lambda_1, \ldots, \lambda_n \in \mathbb{C}$, with $c \neq 0$.

The we claim that $p(T) - \mu$ is invertible $\iff$ all $T - \lambda_i$ are invertible. Then we are done, as $\mu \in \sigma(p(T)) \iff$ some $\lambda_i \in \sigma(T)$, and $\lambda_1, \ldots, \lambda_n$ are precisely all z with $p(z) = \mu$.

Let's prove the claim. Note that $p(T) - \mu = c(T - \lambda_1)\cdots(T - \lambda_n)$.

$\impliedby$ If all the $T - \lambda_i$ are invertible, then $p(T) - \mu$ is invertible (as A, B invertible implies AB is invertible).

$\implies$ Suppose that some $T - \lambda_i$ is not invertible. So $T - \lambda_i$ is either not injective or not surjective. Whence, $(T - \lambda_1)\cdots(T - \lambda_n)$ is not invertible, as:

- If A is non-injective, then BA is non-injective.

- If A is non-surjective, then AB is non-surjective.

- $T - \lambda_1, \ldots, T - \lambda_n$ all commute.

Note that we did not use the false fact that A is non-invertible then AB is non-invertible. A simple example is by letting A be the left-shift, and B the right-shift.

The *spectral radius* of T is

$$r(T) = \sup\{|\lambda| \mid \lambda \in \sigma(T)\}.$$

So $r(T) \leq \|T\|$.

Corollary 7.1. *Let X be a complex Banach space, and $T \in C(X)$. Then*

$$r(T) \leq \inf_{n \geq 1} \|T^n\|^{1/n}.$$

Proof: We have that $r(T^n) \leq \|T^n\|$, and $r(T^n) = r(T)^n$ by spectral mapping theorem.

So $r(T)^n \leq \|T^n\|$.

For example, if T is *nilpotent* (i.e. $T^n = 0$ for some n), then $r(T) = 0$. So we can have $r(T) < \|T\|$. We can also have $r(T) = \|T\|$, as seen with the left or right shift.

Theorem 7.5 (Non-Emptiness of the Spectrum). *Let H be a (non-zero) Hilbert space, and $T \in L(H)$. Then $\sigma(T) \neq \emptyset$.*

Proof: Suppose $\sigma(T) \neq \emptyset$. So $R(\lambda) = (T - \lambda)^{-1}$ is defined for all $\lambda \in \mathbb{C}$, and it is continuous. Also, $\|R(\lambda)\| \to 0$ as $|\lambda| \to \infty$, since

$$R(\lambda) = -\sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}},$$

for all λ with $|\lambda| \geq \|T\|$. Now for $\lambda \neq \mu$,

$$
\begin{aligned}
R(\lambda) - R(\mu) &= (T - \lambda)^{-1} - (T - \mu)^{-1} \\
&= (T - \lambda)^{-1}((T - \mu) - (T - \lambda))(T - \mu)^{-1} \\
&= (\lambda - \mu)R(\lambda)R(\mu),
\end{aligned}
$$

whence we get

$$\frac{R(\lambda) - R(\mu)}{\lambda - \mu} = R(\lambda)R(\mu) \to R(\lambda)^2$$

as $\mu \to \lambda$, since R is continuous. This is good in a Liouville-sort of way.

For $x, y \in H$, define $f : \mathbb{C} \to \mathbb{C}$ by $(\lambda \mapsto (R(\lambda)x, y)$.

Then $f(\lambda) \to 0$ as $|\lambda| \to \infty$, as $|f(\lambda)| \leq \|R(\lambda)\|\|x\|\|y\| \to 0$. Also,

$$
\begin{aligned}
\frac{f(\lambda) - f(\mu)}{\lambda - \mu} = \frac{(R(\lambda)x, y) - (R(\mu)x, y)}{\lambda - \mu} &= (R(\lambda)R(\mu)x, y) \\
&\to (R(\lambda)^2 x, y)
\end{aligned}
$$

as $\mu \to \lambda$. So f is analytic, and defined on $\mathbb{C}$, but goes to 0 as $|\lambda| \to \infty$, whence $f = 0$ by Liouville's theorem.

Therefore $(R(\lambda)x, y) = 0$ for all x, y. Therefore $R(\lambda)x = 0$ for all x, hence $R(\lambda) = 0$, which contradicts $R(\lambda)$ invertible.

Remark.

1. Hence every $T \in L(H)$ has an approximate eigenvalue.

2. For $T \in L(X)$, we would consider $\phi(R(\lambda)x)$, for any fixed $x \in X$ and $\phi \in X^*$ we would let $\phi(R(\lambda)x) = 0$ for all ϕ, x. By Hahn-Banach, this tells us $R(\lambda)x = 0$.

3. We can phrase the above proof in terms of "$L(H)$-valued analytic functions" - we call $g : \mathbb{C} \to L(H)$ *analytic* if for all $\lambda \in \mathbb{C}$, there exists $S \in L(H)$ such that
$$\frac{g(\lambda) - g(\mu)}{\lambda - \mu} \to S,$$
as $\mu \to \lambda$. We proved a Liouville-type theorem for such functions: if g analytic and $g(\lambda) \to 0$ as $|\lambda| \to 0$, then $g = 0$.

4. We know that
$$R(\lambda) = -\sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}.$$

By consider $L(H)$-valued analytic functions, it follows that this Laurent series converges for all $|\lambda| > r(T)$.

But that Laurent series has radius of convergence $\limsup \|T^n\|^{1/n}$. So $r(T) \geq \limsup \|T^n\|^{1/n}$. However, we know that $r(T) \leq \inf \|T^n\|^{1/n}$, hence we can conclude that
$$r(T) = \lim_{n \to \infty} \|T^n\|^{1/n}.$$

7.1   Spectral Theory of Hermitian Operators

Proposition 7.2. *Let H be Hilbert and $T \in L(H)$. Then $\sigma(T^*) = \overline{\sigma(T)}$.*

Proof: We have T is invertible $\iff$ T^* is invertible. Indeed, $ST = TS = I \implies T^*S^* = S^*T^* = I$, and conversely.

So for all $\lambda \in \mathbb{C}$, $T - \lambda I$ invertible $\iff$ $(T - \lambda I)^* = T^* - \overline{\lambda} I$ invertible.

Example 7.5.

On ℓ_2, we have $\sigma(\text{left-shift}) = \triangle$, so $\sigma(\text{right-shift}) = \triangle$.

This holds for ℓ_2, but not ℓ_p in general.

Theorem 7.6. *Let H be Hilbert and $T \in L(H)$ Hermitian. Then $\sigma(T) \subset \mathbb{R}$.*

Proof: For λ an approximate eigenvalue of T, there exists (x_n) with $\|x_n\| = 1$ and $(T - \lambda)x_n \to 0$.

So $(Tx_n - \lambda x_n, x_n) \to 0$, i.e. $(Tx_n, x_n) \to \lambda$. But note (Tx_n, x_n) is real, as $(Tx_n, x_n) = (x_n, Tx_n)$. Thus $\lambda \in \mathbb{R}$.

So $\sigma_{AP}(T) \subset \mathbb{R}$, whence $\partial\sigma(T) \subset \mathbb{R}$. Hence $\sigma(T) \subset \mathbb{R}$.

For Hermitian operators, since $\sigma(T) \subset \mathbb{R}$, we have $\partial\sigma(T) = \sigma(T)$, so every $\lambda \in \sigma(T)$ is an approximate eigenvalue of T.

Remark. Hence the eigenvectors for distinct eigenvalues (when T is Hermitian) are orthogonal. Indeed, let $\lambda \neq \mu$ be eigenvalues, with eigenvectors x, y respectively. Then

$$(Tx, y) = (\lambda x, y) = \lambda(x, y),$$
$$(x, Ty) = (x, \mu y) = \mu(x, y).$$

So $\lambda(x, y) = \mu(x, y)$, giving $(x, y) = 0$.

Theorem 7.7. *Let H be Hilbert and $T \in L(H)$ Hermitian. Then $r(T) = \|T\|$.*

Proof: Without loss of generality, let $\|T\| = 1$. We will show that either 1 or -1 is an approximate eigenvalue of T.

We have a sequence (x_n) with $\|x_n\| = 1$ and $\|Tx_n\| \to 1$. So $(Tx_n, Tx_n) \to 1$, i.e. $(T^2 x_n, x_n) \to 1$. Then

$$\|(T^2 - I)x_n\|^2 = (T^2 x_n - x_n, T^2 x_n - x_n)$$
$$= \|T^2 x_n\|^2 + 1 - 2(T^2 x_n, x_n),$$

but the first term is at most 1, and the last term tends to -2, whence $(T^2 - I)x_n \to 0$, i.e. $(T + I)(T - I)x_n \to 0$.

Now by spectral mapping theorem $0 \in \sigma(T^2 - 1) = \{\lambda^2 - 1 \mid \lambda \in \sigma(T)\}$, so either $1 \in \sigma(T)$ or $-1 \in \sigma(T)$, and hence these are approximate eigenvalues.

Imre prefers the following: if $(T - I)x_n \to 0$, we have 1 is an approximate eigenvalue of T. Otherwise if $(T - I)x_n \nrightarrow 0$, then passing to a subsequence, we have $\|(T - I)x_n\| \geq \delta$, for some $\delta > 0$. But $(T + I)(T - I)x_n \to 0$, so -1 is an approximate eigenvalue of T.

Remark. Alternatively, we have $\|T^2\| = \|T\|^2$. Iterating, we get $\|T^{2^n}\| = \|T\|^{2^n}$.

So the spectral radius formula would tell us

$$r(T) = \lim \|T^n\|^{1/n} = \|T\|.$$

For X Banach and $T \in L(X)$, we say a subspace Y of X is *invariant* if $T(Y) \subset Y$.

> **Example 7.6.**
>
> Examples of invariant subspaces:
>
> (i) The trivial spaces $\{0\}$ and X are always invariant.
>
> (ii) If X is an eigenvector of T with eigenvalue λ, then $T(x) = \lambda x$, so $\langle x \rangle$ is invariant.
>
> (iii) More generally, for any eigenvalue λ, the *eigenspace* $E_\lambda = \{x \mid Tx = \lambda x\}$ is invariant, and closed.
>
> (iv) If T is the right-shift on ℓ_2, then T has no eigenvalues. But $\langle e_2, e_3, \ldots \rangle$ is a closed invariant subspace.

The invariant subspace problem is as follows: does every $T \in L(X)$ have a non-trivial closed invariant subspace? The answer is no, by Read and Enflo.

The example by Read is on ℓ_1. It satisfies that for all $x \neq 0$, $(x, Tx, T^2x, \ldots)$ has dense linear span. However it is hypercyclic, meaning the set itself is dense.

The invariant subspace problem is unsolved for Hilbert spaces.

If $T(Y) \subset Y$, i.e. Y is an invariant subspace, then we sometimes say that T acts on Y.

Proposition 7.3. *Let H be a Hilbert space and Y a subspace of H. Then*

$$T(Y) \subset Y \implies T^*(Y^\perp) \subset Y^\perp.$$

In particular, if T is Hermitian and T acts on Y, then T acts on $Y^\perp$.

> **Proof:** Given $x \in Y^\perp$, we need to show that $T^*x \in Y^\perp$. Indeed, for all $y \in Y$,
> $$(y, T^*x) = (Ty, x) = 0,$$
> as $Ty \in y$ and $x \in Y^\perp$. Hence $T^*x \in Y^\perp$.

Corollary 7.2. *Let H be a Hilbert space of dimension $n \in \mathbb{N}$ and $T \in L(H)$ Hermitian. Then there exists an orthonormal basis $(e_1, \ldots, e_n)$ of eigenvectors of T.*

> **Proof:**   Without loss of generality, assume $H \neq \{0\}$.
>
> Pick an eigenvector e_1 of T with $\|e_1\| = 1$: this is possible as H is over $\mathbb{C}$. Since T acts on $\langle e_1 \rangle$, T acts on $\langle e_1 \rangle^\perp$.
>
> Moreover, $T|_{\langle e_1 \rangle^\perp}$ is Hermitian as T is as well, so by induction we have an orthonormal basis $\langle e_2, \dots, e_n \rangle$ for $\langle e_1 \rangle^\perp$ consisting of eigenvectors of T.
>
> Hence $e_1, \dots, e_n$ is an orthonormal basis for H consisting of eigenvectors of T.

The above is something we have seen many times before. We try to generalise this to infinite dimensions.

7.2   Spectral Theory of Compact Operators

Recall that $T \in L(X)$ is compact if $\overline{T(B_X)}$ is compact, or equivalently if for all bounded $(x_n) \in X^{\mathbb{N}}$, there exists a subsequence (x_{n_j}) for which (Tx_{n_j}) converges.

We have seen that the compact operators form a closed subspace of $L(X)$, hence a limit of finite rank operators is compact.

Proposition 7.4. *Let X be an infinite dimensional Banach space, and $T \in L(X)$ compact.*

Then 0 is an approximate eigenvalue of T.

> **Proof:**   By Riesz's lemma, we can choose $(x_n) \in X^{\mathbb{N}}$ with $\|x_n\| = 1$ for all n, with $\|x_n - x_m\| \geq 1$ for all $n \neq m$.
>
> Since T is compact, we have a subsequence (x_{n_j}) with (Tx_{n_j}) convergent. In particular, $T(x_{n_i} - x_{n_{i+1}}) \to 0$ as $i \to \infty$ and $\|x_{n_i} - x_{n_{i+1}}\| \geq 1$, so 0 is an approximate eigenvalue.

Note that we do not need to have 0 an eigenvalue of T, as we can take $T \in L(\ell_2)$ with

$$T\left(\sum_{i=1}^{\infty} x_i e_i\right) = \sum_{i=1}^{\infty} \frac{x_i}{2^i} e_i.$$

However, we have the following:

Proposition 7.5. *Let X be a Banach space and $T \in L(X)$ compact. Then every $\lambda \neq 0$ that is an approximate eigenvalue of T is an eigenvalue of T.*

Proof: We have $x_1, x_2, \ldots, \in X$ with $\|x_n\| = 1$ and $Tx_n - \lambda x_n \to 0$. As T compact, there exists a subsequence (x_{n_i}) with $Tx_{n_i} \to y$ for some $y \in X$.

But then also $\lambda x_{n_i} \to y$, and $y \neq 0$ since

$$\|y\| = \lim_{i \to \infty} |\lambda| \|x_{n_i}\| = |\lambda|.$$

Hence we have $x_{n_i} \to y/\lambda$. Then from $Tx_{n_i} - \lambda x_{n_i} \to 0$, we have $T(y/\lambda) - y = 0$, i.e. $Ty = \lambda y$.

Proposition 7.6. *Let X be a Banach space, and $T \in L(X)$ compact. Then for any eigenvalue $\lambda \neq 0$, the eigenspace $E(\lambda)$ is finite dimensional.*

Proof: Suppose $E(\lambda)$ is infinite-dimensional. Then

$$T(B_X) \supset T(B_{E(\lambda)}) = \lambda B_{E(\lambda)}.$$

But $B_{E(\lambda)}$ is not totally bounded as it is the unit ball of an infinite-dimensional normed space.

Theorem 7.8. *Let X be a Banach space, and $T \in L(X)$ compact. Then for any $\delta > 0$, T has only finitely many eigenvalues λ with $|\lambda| > \delta$.*

Proof: Suppose not. Say that $\lambda_1, \lambda_2, \ldots,$ are distinct eigenvalues with $|\lambda| > \delta$, and corresponding eigenvectors $x_1, x_2, x_3, \ldots$.

Note that $\{x_1, x_2, \ldots\}$ are linearly independent, as they are eigenvectors for distinct eigenvalues. Write $X_n = \langle x_1, \ldots, x_n \rangle$.

Then by Riesz, there exists $y_n \in X_n$ with $d(y_n, X_{n-1}) \geq 1$ and $\|y_n\| = 1$. For fixed n, say that $y_n = c_1 x_1 + \cdots + c_n x_n$, then $Ty_n = c_1 \lambda_1 x_1 + \cdots + c_n \lambda_n x_n \in X_n$, and $(T - \lambda_n)y_n \in X_{n-1}$.

Now we claim that (Ty_n) has no convergent subsequence. Indeed, we have $Ty_n \in X_n$, and

$$d(Ty_n, X_{n-1}) = d(\lambda_n y_n, X_{n-1}) = |\lambda_n| \geq \delta.$$

Hence for all $m < n$, as $Ty_m \in X_{n-1}$, we have $\|Ty_n - Ty_m\| \geq \delta$, a contradiction to T being compact.

Theorem 7.9 (Spectral Theorem for Compact Operators)**.** *Let X be Banach and*

$T \in L(X)$ *compact. Then:*

(i) $\sigma(T)$ *is finite or* $\sigma(T) = \{\lambda_1, \lambda_2, \ldots\} \cup \{0\}$ *where* $\lambda_n \to 0$.

(ii) *If* $\lambda \in \sigma(T)$, $\lambda \neq 0$ *then* λ *is an eigenvalue of* T.

(iii) *If* $\lambda \in \sigma(T)$ *and* $\lambda \neq 0$, *then* $\dim E(\lambda) < \infty$.

Remark.

1. Part (iii) is a previous proposition.

2. (ii) $\implies$ (i). Indeed, from our previous theorem T has only finitely many eigenvalues with $|\lambda| > \delta$. Hence we can list $\{\lambda \in \sigma(T) \mid \lambda \neq 0\}$ as $\lambda_1, \lambda_2, \ldots$ with $\lambda_n \to 0$, by listing by decreasing $|\lambda|$.

   But now also $0 \in \sigma(T)$, as $\sigma(T)$ is closed.

Proof: We are just required to show (ii).

For any $\delta > 0$, we will show that $\lambda \in \sigma(T)$ and $|\lambda| > \delta$, then λ is an eigenvalue.

We know that if $\lambda \in \partial\sigma(T)$ and $\lambda \neq 0$, then $\lambda \in \sigma_{AP}(T)$, so λ is an eigenvalue as T is compact.

Hence $\partial\sigma(T) \cap \{z \in \mathbb{C} \mid |z| > \delta\}$ is finite, say $(\lambda_1, \ldots, \lambda_n)$. It follows that $\sigma(T) \cap \{z \in \mathbb{C} \mid |z| > \delta\} = \{\lambda_1, \ldots, \lambda_n\}$.

Otherwise, if $\lambda \in \sigma(T)$ and $\lambda \notin \partial\sigma(T)$, we can go outwards from λ in any direction, yielding uncountably many points in $\partial\sigma(T) \cap \{z \mid |z| > \delta\}$.

7.3   Spectral Theory for Compact Hermitian Operators

Putting these together, we have a nice example of compact Hermitian operators:

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \cdots \\ 0 & \lambda_2 & 0 & \ddots \\ 0 & 0 & \lambda_3 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix},$$

where λ_n are real and tend to 0. We will now show this is the only such example.

Theorem 7.10 (Spectral Theorem for Compact Hermitian Operators)**.** *Let H be a separable Hilbert space, and $T \in L(H)$ compact Hermitian. Then:*

(i) *There exists an orthonormal basis (e_n) of eigenvectors of T.*

(ii) *The corresponding eigenvalues $\lambda_n \to 0$.*

In essence, every compact Hermitian operator is diagonalisable.

Note that (i) $\implies$ (ii). Indeed, if there is a subsequence $\lambda_{n_1}, \lambda_{n_2}, \dots$ with $(\lambda_{n_i}) \geq \delta$, then the sequence (e_{n_i}) is a bounded sequence with $(Te_{n_i}) = (\lambda_{n_i} e_{n_i})$ having no convergence subsequence.

Or we can use proposition from before.

> **Proof:** For each eigenvalue λ (of which there may only be finitely many, or none), choose an orthonormal basis for $E(\lambda)$. Then the union of these orthonormal bases, say (e_n) is an orthonormal basis for
>
> $$\overline{\langle E(\lambda) \mid \lambda \text{ an eigenvalue}\rangle} = Y.$$
>
> Note that (e_n) form a sequence, possibly finite, as the set of eigenvalues is countable.
>
> We claim that $Y = H$. Indeed, note that T acts on each $E(\lambda)$, so T acts on each $E(\lambda)^\perp$, hence T acts on
>
> $$\bigcap_{\lambda \text{ eigenvalue}} E(\lambda)^\perp = Y^\perp.$$
>
> Also, $T|_{Y^\perp}$ has no eigenvectors. Now $T|_{Y^\perp}$ is Hermitian and compact (as $Y^\perp$ is closed). Thus no $\lambda \neq 0$ can be in $\sigma(T|_{Y^\perp})$.
>
> In other words, $r(T|_{Y^\perp}) = 0$. But $r(T|_{Y^\perp}) = \|T|_{Y^\perp}\|$ as $T|_{Y^\perp}$ is Hermitian.
>
> So $T = 0$ on $Y^\perp$, whence $Y^\perp = \{0\}$ as 0 is not an eigenvalue in $T|_{Y^\perp}$.

What if H is not separable?

Theorem 7.11. *Let H be a Hilbert space, and $T \in L(H)$ compact Hermitian. Then there exists a closed subspace Y, an orthonormal basis (e_n) for Y, and $\lambda_1, \lambda_2, \dots, \in \mathbb{C}$ such that*

$$T\left(\sum_{n=1}^\infty x_n e_n + z\right) = \sum_{n=1}^\infty \lambda_n x_n e_n,$$

whenever $\sum x_n e_n \in Y$, $z \in Y^\perp$.

Proof: For each eigenvalue $\lambda \neq 0$ (of which there are only countably many) we have a finite-dimensional eigenspace $E(\lambda)$.

Pick an orthonormal basis for each $E(\lambda)$ and let (e_n) be their union. Then (e_n) is an orthonormal basis for

$$\overline{\langle E(\lambda) \mid \lambda \text{ non-zero eigenvalue} \rangle} = Y.$$

Then as before, T acts on $Y^\perp$ with $T|_{Y^\perp}$ having no non-zero eigenvalues, whence

$$0 = r(T|_{Y^\perp}) = \|T|_{Y^\perp}\|.$$

Index