University of
Cambridge

Mathematics Tripos

Part II

Linear Analysis

Michaelmas, 2018

Lectures by
R. Bauerschmidt

Notes by
Qiangru Kuang

Contents

1 Normed spaces and linear operators

Unless otherwise stated, $\mathbb{K}$ stands for $\mathbb{R}$ or $\mathbb{C}$ and all vector spaces are $\mathbb{K}$-vector spaces.

1.1 Normed vector spaces

Definition (normed space)**.** A *normed vector space* $(X, \|\cdot\|)$ is a vector space X with a norm $\|\cdot\| : X \to \mathbb{R}, x \mapsto \|x\|$ satisfying

1. positive-definite: $\|x\| \geq 0$ for all $x \in X$ and $\|x\| = 0$ if and only if $x = 0$,
2. positive homogeneity: $\|\lambda x\| = |\lambda|\|x\|$ for all $\lambda \in K$ and $x \in X$,
3. triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

In particular, every norm induces a *metric* by $d(x, y) = \|x - y\|$.

Fact. Vector space operations and the norm are continuous maps, i.e. the following maps

$$
\begin{aligned}
\mathbb{K} \times X &\to X \\
(\lambda, x) &\mapsto \lambda x \\
X \times X &\to X \\
(x, y) &\mapsto x + y \\
X &\to \mathbb{R} \\
x &\mapsto \|x\|
\end{aligned}
$$

are continuous and the metric is translation invariant: $d(x, y) = d(x + z, y + z)$ for all $x, y, z \in X$.

Proof. We only check scalar multiplication. The others are left as exercises. Since $\mathbb{K}$ and X are both metric spaces, it suffices to check that $\lambda_j \to \lambda$ in $\mathbb{K}$ and $x_j \to x$ in X then $\lambda_j x_j \to \lambda x$.

Indeed,

$$
\begin{aligned}
\|\lambda_j x_j - \lambda x\| &= \|(\lambda_j - \lambda)x_j + \lambda(x_j - x)\| \\
&\leq \|(\lambda_j - \lambda)x_j\| + \|\lambda(x_j - x)\| \\
&= |\lambda_j - \lambda|\|x_j\| + |\lambda|\|x_j - x\| \\
&\to 0
\end{aligned}
$$

□

Example.

1. $\ell_n^2 = (\mathbb{R}^n, \|\cdot\|_2)$ where $\|x\|_2 = (\sum_{i=1}^n |x_i|^2)^{1/2}$, i.e. Euclidean norm.
2. $\ell_n^1 = (\mathbb{R}^n, \|\cdot\|_1)$ where $\|x\|_1 = \sum_{i=1}^n |x_i|$.
3. $\ell_n^\infty = (\mathbb{R}^n, \|\cdot\|_\infty)$ where $\|x\|_\infty = \max_i |x_i|$.

It is often useful to consider the *unit ball* $B = B(X) = \{x \in X : \|x\| \leq 1\}$. (Pictures)

Fact.

1. B determines the norm through $\|x\| = \inf\{t > 0 : x \in tB\}$.
2. B is *convex*: for all $x, y \in B, \lambda \in (0,1), \lambda x + (1-\lambda)y \in B$.

Remark. Any set $B \subseteq \mathbb{R}^n$ which is a closed, bounded, symmetric ($x \in B \implies -x \in B$) neighbourhood of 0 defines a norm by the same formula as above and B is the unit ball of that norm, although we will not use this fact in the course.

1.2 The space l^p

Let $S = \{x = (x_i)_{i=1}^\infty : x_i \in \mathbb{K}\}$ be the set of scalar sequences with

$$\begin{aligned} x + y &= (x_i)_i + (y_i)_i = (x_i + y_i)_i, \\ \lambda x &= \lambda(x_i)_i = (\lambda x_i)_i. \end{aligned}$$

Definition. For $1 \leq p < \infty$, let $\ell^p = \{x \in S : \sum_n |x_n|^p < \infty\}$ with norm $\|x\|_p = (\sum_n |x_n|^p)^{1/p}$. Let $\ell^\infty = \{x \in S : \sup_n |x_n| < \infty\}$ with norm $\|x\|_\infty = \sup_n |x_n|$. Finally, $c_0 = \{x \in S : x_n \to 0\}$ with norm $\|x\|_\infty = \sup_n |x_n|$.

We have yet proved $\|\cdot\|_p$ is a norm for general p. The triangle inequality follows from Minkowski's inequality, discussed next.

Recall that $f : \mathbb{R}^+ \to \mathbb{R}$ is *convex* if

$$f(\lambda t + (1-\lambda)s) \leq \lambda f(t) + (1-\lambda) f(s)$$

for all $s, t \in \mathbb{R}^+, \lambda \in (0,1)$. Graphically, the graph of f lies below the secant between any two points on the graph. f is concave if $-f$ is convex. Note that $\log : \mathbb{R}^+ \to \mathbb{R}$ is a concave function.

Corollary 1.1. *Let $1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$\frac{1}{p}|x|^p + \frac{1}{q}|y|^q \geq |x| \cdot |y|$$

for all $x, y \in \mathbb{K}$.

Proof. Set $t = |x|^p, s = |y|^q, \lambda = \frac{1}{p}$. Then

$$\begin{aligned} &\frac{1}{p}|x|^p + \frac{1}{q}|y|^q \geq |x||y| \\ \iff &\lambda t + (1-\lambda)s \geq t^\lambda s^{1-\lambda} \\ \iff &\log(\lambda t + (1-\lambda)s) \geq \lambda \log t + (1-\lambda)\log s \end{aligned}$$

which holds by concavity of log. $\square$

Theorem 1.2 (Hölder's inequality)**.** *Let $1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, let $x \in \ell^p, y \in \ell^q$. Then $xy = (x_n y_n)_n \in \ell^1$ and*

$$\|xy\|_1 \leq \|x\|_p \|y\|_q.$$

Proof. It suffices to assumes that $\|x\|_p = 1 = \|y\|_q$. By Hölder's inequality,

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

Take $N \to \infty$,

$$\|xy\|_1 \leq \frac{1}{p} + \frac{1}{q} = 1 = \|x\|_p \|y\|_q.$$

□

Theorem 1.3 (Minkowski's inequality)**.** *Let $1 < p < \infty$ and let $x, y \in \ell^p$. Then $x + y \in \ell^p$ and $\|x + y\|_p \leq \|x\|_p + \|y\|_p$.*

Proof. We call the power r. Have

$$\begin{aligned}
&\sum_n |x_n + y_n|^r \\
=&\sum_n |x_n + y_n|^{r-1} |x_n + y_n| \\
\leq&\sum_n |x_n + y_n|^{r-1} |x_n| + \sum_n |x_n + y_n|^{r-1} |y_n|
\end{aligned}$$

Apply Hölder's inequality for $p = \frac{r}{r-1}, q = r$ to the first term and similarly to the second term,

$$\leq \left(\sum_n |x_n + y_n|^r \right)^{\frac{r-1}{r}} \left(\sum_n |x_n|^r \right)^{\frac{1}{r}} + \left(\sum_n |x_n + y_n|^r \right)^{\frac{r-1}{r}} \left(\sum_n |y_n|^r \right)^{\frac{1}{r}}$$

Divide by both sides by a common factor, get

$$\|x + y\|_r \leq \|x\|_r + \|y\|_r.$$

□

1.3 Banach spaces

Definition (Banach space)**.** A normed vector space is a *Banach space* if it is *complete* as a metric space, i.e. every Cauchy sequence converges.

Exercise. For $1 \leq p \leq \infty$, the space ℓ^p is complete.

Example.

1. Any finite dimensional normed space is a Banach space.

2. Let S be a set and let $B(S)$ be the vector space of bounded functions on S. Then $B(S)$ is a Banach space with norm $\|f\|_\infty = \sup_{s\in S}|f(s)|$.

3. Let K be a compact Hausdorff space (for concreteness, take $[0,1]$) and let $C(K)$ be the space of continuous functions on K. Then $C(K) \subseteq B(K)$ as every continuous function on K is bounded. Moreover $C(K) \subseteq B(K)$ is closed as the uniform limit of a sequence of continuous functions is continuous. Therefore $C(K)$ is a Banach space with norm $\|\cdot\|_\infty$.

4. Let $U \subseteq \mathbb{R}^n$ be open bounded and let $C^k(\overline{U})$ be the space of functions $f : \overline{U} \to \mathbb{K}$ k-times continuously differentiable on U. Then $C^k(\overline{U})$ is a Banach space with norm
$$\|f\|_{C^k(\overline{U})} = \max_{|\alpha|\le k}\|D^\alpha f\|_\infty,$$
where
$$D^\alpha(f(x)) = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$$
where $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n, |\alpha| = \sum_{i=1}^n \alpha_i$.

5. Let X be the space of continuous functions on $[0,1]$. Then for $p \in [1,\infty)$,
$$\|f\|_p = \left(\int_0^1 |f(x)|^p dx\right)^{1/p}$$
is a norm on X. However, X is *not* complete in this norm! In IID Probability and Measure, we will show that its completion has a very nice description, namely L^p.

6. Let $D = \{z \in \mathbb{C} : |z| < 1\}$ and let $A(\overline{D})$ be the space of continuous functions $f : \overline{D} \to \mathbb{C}$ that are analytic in D. Then $A(\overline{D})$ is complete with $\|\cdot\|_\infty$ because the uniform limit of a sequence of analytic functions is analytic.

In example 3 above we used the following fact:

Fact. Let X be a normed space and $Y \le X$ a subspace. Then

1. if Y is complete then Y is closed in X.
2. if X is complete and Y is closed then Y is complete.

Proof.

1. Let $x \in \overline{Y}$. Then there is $(y_n) \subseteq Y$ such that $y_n \to x$. In particular, (y_n) is Cauchy so converges to some $y \in Y$ by completeness. Thus by uniqueness of limit $x = y \in Y$.

2. Suppose $(y_n) \subseteq Y$. Then (y_n) is Cauchy in X. By completeness there is $x \in X$ such that $y_n \to x$. Since Y is closed, in fact $x \in Y$.

□

Definition (separable)**.** A topological space is *separable* if it has a countable dense subset.

Exercise.

1. For $1 \le p < \infty$, ℓ^p is separable.
2. ℓ^∞ is not separable.
3. c_0 is separable.

1.4 Bounded operators and the dual space

Proposition 1.4. *Let X, Y be normed spaces, $T : X \to Y$ linear, then TFAE:*

1. *T is continuous.*
2. *T is continuous at 0.*
3. *T is bounded, i.e. there is $C > 0$ such that $\|Tx\| \le C\|x\|$ for all $x \in X$.*

Proof.

- $1 \implies 2$: obvious.
- $2 \implies 3$: since T is continuous at 0 and $\{y \in Y : \|y\| \le 1\}$ is a neighbourhood of $0 = T(0) \in Y$, there is $\delta > 0$ such that $\|x\| < \delta$ implies that $\|Tx\| \le 1$. For any $x \in X, x \neq 0$, by linearity,
$$\|Tx\| = \frac{\|x\|}{\delta}\|T\left(\delta\frac{x}{\|x\|}\right)\| \le \frac{\|x\|}{\delta}.$$
- $3 \implies 1$: let $\varepsilon > 0$. Set $\delta = \frac{\varepsilon}{C}$. Then $\|x - y\| < \delta$ implies that
$$\|Tx - Ty\| = \|T(x - y)\| \le C\|x - y\| \le \varepsilon$$
so T is (uniformly) continuous.

□

The infimum of such C is called

Definition (operator norm)**.** For $T : X \to Y$ bounded linear, the *operator norm* is
$$\|T\| = \|T\|_{\mathrm{op}} = \sup_{\|x\| \le 1} \|Tx\|.$$

Notation. $B(X, Y) = \{T : X \to Y \text{ bounded and linear}\}$.

Fact. $B(X, Y)$ is a normed space with norm given by the operator norm.

Proof. Let $T, S \in B(X, Y)$. Then
$$\|(T + S)x\| = \|Tx + Sx\| \le \|Tx\| + \|Sx\| \le (\|T\| + \|S\|)\|x\|$$
so $\|T + S\| \le \|T\| + \|S\|$.

The other axioms are clear. □

Example. Let $p \in (1, \infty)$.

1. Define
$$\begin{aligned} T : \ell^p &\to \ell^p \\ (x_1, x_2, \dots) &\mapsto (x_1, x_2, \dots, x_r, 0, 0, \dots) \end{aligned}$$
for some fixed $r > 0$. Then $T \in B(\ell^p, \ell^p)$ with $\|T\| = 1$.

2. Define
$$\begin{aligned} T : \ell^p &\to \ell^p \\ (x_1, x_2, \dots) &\mapsto (0, x_1, x_2, \dots), \end{aligned}$$
called the right shift operator. Then $T \in B(\ell^p, \ell^p)$ with $\|T\| = 1$. In fact, $\|Tx\| = \|x\|$ for all $x \in \ell^p$. This means that T is an isometry but not surjective.

3. Similarly define
$$\begin{aligned} S : \ell^p &\to \ell^p \\ (x_1, x_2, \dots) &\mapsto (x_2, x_3, \dots) \end{aligned}$$
with $\|S\| = 1$. Note that S is surjective but not injective. $ST = \mathrm{id} \neq TS$.

4. Let $p, q \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$. Fix $y \in \ell^q$ and define
$$\begin{aligned} \phi_y : \ell^p &\to \mathbb{K} \\ x &\mapsto (x, y) := \sum_n x_y y_n \end{aligned}$$
i.e. $\phi_y = (\cdot, y)$. By Hölder's inequality, this is well-defined and $\|\phi_y\| \leq \|y\|_q$.

5. An unbounded map: let F be the space of finite real sequences with $\|\cdot\|_1$. Define
$$\begin{aligned} T : F &\to \mathbb{R} \\ (x_1, \dots, x_n, 0, 0, \dots) &\mapsto \sum_{i=1}^{n} i x_i \end{aligned}$$
Then T is not bounded (i.e. not continuous) as $\|Te_n\| = n \to \infty$ as $n \to \infty$.

6. Define
$$\begin{aligned} T : \ell^1 &\to \ell^2 \\ x &\mapsto x \end{aligned}$$
which has $\|T\| = 1$ because $\sum_n |x_n| \leq 1$ implies $\sum_n |x_n|^2 \leq 1$. But $T\ell^1 \neq \ell^2$. Since $T\ell^1$ is also dense in ℓ^2, $T\ell^1$ is not closed in ℓ^2 and thus not complete.

Definition (isomorphism, isometric isomorphism)**.** Let X and Y be normed spaces. Then

1. an *isomorphism* from X to Y is a map $T : X \to Y$ that is a linear homeomorphism. Thus $T \in B(X, Y)$ and $T^{-1} \in B(X, Y)$, i.e. there are $C_1, C_2 > 0$ such that
$$C_1\|x\| \leq \|Tx\| \leq C_2\|x\|$$
for all $x \in X$.
2. a bijective linear map $T : X \to Y$ is an *isometric isomorphism* if $\|Tx\| = \|x\|$ for all $x \in X$.

Definition (dual space)**.** Let X be a normed space. Its *dual space* is
$$X^* = B(X, \mathbb{K}).$$
A linear map $X \to \mathbb{K}$ is called a *functional.*

Theorem 1.5. *Let X and Y be normed spaces with Y complete. Then $B(X, Y)$ is also complete. In particular $X^* = B(X, \mathbb{K})$ is complete.*

Proof. Let $(T_n)_n \subseteq B(X, Y)$ be a Cauchy sequence. Then for every $x \in X$, the sequence $(T_nx)_n \subseteq Y$ is Cauchy:
$$\|T_nx - T_mx\| = \|(T_n - T_m)x\| \leq \underbrace{\|T_n - T_m\|}_{<\frac{\varepsilon}{\|x\|}}\|x\| < \varepsilon$$
Since Y is complete, there is $y \in Y$ such that $T_nx \to y$. Set $Tx = y$. Need to check $T \in B(X, Y)$ and $\|T_n - T\| \to 0$.

- T is linear:
$$\begin{aligned} T(\lambda x + \mu y) &= \lim_{n\to\infty} T_n(\lambda x + \mu y) \\ &= \lim_{n\to\infty} (\lambda T_nx + \mu T_ny) \\ &= \lambda Tx + \mu Ty \end{aligned}$$
- T is bounded: for $\|x\| \leq 1$,
$$\begin{aligned} \|Tx\| &\leq \|T_nx\| + \|(T_n - T)x\| \\ &\leq \|T_n\| + \varepsilon \\ &\leq \sup_n \|T_n\| \end{aligned}$$
which is bounded as $(T_n)_n$ is Cauchy.
- $T_n \to T$ in operator norm: for $\|x\| \leq 1$,
$$\begin{aligned} \|T_nx - Tx\| &\leq \|(T_n - T_m)x\| + \|(T_m - T)x\| \\ &\leq \|T_n - T_m\| + \varepsilon \\ &\leq \limsup_{n\to\infty} \|T_n - T_m\| \\ &\to 0 \end{aligned}$$

□

Theorem 1.6. *Let $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then the map*

$$\begin{aligned} \phi : \ell^q &\to (\ell^p)^* \\ y &\mapsto \phi_y = (\cdot, y) \end{aligned}$$

is an isometric isomorphism, i.e. $\ell^q = (\ell^p)^*$.

Proof. Clearly ϕ is linear. We have already seen that $\|\phi_y\| \leq \|y\|_q$. Claim that $\|\phi_y\| \geq \|y\|_q$: note that LHS is a supremum so suffices to find $\|x\| \leq 1$ such that $|\phi_y(x)| \geq \|y\|_q$. Take

$$x_n = \begin{cases} |y_n|^{q/p-1}\overline{y}_n & y_n \neq 0 \\ 0 & y_n = 0 \end{cases}$$

Then

$$\|x\|_p^p = \sum_n |x_n|^p = \sum_n |y_n|^q = \|y\|_q^q < \infty$$

so $x \in \ell^p$. We have

$$\phi_y(x) = (x, y) \geq \sum_n |y_n|^{q/p+1} = \sum_n |y_n|^q = \|y\|_q^q = \|y\|_q \|y\|_q^{q-1}.$$

Note that

$$\|y\|_q^{q-1} = \|x\|_p^{\frac{p}{q}(q-1)} = \|x\|_p^{p(1-\frac{1}{q})} = \|x\|_p$$

so

$$|\phi_y(x)| = \|y\|_q \|x\|_p.$$

Thus ϕ is an isometry. It remains to check that ϕ is surjective. Let $T \in (\ell^p)^*$. Set $y_n = Te_n$. Claim that $y \in \ell^q$ and $\|y\|_q \leq \|T\|$: define

$$x_n = \begin{cases} |y_n|^{q/p-1}\overline{y}_n & n \leq N \text{ and } y_n \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

We want to proceed as before but we don't know if $y \in \ell^q$ this time so we only take the first N terms. Then $\|x\|_p^p = \sum_{n=1}^N |y_n|^q$ so $x \in \ell^p$ and

$$Tx = \sum_{n=1}^N x_n Te_n = \sum_{n=1}^N x_n y_n = \sum_{n=1}^N |y_n|^q$$

Rewrite the equation backward,

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

so

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

so $\|y\|_q \leq \|T\|$.

Finally, claim that $T = \phi_y$: for all n, by construction we know

$$Te_n = \phi_y(e_n) = y_n.$$

Since T and ϕ_y are both continuous and linear, $T = \phi_y$ on the span of $\{e_n : n \geq 1\}$ which is just ℓ^p. $\square$

Remark. Similarly, $(\ell^1)^* = \ell^\infty$ and $c_0^* = \ell^1$ by the same argument. But the argument does *not* show $(\ell^\infty)^* = \ell^1$ since $\{e_n\}$ is not dense in ℓ^∞, i.e. it is not separable.

Corollary 1.7. *For $1 \leq p \leq \infty$, ℓ^p is complete.*

1.5 Finite-dimensional vector spaces

Fact. Any finite-dimensional vector space can be identified with $\mathbb{K}^n$ by choosing a basis. Here n is the dimension.

Definition (equivalent norm)**.** Two norms $\|\cdot\|$ and $\|\cdot\|'$ on a vector space X are *equivalent* if there exists $C > 0$ such that

$$C^{-1}\|x\|' \leq \|x\| \leq C\|x\|',$$

i.e. $\mathrm{id} : (X, \|\cdot\|) \to (X, \|\cdot\|')$ is an isomorphism.

Theorem 1.8. *Let X be a finite-dimensional vector space. Then all norms on X are equivalent.*

Proof. It suffcies to show that any norm $\|\cdot\|$ on $\mathbb{K}^n$ is equivalent to $\|\cdot\|_2$. Claim that $\|x\| \leq C\|x\|_2$ for all $x \in \mathbb{K}^n$:

$$\|x\| = \left\|\sum_{i=1}^n x_i e_1\right\| \leq \sum_{i=1}^n |x_i|\|e_i\| \leq \underbrace{n \max_i \|e_i\|}_{C} \underbrace{\max_i |x_i|}_{\leq \|x\|_2}$$

Also claim that $\|x\|_2 \leq C'\|x\|$ for all x: let

$$S = \{x : \|x\|_2 = 1\}$$

and define $f = \|\cdot\||_S : S \to \mathbb{R}$. Then f is continuous (with respect to $\|\cdot\|_2$):

$$|f(x) - f(y)| = |\|x\| - \|y\|| \leq \|x - y\| \leq C\|x - y\|_2$$

Note also that S is compact (with respect to $\|\cdot\|_2$) as it is closed and bounded. Therefore f assumes its minimum on S, i.e. there exists $\delta > 0$ such that $f(x) \geq \delta$ for all $x \in S$. Then for all $x \in \mathbb{K}^n$, have

$$\|x\| = \left\|\frac{x}{\|x\|_2}\right\| \cdot \|x\|_2 = f(\frac{x}{\|x\|_2}) \cdot \|x\|_2 \geq \delta\|x\|_2.$$

$\square$

Corollary 1.9. *Let X and Y be normed spaces with $\dim X < \infty$. Then every linear map $T : X \to Y$ is continuous.*

Proof. Define a new norm on X by

$$\|x\|' = \|x\| + \|Tx\|.$$

Since all norms on X are equivalent, there is $C > 0$ such that

$$\|x\|' \leq C\|x\|,$$

i.e. $\|Tx\| \leq (C-1)\|x\|$ for all x. Thus T is bounded and thus continuous. $\square$

Corollary 1.10. *Let X and Y be finite-dimensional vector spaces and $T : X \to Y$ is a linear bijection. Then T is an isomorphism.*

In particular for any X and Y if $\dim X = \dim Y < \infty$ then X and Y are isomorphic.

Corollary 1.11.

1. *Every finite-dimensional normed space is complete (as it is true in $\|\cdot\|_2$).*
2. *Every finite-dimensional subspace of a normed space is closed.*

Corollary 1.12. *Let X be a finite-dimensional normed space. Then $\overline{B}(X)$, the closed unit ball, is compact.*

Proof. Closed and bounded in $\|\cdot\|_2$ (because this holds in $\|\cdot\|$). Thus $\overline{B}(X)$ is compact in $\|\cdot\|_2$ so compact in $\|\cdot\|$. $\square$

The converse is also true:

Theorem 1.13. *Let X be a normed space such that $\overline{B}(X)$ is compact. Then X is finite-dimensional.*

Proof. Since $\overline{B}_1(0) = \overline{B}(X)$ is compact, there are $x_1, \dots, x_n \in X$ such that

$$\overline{B}_1(0) \subseteq \bigcup_{i=1}^{n} B_{1/2}(x_i).$$

Let Y be the span of x_i's. Then $\dim Y \leq n$. Also

$$B_1(0) \subseteq Y + B_{1/2}(0)$$

so

$$B_1(0) \subseteq Y + \frac{1}{2}(Y + B_{1/2}(0)) = Y + B_{1/4}(0) \subseteq \cdots \subseteq Y + B_{2^{-m}}(0)$$

for all $m \in \mathbb{N}$. Therefore $B_1(0) \subseteq \overline{Y} = Y$. Since Y is linear, $X \subseteq Y$. Thus $\dim X \leq n$. $\square$

1.6 Completion, products, quotients

1.6.1 Completion

Proposition 1.14. *Let X be a metric space. The* completion *of X is a complete metric space $\tilde{X}$ containing a dense subset that is isometric to X.*

Proof. The construction is as follow. For two Cauchy sequences $x = (x_n), y = (y_n) \subseteq X$, define $x \sim y$ if and only if $d(x_n, y_n) \to 0$. This is an equivalence relation. Denote the equivalence class of a Cauchy sequence x by $\tilde{x}$. Define

$$\tilde{X} = \{\tilde{x} : x \text{ Cauchy in } X\}$$

and define a metric

$$\tilde{d}(\tilde{x}, \tilde{y}) = \lim_{n\to\infty} d(x_n, y_n).$$

The limit exists and is independent of the representatives. Then $\tilde{d}$ is a metric: if $\tilde{d}(\tilde{x}, \tilde{y}) = 0$ then $d(x_n, y_n) \to 0$ so $x \sim y$ so $\tilde{x} = \tilde{y}$. Symmetry and triangle inequality follow from those for d.

Now we show $X \hookrightarrow \tilde{X}$. For $x \in X$, define $j(x) \in \tilde{X}$ to be the equivalence class of $(x, x, \dots)$. Then

$$\tilde{d}(j(x), j(y)) = d(x, y)$$

so j is an isometry. The image of j is dense in $\tilde{X}$ since if (x_n) is Cauchy in X then $(j(x_n))$ in $\tilde{X}$ is Cauchy and $j(x_n) \to \tilde{x}$.

Finally, to show $\tilde{X}$ is complete, let $(\tilde{x}^k) \subseteq \tilde{X}$ be Cauchy. Let $(x_n^k) \subseteq X$ be a representative for $\tilde{x}^k$. Choose n_k such that $d(x_n^k, x_m^k) \le 2^{-k}$ for $n, m \ge n_k$. Define $x_k = x_{n_k}^k \in X$. Claim that $x = (x_k) \subseteq X$ is Cauchy and $\tilde{x}^k \to \tilde{x}$ in $\tilde{X}$. It is left as an exercise. □

Definition (completion)**.** $\tilde{X}$ is called the *completion* of X and we regard $X \subseteq \tilde{X}$.

In the case of normed spaces, the metric completion has more structure:

Theorem 1.15. *Let X be a normed space. Then there is a Banach space $\tilde{X}$ containing X as a dense subspace.*

Proof. Let $\tilde{X}$ be the metric space completion of X. For $\tilde{x}, \tilde{y} \in \tilde{X}$, choose $(x_n), (y_n) \subseteq X$ such that $x_n \to \tilde{x}$ and $y_n \to \tilde{y}$ (in $\tilde{X}$). For any $\lambda, \mu \in \mathbb{K}$, $\lambda x_n + \mu y_n$ is Cauchy. Set $\lambda\tilde{x} + \mu\tilde{y} = \lim_{n\to\infty}(\lambda x_n + \mu y_n)$. This makes $\tilde{X}$ a vector space. Moreover,

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

is a norm on $\tilde{X}$ and since $\tilde{d}$ is complete, this makes $\tilde{X}$ a Banach space. □

Proposition 1.16. *Let X and Y be normed spaces and let $T \in B(X, Y)$.*

Then there is a unique $\tilde{T} \in B(\tilde{X}, \tilde{Y})$ *such that*

$$\tilde{T}|_X = T, \|\tilde{T}\| = \|T\|.$$

Proof. For $\tilde{x} \in \tilde{X}$, choose $(x_n) \subseteq X$ such that $x_n \to \tilde{x}$. Then (x_n) is Cauchy, and since T is bounded, $(Tx_n) \subseteq \tilde{Y}$ is Cauchy as well. By completeness of $\tilde{Y}$, there is $\tilde{y} \in \tilde{Y}$ such that $Tx_n = \tilde{y}$. Set $\tilde{T}\tilde{x} = \tilde{y}$. Note that $\tilde{T}$ is well-defined, linear and $\tilde{T}|_X = T$. Also

$$\|\tilde{T}\tilde{x}\| = \lim_{n\to\infty} \|Tx_n\| \leq \|T\| \lim_{n\to\infty} \|x_n\| = \|T\|\|\tilde{x}\|,$$

so $\|\tilde{T}\| \leq \|T\|$ so equality. Uniqueness follows from continuity and density of X in $\tilde{X}$. □

Remark. The completion $\tilde{X}$ is unique in the sense that if $\tilde{X}'$ is another completion of X then there is an isometric isomorphism $\tilde{X} \to \tilde{X}'$ restricting to identity on X.

1.6.2 Product

Definition (product)**.** Let X and Y be normed spaces. Then $X \times Y$ can be made into a normed space with one of the following equivalent norms:

$$\|(x, y)\| = (\|x\|^p + \|y\|^p)^{1/p}, p \in [1, \infty)$$

or

$$\|(x, y)\| = \max\{\|x\|, \|y\|\}.$$

They are equivalent for precisely the same reason that norms on finite-dimensional spaces are equivalent. Thus henceforth we will just use "norm on $X \times Y$" to mean any of the equivalent norms.

As expected for a product construction, the projections $\pi_X : X \times Y \to X$, and $\pi_Y : X \times Y \to Y$ are continuous.

Fact. If X and Y are complete then $X \times Y$ is complete and $X \cong X \times \{0\} \subseteq X \times Y$ and $Y \cong \{0\} \times Y \subseteq X \times Y$ are closed subspaces.

1.6.3 Quotient

Definition (quotient)**.** Let X be a normed space and let $Y \subseteq X$ be a *closed* subspace. Then $x \sim x'$ if $x - x' \in Y$ defines an equivalence relation with equivalence classes $[x] = x + Y$. Let X/Y be the collection of all equivalence classes and define

$$\|[x]\| = \inf_{y\in Y} \|x + y\|.$$

Proposition 1.17.

1. $\|\cdot\|$ *is a norm on* X/Y.
2. $\pi : X \to X/Y, x \mapsto x + Y$ *is continuous.*

3. If X is complete then X/Y is complete.

Proof.

1. Suffices to show positive definiteness as the other axioms are trivial. Assume that
$$\|\pi(x)\| = \inf_{y\in Y}\|x+y\| = 0,$$
then there exists $(x_n) \subseteq X$ such that $x_n \to 0$ and $\pi(x_n) = \pi(x)$, i.e. $x - x_n \in Y$. Thus $x \in \overline{Y} = Y$.

2. $$\|\pi(x)\| = \inf_{y\in Y}\|x+y\| \le \|x\|$$
so $\|\pi\| \le 1$.

3. Let $(x_n) \subseteq X$ be such that $\pi(x_n) \subseteq X/Y$ is Cauchy. Claim that there exists a subsequence $(n_k) \subseteq \mathbb{N}$ and $(y_n) \subseteq Y$ such that $(x_{n_k} + y_k)_k$ is Cauchy in X: by passing to a subsequence (n_k) we can assume that $\|\pi(x_{n_{k+1}}) - \pi(x_{n_k})\| \le 2^{-k-1}$. Now choose $(z_k) \subseteq Y$ such that
$$\|x_{n_{k+1}} - x_{n_k} + z_k\| \le 2^{-k}.$$
Define $y_1 = 0$ and $y_k = z_1 + \cdots + z_{k-1} \in Y$, then
$$\|(x_{n_{k+1}} + y_{k+1}) - (x_{n_k} + y_k)\| = \|x_{n_{k+1}} - x_{n_k} + z_k\| \le 2^{-k}$$
so $(x_{n_k} + y_k)_k$ is Cauchy.
Claim that $(\pi(x_n)) \subseteq X/Y$ converges: since X is complete, there is $x \in X$ such that $x_{n_k} + y_k \to x$ in X. Then
$$\|\pi(x_{n_k}) - \pi(x)\| = \inf_{y\in Y}\|x - (x_{n_k} + y)\| \le \|x - (x_{n_k} + y_k)\| \to 0$$
which implies convergence of $(\pi(x_n))$ along a subsequence, ergo the original sequence.

□

2 Completeness of the Baire category

2.1 Baire category

Recall that if X is a metric space, then $Y \subseteq X$ is *dense* if $\overline{Y} = X$, i.e. $Y \cap B_r(x) \neq \emptyset$ for all $x \in X, r > 0$.

Theorem 2.1 (Baire category theorem)**.** *Let X be a* complete *metric space. For any sequence of open dense subsets $U_j \subseteq X$, $j \in \mathbb{N}$, the intersection $\bigcap_{j=1}^{\infty} U_j$ is dense in X.*

Proof. Let $U = \bigcap_{j=1}^{\infty} U_j$. Given any $x \in X, r > 0$, we need to show that $B_r(x) \cap U \neq \emptyset$. Since U_1 is dense, there is $x_1 \in X, r \in (0,1)$ such that

$$\overline{B}_{r_1}(x_1) \subseteq B_{2r_1}(x_1) \subseteq U_1 \cap B_r(x).$$

Likewise choose $x_2 \in X, r_2 \in (0, \frac{1}{2})$ such that

$$\overline{B}_{r_2}(x_2) \subseteq U_2 \cap B_{r_1}(x_1)$$

and in general $x_n \in X, r \in (0, \frac{1}{n})$ such that

$$\overline{B}_{r_n}(x_n) \subseteq U_n \cap B_{r_{n-1}}(x_{n-1}).$$

Then $r_n \to 0$ and a nested chain of open balls

$$B_{r_1}(x_1) \supseteq B_{r_2}(x_2) \supseteq \dots$$

so $d(x_n, x_m) < r_n$ if $m \geq n$, i.e. (x_n) is a Cauchy sequence. Since X is complete, there is $y \in X$ such that $x_n \to y$. Note that $y \in \overline{B}_{r_k}(x_k) \cap U_k$ for all k. Thus

$$y \in \bigcap_{j=1}^{\infty} U_j = U$$

and $y \in \overline{B}_{r_1}(x_1) \subseteq B_r(x)$ so $y \in U \cap B_r(x)$. □

The following corollary is equivalent to Baire category theorem is often used in practice:

Corollary 2.2. *Let X be a complete metric space. Let $A_j \subseteq X$ be a sequence of closed subsets such that $\bigcup_j A_j$ has nonempty interior, i.e. it contains some ball, then at least one of the A_j's has nonempty interior.*

Proof. Let $U_j = X \setminus A_j$. Since $\bigcup_j A_j$ has nonempty interior,

$$X \setminus \bigcup_j A_j = \bigcap_j U_j$$

is not dense. Since the U_j's are open, by Theorem 2.1 at least one of the U_j's cannot be dense, say U_k. Thus $A_k = X \setminus U_k$ has nonempty interior. □

Definition (nowhere dense, meagre, residual, set of first/second category)**.** Let X be a metric space.

1. A subset $Y \subseteq X$ is *nowhere dense* if $\mathrm{Int}(\overline{Y}) = \emptyset$, i.e. if Y is not dense in any ball.
2. A subset $Z \subseteq X$ is *meagre* or *of the first category* if there are countably many sets $Y_j \subseteq X$ which are nowhere dense and $Z = \bigcup_j U_j$.
3. A subset $U \subseteq X$ is *nonmeagre* or *of the second category* if it is not meagre.
4. A subset $R \subseteq X$ is *residual* if its complement is meagre.

Remark. TFAE:

- $Y \subseteq X$ is nowhere dense.
- $\overline{Y}$ is nowhere dense.
- $X \setminus \overline{Y}$ is dense.

Example.

1. $\mathbb{Q} = \bigcup_{x \in \mathbb{Q}} \{x\} \subseteq \mathbb{R}$ is meagre in $\mathbb{R}$.
2. Any countable union of meagre sets is meagre.

Remark. There is a similarity of the concepts of meagre, nonmeagre, residual, with those of null sets, sets of positive measure, sets of full measure in measure theory. For metric spaces that are also measure spaces, such as $\mathbb{R}$ with Lebesgue measue, one could ask if there is a closer correspondence. The answer is negative, in general. There exists a meagre set A and a Lebesgue null set B such that $R = A \cup B$.

Yet another formulation of Baire category theorem is

Corollary 2.3. *Let X be a complete metric space. Then X is of the second category.*

Proof. Let $Y_j \subseteq X$ be nowhere dense. It suffices to show that $X \neq \bigcup_j \overline{Y}_j$. But $U_j = X \setminus \overline{Y}_j$ is open dense so by Theorem 2.1

$$\bigcap_j U_j = X \setminus \bigcup_j \overline{Y}_j$$

is dense, in particular not empty. □

Corollary 2.4. *Let X be a complete metric space. Then residual sets are nonmeagre and dense.*

Proof. Let $Z \subseteq X$ be meagre and suppose that $R = X \setminus Z$ was meagre. Then $X = Z \cup R$ would be meagre as a union of two meagre sets. But since X is complete, it is not. So R is nonmeagre.

To show that R is dense, we can suppose $Z = \bigcup_j Y_j$ with Y_j nowhere dense. Then $U_j = X \setminus \overline{Y}$ is open dense. So $R \supseteq \bigcup_j U_j$ is dense by Theorem 2.1. □

Corollary 2.5. *Let X be a complete metric space and $U \subseteq X$ open. Then $U = \emptyset$ or U is of the second category.*

Proof. Assume that U is open and meagre. Then $X \setminus U$ is closed and residual so dense. So $X \setminus U = X$, i.e. $U = \emptyset$. □

2.2 Principle of uniform boundedness

Theorem 2.6 (principle of uniform boundedness)**.** *Let X be a complete metric space. Let $(f_\lambda)_{\lambda\in\Lambda}$ be a family of continuous functions $f_\lambda : X \to \mathbb{R}$. If $(f_\lambda)_{\lambda\in\Lambda}$ is pointwise bounded, i.e. for all $x \in X$, $\sup_{\lambda\in\Lambda} |f_\lambda(x)| < \infty$, then there is a ball $B_r(x_0) \subseteq X$ on which f_λ is uniformly bounded, i.e.*

$$\sup_{\lambda\in\Lambda} \sup_{x\in B_r(x_0)} |f_\lambda(x)| < \infty.$$

Proof. Let

$$A_k = \{x \in X : |f_\lambda(x)| \leq k \text{ for all } \lambda \in \Lambda\} = \bigcap_{\lambda\in\Lambda} \{x \in X : |f_\lambda(x)| \leq k\}.$$

Since f_λ's are continuous, A_k is closed. Since (f_λ) is pointwise bounded,

$$\bigcup_{k\in\mathbb{N}} A_k = X.$$

By Baire category theorem, at least one of the A_k's must contain a ball $B_r(x_0)$. Thus (f_λ) is uniformly bounded on that ball. □

Theorem 2.7 (Banach-Steinhaus)**.** *Let X be a Banach space and let Y be a normed space. Let $(T_\lambda)_{\lambda\in\Lambda} \subseteq B(X,Y)$ be pointwise bounded, i.e. for all $x \in X$, $\sup_{\lambda\in\Lambda} \|T_\lambda x\| < \infty$. Then (T_λ) is uniformly bounded, i.e.*

$$\sup_{\lambda\in\Lambda} \|T_\lambda\| < \infty.$$

Proof. Set $f_\lambda : X \to \mathbb{R}, x \mapsto \|T_\lambda x\|$. Then f_λ is continuous and (f_λ) is pointwise bounded. By the principle of uniform boundedness, there is $B_r(x_0) \subseteq X$ on which

$$\sup_{\lambda\in\Lambda} \sup_{\|x-x_0\|<r} \|T_\lambda x\| < \infty.$$

But since the T_λ's are linear, for any $x \in X$ with $\|x\| \leq 1$,

$$\|T_\lambda x\| = \frac{1}{r}\|T_\lambda(rx + x_0) - T_\lambda(x_0)\| \leq \frac{1}{r} \sup_{\lambda\in\Lambda} \sup_{\|x-x_0\|<r} \|T_\lambda x\| + \frac{1}{r} \sup_{\lambda\in\Lambda} \|T_0 x\|.$$

The second term is bounded since T_λ is pointwise bounded. Thus

$$\sup_{\lambda\in\Lambda} \|T_\lambda\| < \infty.$$

□

The point of the Baire category theorem is not so much of finding a uniform bound on the functionals, as the proof requires axiom of choice and is nonconstructive. Rather it shows that pointwise boundedness implies uniform boundedness so we don't risk losing anything by trying to prove uniform boundedness from onset. As we'll see, in most cases pointwise bound gives uniform bound straightaway.

2.3 Open mapping theorem

Definition (open map)**.** A map between topological spaces is *open* if it maps open sets to open sets.

Example.

1. $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ is continuous but not open.
2. $f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x + \operatorname{sgn}(y)$ is open but not continuous.

Theorem 2.8 (open mapping theorem)**.** *Let X, Y be Banach spaces and $T \in B(X, Y)$. Then*

1. *if T is surjective then it is open.*
2. *if T is bijective then $T^{-1} \in B(Y, X)$.*

Lemma 2.9. *Let X, Y be normed spaces. Then $T : X \to Y$ linear is open if*

$$T(B_1(0)) \supseteq B_r(0)$$

for some $r > 0$.

Proof. Let $U \subseteq X$ be open and $x \in U$. As U is open, choose $\delta > 0$ such that $x + B_\delta(0) \subseteq U$. Then

$$T(U) \supseteq T(x + B_\delta(0)) = Tx + \delta T(B_1(0)) \supseteq Tx + r\delta B_1(0).$$

Thus $T(U)$ contains an open ball around any element $T(x)$, therefore open. □

Lemma 2.10. *Let X be a Banach space, Y a normed space and $T \in B(X, Y)$. If*

$$\overline{T(B_1(0))} \supseteq B_1(0).$$

then

$$T(B_1(0)) \supseteq B_1(0).$$

Proof. Let $y_0 \in B_1(0) \subseteq Y$. We need to find $x \in B_1(0)$ such that $Tx = y_0$. We construct x as the limit of a Cauchy sequence. Let $x_1 \in B_{1/2}(0) \subseteq X$ such that

$$\|Tx_1 - y_0\| < \frac{1}{2}.$$

This is possible since there exists $\tilde{y}_0 \in B_{1/2}(0) \cap B_{1/2}(y_0)$ and we can find $x_1 \in B_{1/2}(0)$ such that $\|Tx_1 - \tilde{y}_0\|$ is arbitrarily small by density of $T(B_{1/2}(0))$ in $B_{1/2}(0)$.

Set $y_1 = y_0 - Tx_1 \in B_{1/2}(0)$. By induction, if $y_1, \dots, y_k$ and $x_1, \dots, x_k$ are such that

$$\|x_i\| < 2^{-i}, y_i = y_{i-1} - Tx_i \in B_{2^{-i}}(0) \subseteq Y,$$

can choose $x_{k+1} \in B_{2^{-k-1}}(0) \subseteq X$ such that

$$y_{k+1} = y_k - Tx_{k+1} \in B_{2^{-k-1}}(0) \subseteq Y$$

so

$$\sum_{k=1}^{\infty} \|x_k\| < 1$$

and $x = \sum_{k=1}^{\infty} x_k \in B_1(0)$ exists since X is complete and

$$y_0 - Tx = \lim_{n\to\infty} \left(y_0 - \sum_{k=1}^{n} Tx_k \right) = \lim_{n\to\infty} \left(y_1 - \sum_{k=2}^{n} Tx_k \right) = \dots = \lim_{n\to\infty} y_n = 0$$

so $y_0 \in T(B_1(0))$ for any $y_0 \in B_1(0)$. Thus $T(B_1(0)) \supseteq B_1(0)$. □

Proof of open mapping theorem.

1. By the previous two lemmas, it suffices to show that

$$\overline{T(B_1(0))} \supseteq B_r(0)$$

for some $r > 0$. We use Baire category theorem to do this. Since T is surjective, $Y = \bigcup_{k\geq 1} \overline{T(B_k(0))}$. Since Y is complete, the Baire category theorem implies that there is $k_0 \in \mathbb{N}$ such that $\overline{T(B_{k_0}(0))}$ has nonempty interior, i.e. there is $r_0 > 0, y_0 = Tx_0$ such that

$$B_{r_0}(y_0) \subseteq \overline{T(B_{k_0}(0))}.$$

By linearity,

$$\begin{aligned} B_{r_0}(0) = B_{r_0}(y_0) - Tx_0 &\subseteq \overline{T(B_{k_0}(0))} - Tx_0 \\ &= \overline{T(B_{k_0}(-x_0))} \subseteq \overline{T(B_{k_0+\ell_0}(0))} = (k_0 + \ell_0)\overline{T(B_1(0))} \end{aligned}$$

where $\ell_0 \geq \|x_0\|$. Now take $r = \frac{r_0}{k_0+\ell_0}$.

2. If T is bijective, then T is open means that T^{-1} is continuous.

□

Remark. The completeness of X and Y are both necessary. See example sheet. We can however do a quick counterexample here. Let

$$F = \{(x_n) : x_n = 0 \text{ except for finitely many } n\}$$

with $\|x\|_\infty = \max_n |x_n|$. Define

$$\begin{aligned} T : F &\to F \\ (x_n) &\mapsto (x_n/n) \end{aligned}$$

Then $\|T\| \leq 1$ so T is continuous and bijective. But $(T^{-1}x)_n = (nx_n)$ is unbounded. In particular F is not complete.

Remark. The basic problem in linear PDE is the following one: given $f \in Y$, for example $Y = L^2(\Omega)$ for some nice $\Omega \subseteq \mathbb{R}^d$, and a linear partial differential operator $L : X \to Y$, say $X = H_0^2(\Omega)$ and $L = \Delta$, is there a unique solution $u \in X$ to $Lu = f$? The typical procedure is to show that for f "nice", say $f \in C^\infty(\Omega)$, spanning a dense subspace of Y, there is a unique solution such that

$$\|u\| \leq C\|f\|.$$

Such an *a priori* estimate allows us to solve $Lu = f$ for general $f \in Y$ by approximation. This implies that L is surjective. The open mapping theorem guarantees that this strategy works if L is surjective.

2.4 Closed graph theorem

Theorem 2.11 (closed graph theorem)**.** *Let X, Y be Banach spaces and $T : X \to Y$ linear. Then T is bounded if and only if the graph*

$$\Gamma = \{(x, Tx) : x \in X\} \subseteq X \times Y$$

is closed.

Proof. Let T be bounded and $(x_k, y_k) \subseteq \Gamma$ be a sequence such that $x_k \to x, y_k = Tx_k \to y$. Since T is continuous,

$$Tx = y,$$

so $(x, y) \in \Gamma$. So Γ is closed.

Conversely, suppose that Γ is closed. We want to show that T is continuous. Since $X \times Y$ is a Banach space with norm

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

and since Γ is closed, it is also a Banach space with the induced norm. The projections

$$\begin{aligned} \pi_X : \Gamma &\to X \\ (x, Tx) &\mapsto x \\ \pi_Y : \Gamma &\to Y \\ (x, Tx) &\mapsto Tx \end{aligned}$$

are continuous and π_X is also a bijection. By the open mapping theorem, $\pi_X^{-1} \in B(X, \Gamma)$. Thus

$$T = \pi_Y \circ \pi_X^{-1} \in B(X, Y).$$

□

Remark. As a consequence, to prove that $T : X \to Y$ is bounded, if X and Y are Banach spaces, it suffices to check if $x_k \to x, Tx_k \to y$ then $Tx = y$, instead of the stronger requirement that if $x_k \to x$ then $Tx_k \to y$ and $y = Tx$.

3 Continuous functions on a compact space

3.1 Normal topological spaces

Recall that a topological space X is *Hausdorff* if for any $x, y \in X, x \neq y$, there exist open neighbourhoods U of x and V of y such that $U \cap V \neq \emptyset$.

Proposition 3.1. *Let X be a Hausdorff space and $K_1, K_2 \subseteq X$ are compact sets with $K_1 \cap K_2 = \emptyset$. Then there exist open $U_1 \supseteq K_1, U_2 \supseteq K_2$ such that $U_1 \cap U_2 = \emptyset$.*

Proof. This is a mundane exercise in general topology. For any $x \in K_1, y \in K_2$, let U_{xy} and V_{xy} be open neighbourhoods such that $x \in U_{xy}, y \in V_{xy}$ and $U_{xy} \cap V_{xy} = \emptyset$. Then $K_1 \subseteq \bigcup_{x \in K_1} U_{xy}$. Since K_1 is compact, there are finitely many points $x_1, \dots, x_n \in K_1$ such that $K_1 \subseteq \bigcup_{i=1}^n U_{x_i y}$. Set $U_y = \bigcup_{i=1}^n U_{x_i y}$ and $Y_y = \bigcap_{i=1}^n V_{x_i y}$. Then $U_y \cap V_y = \emptyset$ and $K_1 \subseteq U_y, y \in V_y$ for all $y \in Y$. Then $K_2 \subseteq \bigcup_{y \in K_2} V_y$. Again by compactness there exist $y_1, \dots y_m \in K_2$ such that $K_2 \subseteq \bigcup_{i=1}^m V_{y_i}$. Set

$$V = \bigcup_{i=1}^m V_{y_i}, U = \bigcap_{i=1}^m U_{y_i}.$$

The sets U and V are open, $U \cap V = \emptyset$ and $U \supseteq K_1, V \supseteq K_2$ by construction. □

Definition (normal)**.** A topological space X is *normal* if for any closed sets $A_1, A_2 \subseteq X$ such that $A_1 \cap A_2 = \emptyset$, there exist open sets $U_1, U_2 \subseteq X$ such that $A_1 \subseteq U_1, A_2 \subseteq U_2$ and $U_1 \cap U_2 = \emptyset$.

Corollary 3.2. *Any compact Hausdorff space is normal.*

Proof. Closed subsets of a compact space are compact. □

Fact. Let X be normal. Then for every closed $A \subseteq X$ and open $U \supseteq A$, then there exists an open set V and closed set B such that

$$A \subseteq V \subseteq B \subseteq U.$$

Proof. Set $A' = X \setminus U$. Then A' and A are closed and disjoint so there exist open sets V and V' such that $V \supseteq A, V' \supseteq A'$ and $V \cap V' = \emptyset$. Take $B = X \setminus V'$, then $A \subseteq V \subseteq B \subseteq U$. □

Proposition 3.3 (Urysohn's lemma)**.** *Let X be normal. For every closed set $A \subseteq X$ and open set $U \supseteq A$, there is a continuous function $f : X \to [0, 1]$ such that*

$$f(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin U \end{cases}$$

Proof. Let $A_1 = A$ and $U_0 = U$. Since $A_1 \subseteq U_0$ there exists an open set $U_{1/2}$ and a closed set $A_{1/2}$ such that

$$A_1 \subseteq U_{1/2} \subseteq A_{1/2} \subseteq U_0.$$

Applying this procedure again, there are open $U_{1/4}, U_{3/4}$ and closed $A_{1/3}, A_{3/4}$ such that

$$A_1 \subseteq U_{3/4} \subseteq A_{3/4} \subseteq U_{1/2} \subseteq A_{1/2} \subseteq U_{1/4} \subseteq A_{1/4} \subseteq U_0.$$

Iterating this procedure, there exist open sets U_q and closed sets A_q for dyadic $q \in \{m2^{-n} : m, n \in \mathbb{N}, 0 < m < 2^n\}$ such that for all $q < q'$,

$$U_{q'} \subseteq A_{q'} \subseteq U_q \subseteq A_q.$$

Define

$$f(x) = \sup\{q : x \in U_q\} = \inf\{q : x \notin A_q\}$$

(where $\inf \emptyset = 1, \sup \emptyset = 0$). Clearly $0 \leq f \leq 1$. If $x \notin U = U_0$ then $f(x) = 0$. If $x \in A = A_1$ then $x \in U_q$ for all q so $f(x) = 1$. To show continuity, note that for any $t \in \mathbb{R}$,

$$\begin{aligned} \{x : f(x) > t\} &= \bigcup_{q>t} U_q \\ \{x : f(x) < t\} &= \bigcup_{q<t} X \setminus A_q \end{aligned}$$

both of which are open. Thus f is continuous. $\square$

Corollary 3.4. *Let X be normal and $A_0, A_1 \subseteq X$ closed and disjoint. Then there exists $f : X \to [0,1]$ continuous such that $f|_{A_0} = 0$ and $f|_{A_1} = 1$.*

Proof. Take $A = A_1$ and $U = X \setminus A_0$ in Urysohn's lemma. $\square$

Corollary 3.5. *Let K be a compact Hausdorff space. Then $C(K)$ separates points, i.e. for all $x, y \in K$, $x \neq y$, there is $f \in C(K)$ such that $f(x) \neq f(y)$.*

Theorem 3.6 (Tietze-Urysohn extension theorem)**.** *Let X be normal, $A \subseteq X$ closed, $g : A \to \mathbb{K}$ continuous. Then there exists a continuous extension $f : X \to \mathbb{K}$ such that $f|_A = g$ and $\|f\|_\infty \leq \|g\|_\infty$.*

Proof. We first assume that g takes values in $[0,1]$. Let $g_0 = g$. Let $A_0 = g^{-1}([0, \frac{1}{3}]), B_0 = g^{-1}([\frac{2}{3}, 1])$ which are disjoint and closed. Thus by Corollary 3.4 there is a continuous $h_0 : X \to [0, \frac{1}{3}]$ such that $h_0|_{A_0} = 0, h_0|_{B_0} = \frac{1}{3}$. Let $g_1 = g_0 - h_0|_A$. Then $g_1(x) \in [0, \frac{2}{3}]$ for all $x \in A$. By induction assume that $g_i : A \to [0, \left(\frac{2}{3}\right)^i]$ is given and set

$$A_i = g_i^{-1}(\left[0, \frac{1}{3}\left(\frac{2}{3}\right)^i\right]), B_i = g_i^{-1}(\left[\frac{2}{3}\left(\frac{2}{3}\right)^i, \left(\frac{2}{3}\right)^i\right])$$

and $h_i : X \to [0, \frac{1}{3}\left(\frac{2}{3}\right)^i]$ a continuous function with $h_i|_{A_i} = 0, h_i|_{B_i} = \frac{1}{3}\left(\frac{2}{3}\right)^i$. Set $g_{i+1} = g_i - h_i|_A$. We find that

$$g = g_0 = g_1 + h_0|_A = g_2 + h_1|_A + h_2|_A = \dots = \sum_{i=0}^{\infty} h_i|_A.$$

Set $\tilde{f} = \sum_{i=0}^{\infty} h_i$. The convergence is uniform by Weierstrass M-test so $\tilde{f}$ is continuous.

If g takes values in $\mathbb{R}$, we can apply the above to the function $\frac{1}{2} + \frac{1}{2\pi} \arctan \circ g$ which takes values in $[\frac{1}{4}, \frac{3}{4}] \subseteq [0,1]$ to obtain an extension $\tilde{f}$. If g takes values in $\mathbb{C}$, we can apply this to the real and imaginary parts to obtain an extension $\tilde{f}$.

Finally define

$$f(x) = \begin{cases} \tilde{f}(x) & |\tilde{f}(x)| \leq \|g\|_\infty \\ e^{i \arg \tilde{f}(x)} \|g\|_\infty & |\tilde{f}(x)| \geq \|g\|_\infty \end{cases}$$

Then f is still a continuous extension. □

3.2 Arzelà-Ascoli theorem

The key object studied in functional analysis is function space. In this section we prove a theorem that answers the important question when a subset of $C(K)$ is compact.

Since we are studying normed spaces and subspaces thereof, which are in particular metric spaces, here are several notions of compactness in metric spaces. Note that in general they are not equivalent.

Definition. A metric space X is compact if any of the following conditions hold:

1. X has the *Heine-Borel property*: any open cover of X has a finite subcover.
2. X is *sequentially compact*, i.e. any sequence in X has a convergent subsequence.
3. X is *complete* and *totally bounded*, i.e. for any $\varepsilon > 0$ there exists a finite ε-net. This is a finite set $M \subseteq X$ such that for any $x \in X$, there exists $m \in M$ such that $d(x, m) < \varepsilon$.

Proof. For 1 $\iff$ 2 see IB Metric and Topological Spaces. 2 $\implies$ 3 easily. We present here only the proof of 3 $\implies$ 2. Let $(x_n) \subseteq X$ be a sequence. We want to find a convergent subsequence. Let M_n be a finite $\frac{1}{n}$-net for X. Let $m_1 \in M_1$ be such that $B_1(m_1)$ contains infinitely many of the x_n's. Let n_1 be the first n such that $x_n \in B_1(m_1)$. Given $m_1 \in M_1, \dots$ and $m_k \in M_k, n_1, \dots, n_k$ such that $B_{1/j}(m_j)$ contains infinitely many points from $(x_n) \cap B_{1/i}(m_i)$ for all $i \leq j$, and $x_{n_\ell} \in \bigcap_{i=1}^{\ell} B_{1/i}(m_i)$ for $\ell \leq k$, let m_{k+1} be such that $B_{1/(k+1)}(m_{k+1})$ contains infinitely many points from $(x_n) \cap \bigcap_{j=1}^{k} B_{1/j}(m_j)$ and n_{k+1} be the first $n > n_k$

such that $x_{n_{k+1}} \in \bigcap_{j=1}^{k+1} B_{1/j}(m_j)$. It follows that for $\ell \geq k$,

$$d(x_{n_k}, x_{n_\ell}) \leq d(x_{n_k}, m_k) + d(m_k, x_{n_\ell}) \leq \frac{2}{k} \to 0$$

so (x_{n_k}) is Cauchy and thus has a convergent subsequence by completeness of X. ☐

Corollary 3.7. *Let X be a complete metric space. Then $Y \subseteq X$ is relatively compact, i.e. has compact closure, if and only if Y is totally bounded.*

Proof. Y is totally bounded if and only if $\overline{Y}$ is totally bounded. ☐

Throughout this chapter, unless otherwise stated, we assume K is compact Hausdorff and equip $C(K)$ with $\|\cdot\|_\infty$ norm, thus making $C(K)$ into a Banach space, which is in particular a complete metric space.

Theorem 3.8 (Arzelà-Ascoli)**.** *Let K be compact Hausdorff and $\mathcal{F} \subseteq C(K)$. Then TFAE:*

1. *$\mathcal{F}$ is relatively compact;*
2. *$\mathcal{F}$ is bounded and equicontinuous, i.e. $\sup_{f\in\mathcal{F}} \|f\|_\infty < \infty$ and for all $\varepsilon > 0, x \in K$ there exists a neighbourhood U of x such that for all $f \in \mathcal{F}$, $|f(x) - f(y)| < \varepsilon$ for all $y \in U$.*

The generalises the fact that a subset of a finite-dimensional space is relative compact if and only if it is bounded, with the additional requirement of equicontinuity.

Proof.

1. $1 \implies 2$: let $\mathcal{F}$ be relatively compact, i.e. totally bounded. Thus for any $\varepsilon > 0$ there exists $f_1, \dots, f_n \in \mathcal{F}$ such that for all $f \in \mathcal{F}$,

   $$\min_i \|f - f_i\| < \varepsilon$$

   so in particular $\|f\| \leq \varepsilon + \max_i \|f_i\|$ for all $f \in \mathcal{F}$ so $\mathcal{F}$ is bounded.

   Let $\varepsilon > 0$, $f_1, \dots, f_n$ as above and $x \in K$. Since the f_i's are continuous, there exist neighbourhoods U_i of x such that

   $$|f_i(x) - f_i(y)| < \varepsilon$$

   for $y \in U_i$. Now let $U = \bigcap_{i=1}^n U_i$, which is again a neighbourhood of x. For all $y \in U$,

   $$|f(x) - f(y)| \leq |f(x) - f_i(x)| + |f_i(x) - f_i(y)| + |f_i(y) - f(y)| < 3\varepsilon$$

   where i is such that for all $f \in \mathcal{F}$, $\|f - f_i\| < \varepsilon$. Thus $\mathcal{F}$ is equicontinuous.

2. 2 $\implies$ 1: Let $\mathcal{F}$ be bounded and equicontinuous. For $\varepsilon > 0$, we construct a finite 3ε-net for $\mathcal{F}$. Let $\varepsilon > 0$. For $x \in K$, let U_x be an open neighbourhood of x such that $|f(x) - f(y)| < \varepsilon$ whenever $f \in \mathcal{F}, y \in U_x$. Since K is compact, there are $x_1, \dots, x_n$ such that
$$K = \bigcup_{i=1}^{n} U_{x_i}.$$
Since $\mathcal{F}$ is (uniformly) bounded, the vector $(f(x_1), \dots, f(x_n)) \in \mathbb{K}^n$ is bounded in any norm on $\mathbb{K}^n$, say $\|\cdot\|_\infty$. Thus
$$F = \{(f(x_1), \dots, f(x_n)) : f \in \mathcal{F}\} \subseteq \mathbb{K}^n$$
is relatively compact in $\mathbb{K}^n$. Thus there are $f_1, \dots, f_m \in \mathcal{F}$ such that
$$F' = \{(f_i(x_1), \dots, f_i(x_n)) : 1 \leq i \leq m\} \subseteq \mathbb{K}^n$$
is a finite ε-net of F. Claim that moreover $f_1, \dots, f_m$ is a finite 3ε-net for $\mathcal{F}$. Indeed, for $x \in U_{x_j}$,
$$|f(x) - f_i(x)| \leq |f(x) - f(x_j)| + |f(x_j) - f_i(x_j)| + |f_i(x_j) - f_i(x)| < 3\varepsilon$$
for some $1 \leq i \leq m$ such since F' is a finite ε-net of F. Thus $\mathcal{F}$ is totally bounded and thus relatively compact.

□

3.3 Aside: compact operator

Definition (compact operator)**.** Let X, Y be normed spaces and $T : X \to Y$ linear. Then T is *compact* if $\overline{T(B_1(0))}$ is compact.

Equivalently, every bounded $(x_n) \subseteq X$ has a subsequence such that (Tx_n) converges along that sequence.

Example.

1. If $T \in B(X, Y)$ of finite rank, i.e. $T(X)$ is finite-dimensional, then T is compact as $\overline{T(B_1(0))}$ is a bounded closed set in the finite-dimensional space $T(X)$. In fact, compact operators are generalisation of matrices in the sense that every compact operator is the uniform limit of finite rank operators.

2. If $\dim X = \infty$ then $\mathrm{id} : X \to X$ is not compact as $\overline{B_1(0)} = \overline{B}_1(0)$ is not compact.

3. Let $K = [0, 1]$. Consider $C^1([0, 1])$ with $\|f\|_{C^1} = \|f\|_\infty + \|f'\|_\infty$ and $C^0([0, 1])$ with $\|f\|_{C^0} = \|f\|_\infty$. Then the embedding $\iota : C^1([0, 1]) \to C^0([0, 1])$ is compact. Indeed, let
$$\mathcal{F} = \{f \in C^1([0, 1]) : \|f\|_\infty + \|f'\|_\infty < 1\} = B_1(0) \subseteq C^1([0, 1])$$
then $\iota(\mathcal{F})$ is bounded in $C^0([0, 1])$ and for any $f \in \mathcal{F}$,
$$|f(x) - f(y)| \leq \|f'\|_\infty |x - y| \leq |x - y| < \varepsilon$$
whenever $|x - y| < \varepsilon$. Thus $\iota(\mathcal{F})$ is equicontinuous. So by Arzelà-Ascoli, $\iota(\mathcal{F})$ is relatively compact. ι is compact.

In general, if Y is a Banach space then $T : X \to Y$ is compact if and only if $T(B_1(0))$ is totally bounded.

Theorem 3.9. *Let X be a normed space and Y a Banach space. Then the compact operators form a closed subspace of the space $B(X, Y)$ of bounded operators.*

Proof. There are two claims in the theorem:

1. if S and T are compact operators then so is $S + T$,
2. if $T_n \to T$ where T_n's are compact and T is bounded then T is compact.

Let $(x_n) \subseteq X$ be bounded. Then there is a subsequence $\Lambda \subseteq \mathbb{N}$ such that $Sx_n \to y$ for some y as $n \in \Lambda, n \to \infty$. Moreover, there is a further subsequence $\Lambda' \subseteq \Lambda$ such that $Tx_n \to z$ as $n \in \Lambda', n \to \infty$. Then

$$(S + T)x_n = Sx_n + Tx_n \to y + z$$

as $n \in \Lambda', n \to \infty$. $S + T$ is compact.

For the second claim, we need to show that $T(B_1(0))$ is totally bounded. Let $\varepsilon > 0$ and $n \in \mathbb{N}$ be such that $\|T - T_n\| < \varepsilon$. Then

$$T_n(B_1(0)) \subseteq \bigcup_{i=1}^{k} B_\varepsilon(T_n x_i)$$

for some $x_1, \dots, x_k \in B_1(0)$ since T_n is compact. Thus

$$T(B_1(0)) \subseteq \bigcup_{i=1}^{k} B_{2\varepsilon}(T_n x_i) \subseteq \bigcup_{i=1}^{k} B_{3\varepsilon}(T x_i)$$

which is a finite 3ε-net for $T(B_1(0))$. $T(B_1(0))$ is totally bounded so T is compact. □

In particular this shows that limits of finite rank operators are compact:

Corollary 3.10. *Any limit in $B(X, Y)$ of finite rank operators is compact.*

3.4 Application: Peano existence theorem

Recall in IB Analysis II

Theorem 3.11 (Picard-Lindelöf). *Let $f : \mathbb{R} \to \mathbb{R}$ be locally Lipschitz continuous. Then for any $x_0 \in \mathbb{R}$ there exists a maximal interval (T_1, T_2), with $T_1 = -\infty$ and/or $T_2 = \infty$ allowed, such that*

$$\begin{cases} x'(t) = f(x(t)) \\ x(0) = x_0 \end{cases} \tag{$*$}$$

has a unique C^1 solution $x : (T_1, T_2) \to \mathbb{R}$ that is maximal, i.e. that is not the restriction of such a solution on a larger interval. Moreover if $T_2 \neq \infty$, for any bounded $K \subseteq \mathbb{R}$ there is $t < T_2$ such that $x([t, T_2)) \cap K = \emptyset$ and similarly if $T_1 \neq -\infty$.

Theorem 3.12 (Peano existence theorem). *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. Then for any $x_0 \in \mathbb{R}$, there is $\varepsilon > 0$ and a solution $x : (-\varepsilon, \varepsilon) \to \mathbb{R}$ to $(*)$.*

Remark. The solution is not necessarily unique. For example take $f(x) = \sqrt{|x|}$.

Lemma 3.13 (a priori bound). *Assume that $b > 0, M > 0$ are such that*

$$|f(x)| \leq M \text{ for } |x - x_0| \leq b \tag{†}$$

Then if $T \leq \frac{b}{M}$ and x is any C^1 solution to $(*)$ for all $|t| \leq T$, it follows that $|x(t) - x_0| \leq b, |x'(t)| \leq M$.

Proof. Assume that $x(t)$ is a C^1 solution for $|t| \leq T' < T$ such that $|x(t) - x_0| \leq b$. Then

$$|x'(t)| = |f(x(t))| \leq M$$

so

$$|x(t) - x_0| = \left| \int_0^t f(x(s)) ds \right| \leq Mt < b$$

for $|t| \leq T'$. This allows us to extend the solution beyond T' by continuous induction. Let

$$I = \{T' \in [0, T] : |x(t) - x_0| \leq b \text{ for } |t| \leq T'\}.$$

Note $I \neq \emptyset$ and that I is closed. Claim that $\sup I = T$: otherwise $|x(t) - x_0| < b$ for $|t| < \sup I$ but by continuity, a neighbourhood of $\sup I$ also has to be contained in I, contradiction. Thus $I = [0, T]$. □

Proof of Peano existence theorem. Let

$$B = \{f + \tilde{g} : \tilde{g} \in C^0, \|\tilde{g}\|_\infty \leq \infty\}$$

and choose $M, b > 0$ such that (†) holds for all $g \in B$. For any $h \in B \cap C^1$ there is a local solution by Picard-Lindelöf. The lemma implies that these solutions are defined on all of $[-T, T]$ with T as in the lemma. Define the solution operator

$$\begin{aligned} S : B \cap C^1 &\to C^1[-T, T] \\ f &\mapsto x \end{aligned}$$

where x is the solution to $(*)$. By the lemma, $S(B \cap C^1)$ is bounded in $C^1[-T, T]$ with norm $\|x\|_\infty + \|x'\|_\infty$. By Arzelà-Ascoli, the embedding $C^1[-T, T] \to C^1[-T, T]$ is compact, i.e. $S(B \cap C^1)$ is relatively compact in $C^0([-T, T])$. Let $f_i \in B \cap C^1$ such that $f_i \to f$ in C^0 (this is not obvious but it will follow from Weierstrass approximation theorem in the next section), i.e. $\|f - f_i\|_\infty \to 0$. By relative compactness there is a subsequence $x_i = Sf_i$ converges to some $x \in C^0[-T, T]$ with $\|\cdot\|_\infty$ norm. Claim that $x \in C^1([-T, T])$ and $(*)$ holds.

Proof. Since $f_i \to f, x_i \to x$ in C^0 (along the subsequence, we also have $f_i \circ x_i \to f \circ x$. Thus $x_i = f_i \circ x_i \to f \circ x$ uniformly in $|t| \leq T$. Thus $x \in C^1$ and $x' = f \circ x$. □

□

3.5 Stone-Weierstrass theorem

Theorem 3.14 (Weierstrass approximation theorem)**.** *The set of polynomials with real coefficients is dense in $C([a,b],\mathbb{R})$ in the uniform topology.*

The theorem follows more or less directly from the approxmation of a single function: the absolute value function, as any continuous function on $[0,1]$ can be approximated uniformly by a piecewise linear function.

Lemma 3.15. *There is a sequence of polynomials $P_n : [-1,1] \to [0,1]$ such that $P_n \to |\cdot|$ uniformly on $[-1,1]$ as $n \to \infty$.*

Proof. We use the Babylonian method to construct square root map, which when composed with square maps gives absolute value. The idea is that if $q : [0,1] \to [0,1]$ is a function with

$$q(t) = \frac{1}{2}(t + q(t)^2)$$

Then

$$(1 - q(t))^2 = 1 - 2q(t) + q(t)^2 = 1 - t$$

so

$$1 - q(t) = \sqrt{1-t}$$

which is square root up to translation and have

$$|t| = 1 - q(1 - t^2).$$

To approximate q, define polynomials

$$\begin{aligned} Q_n &: [0,1] \to [0,1] \\ Q_0(t) &= 0 \\ Q_n(t) &= \frac{1}{2}(t + Q_{n-1}(t)^2) \end{aligned}$$

If Q_n converges to some q then $q(t) \in [0,1]$ and satisfies $q(t) = \frac{1}{2}(t + q(t)^2)$. To show the sequence converges, note that for any $t \in [0,1]$, $Q_{n+1}(t) \geq Q_n(t)$. Indeed

$$Q_{n+1}(t) - Q_n(t) = \frac{1}{2} \underbrace{(Q_n(t) + Q_{n-1}(t))}_{\geq 0} \underbrace{(Q_n(t) - Q_{n-1}(t))}_{\geq 0 \text{ by induction}}.$$

Since Q_n is an increasing function by induction from definition, the last equality implies that $Q_{n+1}(t) - Q_n(t)$ is an increasing function. Thus

$$Q_{n+1}(t) - Q_n(t) \leq Q_{n+1}(1) - Q_n(1)$$

so

$$Q_m(t) - Q_n(t) \leq Q_m(1) - Q_n(1)$$

for all $m > n, t \in [0,1]$. Let $m \to \infty$, get

$$0 \leq 1 - \sqrt{1-t} - Q_n(t) \leq 1 - Q_n(1)$$

by defining properties of q. Thus

$$\|1 - Q_n(t) - \sqrt{1-t}\|_\infty \le 1 - Q_n(1) \to 0$$

as $n \to \infty$.

Now set $P_n(t) = 1 - Q_n(1-t^2)$. Then $\|P_n - |\cdot|\|_\infty \to 0$ as $n \to \infty$. ☐

Proof of Weierstrass approximation theorem. Exercise. ☐

We now state and prove a more abstract and general version of the approximation theorem.

Definition (algebra)**.** A real/complex *algebra* is a real/complex vector space A with a bilinear map

$$\begin{aligned} A \times A &\to A \\ (a, b) &\mapsto ab \end{aligned}$$

called product that is associative, i.e. $(ab)c = a(bc)$ for all $a, b, c \in A$.

If $ab = ba$ for all $a, b \in A$ then A is *commutative.*

If there exists $1 \in A \setminus \{0\}$ such that $1a = a = a1$ for all $a \in A$ then A is *unital.*

Definition (normed/Banach algebra)**.** If an algebra A is a normed vector space such that

$$\|ab\| \le \|a\|\|b\|$$

for all $a, b \in A$ then A is called a *normed algebra.* If A is a Banach space then A is called a *Banach algebra.*

Example.

1. $C(K, \mathbb{R})$ is a commutative unital Banach algebra with product being pointwise multiplication and unit being the constant function 1.

2. $B(X, X)$, where X is a normed vector space, is a normed unital algebra with product being composition and unit being id_X. If X is Banach then so is $B(X, X)$. It is noncommutative.

Theorem 3.16 (Stone-Weierstrass)**.** *Let $A \subseteq C(K, \mathbb{R})$ be a subalgebra that*

1. *separates points: for all $x, y \in K, x \ne y$, there is $f \in A$ such that $f(x) \ne f(y)$,*

2. *vanishes nowhere: for all $x \in K$ there is $f \in A$ such that $f(x) \ne 0$.*

then A is dense in $C(K, \mathbb{R})$.

Example. Let $U \subseteq \mathbb{R}^n$ be open bounded. Let A be the set of polynomials in $x_1, \dots, x_n$. Then A is an algebra, separates points and contains the constant polynomial which vanishes nowhere. Thus Stone-Weierstrass theorem implies that $\overline{A} = C(\overline{U})$. In particular, $C^\infty(\overline{U})$ is dense in $C(\overline{U})$.

We set up some terminologies and intermediate results before we prove the theorem. These definitions will also be useful later in this course and in other areas of maths.

Definition (poset, lattice)**.**

1. A *partially ordered set* or *poset* is a set P with a binary relation $\leq$ such that for all $u, v \in P$, either $u \leq v$ or $u \not\leq v$ and is
   (a) reflective: $u \leq u$,
   (b) transitive: if $u \leq v, v \leq w$ then $u \leq w$,
   (c) antisymmetric: if $u \leq v, v \leq u$ then $u = v$.
2. A *lattice* is a poset L with the property that for any $u, v \in L$, there is a *least upper bound* or *join* $u \vee v$ and a *greatest lower bound* or *meet* $u \wedge v$, i.e.

$$\begin{aligned} &u, v \leq u \vee v \text{ and if } u, v \leq b \text{ then } u \vee v \leq b, \\ &u \wedge v \leq u, v \text{ and if } b \leq u, b \text{ then } v \leq u \wedge v. \end{aligned}$$

Example. $C(K, \mathbb{R})$ is a lattice ($f \leq g$ if and only if $f(x) \leq g(x)$ for all $x \in K$) and

$$\begin{aligned} (f \vee g)(x) &= \max\{f(x), g(x)\} \\ (f \wedge g)(x) &= \min\{f(x), g(x)\} \end{aligned}$$

Lemma 3.17. *Let $A \subseteq C(K, \mathbb{R})$ be a closed subalgebra. Then A is a lattice in $C(K, \mathbb{R})$.*

Proof. We need to show that if $f, g \in A$ then $f \wedge g$ and $f \vee g$ are also in A.

$$\begin{aligned} (f \vee g)(x) &= \frac{1}{2}(f(x) + g(x) + |f(x) - g(x)|) \\ (f \wedge g)(x) &= \frac{1}{2}(f(x) + g(x) - |f(x) - g(x)|) \end{aligned}$$

so suffices to show that if $f \in A$ then $|f| \in A$. Let $f \in A, f \neq 0, \varepsilon > 0$. Replacing f by $f/\|f\|_\infty$ we may assume that f takes values in $[-1, 1]$. By Lemma 3.15 there is a polynomial $P : [-1, 1] \to [0, 1]$ such that $\|P - |\cdot|\|_\infty \leq \varepsilon$. Then $\|P \circ f - |f|\|_\infty \leq \varepsilon$. Since $P \circ f \in A$ and A is closed, have $|f| \in A$. ☐

Lemma 3.18. *Let $L \subseteq C(K, \mathbb{R})$ be a lattice. If $g \in C(K, \mathbb{R})$ is such that for all $\varepsilon > 0$, for all $x, y \in K$, exists $f \in L$ such that*

$$\begin{cases} |f(x) - g(x)| < \varepsilon \\ |f(y) - g(y)| < \varepsilon \end{cases} \tag{$*$}$$

then $g \in \overline{L}$. In particular, if this condition holds for all $g \in C(K, \mathbb{R})$ then $\overline{L} = C(K, \mathbb{R})$.

Proof. Let $g \in C(K, \mathbb{R})$ be as in the assumption and $\varepsilon > 0$. We construct $f \in L$ such that $\|f - g\| < \varepsilon$. For $x, y \in K$, let f_{xy} be f in $(*)$. By continuity, the sets

$$U_{xy} = \{z \in K : f_{xy}(z) < g(z) + \varepsilon\}$$
$$V_{xy} = \{z \in K : f_{xy}(z) > g(z) - \varepsilon\}$$

are open and $\{x, y\} \subseteq U_{xy} \cap V_{xy}$. For any x, $\{U_{xy}\}_y$ is a cover of K so by compactness there are $y_1, \ldots, y_n$ such that $\bigcup_{i=1}^n U_{xy_i} = K$. Define

$$V_x = \bigcap_{i=1}^n V_{xy_i}$$
$$f_x = \bigwedge_{i=1}^n f_{xy_i} \in L$$

then

$$f_x(y) < g(y) + \varepsilon \text{ for all } y \in K$$
$$f_x(y) > g(y) - \varepsilon \text{ for all } y \in V_x$$

Now $\{V_x\}_x$ is an open cover of K. Choose finitely many $x_1, \ldots, x_m$ such that $K = \bigcup_{j=1}^m V_{x_j}$ by compactness. Set

$$f = \bigvee_{j=1}^m f_{x_j} \in L$$

so

$$f(y) < g(y) + \varepsilon \text{ for all } y \in K$$
$$f(y) > g(y) - \varepsilon \text{ for all } y \in K$$

so $|f(y) - g(y)| < \varepsilon$ for all $y \in K$. □

Proof of Stone-Weierstrass. By continuity of addition and multiplication, the closure $\overline{A}$ is a closed subalgebra of $C(K, \mathbb{R})$ so is a lattice. Let $g \in C(K, \mathbb{R}), x, y \in K$. We will find $f \in A$ such that $f(x) = g(x)$ and $f(y) = g(y)$. In particular, $(*)$ in Lemma 3.18 holds.

By assumption A vanishes nowhere and separates points, i.e.

$$\forall x \in K, \exists f_x \in A \text{ such that } f_x(x) \neq 0$$
$$\forall x, y \in K, \exists f_{xy} \in A \text{ such that } f_{xy}(x) \neq f_{xy}(y)$$

Claim that for all $x \neq y$, there are $\alpha, \beta, \gamma \in \mathbb{R}$ such that $h = \alpha f_x + \beta f_y + \gamma f_{xy}$ satisfies

$$h(x) \neq 0, h(y) \neq 0, h(x) \neq h(y).$$

Indeed if $f_{xy}(x) \neq 0$ and $f_{xy}(y) \neq 0$ then we can take $h = f_{xy}$. Otherwise wlog $f_{xy}(y) = 0$ and by rescaling we can assume

$$f_{xy}(x) = 1, f_{xy}(y) = 0, f_y(x) = C, f_y(y) = 1$$

so take

$$\alpha = 0, \beta = 1, \gamma = 2 - C.$$

This gives the claim since

$$\begin{aligned} h(x) &= C + 2 - C = 2 \\ h(y) &= 1 \end{aligned}$$

We find that $(h(x), h(y)), (h(x)^2, h(y)^2) \in \mathbb{R}^2$ are linearly independent. Then there are $s, t \in \mathbb{R}$ such that

$$(g(x), g(y)) = t(h(x), h(y)) + s(h(x)^2, h(y)^2) = (f(x), f(y))$$

with $f = th + sh^2 \in A$. $(*)$ holds for any $g \in C(K, \mathbb{R})$, thus completing the proof. $\square$

Example. Let $K \subseteq \mathbb{R}^n$ be compact. Then $C(K)$ is separable, i.e. there is a countable dense set, given by polynomial with rational coefficients as by Stone-Weierstrass theorem we can approximate continuous functions on K by this set.

In example sheet 3, we will show that given K compact Hausdorff, $C(K)$ is separable if and only if K is metrisable.

Example. Let K and L be compact. Then $A \subseteq C(K \times L)$ consisting of functions of the form

$$\begin{aligned} K \times L &\to \mathbb{R} \\ (x, y) &\mapsto \sum_{i=1}^n f_i(x) g_i(y) \end{aligned}$$

where $(f_i) \subseteq C(K), (g_i) \subseteq C(L)$, is an algebra that separates points and vanishes nowhere. Thus $\overline{A} = C(K \times L)$. In particular if $(f_i) \subseteq C(K), (g_i) \subseteq C(L)$ are dense sequences then functions of the form $(x, y) \mapsto \sum_{i=1}^n f_i(x) g_i(y)$ are dense in $C(K \times L)$.

Corollary 3.19. *For every $f \in C([0,1]^2)$, have*

$$\int_0^1 \int_0^1 f(x, y) dx dy = \int_0^1 \int_0^1 f(x, y) dy dx.$$

3.6 Complex Stone-Weierstrass theorem

Theorem 3.20 (complex Stone-Weierstrass)**.** *Let $A \subseteq C(K, \mathbb{C})$ be a subalgebra such that*

1. *A separates points,*
2. *A vanishes nowhere,*
3. *A is closed under complex conjugation, i.e. $\overline{f} \in A$ if $f \in A$*

then $\overline{A} = C(K, \mathbb{C})$.

This is important in spectral theory, which we'll get back to in the last chapter of the course.

Definition (C^*-algebra)**.** A C^*-*algebra* is a complex unital Banach algebra A with an antilinear involution $a \mapsto a^*$ satisfying

$$\begin{aligned}(ab)^* &= b^* a^* \\ 1^* &= 1 \\ (\lambda a)^* &= \overline{\lambda} a^* \\ \|a^*\| &= \|a\|\end{aligned}$$

Example.

1. $C(K, \mathbb{C})$ is a commutative C^*-algebra with $f^* = \overline{f}$.
2. $B(H, H)$ with H a Hilbert space is a C^*-algebra. We will introduce Hilbert space formally in the next chapter.

Corollary 3.21. *If $A \subseteq C(K, \mathbb{C})$ is a C^*-subalgebra that separates points then $\overline{A} = C(K, \mathbb{C})$.*

Proof. The main observation is that if $f \in A$ then

$$\begin{aligned}\operatorname{Re} f &= \frac{1}{2}(f + \overline{f}) \in A \\ \operatorname{Im} f &= \frac{1}{2i}(f - \overline{f}) \in A\end{aligned}$$

Let $A_{\mathbb{R}}$ be the subalgebra of $C(K, \mathbb{R})$ generated by $\operatorname{Re} f, \operatorname{Im} f$ for $f \in A$. Then $A_{\mathbb{R}}$ vanishes nowhere and separates points since A does, so the real version Stone-Weierstrass theorem implies that $\overline{A}_{\mathbb{R}} = C(K, \mathbb{R})$. Let $f = u + iv \in C(K, \mathbb{C})$ where $u, v \in C(K, \mathbb{R})$. There are $(u_j) \subseteq A_{\mathbb{R}}, (v_j) \subseteq A_{\mathbb{R}}$ such that $u_j \to u, v_j \to v$. Since $u_j + iv_j \in A$, $\overline{A} = C(K, \mathbb{C})$. □

Example (Hardy space)**.** One may wonder if closure under complex conjugation is necessary. Consider $K = \{z \in \mathbb{C} : |z| \leq 1\}$, the closed unit disk. Then

$$A = \{f \in C(K, \mathbb{C}) : f \text{ analytic on } K\}$$

is a subalgebra. It separates points and vanishes nowhere. But $\overline{A} \neq C(K, \mathbb{C})$ since $z \mapsto \overline{z}$ is not in $\overline{A}$.

Example. Let $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ be the circle, i.e. the interval $[-\pi, \pi]$ with end points identified. Let A be the subspace in $C(\mathbb{T}, \mathbb{C})$ spanned by $\{e^{inx}\}_{n \in \mathbb{Z}}$. Its elements are called *trigonometric polynomials*. A is a C^*-subalgebra, separates points and contains the constants so vanishes nowhere. It is also closed under complex conjugation since $\overline{e^{inx}} = e^{-inx}$ so $\overline{A} = C(\mathbb{T}, \mathbb{C})$ by complex Stone-Weierstrass theorem.

Example. On example sheet 3 we'll show that there exists $f \in C(\mathbb{T})$ such that $S_n f(0) \not\to f(0)$ where $S_n f$ is the partial Fourier sum given by the *Dirichlet sum*

$$S_n f = \sum_{k=-n}^{n} \hat{f}_k e^{ikx}$$

where

$$\hat{f}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx.$$

This does *not* contradict the previous observation. The moral is that the trigonometric polynomials that provide a uniform approximation to a given $f \in C(\mathbb{T})$ cannot always be taken to be the partial Fourier sum!

However, we can deduce that the partial Fourier sum of f converges to f in L^2.

Proposition 3.22. *For every $f \in C(\mathbb{T})$,*

$$\lim_{n \to \infty} \int_{-\pi}^{\pi} |f - S_n f|^2 dx = 0.$$

Proof. By complex Stone-Weierstrass theorem, for any $\varepsilon \geq 0$ there is a trigonometric polynomial P such that $\|P - f\| < \varepsilon$. Note that $S_n P = P$ if $n \geq \deg P$, where $\deg P$ is the largest n such that P contains $e^{\pm inx}$. Then

$$|f - S_n f| \leq |f - P| + |S_n f - P| = |f - P| + |S_n f - S_n P|$$

if $\deg P \leq n$ so

$$|f - S_n f|^2 \leq 2|f - P|^2 + 2|S_n f - S_n P|^2$$

since $(a + b)^2 \leq 2a^2 + 2b^2$. Thus

$$\int_{-\pi}^{\pi} |f - S_n f|^2 dx \leq 4 \int_{-\pi}^{\pi} |f - P|^2 dx \leq 8\pi\varepsilon^2$$

where we used *Bessel's inequality*, which we will prove in the next section, that for $g \in C(\mathbb{T})$,

$$\int_{-\pi}^{\pi} |S_n g|^2 dx \leq \int_{-\pi}^{\pi} |g|^2 dx.$$

□

4 Euclidean vector spaces and Hilbert spaces

4.1 Definitions and examples

Definition (inner product). Let X be a vector space (real or complex). Then an *inner product* is a map $(\cdot,\cdot) : X \times X \to \mathbb{K}$ such that

1. (skew-)symmetric: $(x, y) = \overline{(y, x)}$ for all $x, y \in X$,
2. linear in first argument: $(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1(x_1, y) + \lambda_2(x_2, y)$ for all $x_i, y \in X, \lambda_i \in \mathbb{K}$,
3. positive definite: $(x, x) \geq 0$ with $(x, x) = 0$ if and only if $x = 0$.

A vector space X together with an inner product $(\cdot,\cdot)$ is called an *inner product space*.

Remark.

1. In the real case $(\cdot,\cdot)$ is bilinear.
2. In the complex case $(\cdot,\cdot)$ is antilinear in the second argument.
3. There is an opposite convention, for example among physicists, where the role of the first and second argument is interchanged.

Proposition 4.1 (Cauchy-Schwarz). *Let X be an inner product space. Then*

$$|(x, y)| \leq (x, x)^{1/2}(y, y)^{1/2}$$

for all $x, y \in X$ with equality if and only if $x = \lambda y$ for some $\lambda \in \mathbb{K}$.

Proof. We may assume that $(x, x) = 1 = (y, y)$ and $(x, y) \geq 0$. Then for $t > 0$,

$$0 \leq (x - ty, x - ty) = (x, x) - 2t(x, y) + t^2(y, y) = 1 + t^2 - 2t(x, y).$$

Thus

$$(x, y) \leq \inf_{t>0} \frac{1 + t^2}{2t} = 1.$$

□

Corollary 4.2. *Let X be an inner product space. Then $\|x\| = (x, x)^{1/2}$ defines a norm on X.*

Proof. Positive definiteness is immediate from the definition. Positive homogeneity follows from

$$\|\lambda x\| = (\lambda x, \lambda x)^{1/2} = (\lambda\overline{\lambda})(x, x)^{1/2} = |\lambda|\|x\|.$$

For the triangle inequality, note

$$\begin{aligned}\|x+y\|^2 &= (x+y, x+y) \\ &= (x,x) + (y,y) + \underbrace{(x,y) + (y,x)}_{=2\operatorname{Re}(x,y) \le 2|(x,y)|} \\ &\le \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \\ &\le (\|x\| + \|y\|)^2\end{aligned}$$

□

Fact (polarisation identities)**.** Let X be an inner product space, $x, y \in X$. Then

1. real version:
$$(x,y) = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2).$$
2. complex version:
$$(x,y) = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2).$$

Proof. Trivial. □

Corollary 4.3. *The norm determines the inner product.*

Fact (parallelogram law)**.** Let X be an inner product space, $x, y \in X$. Then

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

Proof. Ditto. □

Exercise. If X is a normed vector space satisfying the parallelogram law for any $x, y \in X$, then the polarisation identity defines an inner product on X, so X is also an inner product space.

Definition (Euclidean)**.** A normed space is *Euclidean* if its norm is the norm associated to some inner product.

By polarisation identities, such inner product is unique if exists. The exercise shows that an equivalent characterisation is parallelogram law holds.

Definition (Hilbert space)**.** An inner product space is called a *Hilbert space* if it is complete as a metric space.

Example.

1. $\ell^2 = \{(x_n) \subseteq \mathbb{K} : \sum_{n=1}^\infty |x_n|^2 < \infty\}$ is a Hilbert space with inner product
$$(x,y) = \sum_{n=1}^\infty x_n \overline{y}_n.$$
We will later show that this is the *only* separable Hilbert space up to isometric isomorphism.

2. $C([0,1])$ is an inner product space with
$$(x,y) = \int_0^1 f(x)\overline{g(x)}dx.$$
But it is not complete, so not a Hilbert space. This leads us to completion of inner product spaces.

Proposition 4.4. *Let X be an inner product space. Then the completion of X is also an inner product space, thus a Hilbert space.*

Proof. Let $\tilde{X}$ be the completion of X. For $x, y \in \tilde{X}$, choose $(x_n), (y_n) \subseteq X$ such that $x_n \to x, y_n \to y$. Set
$$(x,y) = \lim_{n\to\infty}(x_n, y_n).$$
Easy to check that this definition is well-defined, that $(\cdot,\cdot)$ is an inner product, and that this iner product induces the completed norm. □

Example. The completion of $C([0,1])$ is a Hilbert space, denoted $L^2([0,1])$, which can be identified with the space of equivalence classes of Lebesgue measurable functions with $f \sim g$ if and only if $f = g$ Lebesgue almost everywhere.

4.2 Orthogonal complements and projections

Definition (orthogonal, orthogonal complement)**.** Let X be an inner product space.

- $x, y \in X$ are *orthogonal* if $(x,y) = 0$, also written as $x \perp y$.
- The *orthogonal complement* of a set $S \subseteq X$ is
$$S^\perp = \{x \in X : (x,y) = 0 \text{ for all } y \in S\}.$$

Fact (Pythagoras)**.** If x, y are orthogonal then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

Fact. $S^\perp$ is a closed subspace of X and $\overline{\operatorname{span} S}^\perp = S^\perp$.

Proof. $S^\perp = \bigcap_{y\in S} f_y^{-1}(0)$ where $f_y(x) = (x,y)$ is continuous. Thus $S^\perp$ is the intersection of closed sets so closed. Clearly
$$S^\perp \supseteq (\operatorname{span} S)^\perp \supseteq \overline{\operatorname{span} S}^\perp.$$
For the other direction, let $x \in S^\perp$ and $y \in \overline{\operatorname{span} S}$, i.e. $y = \lim_{n\to\infty} y_n$ with $y_n \in \operatorname{span} S$. Then
$$(x,y) = \lim_{n\to\infty}(x, y_n) = 0$$
so $x \in \overline{\operatorname{span} S}^\perp$. □

Notation. For $Y \subseteq X$ a subspace, $Y^\perp \cap Y = 0$. Thus the sum $Y + Y^\perp$ is direct and write $Y + Y^\perp = Y \oplus Y^\perp$.

Example. From linear algebra we know if X is finite-dimensional then $X = Y \oplus Y^\perp$. However this is generally false for infinite diemsnional inner product space. Let $X = C[0,1]$ with $(f,g) = \int_0^1 f\bar{g}dx$ and $Y = C^1[0,1] \subseteq C[0,1]$. Then $Y^\perp = 0$ since $\int_0^1 f\bar{g}dx = 0$ for all $g \in C^1$ implies $f = 0$ (as $f \in C[0,1]$).

Theorem 4.5. *Let $Y \subseteq X$ be a complete subspace. Then*

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

Moreover, given $x \in X$, its unique decomposition $x = x_\parallel + x_\perp$ where $x_\parallel \in Y, x_\perp \in Y^\perp$ is characterised by

$$\|x_\perp\| = \|x - x_\parallel\| = \inf_{y\in Y}\|x - y\|.$$

In particular this holds if X is a Hilbert space and $Y \subseteq X$ is a closed subspace.

Proof. Let $x \in X$ and $D = \inf_{y\in Y}\|x - y\|$. Choose any sequence $(y_k) \subseteq Y$ such that $\|y_k - x\| \to D$. Claim that (y_k) is Cauchy: by the parallelogram identity applied to $x - y_j, x - y_k$,

$$\|y_j - y_k\|^2 + \|2x - y_j - y_k\|^2 = 2\|x - y_j\|^2 + 2\|x - y_k\|^2.$$

Rearrange,

$$\begin{aligned}
\|y_j - y_k\|^2 &= 2\underbrace{\|x - y_j\|^2}_{\to D^2} + 2\underbrace{\|x - y_k\|^2}_{\to D^2} - 4\underbrace{\|x - \frac{1}{2}(y_j + y_k)\|^2}_{\geq D^2} \\
&\leq 4(D^2 + \varepsilon) - 4D^2 \\
&= 4\varepsilon
\end{aligned}$$

for all $\varepsilon > 0$ for all j, k sufficiently large. By completeness of Y, the claim implies that $y_j \to x_\parallel$ for some $x_\parallel \in Y$. By continuity of norm $\|x - x_\parallel\| = D$.

Now let $x_\perp = x - x_\parallel$. Have to show that $x_\perp \in Y^\perp$. Suppose not, then there must be $\tilde{y} \in Y$ such that $(\tilde{y}, x_\perp) > 0$. Thus

$$\begin{aligned}
\|x_\perp - t\tilde{y}\|^2 &= \|x_\perp\|^2 - 2t(x_\perp, \tilde{y}) + t^2\|\tilde{y}\|^2 \\
&= D^2 - t(\underbrace{2(x_\perp, \tilde{y}) - t\|\tilde{y}\|^2}_{>0 \text{ for } t>0 \text{ small}}) \\
&< D^2
\end{aligned}$$

But

$$\|x_\perp - t\tilde{y}\|^2 = \|x - (\underbrace{x_\parallel + t\tilde{y}}_{\in Y})\|^2 \geq D^2,$$

contradiction. Thus $x_\perp \in Y^\perp$.

Finally to show that the decomposition is uniquely characterised by the expression, suppose $x = \tilde{x}_\perp + \tilde{x}_\parallel$ for some $\tilde{x}_\perp \in Y^\perp, \tilde{x}_\parallel \in Y$. Then $\tilde{x}_\parallel = x_\parallel + y$ where $y = x_\perp - \tilde{x}_\perp \in Y^\perp$. But $y = \tilde{x}_\parallel - x_\parallel \in Y$ so $y \in Y \cap Y^\perp = 0$. ☐

Theorem 4.6 (Riesz representation theorem for Hilbert space)**.** *Let H be a Hilbert space. Then for any $\ell \in H^*$ there is a unique $x_\ell \in H$ such that*

$$\ell(y) = (y, x_\ell)$$

for all $y \in H$ and $\|\ell\| = \|x_\ell\|$.

This can be seen as a generalisation of $\ell^2 \cong (\ell^2)^*$, as ℓ^2 is an inner product space and thus a Hilbert space.

Proof. Let $\ell \in H^*, \ell \neq 0$. Then $\ker \ell$ is closed and and by Theorem 4.5 we have $H = \ker \ell \oplus (\ker \ell)^\perp$. Since $\ell \neq 0$, $(\ker \ell)^\perp \neq 0$. Claim that there exists $x_0 \in H$ such that $(\ker \ell)^\perp = \operatorname{span}\{x_0\}$ and $\|x_0\| = 1$: let $x_0 \in (\ker \ell)^\perp, \|x_0\| = 1$. Then for any $y \in H$,

$$y = \underbrace{(y - \frac{\ell(y)}{\ell(x_0)} x_0)}_{\in \ker \ell} + \underbrace{\frac{\ell(y)}{\ell(x_0)} x_0}_{\in \operatorname{span}\{x_0\} \subseteq (\ker \ell)^\perp} .$$

Define $x_\ell = \overline{\ell(x_0)} x_0$. Claim that $\ell(x) = (x, x_\ell)$ for all $x \in H$: if $x \in \ker \ell$ then $\ell(x) = 0$ and

$$(x, x_\ell) = \ell(x_0)(x, x_0) = 0.$$

If $x \in (\ker \ell)^\perp$, i.e. $x = \lambda x_0$ where $\lambda \in \mathbb{K}$, x_0 as above, then

$$(x, x_\ell) = \lambda \ell(x_0)(x_0, x_0) = \lambda \ell(x_0) = \ell(\lambda x_0) = \ell(x).$$

Since ℓ and $(\cdot, x_\ell)$ are in H^* and agree on $\ker \ell$ and $(\ker \ell)^\perp$, they also agree on H.

For uniqueness, if $(x, x_\ell) = (x, \tilde{x}_\ell)$ for all $x \in H$ then

$$(x, x_\ell - \tilde{x}_\ell) = 0$$

for all x, in particular

$$(x_\ell - \tilde{x}_\ell, x_\ell - \tilde{x}_\ell) = 0$$

so $x_\ell = \tilde{x}_\ell$.

Finally for isometry,

$$\|\ell\| = \sup_{\|x\| \leq 1} |\ell(x)| = \sup_{\|x\| \leq 1} |(x, x_\ell)| = \|x_\ell\|$$

where the last equality is by taking $x = \frac{x_\ell}{\|x_\ell\|}$ for lower bound and Cauchy-Schwarz for upper bound. ☐

Corollary 4.7. *The map*

$$\begin{aligned} H &\to H^* \\ x &\mapsto (\cdot, x) \end{aligned}$$

is antilinear, bijective and isometric.

Definition (projection, orthogonal projection)**.** Let X be an inner product space.

- A linear operator $P : X \to X$ is a *projection* if $P^2 = P$.
- A projection P is an *orthogonal projection* if $P^2 = P$ and P is self-adjoint, i.e.
$$(Px, y) = (x, Py)$$
for all $x, y \in X$.

Fact. Let P be an orthogonal projection. Then $\|P\| = 1$ or $\|P\| = 0$.

Proof. Suppose $P \neq 0$. For any $x \in X$ such that $Px \neq 0$,

$$\|Px\| = \frac{\|Px\|^2}{\|Px\|} = \frac{(Px, Px)}{\|Px\|} = \frac{(x, P^2x)}{\|Px\|} = \frac{(x, Px)}{\|Px\|} \leq \|x\|$$

so $\|P\| \leq 1$.

On the other hand, since $P \neq 0$ there is x such that $Px \neq 0$. Let $y = Px \neq 0$ and then
$$\|Py\| = \|Px\| = \|y\|$$
so $\|P\| \geq 1$. $\square$

Corollary 4.8. *Let $Y \subseteq X$ be a complete subspace. Then there is an orthogonal projection $P : X \to X$ with*

$$\begin{aligned} \operatorname{im} P &= Y \\ \ker P &= Y^\perp \end{aligned}$$

Proof. Given $x \in X$, let $x = x_\parallel + x_\perp$ with $x_\parallel \in Y$ and $x_\perp \in Y^\perp$ be its orthogonal decomposition. Set $Px = x_\parallel$. Then P is linear since given orthogonal decompositions of $x, y \in X$, have

$$\lambda x + \mu y = \underbrace{\lambda x_\parallel + \mu y_\parallel}_{\in Y} + \underbrace{(\lambda x_\perp + \mu y_\perp)}_{\in Y^\perp}$$

so
$$P(\lambda x + \mu y) = \lambda Px + \mu Py$$
by uniqueness of orthogonal decomposition.

Clearly $P^2 = P$. Also have

$$\begin{aligned} (Px, y) &= (x_\parallel, y_\parallel + y_\perp) \\ &= (x_\parallel, y_\parallel) \\ &= (x_\parallel + x_\perp, y_\perp) \\ &= (x, Py) \end{aligned}$$

so P is orthogonal. $\square$

Example. Let $X = C(\mathbb{T}, \mathbb{C})$ with inner product

$$(f, g) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f\overline{g} dx.$$

Then

$$\begin{aligned} S_n : X &\to X \\ f &\mapsto \sum_{k=-n}^{n} \hat{f}_k e^{ikx} \end{aligned}$$

where $\hat{f}_k = \frac{1}{2\pi} \int f(x) e^{-ikx} dx$, is the orthogonal projection with image $Y = \operatorname{span}\{e^{ikx}\}_{|k| \leq n}$ (which is finite-dimensional so complete).

Proof. Let $e_k(x) = e^{ikx}$. Then $\hat{f}_k = (f, e_k)$ and

$$S_n f = \sum_{k=-n}^{n} e_k (f, e_k).$$

If $f \in Y$, i.e. $f = \sum_{k=-n}^{n} a_k e_k$ then

$$S_n f = \sum_{k=-n}^{n} e_k (\sum_{\ell=-n}^{n} a_\ell e_\ell, e_k) = \sum_{k=-n}^{n} e_k a_k = f$$

as $\{e_k\}$ is an orthonormal basis. If $f \in Y^\perp$ then $(f, e_k) = 0$ for all $|k| \leq n$ so $S_n f = 0$. Thus S_n is a projection.

S_n is also orthogonal since

$$(S_n f, g) = \sum_{k=-n}^{n} \hat{f}_k (e_k, g) = \sum_{k=-n}^{n} \hat{f}_k \overline{(g, e_k)} = \sum_{k=-n}^{n} \hat{f}_k \overline{\hat{g}}_k = (f, S_n g)$$

so S_n is orthogonal. ☐

Corollary 4.9. *Let H be a Hilbert space and $S \subseteq H$. Then*

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

Proof. First equality follows from uniqueness of orthogonal decomposition and second equality follows from fact on a previous remark. ☐

4.3 Orthonormal systems

Definition (orthonormal system)**.** Let X be an inner product space. A set $\{e_\alpha\}_\alpha \subseteq X$ of unit vectors is an *orthonormal system* if $(e_\alpha, e_\beta) = 0$ for all $\alpha \neq \beta$. It is called *maximal* if it cannot be extended to a larger orthonormal system.

Definition (orthonormal basis)**.** Let H be a Hilbert space. Then a maximal orthonormal system is called *complete orthonormal system*, an *orthonormal Hilbert basis* or simply an *orthonormal basis*.

Note that an orthonormal basis is not a basis in the linear algebra sense.

Fact. Let H be a Hilbert space and S an orthonormal system. Then S is an orthonormal Hilbert basis if and only if $\overline{\operatorname{span} S} = H$.

Proof. Let $Y = \overline{\operatorname{span} S}$. Then Y is complete so $H = Y \oplus Y^\perp$. Suppose $Y^\perp \neq 0$. Then there is $x \in Y^\perp = S^\perp$, $\|x\| = 1$, i.e. $S \cup \{x\}$ is an orthonormal system. The converse also holds. □

Example.

1. In ℓ^2, let $e_n = (0, \dots, 0, 1, 0, \dots)$ be the vector with 1 at nth coordinate. Then $\{e_n\}_n$ is an orthonormal basis.
2. In $C(\mathbb{T}, \mathbb{C})$ with the usual inner product, let $e_n(x) = e^{inx}$. Then $\{e_n\}_{n\in\mathbb{Z}}$ is a maximal orthonormal basis. They are orthonormal and their span is dense by the complex Stone-Weierstrass theorem.

Fact (Gram-Schmidt)**.** Let X be an inner product space and $\{x_i\}_{i=1}^N \subseteq X$ be linearly independent, with $N = \infty$ allowed. Then there is an orthonormal system $\{e_i\}$ with $\operatorname{span}\{x_i\}_{i=1}^k = \operatorname{span}\{x_i\}_{i=1}^k$ for all $k \leq N$.

Sketch of proof. Let $e_1 = \frac{x_1}{\|x_1\|}$ and given $e_1, \dots, e_k$, set

$$e_{k+1} = \frac{x_{k+1} - \sum_{i=1}^k e_i(x_{k+1}, e_i)}{\left\| x_{k+1} - \sum_{i=1}^k e_i(x_{k+1}, e_i) \right\|}.$$

□

Example. Let $X = C([-1, 1], \mathbb{R})$ with $(f, g) = \int_{-1}^1 fg dx$. Then $1, t, t^2, \dots$ is a sequence with dense linear span by Stone-Weierstrass. Applying Gram-Schmidt, we obtain

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

These are an example of *orthogonal polynomials* and called the *normalised Legendre polynomials*. In fact, the nth one is a multiple of

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

Corollary 4.10. *Let H be a separable Hilbert space. Then there is a countable orthonormal Hilbert basis.*

Thus from now on, we will always assume orthonormal basis to be countable if H is separable.

Proposition 4.11 (Bessel's inequality)**.** *Let X be an inner product space and $\{e_i\}_{i=1}^N$ an orthonormal system, with $N = \infty$ allowed. Then*

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

for all $x \in X$.

In particular if $N = \infty$ then $(x_i) \in \ell^2$ where $x_i = (x, e_i)$.

Proof. By taking a limit suffice to prove the case $N < \infty$. Define

$$Px = \sum_{i=1}^{N} (x_i, e_i) e_i.$$

Then $P^2 = P$ and

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

Thus P is an orthogonal projection. Thus

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

for all $x \in X$. □

Proposition 4.12 (Riesz-Fisher)**.** *Let H be a separable infinite-dimensional Hilbert space with orthonormal basis $\{e_i\}_{i=1}^\infty$. Then*

1. *for any $x \in X$, set $x_i = (x, e_i) \in \mathbb{K}$. Then $(x_i) \in \ell^2$ and*

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

2. *conversely, if $(x_i) \in \ell^2$ then there is $x \in H$ such that $(x, e_i) = x_i$ for all i.*
3. *Parseval identity: for any $x, y \in H$,*

$$(x, y) = \sum_{i=1}^{\infty} x_i \overline{y}_i.$$

In particular the map

$$\begin{aligned} \phi : H &\to \ell^2 \\ x &\mapsto ((x, e_i))_{i=1}^\infty \end{aligned}$$

is an isometric isomorphism.

In fact, the concrete space ℓ^2 is the one studied by Hilbert and indeed this theorem shows that it is a prototype for a large class of Hilbert spaces. The term "Hilbert space" was coined by von Neumann, in his attempt to formulate quantum mechanics.

Proof.

1. Let $s_n = \sum_{i=1}^n x_i e_i$. Then (s_n) is Cauchy: for $m \geq n$,

$$\|s_m - s_n\|^2 = \left\| \sum_{i=n+1}^m x_i e_i \right\| = \sum_{i=n+1}^m |x_i|^2 \leq \sum_{i=n+1}^\infty |x_i|^2 \to 0$$

   as $n \to \infty$ since $(x_i) \in \ell^2$ by Bessel's inequality. By completeneess of H, there is $s \in H$ such that $s_n \to s$. Claim that $s = x$: for any i,

$$(s - x, e_i) = \lim_{n\to\infty} (s_n - x, e_i) = x_i - x_i = 0$$

   so

$$s - x \in (\text{span}\{e_i\})^\perp = \overline{\text{span}\{e_i\}}^\perp = H^\perp = 0.$$

2. If $(x_i) \in \ell^2$, the sum $x = \sum_{i=1}^\infty x_i e_i$ converges by the same argument. Then

$$(x, e_i) = \lim_{n\to\infty} \left(\sum_{j=1}^n x_j e_j, e_i \right) = x_i.$$

3. Similarly,

$$(x, y) = \lim_{n\to\infty} \lim_{m\to\infty} \left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^m y_j e_j \right) = \lim_{n\to\infty} \sum_{i=1}^n x_i \overline{y}_i$$

   since the infinite sum converges absolutely since $(x_i), (y_i) \in \ell^2$.

□

5 Spectral theory

Roughly speaking spetral theory studies eigenvalues of operators. From now on Banach and Hilbert spaces are complex.

5.1 Spectrum and resolvent

Definition (spectrum, resolvent)**.** Let X be a (complex) Banach space and $T \in B(X) = B(X, X)$.

- The *resolvent set* of T is
  $$\rho(T) = \{z \in \mathbb{C} : T - z = T - z\,\mathrm{id} \text{ is bijective and } (T-z)^{-1} \in B(X)\}.$$
- The *spectrum* of T is
  $$\sigma(T) = \mathbb{C} \setminus \rho(T).$$
- The *resolvent* of T is the map
  $$\begin{aligned} R_T : \rho(T) &\to B(X) \\ z &\mapsto (T-z)^{-1} \end{aligned}$$

Remark. If $T - z$ is bounded (as in our setting) and bijective the condition $(T-z)^{-1} \in B(X)$ is automatic by open mapping theorem. For unbounded operators, which we do not discuss, it does not follow automatically and has to be included in the definition.

Proposition 5.1. *Let $z_0 \in \rho(T)$. Then $\rho(T)$ contains the disk*
$$D = \{z \in \mathbb{C} : |z - z_0| \|R_T(z_0)\| < 1\}.$$
In particular $\rho(T)$ is open, $\sigma(T)$ is closed. Moreover the resolvent map R_T is analytic (can be represented by an absolutely convergent power series in any small enough disk).

Lemma 5.2. *Let $T \in B(X)$ with $\|T\| < 1$. Then the series $\sum_{n=0}^{\infty} T_n$ converges in $B(X)$ and*
$$(1 - T)^{-1} = \sum_{n=0}^{\infty} T^n$$
with
$$\|(1-T)^{-1}\| \leq \frac{1}{1 - \|T\|}.$$

Proof. Basically geometric series. Form the partial sums $S_n = \sum_{k=0}^{n} T^k$. Then (S_n) is a Cauchy sequence in $B(X)$ as for $m \geq n$,
$$\|S_m - S_n\| \leq \sum_{k=n+1}^{\infty} \|T^k\| \leq \sum_{k=n+1}^{\infty} \|T\|^k \to 0$$

as $n \to \infty$ since $\|T\| < 1$. Since $B(X)$ is complete the limit $S = \lim_{n\to\infty} S_n$ exists and by a similar argumet

$$\|S\| \le \sum_{k=0}^{\infty} \|T\|^k = \frac{1}{1-\|T\|}.$$

Moreover,

$$S(1-T) = \sum_{k=0}^{\infty} T^k - \sum_{k=1}^{\infty} T^k = \mathrm{id}\,.$$

□

Proof of Proposition 5.1. For $z \in D$,

$$T - z = (T - z_0) - (z - z_0) = (T - z_0)(1 - \underbrace{R_T(z_0)(z - z_0)}_{\|\cdot\| \le \|R_T(z_0)\||z-z_0|<1})$$

so by the lemma

$$(1 - R_T(z_0)(z - z_0))^{-1} = \sum_{n=0}^{\infty} (z - z_0)^n R_T(z_0)^n \in B(X).$$

Thus

$$(T - z)^{-1} = \sum_{n=0}^{\infty} (z - z_0)^n R_T(z_0)^{n+1} \in B(X)$$

so $z \in \rho(T)$. Thus $D \subseteq \rho(T)$ and R_T is analytic on $\rho(T)$. □

Corollary 5.3. *$\sigma(T) \neq \emptyset$ and*

$$\sigma(T) \subseteq \{z \in \mathbb{C} : |z| \le \|T\|\}.$$

Proof. For any $|z| > \|T\|$,

$$-R_T(z) = \frac{1}{z}\frac{1}{1 - T/z} = \frac{1}{z}\sum_{n=0}^{\infty} \frac{T^n}{z^n} \in B(X).$$

Thus $z \in \rho(T)$ and the second claim follows.

Also $\|R_T(z)\| \to 0$ as $z \to \infty$. Suppose for contradiction $\sigma(T) = \emptyset$. Then $R_T : \mathbb{C} \to B(X)$ would be entire so by Liouville's theorem (which holds when the codomain is a Banach space), it would have to be constant, thus 0. But this is absurd since, for example, $-zR_T(z) \to \mathrm{id}$ as $|z| \to \infty$. □

5.2 Classification of spectrum

We would like to understand why $T - z$ fails to be bijective. One reason, as in finite-dimensional case, is that $\ker(T - z) \neq 0$. But even if the kernel is trivial it may fail to be bijective.

Proposition 5.4. *Let X be a Banach space, Y a normed space and $T \in B(X, Y)$. Then its inverse $T^{-1} \in B(Y, X)$ if and only if $\operatorname{im} T$ is dense in Y and T is bounded below, i.e. exists $\varepsilon > 0$ such that for all $x \in X$, $\|Tx\| \ge \varepsilon\|x\|$.*

Proof. The only if direction is immediate. Thus assume that $T \in B(X,Y)$ is such that $\operatorname{im} T$ is dense and T is bounded below. Since T is bounded below T is injective so bijective onto its image. Let $S : \operatorname{im} T \to X$ be its inverse. Since T is bounded below, S is bounded. Since $\operatorname{im} T$ is dense in Y and X is complete, S extends uniquely to a map $\tilde{S} : Y \to X$. Moreover $\tilde{S} \in B(Y,X)$ and for any sequence $y_k \to y, (y_k) \subseteq \operatorname{im} T$,

$$T\tilde{S}y = \lim_{k\to\infty} TSy_k = \lim_{k\to\infty} y_k = y$$

so $\tilde{S} = T^{-1}$. $\square$

Definition (point spectrum, continuous spectrum, residual spectrum)**.** Let X be a Banach space and $T \in B(X)$.

- The *point spectrum* or set of *eigenvalues* is

$$\sigma_p(T) = \{\lambda \in \sigma(T) : T - \lambda \text{ is not injective}\}.$$

- The *continuous spectrum* is

$$\sigma_c(T) = \{\lambda \in \sigma(T) : T - \lambda \text{ injective and } \operatorname{im}(T - \lambda) \text{ is dense}\}.$$

- The *residual spectrum* is

$$\sigma_r(T) = \{\lambda \in \sigma(T) : T - \lambda \text{ is injective and } \operatorname{im}(T - \lambda) \text{ is not dense}\}.$$

Remark. By previous proposition, if $\lambda \in \sigma_c(T)$ then $T - \lambda$ is not bounded below. Thus there exists a sequence $(x_k) \subseteq X$ with $\|x_k\| = 1$ such that $Tx_k - \lambda x_k \to 0$ as $k \to \infty$. λ is called an *approximate eigenvalue.*

The set

$$\sigma_{ap}(T) = \{\lambda \in \sigma(T) : \lambda \text{ is an approximate eigenvalue}\}$$

is the *approximate point spectrum.*

Example. Let X be a finite-dimensional inner product space. Then $T - \lambda$ is injective if and only if it is surjective. Thus $\sigma(T) = \sigma_p(T)$. Moreover

$$\sigma_p(T) = \{\lambda \in \mathbb{C} : \det(T - \lambda) = 0\}$$

contains at most $n = \dim X$ points. In particular $\rho(T) = \mathbb{C} \setminus \sigma(T)$ is dense in $\mathbb{C}$.

5.3 Adjoints

Definition (dual/adjoint)**.** Let X, Y be normed spaces and $T \in B(X,Y)$. The *dual* or *adjoint* $T^* \in B(Y^*, X^*)$ is defined by

$$(T^*f)(x) = f(Tx)$$

for $f \in Y^*, x \in X$.

Fact. $T^*f \in X^*$ and $\|T^*\| = \|T\|$.

Proof.
$$|(T^*f)x| = |f(Tx)| \leq \|f\|_{Y^*}\|T\| \cdot \|x\|_X$$
so
$$\|T^*f\|_{X^*} \leq \|f\|_{Y^*}\|T\|$$
so $\|T^*\| \leq \|T\|$. In general $\|T^*\| \geq \|T\|$ follows from the Hahn-Banach theorem but for some spaces, such as ℓ^p and Hilbert spaces we may write down an explicit element saturating the bound. □

Definition. Let H be a Hilbert space and $T \in B(H)$. Let $\theta : H \to H^*$ be the isomorphism from Riesz representation theorem. Then we define $\tilde{T}^* \in B(H)$ by
$$\tilde{T}^*x = \theta^{-1}T^*\theta x$$
for $x \in H$.

Fact. $\tilde{T}^*$ is characterised by
$$(Tx, y) = (x, \tilde{T}^*y)$$
for all $x, y \in H$.

Proof. For all $x, y \in H$,
$$(x, \tilde{T}^*y) = (x, \theta^{-1}T^*\theta y) = T^*\theta y(x) = \theta y(Tx) = (Tx, y).$$
□

From now on, we write T^* instead of $\tilde{T}^*$ for a Hilbert space H, which is the only case that we will consider from now on.

Example.

1. $\mathrm{id}^* = \mathrm{id}$. More generally $(\lambda\,\mathrm{id})^* = \overline{\lambda}\,\mathrm{id}$.
2. Let $H = \ell^2$. If T is the left shift operator then T^* is the right shift operator.

Fact. Let H be a Hilbert space and $S, T \in B(H)$. Then

1. $(\lambda S + \mu T)^* = \overline{\lambda}S^* + \overline{\mu}T^*$ for all $\lambda, \mu \in \mathbb{K}$.
2. $(ST)^* = T^*S^*$.
3. $(T^*)^* = T$.
4. $\|T^*T\| = \|T\|^2$.

Note. In general $\|T\|^2 \neq \|T^2\|$, for example $T \neq 0$ such that $T^2 = 0$.

5.4 Normal linear operators

Definition (normal, self-adjoint, unitary)**.** $T \in B(H)$ is

- *normal* if $TT^* = T^*T$.
- *self-adjoint* if $T = T^*$.
- *unitary* if $T^{-1} = T^*$.

In particular self-adjoint and unitary operators are normal.

Exercise. Let $T \in B(H)$ be normal. Then $\|Tx\| = \|T^*x\|$ for all x and

$$\ker T = \ker T^* = (\operatorname{im} T)^\perp = (\operatorname{im} T^*)^\perp.$$

See example sheet 4. It follows that

$$\overline{\operatorname{im} T} = ((\operatorname{im} T)^\perp)^\perp = (\ker T)^\perp = (\ker T^*)^\perp = \overline{\operatorname{im} T^*}.$$

Corollary 5.5. *For T normal, $\sigma_r(T) = \emptyset$.*

Corollary 5.6. *For T normal, if $Tx = \lambda x$ then $T^*x = \overline{\lambda}x$. In particular*

$$\sigma_p(T^*) = \overline{\sigma_p(T)}$$

where the bar denotes conjugation.

Similarly if $Tx_j - \lambda x_j \to 0$ then $T^*x_j - \lambda x_j \to 0$ so

$$\sigma_{ap}(T^*) = \overline{\sigma_{ap}(T)}.$$

Proof. If T is normal then so is $T - \lambda$ and $(T-\lambda)^* = T^* - \overline{\lambda}$. Thus

$$\|(T-\lambda)x\| = \|(T-\lambda)^*x\| = \|(T^* - \overline{\lambda})x\|$$

so $Tx = \lambda x$ if and only if $T^*x = \overline{\lambda}x$. More generally $(T-\lambda)x_j \to 0$, where $\|x_j\| = 1$, if and only if $(T^* - \overline{\lambda})x_j \to 0$ so $\sigma_{ap}(T^*) = \overline{\sigma_{ap}(T)}$. $\square$

Corollary 5.7. *Let T be self-adjoint. Then $\sigma(T) \subseteq \mathbb{R}$.*

Exercise. Let T be unitary. Then $\sigma(T) \subseteq S^1 = \{z \in \mathbb{C} : |z| = 1\}$.

Example. Let $T \in B(H)$ be self-adjoint. Then for any $t \in \mathbb{R}$,

$$e^{itT} = \sum_{n=0}^\infty \frac{(itT)^n}{n!}$$

converges in $B(H)$ and $U(t) = e^{itT}$ is characterised by the ODE

$$-i\frac{\partial}{\partial t}U(t) = TU(t), U(0) = \operatorname{id}.$$

For any $t \in \mathbb{R}$, $U(t) = e^{itT}$ is unitary if T is self-adjoint.

In quantum mechanics, any solution to the Schrödinger equation

$$-i\frac{\partial}{\partial t}\tau(t) = T\psi(t), \psi(0) = \psi_0$$

is given by $\psi(t) = \psi_0 U(t)$. T is called the *Hamiltonian* and ψ is the *wave function*.

Lemma 5.8. *Let $T \in B(H)$ be self-adjoint. Then*

$$\|T\| = \sup_{\|x\|\leq 1} \sup_{\|y\|\leq 1} |(Tx, y)| = \sup_{\|x\|\leq 1} |(Tx, x)|.$$

Proof. Assume $T \neq 0$. By defintion $\|T\| = \sup_{\|x\|\leq 1}(Tx, Tx)^{1/2}$. Let $(x_i) \subseteq H$, $\|x_i\| = 1$ be such that $(Tx_i, Tx_i) \to \|T\|^2$. Thus

$$\begin{aligned}
\|T\| &= \frac{\|T\|^2}{\|T\|} \\
&= \frac{1}{\|T\|} \lim_{i\to\infty} (Tx_i, Tx_i) \\
&= \frac{1}{\|T\|} \lim_{i\to\infty} (x_i, T^2 x_i) \\
&= \lim_{i\to\infty} \frac{1}{\|Tx_i\|} (x_i, T^2 x_i) \\
&= \lim_{i\to\infty} (x_i, Ty_i)
\end{aligned}$$

where $y_i = \frac{Tx_i}{\|Tx_i\|}$. Taking sup,

$$\|T\| \leq \sup_{\|x\|\leq 1} \sup_{\|y\|\leq 1} (x, Ty).$$

On the other hand, $|(Tx, y)| \leq \|T\|$ for $\|x\|, \|y\| \leq 1$ so

$$\|T\| \geq \sup_{\|x\|\leq 1} \sup_{\|y\|\leq 1} |(Tx, y)|$$

so the first equality.

For the second equality,

$$\sup_{\|x\|\leq 1} |(Tx, x)| \leq \sup_{\|x\|\leq 1} \sup_{\|y\|\leq 1} |(Tx, y)|$$

is clear. For the other direction

$$\begin{aligned}
|(x, Ty)| &= \frac{1}{4} |(T(x+y), x+y) - (T(x-y), x-y)| \\
&\leq \frac{1}{4} \sup_{\|z\|\leq 1} |(Tz, z)| (\|x+y\|^2 + \|x-y\|^2) \\
&\leq \frac{1}{4} \sup_{\|z\|\leq 1} |(Tz, z)| \underbrace{(2\|x\|^2 + 2\|y\|^2)}_{\leq 4} \\
&\leq \sup_{\|z\|\leq 1} |(Tz, z)|
\end{aligned}$$

□

Lemma 5.9. *Let T be self-adjoint. Then at least one of $\|T\|$ and $-\|T\|$ must be an approximate eigenvalue.*

Proof. Replacing T by $-T$ if necessary, assume

$$\|T\| = \sup_{\|x\|\leq 1} |(x, Tx)| = \sup_{\|x\|\leq 1} (x, Tx).$$

Then there is $(x_i) \subseteq H$, $\|x_i\| = 1$ such that $(x_i, Tx_i) \to \|T\| = \lambda$. Then

$$\|Tx_i - \lambda x_i\|^2 = \underbrace{\|Tx_i\|^2}_{\leq \lambda^2} - 2\lambda \underbrace{(x_i, Tx_i)}_{\to \lambda} + \lambda^2 \to 0$$

□

5.5 Spectral theorem for compact self-adjoint operators

Recall that given normed spaces X and Y, $T \in B(X, Y)$ is *compact* if $T(B)$ is relatively compact in Y for any bounded set $B \subseteq X$.

> **Lemma 5.10.** *Let T be compact. Then any nonzero approximate eigenvalue is an eigenvalue.*

Proof. Assume that $Tx_i - \lambda x_i \to 0$ with $\|x_i\| = 1, \lambda \neq 0$. By compactness of T there is a subsequence such that $Tx_i \to y$ along that subsequence. Then along that subsequence

$$Ty = T \lim_{i\to\infty} Tx_i = \lim_{i\to\infty} T(\lambda x_i) = \lambda \lim_{i\to\infty} Tx_i = \lambda y$$

so $Ty = \lambda y$. Moreover if $\lambda \neq 0$ then $y \neq 0$. □

> **Corollary 5.11.** *Let H be a Hilbert space. Let $T \in B(H)$ be self-adjoint and compact. Then $\|T\|$ or $-\|T\|$ is an eigenvalue.*

Notation. $E_\lambda = \ker(T - \lambda)$ is the *eigenspace* corresponding to eigenvalue λ.

The strategy is to diagonalise T, construct a sequence of eigenvalues and eigenspaces by applying the corollary to $E_{\lambda_1}^\perp$ where $\lambda_1 = \|T\|$ or $-\|T\|$ and repeat this with H replaced by $E_{\lambda_1}^\perp$.

> **Lemma 5.12.** *Let $T \in B(H)$ be self-adjoint. Then*
>
> 1. *for any eigenvalues $\mu \neq \lambda$ the spaces E_μ and E_λ are orthogonal.*
> 2. *for any nonzero eigenvalues $\{\lambda_i\}_{i\in I}$,*
>
> $$T((\bigoplus_{i\in I} E_{\lambda_i})^\perp) \subseteq (\bigoplus_{i\in I} E_{\lambda_i})^\perp.$$

Proof.

1. Assume $Tx = \mu x$ and $Ty = \lambda y$ and wlog $\lambda \neq 0$. Then

$$(x, y) = \frac{1}{\lambda}(Tx, y) = \frac{1}{\lambda}(x, Ty) = \frac{\mu}{\lambda}(x, y)$$

so $\mu = \lambda$ or $(x, y) = 0$.

2. Let $y \in (\bigoplus_{i \in I} E_{\lambda_i})^\perp$. Then for any $x \in E_{\lambda_i}$ where $\lambda_i \neq 0$, have

$$0 = (x, y) = \frac{1}{\lambda_i}(Tx, y) = \frac{1}{\lambda_i}(x, Ty)$$

so $(x, Ty) = 0$ for all $x \in \bigoplus_{i \in I} E_{\lambda_i}$, i.e. $Ty \in (\bigoplus_{i \in I} E_{\lambda_i})^\perp$.

□

Lemma 5.13. *Let $T \in B(H)$ be self-adjoint and compact. Then for every $\varepsilon > 0$, $\bigoplus_{\substack{\lambda \in \sigma_p(T) \\ |\lambda| > \varepsilon}} E_\lambda$ is finite-dimensional.*

Proof. Assume otherwise. Then there are infinitely many eigenvectors (x_i) such that $\|x_i\| = 1$ and $(x_i, x_j) = 0$ for $i \neq j$ from Gram-Schmidt and

$$\|Tx_i - Tx_j\|^2 = \|Tx_i\|^2 + \|Tx_j\|^2 \geq 2\varepsilon^2,$$

contradicting compactness of T as (Tx_i) does not have a convergent subsequence.

□

Theorem 5.14 (Hilbert-Schmidt)**.** *Let $T \in B(H)$ be self-adjoint and compact. Then there are at most countably many distinct eigenvalues (λ_i) which can accumulate at 0. The eigenspaces E_{λ_i} and E_{λ_j} are orthogonal for $i \neq j$. E_{λ_i} is finite-dimensional for $\lambda_i \neq 0$, and*

$$T = \sum_{j=0}^{\infty} \lambda_j P_{\lambda_j}$$
$$H = (\ker T) \oplus (\overline{\bigoplus_i E_{\lambda_i}})$$

where P_{λ_j} is the orthogonal projection onto E_{λ_j}.

Proof. By the previous two lemmas there is an eigenvalue λ_1 such that $\lambda_1 = \|T\|$ or $-\|T\|$. Given $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_k|$ such that T has no other eigenvalues $> |\lambda_k|$, let

$$H_k = (\bigoplus_{j=1}^{k} E_{\lambda_j})^\perp$$

Then $H_k \subseteq H$ is closed so is itself a Hilbert space. Also T is H_k-stable and $\|T|_{H_k}\| \leq |\lambda_k|$. Thus there is an eigenvalue λ_{k+1} different from the λ_i's for $i < k$, with $|\lambda_{k+1}| = \|T|_{H_k}\| \leq |\lambda_k|$ and there is no other eigenvalue μ with $|\lambda_{k+1}| < |\mu| \leq |\lambda_k|$. This defines a sequence (λ_i) with $|\lambda_{i+1}| \leq |\lambda_i|$ for all i. Since $\bigoplus_{|\lambda_i| \geq \varepsilon} E_{\lambda_i}$ is finite-dimensional, the sequence can only accummulate at 0. In particular (λ_k) is finite or countable with $\lambda_k \to 0$.

Since $T|_{(\bigoplus_{\lambda \neq 0} E_\lambda)^\perp} = T|_{(\overline{\bigoplus_{\lambda \neq 0} E_\lambda})^\perp}$ cannot have a nonzero eigenvalue, must have $\|T|_{(\overline{\bigoplus_{\lambda \neq 0} E_\lambda})^\perp}\| = 0$. Thus

$$H = (\ker T) \oplus (\overline{\bigoplus_{\lambda \neq 0} E_\lambda})$$

and

$$\begin{aligned}
\|Tx - \sum_{i=1}^{n} \lambda_i P_{\lambda_i} x\| &= \|Tx - TP_n x\| \quad \text{where } P_n = \sum_{i=1}^{n} P_{\lambda_i} \\
&= \|T(1-P_n)x\| \quad \text{where } 1 - P_n \text{ is projection onto } (\bigoplus_{i=1}^{n} E_{\lambda_i})^{\perp} = H_n \\
&\le \|T|_{H_n}\|\|x\| \\
&\le |\lambda_n|\|x\|
\end{aligned}$$

so $\|T - \sum_{i=1}^{n} \lambda_i P_{\lambda_i}\| \le |\lambda_n| \to 0$ as $n \to \infty$. Thus

$$T = \sum_{i=1}^{\infty} \lambda_i P_{\lambda_i}$$

as claimed. □

Corollary 5.15. *Let T be self-adjoint and compact. Then*

$$\sigma(T) \cup \{0\} = \sigma_p(T) \cup \{0\}.$$

Proof. Let $T_n = \sum_{i=1}^{n} \lambda_i P_{\lambda_i}$. Let $\mu \notin \sigma_p(T) \cup \{0\}$. Then

$$T_n - \mu = \sum_{i=1}^{n} (\lambda_i - \mu) P_{\lambda_i} - \mu \underbrace{(1 - \sum_{i=1}^{n} P_{\lambda_i})}_{1-P_n}$$

so

$$(T_n - \mu)^{-1} = \sum_{i=1}^{n} (\lambda_i - \mu)^{-1} P_{\lambda_i} - \mu^{-1}(1 - P_n)$$

exists and

$$\|(T_n - \mu)^{-1}\| \le \max\{|\mu|^{-1}, |\lambda_i - \mu|^{-1}\} \le C$$

as $\lambda_i \ne \mu$ and $\mu \ne 0$. Thus

$$T - \mu = T_n - \mu + (T - T_n) = (T_n - \mu)(\mathrm{id} + \underbrace{\underbrace{(T_n - \mu)^{-1}}_{\le C} \underbrace{(T - T_n)}_{\to 0}}_{<1 \text{ eventually}})$$

so $(T - \mu)^{-1} \in B(H)$, thus $\mu \notin \sigma(T)$. □

Corollary 5.16. *Let H be a separable and $T \in B(H)$ self-adjoint and compact. Then there is an orthonormal basis for H of T eigenfunctions.*

Proof. Apply Gram-Schmidt to eigenspaces. □

5.6 Application: boundary value problem

Let T be the 1-dimensional Schrödinger operator acting on $C^2[a,b]$ by

$$Tu(x) = -u''(x) + V(x)u(x), u \in C^2[a,b]$$

with boundary condition

$$u(a) = u(b) = 0$$

and where $V \in C[a,b]$.

We want to apply the spectral theorem but there are two imminent problems: $C^2[a,b]$ is not a Hilbert space, and T is not even a operator on $C^2[a,b]$.

Theorem 5.17. *There exists a continuous function (*Green function*) $k : [a,b] \to \mathbb{R}$ such that the unique solution $u \in C^2[a,b]$ to the boundary value problem*

$$\begin{aligned} Tu(x) &= f(x) \quad f \textit{ continuous} \\ u(a) &= u(b) = 0 \end{aligned}$$

is given by

$$\begin{aligned} u(x) &= \int_a^b k(x,y)f(y)dy \\ k(x,y) &= k(y,x) \end{aligned}$$

Proof. Result in analysis. Omitted. ☐

Lemma 5.18. *Let $k : [a,b]^2 \to \mathbb{R}$ be continuous. Then the integral operator*

$$\begin{aligned} K : (C[a,b], \|\cdot\|_2) &\to (C[a,b], \|\cdot\|_\infty) \\ Kf(x) &= \int_a^b k(x,y)f(y)dy \end{aligned}$$

is bounded and compact.

Proof. By Cauchy-Schwarz,

$$\begin{aligned} \|Kf\|_\infty &\leq \sup_x \int_a^b |k(x,y)||f(y)|dy \\ &\leq \underbrace{\sup_x \left(\int |k(x,y)|^2 dy \right)^{1/2}}_{\leq C} \left(\int |f(y)|^2 dy \right)^{1/2} \\ &\leq C\|f\|_2 \end{aligned}$$

so K is bounded. Also $B = \{Kf : f \in C[a,b], \|f\|_2 \leq 1\}$ is equicontinuous:

$$|Kf(x) - Kf(y)| \leq \underbrace{\left(\int |k(x,z) - k(y,z)|^2 dz \right)^{1/2}}_{\to 0 \text{ as } |x-y| \to 0 \text{ by continuity}} \|f\|_2 \to 0$$

uniform in $f \in B$. Thus compactness follows from Arzelà-Ascoli. ☐

Corollary 5.19. *$K : (C[a,b], \|\cdot\|_2) \to (C[a,b], \|\cdot\|_2)$ is compact.*

Proof.

$$\|Kf\|_2 = \left(\int |Kf(x)|^2 dx\right)^{1/2} \leq \sqrt{|a-b|}\|Kf\|_\infty$$

so the embedding of $(C[a,b], \|\cdot\|_\infty) \to (C[a,b], \|\cdot\|_2)$ is continous. □

Let $H = L^2[a,b]$ be the completion of $(C[a,b], \|\cdot\|_2)$. In particular $C[a,b]$ is dense in $L^2[a,b]$.

Fact. Let X, Y be Banach spaces and $D \subseteq X$ a dense subspace. Then a bounded (compact, respectively) operator $T : D \to Y$ extends uniquely to a bounded (compact, respectively) operator $T : X \to Y$ with the same operator norm.

Corollary 5.20. *K extends uniquely to a compact self-adjoint operator $K : H \to H$. Moreover $Kf \in C[a,b]$ for any $f \in H$.*

Proof. That K is compact follows from the above fact. That K is self-adjoint follows from the symmetry of k:

$$\begin{aligned}
(f, Kg) &= \int f(x)K(x,y)\overline{g(y)}dxdy \\
&= \int f(x)K(y,x)\overline{g(y)}dxdy \\
&= \overline{(g, Kf)} \\
&= (Kf, g)
\end{aligned}$$

That $Kf \in C[a,b]$ for any $f \in H$ follows from the fact that K is also bounded from $(C[a,b], \|\cdot\|_2) \to (C[a,b], \|\cdot\|_\infty)$ and thus by the previous fact from $H \to (C[a,b], \|\cdot\|_\infty)$. Finally embed $(C[a,b], \|\cdot\|_\infty)$ to H. All of these constructions are unique. □

Putting everything together, by the spectral theorem, there exists an orthonormal basis $(f_n) \subseteq H$ with eigenvalues $(\mu_n) \subseteq \mathbb{R}, \mu_n \to 0$ such that

$$Kf = \sum_{n=1}^\infty \mu_n (f, f_n) f_n$$

in H. By the last corollary, if $\mu_n \neq 0$ then

$$f_n = \frac{1}{\mu_n} K f_n \in C[a,b]$$

and in fact then $f_n \in C^2[a,b]$ since $Kf \in C^2[a,b]$ if $f \in C[a,b]$.

Assuming $\ker k = 0$, there is thus an orthonormal basis of C^2 eigenfunctions of K. Write $\lambda_n = \frac{1}{\mu_n}$, have

$$Tf_n = \lambda_n TKf_n = \lambda_n f_n$$

so these eigenfunctions f_n are also eigenfunctions of T_n and $\lambda_n \to \infty$.

In quantum mechanics, the vectors $f \in H$ describe the state of a system and $|f(x)|^2$ the probability density of finding a particle at $x \in [a,b]$.

6 Hahn-Banach theorem*

When can one extend a bounded linear map defined on some $Y \leq X$ onto X? Naturally we require the extension of a bounded map to be bounded. In fact we do things slightly more generally.

> **Definition** (sublinear map). Let X be a real vector space. A map $\rho : X \to \mathbb{R}$ is *sublinear* if
>
> 1. $\rho(\alpha x) = \alpha \rho(x)$ for all $x \in X, \alpha \geq 0$.
> 2. $\rho(x + y) \leq \rho(x) + \rho(y)$ for all $x, y \in X$.

Example. Any norm is sublinear.

> **Theorem 6.1** (Hahn-Banach). *Let X be a real vector space and Y a subspace of X. Let $\rho : X \to \mathbb{R}$ be sublinear, $g : Y \to \mathbb{R}$ linear such that $g(x) \leq \rho(x)$ for all $x, y \in Y$. Then there exists $f : X \to \mathbb{R}$ linear such that $f|_Y = g$ and $f(x) \leq \rho(x)$ for all $x \in X$.*

Proof if Y has codimension 1 *in X.* Suppose Y has codimension 1, i.e. there exists $x_1 \in X \setminus Y$ such that

$$X = \operatorname{span} X \cup \{x_1\} = Y \oplus \mathbb{R}x_1.$$

We'll find $\alpha \in \mathbb{R}$ such that

$$\begin{aligned} f_\alpha : X &\to \mathbb{R} \\ x + tx_1 &\mapsto g(x) + t\alpha \end{aligned}$$

where $x \in Y$, is the required extension. We take

$$\alpha = \sup_{x \in Y}(g(x) - \rho(x - x_1)).$$

and show $f = f_\alpha$ works. Claim first that $\alpha < \infty$: by linearity of g and sublinearity of ρ, for all $x, y \in Y$,

$$g(x) + g(y) = g(x + y) \leq \rho(x + y) \leq \rho(x - x_1) + \rho(y + x_1).$$

Rearrange,

$$g(x) - \rho(x - x_1) \leq -g(y) + \rho(y + x_1)$$

which in particular implies that $\alpha < \infty$.

Note that $f(x - x_1) \leq \rho(x - x_1)$ for all $x \in Y$ by choice of α. Also claim that $f(y + x_1) \leq \rho(y + x_1)$ for all $y \in Y$:

$$f(y + x_1) = g(y) + \alpha \leq \alpha \underbrace{-(g(x) - \rho(x - x_1))}_{\inf = -\alpha} + \rho(y + x_1)$$

so take infimum over x and the result follows.

With this we can extend the boundedness to all scalars: $f(x + tx_1) \leq \rho(x + tx_1)$ for all $t \in \mathbb{R}, x \in Y$: by linearity of positive homogeneity of ρ, for all $t > 0$

$$f(x \pm tx_1) = tf(\frac{x}{t} \pm x_1) \leq t\rho(\frac{x}{t} \pm x_1) = \rho(x \pm tx_1).$$

$\square$

Essentially we're done here as we can keep extending g to a larger space.. However when the codimension is not finite, there is a slight (or enormous, depending on how seriously you treat axiom of choice) issue with tethe termination of this process.

Definition (total order, maximal element)**.** Let P be a poset.

- A subset $T \subseteq P$ is *totally ordered* if for all $x, y \in T$, either $x \leq y$ or $y \leq x$.
- An element $m \in P$ is *maximal* if for all $x \in P$, $m \leq x$ implies $x = m$.

Proposition 6.2 (Zorn's lemma)**.** *Let $P \neq \emptyset$ be a poset such that whenever T is a totally ordered subset then there exists a least upper bound for T. Then there exists a maximal element of P.*

Note that this easily follows from axiom of choice. For more discussion see IID Logic and Set Theory.

Proof of Hahn-Banach. Let

$$P = \{(N, h) : N \leq X, h : N \to \mathbb{R} \text{ linear}, h|_Y = g, h(x) \leq \rho(x) \text{ for all } x \in N\}.$$

Equip P with partial order $\leq$ where $(N, h) \leq (N', h')$ if and only if $N \subseteq N'$ and $h'|_N = h$. As $(Y, g) \in P$, P is nonempty.

To apply Zorn's lemma we need to check that every totally ordered subset has an upper bound. Let $\{(N_i, h_i)\}_{i \in I} \subseteq P$ be totally ordered. Then let $N = \bigcup_{i \in I} N_i$ and $h(x) = h_i(x)$ if $x \in N_i$. h is well-defined. Clearly N is a subspace of X and $h(v) \leq \rho(v)$ for all N and $h_N = g$ so $(N, h) \in P$. It is also an upper bound. Thus by Zorn's lemma, P has a maximal element (M, f). We must have $M = X$ as otherwise (M, f) would not be maximal by codimension 1 argument. □

Corollary 6.3. *Let V be a normed vector space and $W \leq V$. Let $g \in W^*$. Then there exists $f \in V^*$ such that $f|_W = g$ and $\|f\| \leq \|g\|$.*

Proof. If V is real then apply Hahn-Banach with ρ being norm on V. The complex case is similar. □

Corollary 6.4. *Let V be a normed vector space and $v \in V$. Then there exists $f_v \in V^*$ such that $\|f_v\| = 1, f_v(v) = \|v\|$.*

Such an f_v is called a support functional *for f.*

In particular this implies that the dual of any nontrivial normed space is nontrivial.

Proof. Let W be the span of v and define

$$\begin{aligned} g : W &\to \mathbb{R} \\ tv &\mapsto t\|v\| \end{aligned}$$

Then $\|g\| = 1, g(v) = \|v\|$. Extend g to f_v by Hahn-Banach gives desired element of V^*. □

Corollary 6.5. *Let V be a normed vector space and $v \in V$. If $f(v) = 0$ for all $f \in V^*$ then $v = 0$.*

Index