Part II – Linear Analysis

Based on lectures by Dr R. Bauerschmidt
Notes taken by Bhavik Mehta

Michaelmas 2017

Contents

0   Introduction

As the name suggests, Linear Analysis is the study of linear spaces of functions, mostly infinite dimensional. In particular, properties like convexity, completeness, closedness are of interest here. Like any pure course, we start with a lot of definitions which come out of nowhere, and then clear them up, but functional analysis is not devoid of interesting theorems. In particular, in the field of differential equations both ordinary and partial it is often useful to view the differential operators as a linear operator on a space of functions. Markov processes can also be seen using a transition operator, and dynamical processes are given by a measure preserving map, all fitting into the realm of linear maps. Quantum mechanics to a certain extent is the study of the spectrum of certain self-adjoint linear operators on a Hilbert space, and so requires functional analysis. As much as possible, examples of applications will be given briefly.

1   Normed vector spaces

Unless stated, vector spaces will be either over the real numbers or the complex numbers, denoted by $\mathbb{K}$ to represent $\mathbb{R}$ or $\mathbb{C}$.

1.1   Topology of vector spaces

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

i. $\|v\| \geq 0$ for all $v \in V$ and $\|v\| = 0$ if $v = 0$. (Positive definite)

ii. $\|\lambda v\| = |\lambda| \|v\|$ for every $v \in V$ and $\lambda \in K$. (Positive homogeneous)

iii. $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in V$. (Triangle inequality)

In particular, a metric on V is defined by $d(v, w) = \|v - w\|$.

Fact. The vector space operations of scalar multiplication and vector addition are continuous.

$$K \times V \to V \qquad\qquad (\lambda, v) \mapsto \lambda v$$
$$V \times V \to V \qquad\qquad (v, w) \mapsto v + w$$

Proof. We only check that scalar multiplication is continuous. Since K and V are metric spaces, it suffices to show that $\lambda_j \to \lambda$ and $v_j \to v$ implies $\lambda_j v_j \to \lambda v$. But

$$\|\lambda_j v_j - \lambda v\| = \|(\lambda_j - \lambda)v_j + \lambda(v_j - v)\|$$
$$\leq \underbrace{|\lambda_j - \lambda|}_{\to 0} \underbrace{\|v_j\|}_{\text{bounded}} + |\lambda| \underbrace{\|v_j - v\|}_{\to 0}$$

as required. $\square$

Corollary. Translations $(v \mapsto v + v_0)$ and dilations $(v \mapsto \lambda v, \lambda \neq 0)$ are homomorphisms.

Definition (Topological vector space). A **topological vector space** is a vector space together with a topology that makes the vector space operations continuous and in which points are closed.

Notation. For a subset C of a vector space V over $\mathbb{K}$ and $t \in \mathbb{K}$, we write tC for the following subset:

$$tC := \{\, tv \mid v \in C \,\}$$

Definition (Convex subset)**.** Let V be a vector space and $C \subset V$ a subset. We say that C is **convex** if and only if $tC + (1-t)C \subset C$ for all $t \in [0,1]$. Specifically, this means $tv + (1-t)w \in C$ for all $v, w \in C$ and $t \in [0,1]$.

Fact. Let V be a normed vector space. Then $B_1(0)$ is convex.

Fact. If C is convex, then $v + \lambda C$ is convex for all $\lambda \in K$ and $v \in V$.

Definition (Locally convex space)**.** A topological vector space is **locally convex** if its topology has a basis of convex sets.

Definition (Bounded subset)**.** Let V be a topological vector space and $B \subset V$. We say that B is **bounded** if for every open neighbourhood U of 0, there exists $t > 0$ such that $sU \supset B$ for all $s \geq t$.

Definition (Balanced subset)**.** Let V be a vector space, and $C \subset V$ a subset. Call C **balanced** if for all $|\lambda| \leq 1$, we have $\lambda C \subset C$.

Example.

(i) Balanced sets in $\mathbb{R}$ are sets of the form $[-t, t]$, $(-t, t)$, or $\{0\}$ and all of $\mathbb{R}$, and $\mathbb{Q}$ is not balanced in $\mathbb{R}$.

(ii) In $\mathbb{C}$, the only balanced sets are $\{0\}$, $\mathbb{C}$, and the open or closed balls centred at 0.

(iii) There are more interesting examples of balanced sets in $\mathbb{R}^2$, for instance the open disk or any ellipse centred at 0.

Lemma. Let V be a topological vector space and $C \subset V$ be a bounded, convex neighbourhood of 0. Then there exists a bounded, balanced, convex neighbourhood $\widetilde{C}$ of 0.

Proof. Exercise (on example sheet). $\qquad\square$

Proposition. Let V be a topological vector space and $C \subset V$ be a bounded, convex neighbourhood of 0. Then the topology on V is induced by a norm.

Proof. Use the previous lemma to construct $\widetilde{C}$. Let

$$\mu_{\widetilde{C}}(v) = \inf \{\, t > 0 \mid v \in t\widetilde{C} \,\}$$

referred to as the Minkowski functional of $\widetilde{C}$. We claim that $\|v\| = \mu_{\widetilde{C}}(v)$ is a norm on V and that the topology induced by it is the same as the original topology. Check the norm axioms in turn:

i. We clearly have positivity, and $\mu_{\widetilde{C}}(v) = 0$ if and only if $v = 0$ since $\widetilde{C}$ is bounded.

ii. Since $\widetilde{C}$ is balanced,

$$
\begin{aligned}
\mu_{\widetilde{C}}(\lambda v) &= \inf\left\{ t > 0 \,\Big|\, \lambda v \in t\widetilde{C} \right\} \\
&= \inf\left\{ t > 0 \,\Big|\, v \in \frac{t}{|\lambda|}\widetilde{C} \right\} \\
&= \inf\left\{ |\lambda|\frac{t}{|\lambda|} > 0 \,\Big|\, v \in \frac{t}{|\lambda|}\widetilde{C} \right\} \\
&= |\lambda|\,\mu_{\widetilde{C}}(v)
\end{aligned}
$$

iii. Given $v, w \in V$, write $v = \lambda v_0$ and $w = \mu w_0$ with $\lambda, \mu > 0$, $v_0, w_0 \in \widetilde{C}$. Since $\widetilde{C}$ is convex,

$$
\frac{\lambda v_0 + \mu w_0}{\lambda + \mu} \in \widetilde{C}
$$
$$
\implies \mu_{\widetilde{C}}\left(\frac{\lambda v_0 + \mu w_0}{\lambda + \mu} \right) \leq 1
$$

Therefore,

$$
\begin{aligned}
\mu_{\widetilde{C}}(v + w) &= (\lambda + \mu)\,\mu_{\widetilde{C}}\left(\frac{\lambda v_0 + \mu w_0}{\lambda + \mu} \right) \\
&\leq \lambda + \mu \\
&\leq \mu_{\widetilde{C}}(v) + \mu_{\widetilde{C}}(w)
\end{aligned}
$$

$\square$

Corollary. A topological vector space is normable if and only if it is locally convex and locally bounded (that is, there exists a bounded convex neighbourhood of 0).

Definition (Banach space). A **Banach** space is a normed vector space that is *complete* as a metric space, that is, any Cauchy sequence converges.

Example.

(i) Any finite dimensional vector space is a Banach space (with any norm).

(ii) Let X be a set, and $\mathcal{B}(X)$ the set of bounded $\mathbb{K}$-valued functions on X. Then $\mathcal{B}(X)$ is a Banach space with norm

$$
\|f\|_\infty = \sup_{x \in X} |f(x)|\,, \quad f \in \mathcal{B}(X)
$$

(iii) Let X be a *compact* Hausdorff space (e.g. $X = [0,1]$) and $\mathcal{C}(X)$ be the space of continuous functions on X. Then $\mathcal{C}(X) \subset \mathcal{B}(X)$ since every continuous function on a compact space is bounded. In addition, $\mathcal{C}(X)$ is a Banach space, as the uniform limit of a sequence of continuous functions is continuous.

(iv) Let $U \subset \mathbb{R}^n$ be an open, bounded subset, and let $\mathbb{C}^{\mathbb{K}}(\bar{U})$ be the space of k–times continuously differentiable functions $f : \bar{U} \to \mathbb{K}$ with norm defined as follows:

$$D^\alpha f(x) = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} f(x)$$

for any multi-index $\alpha = (\alpha_1, \dots \alpha_n) \in \mathbb{N}_0^n$, and where $|\alpha| = \sum_{i=1}^n \alpha_i$. From here, we use the norm

$$\|f\|_{C^k(\bar{U})} = \max_{|\alpha| \le k} \|D^\alpha f\|_\infty$$

Then $C^{\mathbb{K}}(\bar{U})$ is a Banach space.

(v) For a sequence $x = (x_1, x_2, \dots) \subset \mathbb{K}$, define

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

for $p \in [1, \infty)$, and $\|x\|_\infty = \sup_i |x_i|$. Then $\ell^p = \{ x \mid \|x\| < \infty \}$ for $p \in [1, \infty]$ is a Banach space. Note for $p < 1$, this does not generate a norm.

(vi) Let $U \subset \mathbb{R}^n$ be open (not necessarily bounded) and denote by $\mathcal{C}(U)$ the space of continuous functions on U (not necessarily bounded). Then $\mathcal{C}(U)$ is a topological vector space with topology generated as follows.

Let $K_i \subset U$ be compact subsets such that $K_i \subset K_{i+1}$ and $\bigcup_{i=1}^\infty K_i = U$.

Let $V(i, n) = \left\{ f \mid \|f\|_{\mathcal{C}(K_i)} < \frac{1}{n} \right\}$ where we use $\|f\|_{\mathcal{C}(K_i)}$ to denote $\sup_{x \in K_i} |f(x)|$.

Then the topology generated by the $V(i, n)$ and their translates makes $\mathcal{C}(U)$ a locally convex topological vector space, and the topology on $\mathcal{C}(U)$ is generated by the metric

$$d(f, g) = \sum_{i=1}^\infty 2^{-i} \frac{\|f - g\|_{\mathcal{C}(K_i)}}{1 + \|f - g\|_{\mathcal{C}(K_i)}}$$

This is not a Banach space, but it is a Frechet space.

(vii) Let $X = \{ f : [0, 1] \to \mathbb{K} \text{ continuous} \}$. Then

$$\|f\|_p = \left(\int_0^1 |f(x)|^p \, dx \right)^{\frac{1}{p}}$$

for $p \in [1, \infty)$ is a norm on X. However, X is not complete with this norm.

1.2 Bounded linear maps and the dual space

Fact. In any topological vector spaces V, W, a linear map $T : V \to W$ is continuous if and only if it is continuous at 0.

Proof. Let T be continuous at 0 and $v \in V$. Let $w = Tv$ and $U \subset W$ an open neighbourhood of w. Then $U - w$ is an open neighbourhood of $0 \in W$. Since T is continuous at 0, $T^{-1}(U - w)$ contains an open neighbourhood $U' \subset V$ of 0. By linearity, $T(v + U') = Tv + T(U') \subset Tv + U - w = U$. Since $v + U'$ is an open neighbourhood of v this means that T is continuous at v. □

Definition (Bounded linear map). Let V, W be topological vector spaces and $T : V \to W$ a linear map. Then T is **bounded** if $T(B)$ is bounded for any bounded $B \subset V$.

Fact. If V, W are normed vector spaces, a linear map $T : V \to W$ is bounded if and only if there is $\lambda > 0$ such that

$$T(B_1(0)) \subset B_\lambda(0) \quad \text{i.e.} \quad \|Tv\| < \lambda \text{ if } \|v\| \leq 1$$

Definition (Operator norm). Let V, W be normed vector spaces. The **operator norm** of a linear map $T : V \to W$ is

$$\|T\| = \sup_{\|v\|=1} \|Tv\| = \sup_{\|v\|\leq 1} \|Tv\|$$

Denote by $\mathcal{L}(V, W)$ the space of linear maps $V \to W$ and by $\mathcal{B}(V, W)$ the space of bounded linear maps $V \to W$.

Fact. The operator norm $\|\cdot\|$ is a norm on $\mathcal{B}(V, W)$.

Proposition. Let V, W be normed vector spaces. Then a linear map $T : V \to W$ is bounded if and only if it is continuous.

Proof.

- T bounded $\implies T$ continuous

  Assume $\|v_k - v\| \to 0$. Then

  $$\|Tv_k - Tv\| \leq \underbrace{\|T\|}_{<\infty} \underbrace{\|v_k - v\|}_{\to 0} \to 0$$

- T continuous $\implies T$ bounded

  Since T is continuous, $T^{-1}(B_1(0))$ contains an open ball $B_\epsilon(0) \subset V$. Thus

  $$T(B_1(0)) = \epsilon^{-1} T(B_\epsilon(0)) \subset \epsilon^{-1} B_1(0) = B_{\epsilon^{-1}}(0)$$

  so T is bounded.

$\square$

Definition (Dual space). Let V be a topological vector space. The (topological) **dual space** of V is the space of continuous linear maps $V \to \mathbb{K}$, denoted by V^*. In a normed space, $V^* = \mathcal{B}(V, \mathbb{K})$. (We call $\mathcal{L}(V, \mathbb{K})$ the algebraic dual of V.)

Proposition. Let V be a normed vector space and W a Banach space. Then $\mathcal{B}(V, W)$ is a Banach space.

Proof. Let $(T_i) \subset \mathcal{B}(V, W)$ be a Cauchy sequence. We aim to show that there is some $T \in \mathcal{B}(V, W)$ such that $T_i \to T$. For any vector $v \in V$, the sequence $(T_i v) \subset W$ is Cauchy:

$$\|T_i v - T_j v\| \leq \underbrace{\|T_i - T_j\|}_{\leq \epsilon} \cdot \|v\|$$

Since W is complete, $T_i v$ converges, and define Tv as this pointwise limit.

- T is linear:

$$T(\lambda v + \mu w) = \lim_{i \to \infty} T_i(\lambda v + \mu w)$$
$$= \lambda \lim_{i \to \infty} T_i v + \mu \lim_{i \to \infty} T_i w$$
$$= \lambda T v + \mu T w$$

- T is bounded: For $\|v\| \leq 1$,

$$\|Tv\| \leq \underbrace{\|T_i v\|}_{\leq \|T_i\|} + \underbrace{\|T_i v - Tv\|}_{<1 \text{ for } i > N(v)}$$
$$\leq \sup_i \|T_i\| + 1$$
$$\leq \infty$$

- $T_i \to T$ in norm:

$$\|T_i v - Tv\| \leq \underbrace{\|T_i v - T_j v\|}_{\leq \|T_i - T_j\|} + \underbrace{\|T_j v - Tv\|}_{<\epsilon \text{ for } j > N(\epsilon, v), \, j \geq i}$$
$$\implies \|T_i - T\| \to 0 \text{ as } i \to \infty$$

$\square$

Corollary. Let V be a normed vector space. Then V^* is a Banach space.

Definition (Dual Map)**.** Let V, W be normed vector space and $T \in \mathcal{B}(V, W)$. Then the **adjoint** or **dual map** $T^* : W^* \to V^*$ is defined by

$$[T^* f]v = f(Tv) \quad f \in W^*, \, v \in V$$

Fact. We can easily see $T^* f$ is in $V^* = \mathcal{B}(V, K)$ and we can check $\|T^*\| \leq \|T\|$:

Proof.

$$\|(T^* f)v\| = \|f(Tv)\|$$
$$\leq \|f\| \, \|Tv\|$$
$$\leq \|f\| \, \|T\| \, \|v\|$$
$$\implies \|T^* f\| \leq \|f\| \, \|T\|$$
$$\implies \|T^*\| \leq \|T\|$$

Later we will see this last inequality is in fact equality. $\square$

Definition (Double dual)**.** Let V be a normed vector space. The **double dual** of V is the dual space of V^*, that is $V^{**} = (V^*)^*$.

Fact. The map $\phi : V \to V^{**}, v \mapsto \widetilde{v}$ where $\widetilde{v}(f) = f(v)$, $f \in V^*$ is bounded and linear.

Remark. The Hahn-Banach Theorem implies that ϕ is isometric (injective), so V can be considered a subspace of V^{**}. In general, ϕ is not surjective.

Definition (Reflexive space). A Banach space is **reflexive** if ϕ is a bijection.

Example.

- Let V, W be finite-dimensional vector spaces. Then any linear map $T : V \to W$ is bounded with respect to any norms on V, W, and T can be represented by a matrix while T^* is represented by the transpose of T. Finite dimensional normed vector spaces are reflexive.

- The map $T : \ell^p \to \ell^p$, $(x_1, x_2, \dots) \mapsto (0, x_1, x_2, \dots)$ is bounded, $\|T\| = 1$, injective but not surjective.

- The derivative map $D : C^1[0,1] \to C^0[0,1]$ is

    bounded as a map $C^1[0,1]$, $\|f\|_\infty + \|f'\|_\infty$ to $C^0[0,1]$ $\|f\|_\infty$

    unbounded as a map $C^1[0,1]$, $\|f\|_\infty$ to $C^0[0,1]$, $\|f\|_\infty$

- The map $\mathrm{id} : C^1[0,1]$, $\|\cdot\|_\infty \to C^1[0,1]$, $\|\cdot\|_\infty + \|\cdot\|'_\infty$ is unbounded.

1.3 Finite-dimensional normed spaces

Fact. Any finite dimensional vector space can be identified with $\mathbb{K}^n$ for some n by choosing a basis.

Definition (Norm equivalence). Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a vector space V are **equivalent** if there exists $C > 0$ such that

$$C^{-1} \|v\|_2 \le \|v\| \le C \|v\|_2$$

for $v \in V$.

Proposition. All norms on a finite dimensional space are equivalent.

Proof. If suffices to check that any norm $\|\cdot\|$ on $\mathbb{K}^n$ is equivalent to $\|\cdot\|_1$ where

$$\|v\|_1 = \sum_{i=1}^n |v_i|$$

- $\|v\| \le C \|v\|_1$: Let $e_i = (0, \dots, 0, 1, 0, \dots) \in \mathbb{K}^n$. Then

$$\|v\| = \left\| \sum_{i=1}^n v_i e_i \right\| \le \sum_{i=1}^n \|v_i\| \|e_i\| \le \left(\max_{i=1,\dots,n} \|e_i\| \right) \|v\|_1$$

- $\|v\|_1 \le \|v\|$: Let $S = \{ v \in \mathbb{K}^n \mid \|v\|_1 = 1 \}$ and $f = S \to \mathbb{R}$, given by $v \mapsto \|v\|$.

    1. f is continuous with respect to $\|\cdot\|_1$ topology:

    $$|f(v) - f(w)| = |\|v\| - \|w\|| \le \|v - w\| \le C \|v - w\|_1$$

    2. S is compact with respect to $\|\cdot\|_1$ topology: It is closed as the inverse image of $\{1\}$ under $\|\cdot\|_1$, and it is bounded by definition.

So, f is a continuous function on a compact set, and hence attains its minimum. Since $\|\cdot\|$ and $\|\cdot\|_1$ are norms (positive definite), this minimum must be strictly positive. So,

$$\|v\| = f(v) \geq f(v_0) = c$$

for all $\|v\|_1 = 1$. For arbitrary $v \in \mathbb{K}^n$, we have

$$\|v\| = \|v\|_1 \left\| \frac{v}{\|v\|_1} \right\| = \|v_1\| f\left(\frac{v}{\|v\|_1} \right) \geq c \|v\|_1$$

$\square$

Corollary. In any finite-dimensional normed space, the closed unit ball is compact.

Corollary. Every finite-dimensional normed space is a Banach space.

Proof. Let (v_i) be Cauchy, then (v_i) is bounded, that is $(v_i) \subset \overline{B_R(0)}$ for some $R > 0$. But then $\overline{B_R(0)}$ is compact, so (v_i) converges. $\square$

Corollary. Let V be a normed space, $W \subset V$ a finite-dimensional subspace. Then W is closed.

Corollary. Let V be a normed space, W be a finite-dimensional normed space and $T : W \to V$ is linear, then T is bounded.

Proof. Since $\operatorname{Im} T$ is finite dimensional, we can assume without loss of generality that V is finite dimensional. Moreover, we can assume $V = \mathbb{K}^m, \|\cdot\|_\infty$ and $W = \mathbb{K}^n, \|\cdot\|_1$. Let (T_{ij}) be the matrix associated to T.

$$T(w_1, \ldots, w_n) = \left(\sum_{i=1}^n T_{1i} w_i, \ldots, \sum_{i=1}^n T_{mi} w_i \right) \in V$$

$$\implies \|Tw\|_\infty = \max_{j=1,\ldots,m} |(Tw)_j|$$

$$= \max_{j=1,\ldots,m} \left| \sum_{i=1}^n T_{ji} w_i \right|$$

$$\leq \max_{i,j} |T_{ji}| \sum_{i=1}^n |w_i|$$

$$= C(T) \|w\|_1$$

$\square$

Theorem. Let V be a normed vector space such that $\overline{B_1(0)}$ is compact. Then V is finite-dimensional.

Proof. Since $\overline{B_1(0)}$ is compact, there are $w_1, \ldots, w_n \in V$ such that

$$\overline{B_1(0)} \subset \bigcup_{i=1}^n B_{\frac{1}{2}}(w_i)$$

Let $W = \text{span}\,\{\,w_1, \dots, w_n\,\}$. Note that $\dim W \leq n$. Our goal is to show $W = V$.

$$B_1(0) \subset W + B_{\frac{1}{2}}(0)$$

$$B_1(0) \subset W + \frac{1}{2}\left(W + B_{\frac{1}{2}}(0)\right) = W + B_{\frac{1}{4}}(0)$$

By induction, $B_1(0) \subset W + B_{2^{-i}}(0)$ for any $i \in \mathbb{N}$, hence $B_1(0) \subset \overline{W} - W$. Since V is a vector space, $V \subset W$, and so $\dim V \leq n$. $\qquad\square$

2 The Hahn-Banach Theorem

Definition (Extends). Given vector spaces $W \subset V$, and linear maps $g : W \to \mathbb{K}$, $f : V \to \mathbb{K}$, we say that f **extends** g if $f \mid_W = g$, that is $f(v) = g(v)$ for every $v \in W$.

When can one extend a linear map in a continuous way?

2.1 Finite codimension

Let V be a *real* vector space.

Definition (Sublinear). A map $p : V \to \mathbb{R}$ is **sublinear** if

1. $p(\alpha v) = \alpha p(v)$ for all $v \in V$, $\alpha \geq 0$

2. $p(v + w) \leq p(v) + p(w)$ for all $v, w \in V$.

Example. Any norm is sublinear.

Lemma (Codimension 1 case of the Hahn-Banach Theorem). Let $W \subset V$ be a subspace of codimension 1, that is, there exists $v_1 \in V \setminus W$ such that

$$V = \{\, v + tv_1 \mid v \in W,\, t \in \mathbb{R} \,\}$$

Let $p : V \to \mathbb{R}$ be sublinear, and $g : W \to \mathbb{R}$ linear with $g(v) \leq p(v)$ for all $v \in W$. Then there exists a linear map $f : V \to \mathbb{R}$ that extends g and $f(v) \leq p(v)$ for all $v \in V$.

Proof. We will find $\alpha \in \mathbb{R}$ such that $f_\alpha : V \to \mathbb{R}$ defined by

$$f_\alpha(v + tv_1) = g(v) + t\alpha$$

for all $v \in W$, $t \in \mathbb{R}$ is the asserted linear extension of g. Let

$$\alpha = \sup_{v \in W} (g(v) - p(v - v_1))$$

We proceed with a sequence of claims which will complete the proof.

Claim $\alpha < \infty$. By linearity of g and sublinearity of p,

$$g(v) + g(w) = g(v + w) \leq p(v + w) \leq p(v - v_1) + p(w + v_1)$$

for all $v, w \in W$.

So, $g(v) - p(v - v_1) \leq p(w + v_1) - g(w)$ and hence $\alpha \leq p(w + v_1) - g(w) < \infty$.

Claim $f_\alpha(v - v_1) \leq p(v - v_1)$ for all $v \in W$.

$$
\begin{aligned}
f_\alpha(v - v_1) &= g(v) - \alpha \\
&= \underbrace{g(v) - p(v - v_1)}_{\leq \alpha} - \alpha + p(v - v_1) \\
&\leq p(v - v_1)
\end{aligned}
$$

12

Claim $f_\alpha(w + v_1) \leq p(w + v_1)$ for all $w \in W$. By (*),

$$
\begin{aligned}
f_\alpha(w + v_1) &= g(w) + \alpha \\
&\leq \alpha - \underbrace{(g(v) - p(v - v_1))}_{\sup = 0} + p(w + v_1)
\end{aligned}
$$
$$
\implies f_\alpha(w + v_1) \leq p(w + v_1)
$$

Claim $f_\alpha(v + tv_1) \leq p(v + tv_1)$ for all $v \in W$, $t \in \mathbb{R}$. By linearity of f and positive homogeneity of p, for any $t > 0$,

$$
\begin{aligned}
f(v \pm tv_1) &= tf\left(\frac{v}{t} \pm v_1\right) \\
&\leq tp\left(\frac{v}{t} \pm v_1\right) \\
&= p(v \pm tv_1) \qquad\qquad\qquad \square
\end{aligned}
$$

Corollary. The same statement holds if $W \subset V$ is a subspace of *finite* codimension.

Proof. By induction. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

2.2   Zorn's Lemma

Definition (Poset).

(i) A partially ordered set (**poset**) is a set p with a binary relation $\leq$ such that for all $x, y \in P$ either $x \leq y$ or $x \not\leq y$, and

$$
\begin{aligned}
&x \leq x && \text{(reflexive)} \\
&x \leq y, y \leq z \implies x \leq z && \text{(transitive)} \\
&x \leq y, y \leq x \implies x = y && \text{(antisymmetric)}
\end{aligned}
$$

(ii) Let P be a poset. A subset $T \subset P$ is called **totally ordered** (or a chain) if $x \not\leq y \implies y \leq x$, that is, at least one of $x \leq y$ or $y \leq x$ must hold.

(iii) Let P be a poset and $U \subseteq$ a subset. Then

   - $b \in P$ is an **upper bound** for U if $x \leq b$ for all $x \in U$
   - $l \in P$ is a **least upper bound** for U if l is an upper bound and any other upper bound b for U satisfies $l \leq b$.

(iv) Let P be a poset. An element $m \in P$ is **maximal** if $m \leq x \implies m = x$.

Lemma (Zorn's Lemma)**.** Let P be a poset with the property that every non-empty totally ordered subset has a least upper bound. Then P has at least one maximal element.

Remark.

(i) Zorn's Lemma is trivial if P is finite.

(ii) In infinite posets, it is equivalent to the Axiom of Choice (assuming the usual other axioms of set theory).

Recall that in a vector space V, elements $v_1, \ldots, v_k$ are linearly independent if

$$\sum_{i=1}^{k} \alpha_i v_i = 0 \implies \alpha_1 = \alpha_2 = \cdots \alpha_k = 0$$

A set $S \subset V$ is linearly independent if any *finite* subset is. A basis of V is a set $B \subset V$ that is linearly independent and such that every element of V is a *finite* linear combination of elements of B.

Proposition. Let $V \neq \{0\}$ be a vector space and $S \subset V$ linearly independent. Then V has a basis B containing S.

Proof. Let $P = \{\, T \subset V \mid T \supset S, T \text{ linearly independent}\,\}$. Then P is a poset with partial order $\subseteq$, i.e. $T_1 \leq T_2$ if and only if $T_1 \subseteq T_2$. For any $L \subseteq P$ that is totally ordered, set $T_b = \bigcup_{T \in L} T$.

We claim $T_b \in P$ and T_b is a least upper bound for L. Clearly, $S \subseteq T_b$. Let $\alpha_i \in K$ and $v_i \in T_b$ such that $\sum_{i=1}^{m} \alpha_i v_i = 0$. Since L is totally ordered and $m < \infty$, there must be $T \in L$ such that $v_1, \ldots, v_m \in T$. Since T is linearly independent, we have $\alpha_1 = \cdots = \alpha_m = 0$. Thus T_b is linearly independent. Thus, $T_b \in P$ and clearly a least upper bound for L.

By Zorn's Lemma, there exists a maximal element B in P.

Claim: B is a basis for V. By definition of P, B is linearly independent. We need to check that B spans V. Suppose otherwise. Then there exists $v \in V \setminus \operatorname{span} B$ and $B \cup \{v\}$ is linearly independent. This contradicts the maximality of B, so B is a basis for V. $\qquad\square$

2.3   The Hahn-Banach Theorem

Theorem (Hahn-Banach)**.** Let V be a *real* vector space, $W \subset V$ a subspace, $p : V \to \mathbb{R}$ sublinear, and $g : W \to \mathbb{R}$ linear such that $g(v) \leq p(v)$ for all $v \in W$.

Then there exists $f : V \to \mathbb{R}$ linear such that $f|_W = g$ and $f(v) \leq p(v)$ for all $v \in V$.

Proof. Let

$$P = \{\, (N, h) \mid N \subset V \text{ subspace}, W \subset N, h : N \to \mathbb{R} \text{ linear}, h(v) \leq p(v)\, \forall v \in N \,\}$$

For $(N, h), (N', h') \in P$, set $(N, h) \leq (N', h')$ if and only if $N \subset N'$ and $h'|_N = h$. Then P is a poset, $P \neq \varnothing$ because $(W, g) \in P$. We want to apply Zorn's Lemma.

Let $(N_i, h_i)_{i \in I} \subseteq P$ be a totally ordered subset of P. Set

$$N = \bigcup_{i \in I} N_i, \quad h(v) = h_i(v) \text{ if } v \in N_i$$

This is well-defined since $(N_i, h_i)_{i \in I}$ is totally ordered. Moreover, $N \subseteq V$ is a subspace and $h : N \to \mathbb{R}$ is linear. Also $h(v) \leq p(v)$ for all $v \in \mathbb{N}$. Thus, $(N, h) \in P$. Clearly, (N, h) is a least upper bound for $(N_i, h_i)_{i \in I}$.

By Zorn's Lemma, there is a maximal element (M, f) in P.

Claim: $M = V$. Suppose not, then the codimension 1 case gives (M', f') with $(M, f) \lneq (M', f')$, a contradiction of maximality of (M, f). $\qquad\square$

Corollary. Let V be a normed vector space (real or complex), $W \subseteq V$ a subspace. For any $g \in W^*$, there exists $f \in V^*$ such that $f|_W = g$, $\|f\| \leq \|g\|$.

Proof. Assume that V is real. The complex case is treated on an example sheet. Then $p(v) = \|v\| \, \|g\|$ is sublinear and the claim follows directly from the general Hahn-Banach Theorem. $\qquad \square$

Corollary. Let V be a normed vector space, and $v \in V$. Then there exists $f_v \in V^*$ such that $\|f_v\| = 1$ and $f_v(v) = \|v\|$. Such an f_v is called a **support functional** for v.

Proof. Let $W = \{\, tv \mid t \in K \,\}$. Define $g \in W^*$ by $g(tv) = t \|v\|$. Then $\|g\| = 1$ and $g(v) = \|v\|$. By Hahn-Banach, there is $f = f_v$ as desired. $\qquad \square$

Remark. In concrete examples, one can often construct f_v by hand.

Corollary. Let V be a normed vector space and $v \in V$. Then $v = 0 \iff f(v) = 0$ for all $f \in V^*$. In particular, $V^* \neq \{0\}$.

Corollary. Let V be a normed space, $v, w \in V$, $v \neq w$. Then $\exists f \in V^*$ such that $f(v) \neq f(w)$.

Proof. Take $f = f_{v-w}$. $\qquad \square$

Corollary. The map $\phi : V \to V^{**}$ given by $\phi(v)f = f(v)$ is an isometry:

$$\forall v \in V, \|\phi(v)\| = \|v\|$$

Proof. First check $\|\phi(v)\| \leq \|v\|$:

$$\begin{aligned}
\|\phi(v)\| &= \sup_{\|f\|=1} |\phi(v)f| \\
&= \sup_{\|f\|=1} |f(v)| \\
&\leq \|v\| \, .
\end{aligned}$$

On the other hand,

$$|\phi(v)f_v| = |f_v(v)| = \|v\| \implies \|\phi(v)\| \geq \|v\| \, . \qquad \square$$

Corollary. Let V, W be normed spaces. For any $T \in \mathcal{B}(V, W)$, the dual map $T^* \in \mathcal{B}(W^*, V^*)$ satisfies $\|T^*\| = \|T\|$.

Proof. We have already seen $\|T^*\| \leq \|T\|$. Now, with $f = f_{T_v} \in W^*$, $[T^*f]v = f(Tv) = \|Tv\|$ so

$$\begin{aligned}
\|T^*\| &= \sup_{\|f\|\leq 1} \|T^*f\| \\
&\geq \|T^*f_{T_v}\| \\
&= \sup_{\|v\|\leq 1} \|[T^*f_{T_v}]v\| \\
&\geq \sup_{\|v\|\leq 1} \|Tv\| \\
&= \|T\| \qquad\qquad\qquad\qquad\qquad \square
\end{aligned}$$

3 Completeness and the Baire Category Theorem

3.1 Baire Category

Definition (Dense)**.** Recall that if X is a metric space, a subset $Y \subset X$ is **dense** if $\overline{Y} = X$. Equivalently $Y \cap B_r(X) \neq \varnothing$ for any $x \in X$, $r > 0$.

Theorem (Baire Category Theorem)**.** Let X be a complete metric space. For any sequence of open and dense sets $U_j \subset X$, $\bigcap_j U_j$ is dense in X.

Proof. Let $U = \bigcap_j U_j$. Given any $x \in X$, $r > 0$, we need to show that $B_r(x) \cap U \neq \varnothing$. Since U_1 is dense and open, there exists $x_1 \in X$, $r_1 \in (0, 1)$ such that

$$B_r(x) \cap U_1 \supset B_{2r_1}(x_1) \supset \overline{B_{r_1}(x_1)}$$

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

$$\overline{B_{r_2}(x_2)} \subset B_{r_1}(x_1) \cap U_2$$

and in general $x_n \in X$, $r_n \in (0, 2^{-n})$ such that

$$\overline{B_{r_n}(x_n)} \subset B_{r_{n-1}}(x_{n-1}) \cap U_n$$

Then $r_n \to 0$ and $B_{r_1}(x_1) \supset B_{r_2}(x_2) \supset \cdots$ so $d(x_n, x_m) < r_n \ \forall m \geq n$. So, (x_n) is a Cauchy sequence, and since X is complete, $\exists y \in X$ such that $x_n \to y$.
Note that $y \in \overline{B_{r_k}(x_k)} \cap U_{k+1}$ for all $k \implies y \in \bigcap_k U_k$ and $y \in B_r(x)$. $\qquad \square$

Corollary. Let X be a complete metric space. Let $A_j \subset X$ be a sequence of closed subsets of such that $\bigcup_j A_j$ has nonempty interior, that is, contains some ball. Then at least one of the A_j 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 are open, the Baire Category Theorem implies that there must be at least one U_j that is not dense. Thus $A_j = X \setminus U_j$ has nonempty interior. $\qquad \square$

Definition (Meagre sets)**.** Let X be a metric space.

(i) A subset $Y \subset X$ is **nowhere dense** if $\mathrm{int}(\overline{Y}) = \varnothing$, equivalently Y is not dense in any ball.

(ii) A subset $Z \subset X$ is **meagre** or of the **first category** if there are countably many sets $Y_j \subset X$ that are nowhere dense and $Z = \bigcup_j Y_j$.

(iii) A subset $U \subset X$ is **nonmeagre** or of the **second category** if it is not meagre.

(iv) A subset $R \subset X$ is **residual** if $X \setminus R$ is meagre.

Fact. $Y \subset X$ is nowhere dense $\iff \overline{Y}$ is nowhere dense $\iff X \setminus \overline{Y}$ is open and dense.

Example.

- $\mathbb{Q}$ is meagre in $\mathbb{R}$

- Any countable union of meagre sets is meagre.

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

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

Proof. Let $Y_j \subset X$ be nowhere dense sets. It suffices to show that $X \neq \bigcup_j \overline{Y_j}$. But $U_j = X \setminus \overline{Y_j}$ is open dense. So $\bigcap_j U_j = X \setminus \bigcup_j \overline{Y_j}$ is dense by the Baire Category Theorem. So in particular, $\bigcap_j U_j \neq \varnothing$, that is, $\bigcup_j Y_j \neq X$. $\qquad\square$

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

Proof. Let $Z \subset X$ meagre and suppose that $R = X \setminus Z$. Assume R was meagre. Then $X = R \cup Z$ would be the union of two meagre sets, so meagre. But since X is complete, it is not meagre, a contradiction.

To show R is dense, we can suppose that $Z = \bigcup_j Y_j$ with Y_j nowhere dense. Then $U_j = X \setminus \overline{Y_j}$ is dense and open. So, $\bigcap_j U_j$ is dense. But $R \supset \bigcap_j U_j$, so R is dense. $\qquad\square$

Corollary. Let X be a complete metric space and $U \subset X$ open. Then $U = \varnothing$ 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 $X \setminus U = X$, so $U = \varnothing$. $\qquad\square$

3.2   Principle of Uniform Boundedness

Theorem. Let X be a complete metric space and 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:

$$\forall x \in X \ \sup_{\lambda \in \Lambda} |f_\lambda(x)| < \infty$$

then there is a ball $B_r(x_0) \subset X$ such that (f_λ) is uniformly bounded on $B_r(x_0)$:

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

Proof. Let

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

Since the f_λ are continuous, A_k is closed as the intersection of closed sets. Since (f_λ) is pointwise bounded,

$$\bigcup_{n \in \mathbb{N}} A_n = X.$$

By the Baire Category Theorem, at least one of the A_λ must contain a ball $B_r(x_0)$. Therefore, (f_λ) is uniformly bounded (by k) on that ball. $\qquad\square$

Theorem (Banach-Steinhaus). Let V be a Banach space, and W a normed space. Let $(T_\lambda)_{\lambda \in \Lambda} \subset \mathcal{B}(V, W)$ be pointwise bounded:

$$\sup_{\lambda \in \Lambda} \|T_\lambda v\| < \infty \quad \forall v \in V.$$

Then (T_λ) is uniformly bounded:

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

Proof. Let $f_\lambda : V \to \mathbb{R}$, given by $v \mapsto \|T_\lambda v\|$. Then f_λ is continuous and (f_λ) is pointwise bounded. By the principle of uniform boundedness, there is a ball $B_r(v_0) \subset V$ such that (f_λ) is uniformly bounded on $B_r(v_0)$, i.e.

$$\sup_{\lambda \in \Lambda} \sup_{\|v - v_0\| < r} \|T_\lambda v\| < \infty.$$

Since the T_λ are linear, for any $v \in V$ with $\|v\| = 1$,

$$\begin{aligned}
\|T_\lambda v\| &= \|T_\lambda(v_0 + v) - T_\lambda v_0\| \\
&\leq \frac{1}{r} \sup_{\lambda \in \Lambda} \sup_{\|v - v_0\| < r} \|T_\lambda v\| + \frac{1}{r} \sup_\lambda \|T_\lambda v_0\| \\
&=: M < \infty \\
&\text{independent of } v, \|v\| = 1 \text{ and } \lambda
\end{aligned}$$

$$\implies \sup_{\lambda \in \Lambda} \|T_k\| \leq M \qquad \square$$

3.3 Open Mapping Theorem

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

Example.

(i) $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ is continuous but not open, as e.g. $f((-1, 1)) = [0, 1)$ is not open.

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

Theorem (Open Mapping Theorem). Let V, W be Banach spaces and $T \in \mathcal{B}(V, W)$.

(i) If T is surjective, then T is open.

(ii) If T is bijective, then $T^{-1} \in \mathcal{B}(W, V)$.

Lemma. Let V, W be normed spaces. Then $T : V \to W$ is open if $T(B_1(0)) \supset B_r(0)$ for $r > 0$.

Proof. Let $U \subset V$ be open and $v \in U$. Using that U is open, choose $\delta > 0$ such that $B_\delta(v) \subset U$. Then

$$\begin{aligned}
T(U) &\supset T(v + B_\delta(0)) \\
&= Tv + \delta T(B_1(0)) \\
&\supset Tv + \delta r B_1(0).
\end{aligned}$$

Therefore $T(U)$ contains a ball around any element Tv, so T is open. $\qquad \square$

Lemma. Let V be a Banach space, W a normed space, $T \in \mathcal{B}(V, W)$. Then

$$\underbrace{\overline{T(B_1(0))}}_{\subset V} \supset \underbrace{B_1(0)}_{\subset W} \implies T(B_1(0)) \supset B_1(0).$$

Proof. Let $w_0 \in B_1(0) \subset W$. We need to find $v \in B_1(0) \subset V$ such that $Tv = w_0$. We construct v as the limit of a Cauchy sequence. Let $v_1 \in B_{\frac{1}{2}}(0) \subset V$ such that

$$\|Tv_1 - w_0\| < \frac{1}{2}.$$

This is possible since there is $\widetilde{w}_0 \in B_{\frac{1}{2}}(0)$ with $\|w_0 - \widetilde{w}_0\| < \frac{1}{2}$ and we can find $v_1 \in B_{\frac{1}{2}}(0)$ such that $\|Tv_1 - \widetilde{w}_0\|$ arbitrarily small by $\overline{T(B_{\frac{1}{2}}(0))} \supset B_{\frac{1}{2}}(0)$ (assumption rescaled).

Set $w_1 = w_0 - Tv \in B_{\frac{1}{2}}(0)$. By induction, if $w_1, \dots, w_k$ and $v_1, \dots, v_k$ are given such that $\|v_i\| < 2^{-i}$,

$$w_i = w_{i-1} - Tv_i \in B_{2^{-i}}(0) \subset W, \ 1 \leq i \leq k$$

then choose $v_{k+1} \in B_{2^{-k-1}}(0) \subset V$ such that

$$w_{k+1} = w_k - Tv_{k+1} \in B_{2^{-k-1}}(0).$$

This gives

$$\sum_{k=1}^{\infty} \|v_k\| < 1 \text{ and } v := \sum_{k=1}^{\infty} v_k \in B_1(0) \text{ exists}$$

since V is complete.

$$\implies w_0 - Tv = \lim_{n \to \infty} \left(w_0 - \sum_{k=1}^{n} Tv_k \right)$$

$$= \lim_{n \to \infty} \left(w_1 - \sum_{k=2}^{n} Tv_k \right)$$

$$\vdots$$

$$= \lim_{n \to \infty} w_n = 0$$

$$\implies w_0 \in T(B_1(0)) \text{ for any } w_0 \in B_1(0)$$

$$\implies T(B_1(0)) \supset B_1(0). \qquad \square$$

We now return to prove the Open Mapping Theorem.

Proof.

(i) By the previous two lemmas, it suffices to prove that $\overline{T(B_1(0))} \supset B_r(0)$ for some $r > 0$. We use the Baire Category Theorem to do this. Since T is surjective,

$$W = \bigcup_{k=1}^{\infty} \overline{T(B_k(0))}.$$

Since W is complete, the Baire Category Theorem implies that there is a $k_0 \in \mathbb{N}$ such that $\overline{T(B_{k_0}(0))}$ has nonempty interior, i.e. there is $r_0 > 0$, $w_0 = Tv_0$ such that

$$B_{r_0}(w_0) \subset \overline{T(B_{k_0}(0))}.$$

By linearity,

$$\begin{aligned}
B_{r_0}(0) &= B_{r_0}(w_0) - Tv_0 \\
&\subset \overline{T(B_{r_0}(0))} - Tv_0 \\
&= \overline{T(B_{r_0}(-v_0))} \\
&\subset \overline{T(B_{k_0+l_0}(0))} \\
&= (k_0 + l_0)\overline{T(B_1(0))} \quad \text{(if } l_0 \geq \|v_0\|) \\
\implies B_r(0) &\subset \overline{T(B_r(0))} \text{ for } r = \frac{r_0}{k_0 + l_0}.
\end{aligned}$$

(ii) If T is bijective, that T is open means that T^{-1} is continuous, that is, $T^{-1} \in \mathcal{B}(W, V)$.

$\square$

Example. Let
$$c_c = \{ (x_n) \mid x_n = 0 \text{ except for finitely many } n \},$$
with norm $\|x\|_\infty = \max_n |x_n|$. Define $T : c_c \to c_c$ by $(Tx)_n = \frac{x_n}{n}$. Then T is bounded and bijective, but $(T^{-1}x)_n = nx_n$ is unbounded.

Remark. The basic problem in linear PDEs is the following one. Given $f \in W = L^2(\Omega)$ for some bounded nice domain $\Omega \subset \mathbb{R}^n$, and a linear partial differential operator $L : V \to W$, say $V = H_0^2(\Omega)$ and $L = \Delta$, is there a unique solution to $Lu = f$? The typical procedure is to show that for f 'nice', say $f \in C^\infty(\overline{\Omega})$, spanning a dense subspace of U, there exists a unique solution with
$$\|u\| \leq C \|f\|.$$

Such an *a priori* bound allows to solve $Lu = f$ for general $f \in W$ by approximation. This implies that L is surjective. The Open Mapping Theorem guarantees that this strategy works, provided that L is surjective.

3.4   Closed Graph Theorem

Theorem. [Closed Graph Theorem] Let V, W be Banach spaces, $T : V \to W$ linear. Then T is bounded if and only if the graph $\Gamma = \{ (v, Tv) \mid v \in V \} \subset V \times W$ is closed.

Proof. Let T be bounded and $(v_k, w_k) \subset \Gamma$ be a sequence such that

$$v_k \to v, \quad w_k = Tv_k \to w.$$

Since T is continuous, $w = Tv$ so $(v, w) \in \Gamma$, so Γ is closed.

Conversely, assume that Γ is closed. Since $V \times W$ is a Banach space with norm $\|(v,w)\| = \|v\| + \|w\|$, and since Γ is a closed subspace, Γ itself is also a Banach space with induced norm. The projections

$$
\begin{aligned}
\pi_v &: \Gamma \to V, & (v, Tv) &\mapsto v \\
\pi_w &: \Gamma \to W, & (v, Tv) &\mapsto Tv
\end{aligned}
$$

are continuous, and π_v is a bijection. By the Open Mapping Theorem, $\pi_v^{-1} \in \mathcal{B}(V, \Gamma)$. Thus $T = \pi_w \circ \pi_v^{-1} \in \mathcal{B}(V, W)$. $\qquad\square$

Remark. As a consequence of the Closed Graph Theorem, to prove that $T : V \to W$ is bounded, if V, W are Banach spaces, it suffices to check

$$
\left.\begin{aligned}
v_k &\to v \\
Tv_k &\to w
\end{aligned}\right\} \implies Tv = w
$$

instead of

$$
v_k \to v \implies \begin{cases} Tv_k \to w \\ Tv = w. \end{cases}
$$

4   The space of continuous functions on a compact space

4.1   Normal topological space

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

Proposition. Let X be Hausdorff and $K_1, K_2 \subset X$ compact sets with $K_1 \cap K_2 = \varnothing$. Then there exist open $U_1 \supset K_1$ and $U_2 \supset K_2$ such that $U_1 \cap U_2 = \varnothing$.

Proof. For any $x \in K_1$ and $y \in K_2$, let U_{xy} and V_{xy} be open neighbourhoods such that $x \in U_{xy}$, $y \in V_{xy}$ and $U_{xy} \cap V_{xy} = \varnothing$ (using that X is Hausdorff).

Then $\bigcup_{x \in K_1} U_{xy} \supset K_1$. Since K_1 is compact, there are finitely many points $x_1, \ldots, x_n \in K_1$ such that $\bigcup_{i=1}^{n} U_{x_i y} \supset K_1$. Set

$$U_y = \bigcup_{i=1}^{n} U_{x_i y} \text{ and } V_y = \bigcap_{i=1}^{n} V_{x_i y}.$$

Then $U_y \cap V_y = \varnothing$ and $U_y \supset K_1$ and $y \in V_y$. Thus $\bigcup_{y \in K_2} V_y \supset K_2$.

Again by compactness, there are $y_1, \ldots, y_n \in K_2$ such that $\bigcup_{i=1}^{n} V_{y_i} \supset K_2$. Set

$$V = \bigcup_{i=1}^{n} V_{y_i} \text{ and } U = \bigcap_{i=1}^{n} U_{y_i}.$$

These sets are open, $V \cap U = \varnothing$ and $U \supset K_1$ and $V \supset K_2$. $\qquad\square$

Definition (Normal space)**.** A Hausdorff topological space is **normal** if for any closed sets $A_1, A_2 \subset X$ such that $A_1 \cap A_2 = \varnothing$, there exist open sets $U_1 \supset A_1$, $U_2 \supset A_2$ such that $U_1 \cap U_2 = \varnothing$.

Corollary. Any compact Hausdorff space is normal.

Proof. Closed subsets of a compact space are compact. $\qquad\square$

Fact. Let X be normal. Then for every closed $A \subset X$ and open $U \supset A$, there exist an open set V and a closed set B such that $A \subset V \subset B \subset U$.

Proof. Set $A' = X \setminus U$. Then A and A' are disjoint. So, since X is normal, there are open sets $V, V' \subset X$ such that
$$V \supset A, \ V' \supset A', \ V \cap V' = \varnothing.$$
Take $B = X \setminus V'$. Then $A \subset V \subset B \subset U$. $\qquad\square$

Proposition (Urysohn's Lemma)**.** Let X be normal. For every closed $A \subset X$ and open $U \supset A$, there exists a continuous $f : X \to [0, 1]$ 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 $A_1 \subset U_0$ there exists an open set $U_{\frac{1}{2}}$ and a closed set $A_{\frac{1}{2}}$ such that

$$A_1 \subset U_{\frac{1}{2}} \subset A_{\frac{1}{2}} \subset U_0.$$

Applying this procedure again with $A_1 \subset U_{\frac{1}{2}}$ and with $A_{\frac{1}{2}} \subset U_0$,

$$A_1 \subset U_{\frac{3}{4}} \subset A_{\frac{3}{4}} \subset U_{\frac{1}{2}} \subset A_{\frac{1}{2}} \subset U_{\frac{1}{4}} \subset A_{\frac{1}{4}} \subset U_0.$$

Iterating this procedure, there are open sets U_q and A_q for

$$q \in \{\, m2^{-n} \mid n, m \in \mathbb{N}_0, 0 < m < 2^n \,\} = D$$

such that

$$U_q \supset U'_q \text{ and } A_q \supset A'_q \text{ for } q' > q,$$
$$U_q \subset A_q \text{ for all } q \in D,$$
$$A'_q \subset U_q \text{ for } q' > q.$$

Define $f(x) = \sup \{\, q \in D \mid x \in U_q \,\} = \inf \{\, q \in D \mid x \notin A_q \,\}$, where $\inf \varnothing = 1$ and $\sup \varnothing = 0$.

Clearly $0 \leq f \leq 1$. If $x \notin U = U_0$ then $f(x) = 0$. If $x \in A = A_1$, then also $x \in U_q$ for $q \in D$ and $f(x) = 1$. To show continuity, note that, for $t \in R$,

$$\{\, x \mid f(x) > t \,\} = \bigcup_{q > t} U_q \text{ is open}$$
$$\{\, x \mid f(x) < t \,\} = \bigcup_{q > t} X \setminus A_q \text{ is open.}$$

Thus, the preimage of any open interval is open. Thus f is continuous. $\qquad\square$

Corollary. Let X be normal and $A_0, A_1 \subset X$ closed and disjoint. Then there is a continuous $f : X \to [0, 1]$ such that

$$f|_{A_0} = 0 \,, \ f|_{A_1} = 1.$$

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

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

Proof. We first assume that g takes values in $[0, 1]$. Let $g_0 = g$. Let $A_0 = g^{-1}([0, \frac{1}{3}])$ and $B_0 = g^{-1}([\frac{2}{3}, 1])$. Then A_0, B_0 are closed and disjoint. By the corollary to Urysohn's Lemma, there is $h_0 : X \to [0, \frac{1}{3}]$ continuous 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 $x \in A$. By induction, assume that $g_i : A \to [0, (\frac{2}{3})^i]$ is given and set

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

23

and a continuous function $h_i|A_i = 0$ and $h_i|B_i = \frac{1}{3}(\frac{2}{3})^i$. Set $g_{i+1} = g_i - h_i|_A$, so

$$g = g_0 = \sum_{i=0}^{\infty} h_i|_A \text{ uniformly on } A.$$

Set $\widetilde{f} = \sum_{i=0}^{\infty} h_i$ uniformly on X, so $\widetilde{f}$ is continuous and $\|f\|_\infty \le 1$.

If g takes values in $\mathbb{R}$ but is not bounded, we can apply the above to $\frac{1}{2} + \frac{1}{\pi} \arctan \circ g$ which does take values in $[0, 1]$ to obtain an extension $\widetilde{f}$.

If g takes values in $\mathbb{C}$, we apply this separately to the real and imaginary parts to obtain an extension $\widetilde{f} : X \to \mathbb{C}$. Finally, define

$$\begin{cases} \widetilde{f}(x) & \text{if } \left|\widetilde{f}(x)\right| \le \|g\|_\infty \\ e^{i \arg \widetilde{f(x)}} \|g\|_\infty & \text{otherwise.} \end{cases}$$

This f is still an extension of g and $\|f\| \le \|g\|$. $\qquad\square$

4.2   Arzelá-Ascoli Theorem

When is a subset of $C(K)$ compact?

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

 (i) X has the Heine-Borel property: any open cover of X has a finite subcover.

 (ii) X is sequentially compact: any sequence in X has a convergent subsequence.

 (iii) X is complete and totally bounded: for $\epsilon > 0$ there is a finite ϵ-net for X. This is a finite set $M \subset X$ such that for any $x \in X$ there exists $m \in M$ with $d(x, m) < \epsilon$.

Proof. The first two are equivalent from Analysis II and Metric and Topological spaces. It is easy to check (i) and (ii) give (iii), so we only check (iii) $\Rightarrow$ (ii).

Let $(x_n) \subset X$ be a sequence. We need 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). Let n_1 be the first n such that $x_n \in B_1(m_1)$. Given $m_1 \in M_1, \dots, m_k \in M_k$, $n_1, \dots, n_k$ such that $B_{\frac{1}{k}}(M_i)$ contains infinitely many points from $(x_n) \cap B_{\frac{1}{i}}(m_i)$ for all $i \le j$ and $x_{n_l} \in \bigcap_{j=1}^{l} B_{\frac{1}{j}}(m_j) \ \forall l < k$.

Let m_{k+1} be such that $B_{\frac{1}{k+1}}(m_{k+1})$ contains infinitely many points from $(x_n) \cap \bigcap_{j=1}^{n} B_{\frac{1}{j}}(m_j)$ and n_{k+1} be the first $n > n_k$ such that $x_{n_{k+1}} \in \bigcap_{i=1}^{k+1} B_{v_j}(m_j)$.

It follows that

$$d(x_{n_k}, x_{n_l}) \le d(x_{n_k}, m_k) + d(x_{n_l}, m_k) \le \frac{1}{k} + \frac{1}{k} \text{ for } l \ge k$$

so (x_{n_k}) is a Cauchy sequence. X is complete, so $(x_{n_k})_k$ converges. $\qquad\square$

Corollary. Let X be a complete metric space. Then $Y \subset X$ is relatively compact (i.e. has compact closure in X) if and only if Y is totally bounded.

For K compact, we always assume that $C(K)$ has $\|\cdot\|_\infty$-norm, making it a Banach space.

Theorem (Arzelá-Ascoli). Let K be compact Hausdorff and $\mathcal{F} \subset C(K)$. Then the following are equivalent:

(i) $\mathcal{F}$ is relatively compact

(ii) $\mathcal{F}$ is bounded and equicontinuous, i.e. $\sup_{f \in \mathcal{F}} \|f\|_\infty < \infty$ and $\forall \epsilon > 0 \; \forall x \in K, \; \exists U$ a neighbourhood of x such that

$$\forall f \in \mathcal{F}, \; |f(x) - f(y)| < \epsilon \quad \text{for all } y \in U.$$

Proof. (i) $\implies$ (ii): Let $\mathcal{F}$ be relatively compact, i.e. totally bounded. Thus for any $\epsilon > 0$, there is $f_1, \ldots, f_n \in \mathcal{F}$ such that $\forall f \in \mathcal{F}$

$$\min_{i=1,\ldots,n} \|f - f_i\|_\infty < \epsilon.$$

Thus $\|f\|_\infty \leq \epsilon + \max_{i=1,\ldots,n} \|f_i\|_\infty = M$ for all $f \in \mathcal{F}$, so $\mathcal{F}$ is bounded. Let $\epsilon > 0$ and $x \in K$. Since the f_i are continuous, there exist neighbourhoods U_i of x such that $|f_i(x) - f_i(y)| < \epsilon$ for $y \in U_i$. Let $U = \bigcap_{i=1}^k$. Then U is a neighbourhood of x and for $y \in U$,

$$|f(x) - f(y)| \leq \min_i (|f(x) - f(y)| + |f_i(x) - f_i(y)| + |f(y) - f_i(y)|) < 3\epsilon \quad \text{for all } f \in \mathcal{F}.$$

Thus f is equicontinuous.

(ii) $\implies$ (i): Let $\mathcal{F}$ be bounded and equicontinuous. For every $\epsilon > 0$, we construct a finite ϵ-net for $\mathcal{F}$. Let $\epsilon > 0$. For $x \in K$, let U_x be an open neighbourhood of x such that $|f(x) - f(y)| < \epsilon$ if $f \in \mathcal{F}, \; y \in U_x$ (by equicontinuity). Since K is compact, there are $x_1, \ldots x_n \in K$ such that

$$K = \bigcup_{i=1}^n U_{x_i}.$$

Since $\mathcal{F}$ is uniformly bounded, the vectors $(f(x_1), \ldots, f(x_n)) \in \mathbb{K}^n, \; f \in \mathcal{F}$ are bounded in any norm on $\mathbb{K}^n$. We take the $\|\cdot\|_\infty$ on $\mathbb{K}^n$. Thus

$$F = \{\, (f(x_1), f(x_2), \ldots, f(x_n)) \mid f \in \mathcal{F} \,\}$$

is relatively compact in $\mathbb{K}^n$. Thus there are $f_1, \ldots f_m \in \mathcal{F}$ such that

$$\{\, (f_i(x_1), \ldots, f_i(x_n)) \mid 1 \leq i \leq m \,\} \subset \mathbb{K}^n$$

is a finite ϵ-net for F.

Now claim $f_1, \ldots, f_m$ are a finite 3ϵ-net for $\mathcal{F}$.

$$|f(x) - f_i(x)| \leq \underbrace{|f(x) - f(x_j)|}_{<\epsilon \text{ if } x \in U_{x_j}} + \underbrace{|f(x_j) - f_i(x_j)|}_{<\epsilon \text{ for some } i \in \{1,\ldots,m\}} + \underbrace{|f_i(x_i) - f_i(x)|}_{<\epsilon \text{ if } U_{x_i} \ni x}$$

$$< 3\epsilon$$

for some choice of j. $\qquad\square$

Example. Let $K = [0, 1]$ and $\mathcal{F} = \{\, f \in C^1[0,1] \mid \|f\|_\infty + \|f'\|_\infty < 1 \,\}$. Then $\mathcal{F}$ is relatively compact in $C^0[0, 1]$. Indeed $\mathcal{F}$ is bounded (in C^0) and

$$|f(x) - f(y)| \le \|f'\|_\infty \, |x - y| \le |x - y| \le \epsilon \quad \text{whenever } |x - y| \le \epsilon$$

for any $f \in \mathcal{F}$.

Since $\mathcal{F}$ is the unit ball of $C^1[0, 1]$, this means that the identity map $C^1[0, 1] \to C^0[0, 1]$, is a compact embedding: it is continuous and maps bounded sets to relatively compact sets.

4.3 Application: Peano Existence Theorem

Theorem (Picard-Lindelöf). Let $f : \mathbb{R} \to \mathbb{R}$ be locally Lipschitz continous. Then for any $x_0 \in R$ there exists a maximal interval $(T_1, T_2) \subset \mathbb{R}$ with $T_1 = -\infty$ and $T_2 = +\infty$ 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 not the restriction of such a solution to a larger interval). Moreover, if $T_2 \ne +\infty$, for any bounded $K \subset \mathbb{R}$ there exists $t < T_2$ such that $x([t, T_2)) \cap K = \varnothing$ and similarly if $T_1 \ne -\infty$.

Theorem (Peano). Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. Then for any $x_0 \in \mathbb{R}$, $\exists T > 0$ and a solution $x : (-T, T) \to \mathbb{R}$ for $(*)$.

Remark. The solution is not necessarily unique, e.g. $f(x) = \sqrt{|x|}$.

Lemma. Assume that $b > 0$, $m > 0$ such that

$$|f(x)| \le M \quad \text{for } |x - x_0| \le b. \tag{@}$$

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

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

$$|x'(t)| = |f(x(t))| \le M, \ |x(t) - x_0| = \left| \int_0^t f(x(s)) \, ds \right| \le Mt \lneqq b \text{ for } |t| \le T'. \tag{$\dagger$}$$

This allows us to extend the solution beyond T' by 'continuous induction': Let $I = \{\, T' \in [0, T] \mid |x(t) - x_0| \le b \text{ for } |t| \le T' \,\}$. Note $I \ne \varnothing$ and that I is closed (both by continuity of x).

We now claim that $\sup I = T$. Otherwise, $|x(t) - x_0| < b$ by $(\dagger)$ for $|t| \le \sup I$. But by continuity, a neighbourhood of $\sup I$ must also then be in T, a contradiction. Then, $I = [0, T]$. $\qquad \square$

Proof of Peano's Theorem. Let $B = \{\, f + \tilde{g} \mid \tilde{g} \in C^0, \|\tilde{g}\| \le 1 \,\}$ and choose $M, b > 0$ such that (@) holds for all $g \in B$. For any $f \in B \cap C^1$ there is a local solution to $(*)$ by the Picard-Lindelöf Theorem. The lemma implies that these solutions are defined on $[-T, T]$ with T as in the lemma. Define

$$S : B \cap C^1 \longrightarrow C^1[-T, T]$$
$$f \longmapsto x \quad \text{where } x \text{ is the solution to } (*).$$

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 $C^1[-T, T] \to C^0[-T, T]$ is compact so $S(B \cap C^1)$ is relatively compact in $C^0[-T, T]$. Let $g_i \in B \cap C^1$ be such that $g_i \to f$ is C^0, i.e. $\|f - g_i\|_\infty \to 0$. (The existence of such g_i follows from the Weierstrass approximation theorem in the next section.)

By relative compactness of $S(B \cap C')$ in C^0 there is a subsequence such that $x_i = Sg_i$ converges to some $x \in C^0[-T, T]$.

Claim: $x \in C^1[-T, T]$ and $(*)$ holds.

Since $g_i \xrightarrow{C^0} f$ and $x_i \xrightarrow{C^0} x$ (along the subsequence), also $g_i \circ x_i \xrightarrow{C^0} f \circ x$

$$\implies x_i' = g_i \circ x_i \to f \circ x \text{ uniformly in } |t| \le T \implies x \in C^1 \text{ and } x' = f \circ x.$$

<div align="right">□</div>

4.4   Stone-Weierstrass Theorem

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

Lemma. There is a sequence of polynomials $P_n : [-1, 1] \to [0, 1]$ such that $P_n(s) \to |s|$ uniformly in $s \in [-1, 1]$ as $n \to \infty$.

Proof. We use the Babylonian method to construct square roots. Note that if $q : [0, 1] \to [0, 1]$ is a function that satisfies

$$q(t) = \frac{1}{2}(t + q(t)^2) \qquad (*)$$
$$\implies (1 - q(t))^2 = 1 - t$$
$$\implies 1 - q(t) = \sqrt{1 - t}$$
$$\implies |t| = 1 - q(1 - t^2).$$

To approximate q, define polynomials $Q_n : [0, 1] \to [0, 1]$ by

$$Q_0(t) = 0, \ Q_{n+1}(t) = \frac{1}{2}(t + Q_n(t))^2.$$

If $Q_n(t)$ converges to some $q(t)$ then $q(t) \in [0, 1]$ and $(*)$ must hold. To show that Q_n does converge, note that for any $t \in [0, 1]$,

$$Q_{n+1}(t) \ge Q_n(t).$$

Indeed,

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

Also, $Q_n(t)$ is increasing in $t \in [0, 1]$, and the last equality implies that also $Q_{n+1}(t) - Q_n(t)$ is increasing in $t \in [0, 1]$.

$$\implies Q_{n+1}(t) - Q_n(t) \le Q_{n+1}(1) - Q_n(1)$$
$$\implies Q_m(t) - Q_n(t) \le Q_m(1) - Q_n(1) \quad \text{for } m \ge n, \quad t \in [0, 1].$$

<div align="center">27</div>

Let $m \to \infty$:

$$0 \leq 1 - \sqrt{1 - t^2} - Q_n(t) \leq 1 - Q_n(1)$$
$$\implies \left\| 1 - Q_n - \sqrt{1 - t} \right\|_\infty \leq 1 - Q_n(1) \to 0.$$

Set $P_n(t) = 1 - Q(1 - t^2)$. Then $\| P_n - |\cdot| \|_\infty \to 0$. $\qquad\qquad\square$

Exercise. Using the lemma, prove the Weierstrass Approximation Theorem for piecewise linear functions. Then deduce the general case.

Definition.   (i) A real (complex) **algebra** is a real (complex) vector space A with a bilinear map $A \times A \to A$, $(a, b) \mapsto ab$ called product such that

$$(ab)c = a(bc) \text{ for all } a, b, c \in A. \qquad\qquad ((\text{associativity}))$$

(ii) If A is a normed vector space and

$$\|ab\| \leq \|a\| \, \|b\| \text{ for all } a, b \in A$$

then A is called a **normed algebra**. If A is a Banach space then A is called a **Banach algebra**.

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

(iv) If there exist an element $1 \in A \setminus \{0\}$ such that $1a = a1 = a$ for all $a \in A$ then A is **unital** (and 1 is called a **unit**).

Example.   • $C(K, \mathbb{R})$ is commutative unital Banach algebra with product

$$(fg)(x) = f(x)g(x) \quad \forall x \in K$$

and unit $1(x) = 1$ for all $x \in K$.

• $\mathcal{B}(V, V)$ where V is a normal vector space is a normed unital algebra with product given by composition and unit by the identity operation. Then if V is a Banach space, $\mathcal{B}(V, V)$ is a Banach algebra.

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

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

• vanishes nowhere: for all $x \in K$, there is $f \in A$ such that $f(x) \neq 0$.

Then A is dense in $C(K, \mathbb{R})$, i.e. $\overline{A} = C(K, \mathbb{R})$.

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

Definition. A **lattice** is a poset L with the property that for any $u, v \in L$, there is a least upper bound (join) $u \vee v$ for $\{u, v\}$ and a greatest lower bound (meet) $u \wedge v$.

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

$$(f \vee g)(x) = \max\{f(x), g(x)\}$$
$$(f \wedge g)(x) = \min\{f(x), g(x)\}.$$

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

Proof.

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

Thus it suffices to show that if $f \in \overline{A}$ then $|f| \in \overline{A}$.

Let $f \in \overline{A} \setminus \{0\}$ and $\epsilon > 0$. Replacing f by $\frac{f}{\|f\|_\infty}$ we may assume that f takes values in $[-1, 1]$. By the lemma at the beginning of this section (link?), there is a polynomial $P : [-1, 1] \to [0, 1]$ such that

$$\sup_{|t| \leq 1} |P(t) - |t|| < \epsilon$$
$$\implies \|P \circ f - |f|\| < \epsilon.$$

Since $P \circ f \in \overline{A}$ and since $\overline{A}$ is closed therefore $|f| \in \overline{A}$. $\qquad\square$

Lemma. Let $L \subset C(K, \mathbb{R})$ be a sublattice, i.e. $f, g \in L \implies f \wedge g, f \vee g \in L$. Let $g \in C(K, \mathbb{R})$ such that

$$\forall \epsilon > 0 : \forall x, y \in K : \exists f_{xy} \in L, \text{ such that } |f_{xy}(x) - g(x)| < \epsilon \text{ and } |f_{xy}(y) - g(y)| < \epsilon. \quad (*)$$

Then $g \in \overline{L}$. In particular, if $(*)$ 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 $\epsilon > 0$. We will construct $f \in L$ such that $\|f - g\| < \epsilon$.

For $x, y \in K$, let f_{xy} be as in $(*)$. By continuity, there is some open U_{xy} containing y such that $|f_{xy}(z) - g(z)| < \epsilon \; \forall z \in U_{xy}$.

Then for fixed x, these form an open cover of K so $\exists y_1, \ldots, y_n$ such that $\bigcup_{i=1}^{n} U_{x,y_i} \supseteq K$.

Now let $f_x = f_{xy_1} \wedge f_{xy_2} \wedge \cdots \wedge f_{xy_n}$. Since $f_{xy_i}(z) < g(z) + \epsilon \forall z$, we have $f_x(z) < g(z) + \epsilon$.

On the other hand, $|f_x(x) - g(x)| < \epsilon$. By continuity of f_x, $\exists V_x \ni x$ open such that $|f_x(w) - g(w)| < \epsilon \quad \forall w \in V_x$. Then (V_x) is an open cover of K so

$$x_1, \ldots, x_m \text{ such that } \bigcup_{j=1}^{m} V_{x_j} \supseteq K.$$

Then if $f = f_{x_1} \vee f_{x_2} \vee \cdots \vee f_{x_m}$, we have $f(z) > g(z) - \epsilon \quad \forall z$, and our first bound $f(z) < g(z) + \epsilon$ still holds, so $\|f - g\| < \epsilon$. $\qquad\square$

Proof of Stone-Weierstrass Theorem. By continuous of addition and multiplication, the closure $\overline{A}$ is a sublattice of $C(K, \mathbb{R})$. We apply the last lemma. 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, $(*)$ holds.

By assumption, A vanishes nowhere and separates points:

$$\forall x \in K \quad \exists f_x \in A \text{ s.t. } f_x(x) \neq 0$$
$$\forall x, y \in K, \ x \neq y, \ f_{xy} \in A \text{ s.t. } f_{xy}(x) \neq f_{xy}(y).$$

If $x = y$, we can take $f = \frac{g(x)}{f_x(x)} f_x \in A$. Thus assume $x \neq y$.

Claim: There are $\alpha, \beta, \gamma \in \mathbb{R}$ such that

$$h = \alpha f_x + \beta f_y + \gamma f_{xy} \tag{$\dagger$}$$

satisfies

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

Indeed, if $f_{xy}(x) \neq 0$ and $f_{xy}(y) \neq 0$, we can take $h = f_{xy}$. Otherwise, without loss of generality $f_{xy}(y) = 0$, and by rescaling assume that

$$\begin{aligned} f_{xy}(x) &= 1 & f_y(x) &= C \\ f_{xy}(y) &= 0 & f_y(y) &= 1. \end{aligned}$$

Take $\alpha = 0$, $\beta = 1$, $\gamma = 2 - C$ in $(\dagger)$. This gives the claim since

$$h(x) = C + \gamma = 2, h(y) = 1.$$

The claim implies that $(h(x), h(y))$ and $(h(x)^2, h(y)^2)$ are linearly independent in $\mathbb{R}^2$. Thus there exist $t, s \in \mathbb{R}$ such that

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

with $f = th + sh^2 \in A$.

Thus $(*)$ holds for all $g \in C(K, \mathbb{R})$. So we have shown $\overline{A} = C(K, \mathbb{R})$. $\qquad\square$

Corollary. Let $U \subset \mathbb{R}^n$ be a bounded open domain. Then $C(\overline{U})$ is separable, i.e. there is a countable dense subset.

Proof. By the Weierstrass approximation theorem, the set of polynomials with real coefficients is dense in $C(U)$. But every polynomial with real coefficients can be approximated by one with rational coefficients. These are countable. $\qquad\square$

Remark. For $p \in [1, \infty)$, ℓ^p is separable, but ℓ^∞ is not separable.

Proof that l^∞ is not separable. Let $X = \{ (x_n) \in \ell^\infty \mid x_n \in \{0, 1\} \text{ for all } n \}$. Then X is not countable.

Note that $|x - y|_\infty = 1$, $\forall x \neq y$, $x, y \in X$. Assume ℓ^∞ was separable, i.e. there is a dense countable subset $(y^k)_{k \in \mathbb{N}}$. Then for any $x \in X$, there is $k \in \mathbb{N}$ such that $\|x - y^k\| \leq \frac{1}{2}$ and $\|y - y^k\| \geq \frac{1}{2}$ for all $y \in X$, $y \neq x$. This gives a bijection $\mathbb{N} \to X$, a contradiction. $\qquad\square$

4.5 Complex Stone-Weierstrass

Theorem. Let $A \subset C(K, \mathbb{C})$ be a subalgebra such that

- A separates points

- A vanishes nowhere

- A is closed under complex conjugation, i.e., $\overline{f} \in A$ if $f \in A$.

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

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

$$(ab)^* = b^* a^*, \quad 1^* = 1, \quad a^{**} = a, \quad (\lambda a)^* = \overline{\lambda} a^*, \quad \|a^*\| = \|a\|$$

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

Example.

- $C(K, \mathbb{C})$ is a commutative C^*-algebra with $f^* = \overline{f}$.

- $\mathcal{B}(H, H)$ where H is a Hilbert space is a C^* algebra (later).

Proof of theorem. Note

$$f \in A \implies \operatorname{Re} f = \frac{1}{2}(f + \overline{f}) \in A, \quad \operatorname{Im} f = \frac{1}{2i}(f - \overline{f}) \in A$$

Let $A_\mathbb{R}$ be the subalgebra of $C(K, \mathbb{R})$ generated by $\operatorname{Re} f$, $\operatorname{Im} f$, $f \in A$. Then $A_\mathbb{R}$ vanishes nowhere and separates points, since A has these properties. So the real version of the Stone-Weierstrass Theorem implies $\overline{A_\mathbb{R}} = C(K, \mathbb{R})$.

Let $f = u + iv \in C(K, \mathbb{C})$ with $u, v \in C(K, \mathbb{R})$. There are $u_j, v_j \in A_\mathbb{R}$ such that $u_j + iv_j \to f$ (with norm $\|\cdot\|_\infty$). Since $u_j + iv_j \in A$, hence $\overline{A} = C(K, \mathbb{C})$. $\square$

Example (Hardy Space)**.** The assumption that A is closed under complex conjugation is necessary. Let $K = \overline{\mathbb{D}} = \{ z \in C \mid |z| \le 1 \}$. Then

$$A = \{ f \in C(\overline{\mathbb{D}}, \mathbb{C}) \mid f \text{ analytic in } \mathbb{D} \}$$

is an algebra. It separates points and vanishes nowhere. But $\overline{A} \ne C(\overline{\mathbb{D}}, \mathbb{C})$ since $\overline{z} \notin A$.

4.6 Application: Convergence of Fourier Series

Let $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ be the circle, ie $[\pi, \pi]$ with its endpoints identified. Let A be the vector space spanned by $\{e^{inx}\}_{n \in \mathbb{Z}}$. Its elements are called trigonometric polynomials. It is an algebra, it separates points, vanishes nowhere (contains constants), and is closed under complex conjugation since $\overline{e^{inx}} = e^{-inx}$. Thus $\overline{A} = C(\mathbb{T}, \mathbb{C})$ by Stone-Weierstrass.

Example (Example sheet)**.** There exists $f \in C(\mathbb{T})$ such that $S_n f(0) \nrightarrow f(0)$ as $n \to \infty$, where

$$S_n f = \sum_{k=-n}^{n} \hat{f}_k e^{ikx}, \qquad \hat{f}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} \, dx$$

Thus 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 $S_n f$. However, we can show that $S_n f \to f$ in L^2.

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

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

Proof. By the complex Stone-Weierstrass Theorem, for any $\epsilon > 0$, there is a trig poly $P \in A$ such that $\|f - P\|_{\infty} < \epsilon$. Note that $S_n P = P$ if n is larger than the degree of P.

$$\begin{aligned}
\implies & |f - S_n f| \leq |f - P| + |S_n f - S_n p| \qquad \text{for } n \text{ large enough} \\
\implies & |f - S_n f|^2 \leq 2 |f - P|^2 + 2 |S_n f - S_n P|^2 \quad \text{since } (a+b)^2 \leq 2a^2 + 2b^2 \\
\implies & \int_{-\pi}^{\pi} |f - S_n f|^2 \, dx \leq 4 \int_{-\pi}^{\pi} \pi |f - P|^2 \, dx \leq 8\pi\epsilon^2
\end{aligned}$$

The first inequality here is called Bessel's inequality:

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

for any $g \in C(\mathbb{T})$. $\qquad \square$

4.7 Aside: Radon measures and the weak-$*$ topology

Definition (Positive functional)**.** An element $\mu \in C(K)^*$ is called positive if $\mu(f) \geq 0$ for every $f \in C(K)$ with $f(x) \geq 0$ for all $x \in K$.

Example. Let $K = [0,1]$. Then,

- $\mu(f) = \int_0^1 f(x) \, dx$ defines a positive functional.

- for any $x \in K$, $\delta_x(f) = f(x)$ defines a positive functional.

Theorem (Riesz-Markov)**.** Let K be compact Hausdorff. Then for any positive $\mu \in C(K)^*$, there is a unique regular Borel measure $d\mu$ such that

$$\mu(f) = \int_K f(x) \, d\mu(x) \qquad \text{for all } f \in C(K).$$

Proof. Not proved in this course. $\qquad \square$

Example. Let $x, y \in K$, $x \neq y$. Then $\|\delta_x - \delta_y\|_{C(K)^*} = 2$ using Urysohn's Theorem. In particular, the map $x \in K \mapsto \delta_x \in C(K)^*$ is not continuous in the norm topology.

Definition. Let V be a normed vector space.

- The topology on V induced by the sets $\Omega_{l,U} = l^{-1}(U)$, $l \in V^*$, $U \subset \mathbb{R}$ open is called the **weak topology** on V.

- The topology on V^* by $\Omega_{v,U} = f_v^{-1}(U)$, $v \in V$, $U \subset \mathbb{R}$ open, $f_v \in V^{**}$, $f_v(l) = l(v)$ is called the **weak-$*$ topology** on V^*.

The weak topology makes V be a locally convex topological vector space (and it does not come from a norm). On V^*, we now have three topologies.

The weak-$*$ topology is 'weaker' than the weak topology on V^*, as we are restricting our attention to the subset of V^{**} induced by the includsion map $V \hookrightarrow V^{**}$.

Equivalently, we can define the weak topology on V^* as the weakest topology that makes every bounded linear $l : V \to \mathbb{K}$ continuous. Similarly, the weak-$*$ topology on V^* is the weakest topology that makes the maps $f_v \in V^{**}$ continuous for all $v \in V$, where $f_v \in V^{**}$ is defined by $f_v(l) = l(v)$.

Example. Let K be compact Hausdorff. Then the map $x \in K \mapsto \delta_x \in C(K)^*$ is continuous with respect to the weak topology on $C(K)^*$:

$$x \to y \implies \delta_x(f) - \delta_y(f) = f(x) - f(y) \to 0 \quad \text{if } f \in C(K)$$

Theorem. A normed vector space V is separable iff the weak-$*$ topology on the closed unit ball in V^* is metrisable.

Proof. See example sheet 3, question 9. $\qquad\qquad\square$

Theorem (Banach-Alaogla). Let V be a (separable) normed space. Then the closed unit ball in V^* is compact with respect to the weak-$*$ topology. (Similar to Arzelà-Ascoli).

Theorem. Let V be a (separable) normed space. Let K be the closed unit ball in V^* with the weak-$*$ topology. Then

$$\Phi : V \to C(K), \qquad v \mapsto f_v \quad \text{where} \quad f_v(l) = l(v) \text{ for } l \in V^*$$

is an isometric isomorphism onto its image.

Proof. That $f_v \in C(K)$ holds exactly by definition of the weak-$*$ topology. Linearity is clear. That Φ is isometric follows from Hahn-Banach:

$$\|f_v\| = \sup_{l \in \mathbb{K}} |f_v(l)| = \sup_{\|l\| \leq 1} |l(v)| = \|v\| . \qquad\qquad\square$$

Example. Let K be a compact metric space and

$$P(K) = \{\, \mu \in C(K)^* \mid \mu \text{ is positive, } \mu(\mathbb{1}) = 1 \,\}$$

(where $\mathbb{1}$ denotes the constant function $\mathbb{1}(x) = 1$) be the set of Borel probability measures. Let $\phi : K \to K$ be a homeomorphism. Then there exists a ϕ-invariant probability measure, i.e., $\mu \in P(K)$ such that

$$\mu(f) = \mu(f \circ \phi) \text{ for any } f \in C(K).$$

Proof. Fix $x_0 \in K$ and define $\mu_n \in P(K)$ by

$$\mu_n(f) = \frac{1}{n} \sum_{k=0}^{n-1} f(\phi \circ \cdots \circ \phi(x_0)) \text{ for } f \in C(K).$$

Since K is metric, $C(K)$ is separable (example sheet 3). Also by example sheet 3, $\|\mu_n\| \leq 1$. By the Banach-Alaoglu Theorem, there exists a weak-$*$ convergent subsequence $(\mu_n)_{n \in \Lambda}$. Its limit satisfies

$$\mu(1) = \lim_{\substack{n \to \infty \\ n \in \Lambda}} \mu_n(1) = 1, \quad \mu(f) \geq 0 \text{ if } f \geq 0.$$

Thus $\mu \in P(K)$. Moreover, for any $f \in C(K)$,

$$\mu(f \circ \phi) = \lim_{\substack{n \to \infty \\ n \in \Lambda}} \frac{1}{n} \sum_{k=1}^{n} f(\underbrace{\phi \circ \cdots \circ \phi}_{k \text{ times}}(x_0))$$

$$= \lim_{\substack{n \to \infty \\ n \in \Lambda}} \frac{1}{n} \sum_{k=0}^{n-1} f(\underbrace{\phi \circ \cdots \circ \phi}_{k \text{ times}}(x_0))$$

$$= \mu(f).$$

$\square$

5 Euclidean vector spaces and Hilbert spaces

5.1 Definitions and Examples

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

   (i) $(v, w) = \overline{(w, v)}$ for all $v, w \in V$ (skew-symmetry)

   (ii) $(\lambda_1 v_1 + \lambda_2 v_2, w) = \lambda_1 (v_1, w) + \lambda_2 (v_2, w)$ for all $v_i, w \in V$, $\lambda_i \in \mathbb{K}$ (linear in first argument)

   (iii) $(v, v) \geq 0$ with $(v, v) = 0$ iff $v = 0$ (positive definite).

A vector space together with an inner product is called an **inner product space**.

Proposition (Cauchy-Schwarz inequality)**.** Let V be an inner product space. Then

$$|(v, w)| \leq (v, v)^{\frac{1}{2}} (w, w)^{\frac{1}{2}} \quad \forall v, w \in V$$

with equality iff $v = \lambda w$ for some $\lambda \in \mathbb{K}$ or $w = 0$.

Corollary. Let V be an inner product space. Then $\|v\| = (v, v)^{\frac{1}{2}}$ is a norm.

Definition. A normed vector space whose norm is induced by an inner product is called a **Euclidean space**. If it is also Banach, then it is a **Hilbert space**.

Fact. Let V be Euclidean. Then there is a unique inner product such that $\|v\| = (v, v)^{\frac{1}{2}}$.

Proof. $\|v\|^2 = (v, v)$ implies $\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2 \operatorname{Re}(v, w)$. If V is a real vector space, then $(v, w) = \frac{1}{2}(\|v + w\|^2 - \|v\|^2 - \|w\|^2)$. If V is a complex vector space, then also $\|v + iw\|^2 = \|v\|^2 + \|w\|^2 + 2 \operatorname{Re}(v, iw)$. But $\operatorname{Re}(v, iw) = -\operatorname{Re} i(v, w) = +\operatorname{Im}(v, w)$.

$$(v, w) = \operatorname{Re}(v, w) + i \operatorname{Im}(v, w)$$
$$= \frac{1}{2}(\|v + w\|^2 + i \|v + iw\|^2 - (i + 1)(\|v\|^2 + \|w\|^2)).$$

$\square$

Fact. Let V be Euclidean. Then the unique inner product $(\cdot, \cdot)$ inducing the norm is continuous.

Example.

- $\ell^2 = \{ (x_n) \in K \mid \sum_{n=1}^{\infty} |x_n|^2 < \infty \}$ is a Hilbert space with inner product $(x, y) = \sum_{n=1}^{\infty} x_n \overline{y_n}$.

- $C[a, b]$ iwth norm

$$\|f\|_2 = (\int_a^b |f(x)|^2 \, dx)^{\frac{1}{2}}$$

  (and inner product $\int_a^b f(x) g(x) \, dx$) is a Euclidean vector space, but it is not complete. Its completion can be identified with the Lebesgue space $L^2[a, b]$.

Remark. Any metric space X can be completed. For two Cauchy sequences (x_n), $(y_n) \subset X$, set $(x_n) \sim (y_n)$ if $d(x_n, y_n) \to 0$. This is an equivalence relation. Denote the equivalence class of a Cauchy sequence $x = (x_n)$ by $\overline{x}$. Define $\overline{X} = \{\, \overline{x} \mid x \text{ is a Cauchy sequence in } X \,\}$. Then $\overline{d}(\overline{x}, \overline{y}) = \lim_{n \to \infty}$ exists (exercise), it is well-defined, i.e. independent of the sequences representing $\overline{x}$, $\overline{y}$ and one can check that $\overline{d}$ is a metric on $\overline{X}$.

The metric space $\overline{X}$ is complete: if $(\overline{x}^k) \subset \overline{X}$ is a Cauchy sequence in $\overline{X}$ and $(x_n^k)_n \subset X$ is a representative for $\overline{x}^k$, choose n_k such that $d(x_i^k, x_j^k) \leq 2^{-k}$ for $i, j \geq n_k$. Define $(x_k) \subset X$ by $x_k = x_{n_k}^k$. Then $x = (x_k) \subset X$ is a Cauchy sequence and $\overline{x}^k \to \overline{x}$ in $\overline{X}$.

The space X is isometrically embedded in $\overline{X}$ by the map that sends $x \in X$ to the equivalence class of the constant sequence $(x, x, \dots)$. Thus we can regard X as a subspace of $\overline{X}$. One can check that X is dense in $\overline{X}$.

Exercise. Let X be a normed vector space (inner product space). Then the norm (inner product) extends uniquely from X to its completion, making it a Banach (Hilbert) space.

Example.

- $L^2[a, b]$ is the completion of $C[a, b]$ with inner product $(f, g) = \int_a^n f\overline{g}\, dx$.

- The Sobolev space $H^1[a, b]$ is the completion of C^∞ with inner product $(f, g) = \int_a^b f(x)\overline{g(x)}\, dx + \int_a^b f'(x)\overline{g'(x)}\, dx$.

5.2 Orthogonal complements and projections

Definition. Let V be a Euclidean vector space.

- $v, w \in V$ are **orthogonal** if $(v, w) = 0$.

- The **orthogonal complement** of a set $S \subset V$ is $S^\perp = \{\, v \in V \mid (v, w) = 0 \forall w \in S \,\}$.

Fact. If $v, w \in V$ are orthogonal, then $\|v + w\|^2 = \|v\|^2 + \|w\|^2$.

Fact. $S^\perp$ is a closed subspace of V, and $(\overline{\operatorname{span} S})^\perp = S^\perp$.

Proof. $S^\perp = \bigcup_{w \in S} f_w^{-1}(0)$ where $f_w(v) = (v, w)$ is continuous. Thus $S^\perp$ is the intersection of closed sets, so closed. It is clear that $S^\perp$ is a linear subspace. Clearly,

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

On the other hand, let $v \in S^\perp$ and $w \in \overline{\operatorname{span} S}$, i.e. $w = \lim_{k \to \infty} w_k$ with $w_k \in \operatorname{span} S$. Then $(v, w) = \lim_{k \to \infty}(v, w_k) = 0$, so $v \in (\operatorname{span} S)^\perp$. $\qquad\square$

For $W \subset V$ a linear subspace, $W^\perp \cap W = \{0\}$. Thus the sum $W + W^\perp$ is direct, and we write $W + W^\perp = W \oplus W^\perp$.

Example.

- If X is finite-dimensional, then $V = W \oplus W^\perp$ for any linear subspace $W \subset V$.

- Let $V = C[0, 1]$ with $(f, g) = \int_0^1 f\overline{g}\, dx$ and $W = C^1[0, 1] \subset C[0, 1]$. Then $W^\perp = \{0\}$, since $\int_0^1 f\overline{g}\, dx = 0 \forall g \in C^1$ implies $f = 0$. Thus $V \neq W \oplus W^\perp$.

Theorem. Let $W \subset V$ be a complete subspace. Then $W \oplus W^\perp = V$. Moreover, given $v \in V$, its unique decomposition $v = v_\parallel + v_\perp$ with $v_\parallel \in W$ and $v_\perp \in W^\perp$ is characterised by

$$\|v_\perp\| = \|v - v_\parallel\| = \inf_{w \in W} \|v - w\|. \tag{$*$}$$

The assumption that W is complete holds in particular if V is a Hilbert space and $W \subset V$ is closed.

Proof. Let $v \in V$, and $D = \inf_{w \in W} \|v - w\|$. Choose any sequence $(w_k) \subset W$ such that $\|v - w_k\| \to D$.

Claim: (w_k) is Cauchy. By the parallelogram identity, $\|x - y\|^2 + \|x + y\|^2 = 2\|x\|^2 + 2\|y\|^2$, with $x = v - w_j$ and $y = v + w_k$, we get

$$\|w_j - w_k\|^2 + \|2v - w_j - w_k\|^2 = 2\|v - w_j\|^2 + 2\|v - w_k\|^2 \tag{1}$$

$$\implies \|w_j - w_k\|^2 = 2\underbrace{\|v - w_j\|^2}_{\to D^2} + 2\underbrace{\|v - w_k\|^2}_{\to D^2} - 4\underbrace{\left\|v - \frac{1}{2}(w_j + w_k)\right\|^2}_{\geq D^2} \tag{2}$$

$$\leq 4(D^2 + \epsilon) - 4D^2 \text{ for } i, k \text{ sufficiently large} \tag{3}$$

$$= 4\epsilon \tag{4}$$

So, (w_j) is Cauchy, and by completeness of W, $w_j \to v_\parallel$ for some $v_\parallel$ in W.

Claim: $v_\perp = v - v_\parallel \in W^\perp$. If not, $\exists \tilde{w} \in W$ such that $(\tilde{w}, v_\perp) > 0$.

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

But also,

$$\|v_\perp - t\tilde{w}\|^2 = \|v - (v_\parallel + t\tilde{w})\|^2 \geq D^2$$

by definition of D, a contradiction.

Thus, $v_\perp \in W^\perp$ and $v = v_\parallel + v_\perp$, $v_\parallel \in W$, $v_\perp \in W^\perp$.

To show that the decomposition is characterised by ($*$), suppose

$$v = \tilde{v}_\perp + \tilde{v}_\parallel \text{ for } \tilde{v}_\perp \in W^\perp, \tilde{v}_\parallel \in W.$$

Then $\tilde{v}_\parallel = v_\parallel + w$, $w = v_\perp - \tilde{v}_\perp \in W^\perp$ but also $w = \tilde{v}_\parallel - v_\parallel \in W$. Thus $w \in W \cap W^\perp = \{0\}$. $\square$

Corollary. Let H be a Hilbert space and $S \subset H$. Then $\overline{\operatorname{span} S} = ((\overline{\operatorname{span} S})^\perp)^\perp = (S^\perp)^\perp$.

Theorem (Riesz Representation Theorem)**.** let H be a Hilbert space. Then for any $l \in H^*$ there is a unique $v_l \in H$ such that $l(w) = (w, v_l)$ for all $w \in H$. Moreover, $\|l\| = \|v_l\|$.

Corollary. The antilinear map $H \to H^*$, $v \mapsto (\cdot, v)$ is bijective and isometric. Thus H and H^* can be identified. In particular, H^* is a Hilbert space. For $l, l' \in H^*$ the inner product on H^* is $(l, l') = (v'_l, v_l)$.

Proof of Riesz Representation Theorem. Let $l \in H^*$, $l \neq 0$. Then $\ker L$ is closed and the previous theorem implies the existence of the orthogonal decomposition $H = \ker l \oplus (\ker l)^\perp$. Since $l \neq 0$, $(\ker l)^\perp \neq \{0\}$.

Claim: There is $v_0 \in H$ such that $(\ker l)^\perp = \mathrm{span}\{v_0\}$ and $\|v_0\| = 1$. Let $v_0 \in (\ker l)^\perp$, $\|v_0\| = 1$. Then for any $w \in H$,

$$w = \underbrace{\left(w - \frac{l(w)}{l(v_0)} v_0 \right)}_{\in \ker l} + \underbrace{\frac{l(w)}{l(v_0)} v_0}_{\in \mathrm{span}\{v_0\} \subset (\ker l)^\perp}$$

Claim: Define $v_l = \overline{l(v_0)} v_0$. Then $l(v) = (v, v_l)$ for any $v \in H$. If $v \in \ker l$, then $l(v) = 0$ and $(v, v_l) = l(v_0)(v, v_0) = 0$. If $v \in (\ker l)^\perp$, i.e. $v = \lambda v_0$, $\lambda \in \mathbb{K}$, then

$$(v, v_l) = \lambda l(v_0)(v_0, v_0) = \lambda l(v_0) = l(v).$$

Since l and $(\cdot, v_l)$ are linear and agree on $\ker l$ and $(\ker l)^\perp$, they must also agree on $\ker l \oplus (\ker l)^\perp = H$.

Norm: $\|l\| = \sup_{\|v\| \leq 1} |l(v)| = \sup_{\|v\| \leq 1} |(v, v_l)| = \|v_l\|$ where the last equality follows from taking $v = \frac{v_l}{\|v_l\|}$ and Cauchy-Schwarz.

Uniqueness: If $(v, v_l) = (v, \tilde{v}_l)$ for all $v \in H$, then

$$(v, v_l - \tilde{v}_l) = 0 \quad \forall v$$
$$\implies (v_l - \tilde{v}_l, v_l - \tilde{v}_l) = 0 \implies \tilde{v}_l = v_l.$$

$\square$

Definition. Let V be a Euclidean vector space.

- A linear operator $P : V \to V$ is a **projection** if $P^2 = P$.

- P is an **orthogonal projection** if $P^2 = P$ and $(Pv, w) = (v, Pw)$ for all $v, w \in V$ (P is self-adjoint)

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

Proof. Let $v \in V$, $Pv \neq 0$. Then

$$\|Pv\| = \frac{(Pv, Pv)}{\|Pv\|} = \frac{(v, P^2 v)}{\|Pv\|} = \frac{(v, Pv)}{\|Pv\|} \leq \|v\| \cdot \|P\| \leq 1.$$

On the other hand, if $\|P\| \neq 0$, there is v such that $Pv \neq 0$. Let $w = Pv$. Then $\|Pw\| = \|Pv\| = \|w\| \implies \|P\| \geq 1$. $\square$

Corollary. Let $W \subset V$ be complete (i.e. closed if V is a Hilbert space). Then there is an orthogonal projection $P : V \to V$ with

$$\mathrm{im}\, P = W, \ker P = W^\perp$$

Proof. Let $Pv = v_\parallel$ with $v_\parallel \in W$ and $v_\perp = v - v_\parallel \in W^\perp$ as in the theorem. Then P is linear since

$$v = v_\parallel + v_\perp, \quad w = w_\parallel + w_\perp, \quad v_\parallel, w_\parallel \in W, \quad v_\perp, w_\perp \in W^\perp \implies \lambda v + \mu w = \underbrace{(\lambda v_\parallel + \mu w_\parallel)}_{\in W} + \underbrace{(\lambda v_\perp + \mu_\perp)}_{\in W^\perp} \implies P(\lambda v$$

by uniqueness of the orthogonal decomposition. Clearly, $P^2 = P$ and $(Pv, w) = (v_\parallel, w_\parallel + w_\perp) = (v_\parallel, w_\parallel) = (v_\parallel + v_\perp, w_\parallel) = (v, Pw)$ so P is orthogonal. $\square$

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

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

Then $S_n : V \to V$ defined by the partial Fourier sum

$$S_n f(x) = \sum_{k=-n}^{n} \hat{f}_k e^{ikx}, \quad \hat{f}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} \, dx$$

is the orthogonal projection with image $W = \text{span}\{e^{ikx}\}_{|k| \leq n}$ (which is finite dimensional and thus complete).

Proof. Let $e_k(x) = e^{ikx}$. Then $\hat{f}_k = (f, e_k)$ so $S_n f = \sum_{k=-n}^{n} e_k(f, e_k)$. If $f \in W$, i.e., $f = \sum_{k=-n}^{n} a_k e_k$, $a_k \in \mathbb{C}$, then

$$S_n f = \sum_{|k| \leq n} e_k \left(\sum_{|k| \leq n} a_l e_l, e_k \right) = f \text{ since } (e_k, e_l) = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{if } k \neq l \end{cases}$$

If $f \in W^\perp$ then $(f, e_k) = 0$ for all $|k| \leq n$, so $S_n f = 0$.

$\implies$ S_n is a projection with image W. Also, $(S_n f, g) = (f, S_n g)$ for all f, g so S_n is orthogonal. $\square$

5.3 Orthonormal systems

Definition (Orthonormal system).

- Let V be a Euclidean vector space. A set $\{e_\alpha\}_\alpha \subset V$ of unit vectors is an **orthonormal system** (ONS) if $(e_\alpha, e_\beta) = 0$ for all $\alpha \neq \beta$.

- An orthonormal system is **maximal** if it cannot be extended to a larger orthonormal system.

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

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

Proof. Let $W = \overline{\text{span} \, S}$. Then W is complete, so $H = W \oplus W^\perp$. Suppose that $W^\perp \neq \{0\}$. Then there exists $0 \neq v \in W^\perp = S^\perp$, so $S \cup \{\frac{v}{\|v\|}\}$ is an orthonormal system. Thus

$$\overline{\text{span} \, S} \neq H \iff W^\perp \neq \{0\} \iff S \text{ is not maximal.} \qquad \square$$

Fact (Gram-Schmidt). Let V be Euclidean and $\{v_i\}_{i=1}^N$ linear independent (with $N = \infty$ permitted). Then there is an orthonormal system $\{e_i\}_{i=1}^N$ with $\text{span}\{e_i\}_{i=1}^N$ with $\text{span}\{v_i\}_{i=1}^k$ for all $k \leq N$.

Proof. Let $e_1 = \frac{v_1}{\|v_1\|}$ and given $e_1, \ldots, e_k$, set

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

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

Thus from now on, we will always take Hilbert bases to be countable if H is separable.

Proposition (Bessel inequality)**.** Let V be a Euclidean vector space and $\{e_i\}_{i=1}^N$ an orthonormal system (with $N = \infty$ permitted). Then

$$\sum_{i=1}^N |(v, e_i)|^2 \leq \|v\|^2 \quad \text{for all } v \in V.$$

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

Proof. By taking a limit, we may assume that $N < \infty$. Define $Pv = \sum_{i=1}^N (v, e_i) e_i$. Then $P^2 = P$,

$$(Pv, w) = \sum_{i=1}^N (v, e_i)\overline{(v, e_i)} = (v, Pw) \quad \text{for all } v, w \in V.$$

Thus P is an orthogonal projection $\implies \sum_{i=1}^N |(v, e_i)|^2 = \|Pv\|^2 \leq \|v\|^2$ for all $v \in V$. $\quad\square$

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

(i) For any $x \in H$, set $x_i = (x, e_i)$. Then $(x_i) \in \ell^2$ and

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

(ii) Conversely, if $(x_i) \in \ell^2$ then there is $x \in H$ such that $(x, e_i) = x_i$ for all i.

(iii) For any $x, y \in H$,

$$(x, y) = \sum_{i=1}^\infty \qquad \text{(Parseval identity)}.$$

In particular, the map

$$\phi : H \longrightarrow \ell^2$$
$$x \longmapsto ((x, e_i))_{i=1}^\infty \in \ell^2$$

is an isometric isomorphism. This is sometimes called the Riesz-Fischer Theorem.

Proof.

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

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

since $(x_i) \in \ell^2$ by Bessel's inequality. By completeness of H, there is $s \in H$ such that $s_n \to s$.

Now claim $s = x$. For any i,

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

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

40

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

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

(iii) Similarly.

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

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

6 Spectral theory

From now on, Banach and Hilbert spaces are always complex.

6.1 Spectrum and resolvent

Definition. Let X be a (complex) Banach space and $T \in \mathcal{B}(X)$ where from now on $\mathcal{B}(X) = \mathcal{B}(X, X)$.

- The **resolvent set** of T is

$$\rho(T) = \{\, z \in \mathbb{C} \mid T - z \text{ is bijective and } (T - z)^{-1} \in \mathcal{B}(X) \,\}.$$

  $T - z$ denotes the operator $T - z\,\mathrm{id}$.

- The **spectrum** of T is

$$\sigma(T) = \mathbb{C} \setminus \rho(T).$$

- The **resolvent** of T is the map $R_T : \rho(T) \to \mathcal{B}(X)$,

$$R_T(z) = (T - z)^{-1}.$$

Remark. If T is bounded above and $T - z$ is bijective, then $T - z \in \mathcal{B}(X)$ is automatic by the [open mapping theorem].

Fact. Let $T \in \mathcal{B}(X)$ with $\|T\| \leq 1$. Then the series $\sum_{n=0}^{\infty} T^n$ converges in $\mathcal{B}(X)$ and

$$(\mathrm{id} - T)^{-1} = \sum_{n=0}^{\infty} T^n, \quad \left\| (\mathrm{id} - T)^{-1} \right\| \leq \frac{1}{1 - \|T\|}.$$

Proof. Let $S_n = \sum_{k=0}^{n} T^k$. Then (S_n) is Cauchy in $\mathcal{B}(X)$:

$$\|S_n - S_m\| \leq \sum_{k=n}^{\infty} \|T\|^k \to 0 \text{ as } n \to \infty.$$

Thus $S = \lim_{n \to \infty} S_n \in \mathcal{B}(X)$ exists and $\|S\| \leq \lim_{n \to \infty} \sum_{k=0}^{n} \|T\|^k \leq \frac{1}{1 - \|T\|}$ and

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

$\square$

Proposition. Let $z_0 \in \rho(T)$. Then $\rho(T)$ contains the disk

$$D = \{\, z \in \mathbb{C} \mid |z - z_0| \, \|R_T(z_0)\| < 1 \,\}.$$

In particular, $\rho(T)$ is open, $\sigma(T)$ is closed. Moreover, the map $R_T : \rho(T) \to \mathcal{B}(X)$ is analytic (can be represented by an absolutely convergent power series in any small enough disk).

Proof. For $z \in D$,

$$T - z = (T - z_0) - (z - z_0)$$
$$= (T - z_0)(\text{id} - \underbrace{R_T(z_0)(z - z_0)}_{\|\cdot\| < 1 \text{ by definition of } D.})$$

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

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

$$\implies z \in \rho(T).$$

Therefore, $D \subset \rho(T)$. Moreover, R_T is analytic on $\rho(T)$. $\square$

Corollary. $\sigma(T) \neq \varnothing$ and $\sigma(T) \subseteq \{ z \in \mathbb{C} \mid |z| \leq \|T\| \}$.

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

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

Thus $\sigma(T) \subseteq \{ z \in \mathbb{C} \mid |z| \leq \|T\| \}$ and $\|R_T(z)\| \to 0$ as $|z| \to \infty$. If $\sigma(T) = \varnothing$ then $R_T : \mathbb{C} \to \mathcal{B}(X)$ would be entire. By Liouville's Theorem it would have to be constant, thus 0 (since it tends to 0 as $|z| \to \infty$). But $R_T(z)$ is not identically 0, since e.g. $z R_T(z) \xrightarrow{|z| \to \infty}$ id. $\square$

6.2   Classification of the spectrum

Proposition. Let X be a Banach space, Y a normed space, $T \in \mathcal{B}(X, Y)$. Then $T^{-1} \in \mathcal{B}(Y, X)$ iff $\text{im } T$ is dense in Y and T is bounded below:

$$\exists \epsilon > 0 \text{ s.t. } \forall x \in X \setminus \{0\}, \; \|Tx\| > \epsilon \|x\|.$$

Proof. The direction $T^{-1} \in \mathcal{B}(Y, X) \implies (\text{im } T \text{ dense and } T \text{ bounded below})$ is clear. Thus assume that $T \in \mathcal{B}(X, Y)$ is such that $\text{im } T$ is dense and T bounded below. Then T is injective, so bijective onto its image. Let $S : \text{im } T \to X$ be its inverse. Since T is bounded below, S is bounded. Since $\text{im } T$ is dense in Y, for every $y \in Y$, there are $(y_k) \subset \text{im } T$ such that $y_k \to y$. Define $T^{-1}y = \lim_{k \to \infty} Sy_k$. The limit exists since S is bounded and X is complete, $T^{-1}y$ is linear in y and independent of the approximating sequence. Moreover, $T^{-1} \in \mathcal{B}(Y, X)$,

$$TT^{-1}y = \lim_{k \to \infty} TSY_k = \lim_{k \to \infty} y_k = y,$$

as claimed. $\square$

Definition.   • $\sigma_p(T) = \{ \lambda \in \mathbb{C} \mid T - \lambda \text{ is not injective} \}$ is the **point spectrum** or the set of eigenvalues.

- $\sigma_c(T) = \{\, \lambda \in \sigma(T) \mid T - \lambda \text{ is injective and } \operatorname{im}(T - \lambda) \text{ is dense}\,\}$ is the **continuous spectrum**.

- $\sigma_r(T) = \{\, \lambda \in \sigma(T) \mid T - \lambda \text{ is injective and } \operatorname{im}(T - \lambda) \text{ is not dense}\,\}$ is the **residual spectrum**.

Remark. By the previous proposition, if $\lambda \in \sigma_c(T)$, then $T - \lambda$ is not bounded below. Then there exists a sequence $(x_k) \subset X$ with $\|x_k\| = 1$ such that

$$Tx_k - \lambda x_k \to 0 : \quad \lambda \text{ is an approximate eigenvalue.}$$

The set $\sigma_{ap}(T) = \{\, \lambda \in \sigma(T) \mid \lambda \text{ is an approximate eigenvalue}\,\}$ is the approximate point spectrum.

Example. Let X be finite dimensional. Then equivalently

- $T - \lambda$ is injective

- $T - \lambda$ is surjective

so $\sigma(T) = \sigma_p(T)$. Moreover, $\sigma_p(T) = \{\, \lambda \in \mathbb{C} \mid \det(T - \lambda) = 0\,\}$ contains at most $\dim X$ points.

6.3   Normal linear operators on a Hilbert space

Recall that given $T : H \to H$ we defined the adjoint $T^* : H^* \to H^*$ by $(T^* l)(x) = l(Tx)$ for any $l \in H^*$ and $x \in H$. In a Hilbert space, $H \cong H^*$ and thus regard T^* as a map $T^* : H \to H$. Explicitly,

$$(Tx, y) = (x, T^* y) \quad \text{for } x, y \in H.$$

Definition. $T \in \mathcal{B}(H)$ is

- **normal** if $TT^* = T^* T$

- **self-adjoint** if $T = T^*$

- **unitary** if $TT^* = T^* T = \operatorname{id}$

In particular, self-adjoint and unitary operators are normal.

Exercise. Let $T \in \mathcal{B}(H)$ be normal. Then $\|Tx\| = \|T^* x\| \ \forall x \in H$ and $\ker T = \ker T^* = (\operatorname{im} T)^\perp = (\operatorname{im} T^*)^\perp$. Also

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

Corollary. For T normal, $\sigma_r(T) = \varnothing$ and $\sigma(T) = \sigma_{ap}(T)$.

Corollary. For T normal, if $Tx = \lambda x$ then $T^* x = \bar{\lambda} x$. In particular, $\overline{\sigma_p(T)} = \sigma_p(T^*)$ and similarly $\overline{\sigma_c(T)} = \sigma_c(T^*)$.

Proof. If T is normal, then so is $T - \lambda$ and $(T - \lambda)^* = T^* - \bar{\lambda}$. Thus $\|(T - \lambda)x\| = \|(T - \lambda)^* x\| = \|(T^* - \bar{\lambda})x\|$ so $Tx = \lambda x \iff T^* x = \bar{\lambda} x$ and more generally

$$(T - \lambda)x_j \to 0 \iff (T^* - \bar{\lambda})x_j \; to0.$$

$\square$

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

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

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

$$e^{itT} = \sum_{n=0}^{\infty} \frac{1}{n!}(itT)^n \in \mathcal{B}(H)$$

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

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

Any solution to the Schrödinger equation

$$-i\frac{\partial}{\partial t}\Psi(t) = t\Psi(t), \quad \Psi(0) = \psi_0$$

is given by $\Psi(t) = U(t)\psi_0$. For any $t \in \mathbb{R}$, e^{itT} is unitary. In Quantum Mechanics, T is called the Hamiltonian and Ψ the wave function.

Lemma. Let $T \in \mathcal{B}(H)$ 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 definition,

$$\|T\| = \sup_{\|x\|\leq 1} (Tx, Tx)^{\frac{1}{2}}\,.$$

Let $(x_i) \subset H$, $\|x_i\| = 1$ be such that $(Tx_i, Tx_i) \to \|T\|^2$.

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

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

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

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

On the other hand $|(x, Ty)| \leq \|T\|$ for $\|x\| \leq 1$, $\|y\| \leq 1$, showing the first equality.

To show the second equality, first note $\sup_{\|x\|\leq 1}|(Tx, x)| \leq \sup_{\|x\|leq1}\sup_{\|y\|\leq 1}|(Tx, y)|$ is clear. On the other hand,

$$|(x, Ty)| = \frac{1}{4}|((T(x + y), x + y) - (T(x - y), x - y))|$$

$$\leq \frac{1}{4}\sup_{\|x\|\leq 1}|(Tx, x)|\,(\underbrace{\|x+y\|^2 + \|x-y\|^2}_{\leq 2\|x\|^2 + 2\|y\|^2})$$

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

as required. $\qquad\square$

Lemma. 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

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

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

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

$\square$

6.4 Spectral Theorem for compact self-adjoint operators

Definition. Let X, Y be normed vector spaces. Then $T \in \mathcal{B}(X, Y)$ is compact if $T(B)$ is relatively compact for any bounded $B \subset X$.

Lemma. Let T be compact. Then any nonzero approximate eigenvalue of T is an eigenvalue.

Proof. Assume that $Tx_i - \lambda x_i \to 0$ with $\|x_i\| = 1$, $\lambda \ne 0$. By compactness there is a subsequence Λ such that $Tx_i \to y$ ($i \in \Lambda$, $i \to \infty$)

$$\implies Ty = T(\lim Tx_i) = \lim T(\lambda x_i) = \lambda y.$$

Moreover if $\lambda \ne 0$ then $y \ne 0$. $\square$

Corollary. Let $T \in \mathcal{B}(H)$ be self-adjoint and compact. Then $\|T\|$ or $-\|T\|$ is an eigenvalue.

Notation. $E_\lambda = \ker(T - \lambda)$.

Strategy: repeat corollary with T replaced by $T|_{E_{\lambda_1}^\perp}$, $|\lambda_i| = \|T\|$.

Lemma. Let $T \in \mathcal{B}(H)$ be self-adjoint. Then for eigenvalues $\lambda \ne \mu$, the spaces E_λ and E_μ are orthogonal.

Proof. Assume $Tx = \lambda x$ and $Ty = \mu y$, for some x, y, $\lambda \ne \mu$ and say $\lambda \ne 0$. Then

$$(x, y) = \frac{1}{\lambda}(Tx, y) = \frac{1}{\lambda}(x, Ty) = \frac{\mu}{\lambda}(x, y)$$
$$\implies \mu = \lambda \text{ or } (x, y) = 0.$$

$\square$

Lemma. Let $T \in \mathcal{B}(H)$ be self-adjoint. For any non-zero eigenvalues $\{\lambda_i\}_{i \in I}$,

$$T\left(\left(\bigoplus_{i \in I} E_{\lambda_i}\right)^\perp\right) \subset \left(\bigoplus_{i \in I} E_{\lambda_i}\right)^\perp.$$

Proof. Let $y \in (\bigoplus_{i \in I} E_{\lambda_i})^\perp$. Then for any $x \in E_{\lambda_i}$,

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

$$\implies (x, Ty) = 0 \quad \forall x \in \bigoplus_{i \in I} E_{\lambda_i}$$

$$\implies T((\bigoplus_{i \in I} E_{\lambda_i})^\perp) \subset (\bigoplus_{i \in I} E_{\lambda_i})^\perp. \qquad \square$$

Lemma. Let $T \in \mathcal{B}(H)$ be self-adjoint and compact. Then, for any $\epsilon > 0$,

$$\bigoplus_{\substack{\lambda \in \sigma_p(T) \\ |\lambda| > \epsilon}} E_\lambda$$

is finite dimensional.

Proof. Otherwise there are infinitely many eigenvectors (x_i) such that $\|x_i\| = 1$ and $(x_i, x_j) = 0$ for $i \neq j$ and

$$\|Tx_i - Tx_j\|^2 = \|Tx_i\|^2 + \|Tx_j\|^2 \geq 2\epsilon.$$

This is a contradiction to compactness of T. $\qquad \square$

Theorem (Hilbert-Schmidt)**.** Let $T \in \mathcal{B}(H)$ be self-adjoint and compact. Then there are at most countably many eigenvalues (λ_i) which can accumulate to 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

$$T = \sum_{j=1}^\infty \lambda_j P_{\lambda_j} \text{ in } \mathcal{B}(H), \quad H = \ker T \oplus E_{\lambda_1} \oplus \cdots$$

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

Proof. By the previous lemmas, there is an eigenvalue λ_1 such that $|\lambda_1| = \|T\|$. Given $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_k|$ such that T has no eigenvalue with magnitude $> |\lambda_k|$, let $H_k = (\bigoplus_{i=1}^k E_{\lambda_i})^\perp$. Then $H_k \subset H$ is closed, so itself a Hilbert space. Also,

$$T|_{H_k} : H_k \to H_k, \quad \|T|_{H_k}\| \leq |\lambda_k|.$$

Thus there is an eigenvalue λ_{k+1} different from the λ_i, $i \leq k$, with $|\lambda_{k+1}| = \|T|_{H_k}\| \leq |\lambda_k|$ and there is no other eigenvalue μ with $|\lambda_{k+1}| < |\mu| \geq |\lambda_k|$. This defines a sequence (possibly finite) (λ_i) with $|\lambda_{k+1}| \leq |\lambda_k|$. Since $\bigoplus_{|\lambda_j| > \epsilon} E_{\lambda_j}$ is finite dimensional, this sequence can only accumulate at 0. In particular, (λ_k) is finite or countable with $\lambda_k \to 0$.

Since $T|_{(\bigoplus_{i=1}^\infty E_{\lambda_i})^\perp}$ cannot have a nonzero eigenvalue, $\left\|T|_{(\bigoplus_{i=1}^\infty E_{\lambda_i})^\perp}\right\| = 0$ and $H = \ker T \oplus (\overline{\bigoplus_{i=1}^\infty E_{\lambda_i}})$.

It follows that

$$\left\|Tx - \sum_{i=1}^n \lambda_i P_{\lambda_i} x\right\| = \|Tx - TP_n x\| \leq |\lambda_n| \, \|x\|$$

$$\implies \left\|T - \sum_{i=1}^n \lambda_i P_{\lambda_i}\right\| \leq |\lambda_n| \to 0 \implies T = \sum_{i=1}^\infty \lambda_i P_{\lambda_i} \in \mathcal{B}(H)$$

where $P_n = \sum_{i=1}^n P_{\lambda_i}$. $\qquad \square$

47

Corollary. Let H be separable and $T \in \mathcal{B}(H)$ self-adjoint and compact. Then there is an orthonormal (Hilbert) bases for H of T eigenvectors.

Proof. Apply Gram-Schmidt to the eigenspaces and the kernel. $\qquad\qquad\square$

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

Proof. Let $T_n = \sum_{i=1}^{n} \lambda_i P_{\lambda_i}$. We will show that if $\mu \notin \sigma_p(T) \cup \{0\}$ then $\mu \notin \sigma(T)$.

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

$$\implies (T_n - \mu)^{-1} = \sum_{i=1}^{n}(\lambda_i - \mu)^{-1}P_{\lambda_i} - \mu^{-1} \sum_{i=n+1}^{\infty} P_{\lambda_i} \text{ exists and}$$

$$\left\|(T_n - \mu)^{-1}\right\| \leq \max\{|\mu|^{-1}, |\lambda_i - \mu|^{-1}\}. \implies T - \mu = T_n - \mu + T - T_n = (T_n - \mu)(1 + \underbrace{(T}$$

$$\implies (T - \mu)^{-1} \in \mathcal{B}(H) \implies \mu \notin \sigma(T)$$

$$\square$$

6.5 Application: Boundary Value Problem

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

$$Tu(x) = -u''(x) + V(x)u(x)$$

with boundary condition $u(a) = u(b) = 0$ and V is a continuous function.

More generally, we could consider a Sturm-Liouville operator

$$Tu(x) = -\frac{d}{dx}\left(a(x)\frac{d}{dx}u(x)\right) + V(x)$$

where $a(x) > 0$ is in C^1, V is continuous, and

$$A_1 u(a) + A_2 u'(a) = 0, \qquad\qquad (A_1, A_2) \neq (0, 0)$$
$$B_1 u(b) + B_2 u'(b) = 0, \qquad\qquad (B_1, B_2) \neq (0, 0).$$

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

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

is given by $u(x) = \int_a^b k(x, y)f(y)\, dy$ and $k(x, y) = k(y, x)$.

Proof. By the theory of linear ODEs, there are C^2 functions $u_a : [a, b] \to \mathbb{R}$ and $u_b : [a, b] \to \mathbb{R}$ such that any solution to

$$-u''(x) + V(x)u(x) = 0, \quad u(a) = 0$$

48

is a multiple of u_a and any solution to

$$-u''(x) + V(x)u(x) = 0, \quad u(b) = 0$$

is a multiple of u_b. Define

$$k(x,y) = \begin{cases} u_a(y)u_b(x) & \text{if } a \leq y \leq x \leq b \\ u_a(x)u_b(y) & \text{if } a \leq x \leq y \leq b \end{cases}$$

Claim: $u(x) = \int k(x,y)f(y)\,dy$ solves $(*)$.

$$u(x) = \int_a^x u_a(y)u_b(x)f(y)\,dy + \int_x^b u_a(x)u_b(y)f(y)\,dy$$

$$\implies u'(x) = \underline{u_a(x)u_b(x)f(x)} + \int_a^x u_a(y)u_b'(x)f(y)\,dy$$

$$- \underline{u_a(x)u_b(x)f(x)} + \int_x^b u_a'(x)u_b(y)f(y)\,dy$$

$$\implies u''(x) = u_a(x)u_b'(x)f(x) + \int_a^x u_a(y)u_b''(x)f(y)\,dy$$

$$- u_a'(x)u_b(x)f(x) + \int_x^b u_a''(x)u_b(y)f(y)\,dy$$

$$= \underbrace{(u_a u_b' - u_a' u_b)(x)}_{=:W(x)} f(x) + V(x)u(x)$$

$$W'(x) = (u_a u_b'' - u_a'' u_b)(x)$$
$$= V(x)(u_a u_b - u_a u_b)(x) = 0 \implies W(x) = C$$

Thus,

$$u''(x) = Cf(x) + V(x)u(x).$$

Multiplying u_a by a constant, we can assume $C = -1$. Then

$$-u''(x) + V(x)u(x) = f(x).$$

Also,

$$u(a) = \int_a^b k(a,y)f(y)\,dy = u_a(a)\int_a^b u_b(y)f(y)\,dy = 0$$

$$u(b) = \int_a^b k(b,y)f(y)\,dy = u_b(b)\int_a^b u_a(y)f(y)\,dy = 0$$

and $k(x,y) = k(y,x)$. $\qquad\square$

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

$$Kf(x) = \int_a^b k(x,y)f(y)\,dy$$

is bounded and compact from $(C[a,b], \|\cdot\|_2)$ to $(C[a,b], \|\cdot\|_\infty)$.

49

Proof. By Cauchy-Schwarz,

$$\|Kf\|_\infty \le \sup_x \int |k(x,y)|\, |f(y)|\ dy \le \underbrace{\sup_x (\int |k(x,y)|^2\ dy)^{\frac{1}{2}}}_{\le C} \|f\|_2$$

so K is bounded from $(C[a,b], \|\cdot\|_2)$ to $(C[a,b], \|\cdot\|_\infty)$. Also $\{\, Kf \mid f \in C[a,b], \|f\|_2 \le 1 \,\}$ is equicontinuous:

$$|Kf(x) - Kf(y)| \le \underbrace{(\int_a^b |k(x,z) - k(y,z)|^2\ dz)^{\frac{1}{2}}}_{\to 0 \text{ as } x \to y} \|f\|_2 \to 0$$

uniformly in f.

Thus compactness follows from the Arzelà-Ascoli Theorem. □

Corollary. K is bounded and compact from $(C[a,b], \|\cdot\|_2)$ to $(C[a,b], \|\cdot\|_2)$.

Proof. This follows from

$$\|Kf\|_2 \le (\int_a^b |Kf(x)|^2\ dx)^{\frac{1}{2}} \le \sqrt{|a-b|}\, \|Kf\|_\infty\,. \qquad\qquad □$$

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 (Example sheet 4)**.** Let X, Y be Banach spaces and $D \subset X$ a dense subspace. Then a bounded (compact) operator $T : D \to Y$ extends uniquely to a bounded (compact) operator $T : X \to Y$ with the same norm (BLT thoerem).

Corollary. 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 previous fact. That K is self-adjoint follows from the symmetry of k.

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 from H to $(C[a,b], \|\cdot\|_\infty)$ by the previous fact.

By the spectral theorem, there exists an orthonormal system $(f_n) \subset H$, $(\mu_n) \subset \mathbb{R}$ with $\mu_n \to 0$ such that

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

By the last corollary, if $\mu_n \ne 0$, $f_n = \frac{1}{\mu_n} K f_n \in C[a,b]$ and in fact then $f_n \in C^2[a,b]$ since $Kf \in C^2$ if $f \in C^0$. □

Assuming $\ker(K) = 0$, there is thus a Hilbert basis of C^2 eigenfunctions of K. With $\lambda_n = \frac{1}{\mu_n}$, then

$$Tf_n = \lambda_n T K f_n = \lambda_n f_n,$$

so these (f_n) are also eigenfunctions of T, and $\lambda_n \to \infty$.

In QM, the vectors $f \in H$ describe the state of a system and $|f(x)|^2$ the probability density of finding a particle at $x \in [a,b]$. The (λ_n) are the quantized energy levels of the system.

6.6 Continuous functional calculus

Let $T \in \mathcal{B}(H)$ be self-adjoint compact with spectral resolution

$$T = \sum_{\lambda \in \sigma(T)} \lambda P_\lambda.$$

For $f \in C(\sigma(T))$ define

$$f(T) = \sum_{\lambda \in \sigma(T)} f(\lambda) P_\lambda \in \mathcal{B}(H).$$

Theorem. The map $\Phi : C(\sigma(T)) \to \mathcal{B}(H)$, $f \mapsto f(T)$ is a $*$-homomorphism (a homomorphism of C^* algebras):

$$\Phi(fg) = \Phi(f)\Phi(g), \quad \Phi(\lambda f) = \lambda \Phi(f), \quad \Phi(f)^* = \Phi(\bar{f}), \quad \Phi(1) = \mathrm{id} \qquad (*)$$

and it has the following properties:

    a $\|\Phi(f)\| \leq \|f\|$ (in particular Φ is continuous)

    b if $f(t) = t$ for all t then $f(T) = T$

    c if $Tx = \lambda x$ then $f(T)x = f(\lambda)x$

    d $\sigma(f(T)) = f(\sigma(T)) = \{\, f(\lambda) \mid \lambda \in \sigma(T) \,\}$

    e if $f \geq 0$ then $f(T) \geq 0$, i.e. $(x, f(T)x) \geq 0$ for all $x \in H$.

Remark. For any $x \in H$, $f \in C(\sigma(T)) \mapsto \mu_x(f) = (x, f(T)x)$ defines a positive linear functional on $C(\sigma(T))$, i.e. a measure. This is the spectral measure of $x \in H$.

Theorem. Let $T \in \mathcal{B}(H)$ be self-adjoint. Then there is a unique $*$-homomorphism $\Phi : C(\sigma(T)) \to \mathcal{B}(H)$ with all properties (a)-(e).

Idea. $(*)$ and (b) determine Φ uniquely for any polynomial f. Since the set of polynomials is dense in $C(\sigma(T))$, the main task is to prove (a). Assuming (a), the existence and uniqueness follow by approximation (BLT theorem).

Lemma. Let $P(t) = \sum_{n=0}^{N} a_n t^n$, $P(T) = \sum_{n=0}^{N} a_n T^n$. Then (d) holds for $f = P$:

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

Proof. Let $\lambda \in \sigma(T)$ and write $P(t) - P(\lambda) = (t - \lambda)Q(t)$ for a polynomial Q.

$$\implies P(T) - P(\lambda) = \underbrace{(T - \lambda)}_{\text{not invertible since } \lambda \in \sigma(T)} Q(T) = Q(T)(T - \lambda)$$

$$\implies P(T) - P(\lambda) \text{ is not invertible}$$

$$\implies P(\lambda) \in \sigma(P(T))$$

Let $\mu \in \sigma(P(T))$, let $\lambda_1, \ldots, \lambda_n$ be roots of $P(\lambda) - \mu$ (with multiplicity).

$$\implies P(t) - \mu = a_N(t - \lambda_1) \cdots (t - \lambda_n).$$

If $\lambda_1, \ldots, \lambda_n \notin \sigma(T)$ then

$$(P(T) - \mu)^{-1} = a_N^{-1}(T - \lambda_1)^{-1} \cdots (T - \lambda_n)^{-1} \in \mathcal{B}(H) \text{ i.e. } \mu \notin \sigma(P(T)).$$

Thus some $\lambda \in \sigma(T)$, so $\mu = P(\lambda_i)$ for some λ_i, i.e. $\mu \in P(\sigma(T))$. $\qquad \square$

Lemma. Let $T \in \mathcal{B}(H)$ be self-adjoint. Then

$$\|T\| = \sup_{\lambda \in \sigma(T)} |\lambda|. \tag{\dagger}$$

In particular, together with the previous lemma,

$$\|P(T)\| = \sup_{\lambda \in \sigma(T)} |P(\lambda)|.$$

Proof. We have proved before that $\|T\| \in \sigma(T)$ or $-\|T\| \in \sigma(T)$. Thus $\|T\| \leq \sup_{\lambda \in \sigma(T)} |\lambda|$. We also saw that $\sigma(T) \subset \{\, z \in \mathbb{C} \mid |z| \leq \|T\| \,\}$. Thus $\|T\| \geq \sup_{\lambda \in \sigma(T)} |\lambda|$.

Now:

$$\|P(T)\|^2 = \sup_{\|x\| \leq 1} \|P(T)x\|^2 = \sup_{\|x\| \leq 1} (x, P(T)^* P(T) x)$$
$$= \|P(T)^* P(T)\|$$
$$= \|(\overline{P}P)(T)\|$$
$$\stackrel{\dagger}{=} \sup_{\lambda \in \sigma(\overline{P}P(T))} |\lambda|$$
$$\stackrel{\text{lemma}}{=} \sup_{\lambda \in \sigma(T)} |\overline{P}P(\lambda)| = (\sup_{\lambda \in \sigma(T)} |P(\lambda)|)^2.$$

$\square$

Proof of theorem. Let $A \subset C(\sigma(T), \mathbb{C})$ be the algebra of complex polynomials. For $P(A)$, set $\Phi(P) = P(T)$. Then

$$\|\Phi(P)\|_{\mathcal{B}(H)} = \|P\|_{C(\sigma(T))}.$$

By the complex Stone-Weierstrass Theorem, A is dense in $C(\sigma(T), \mathbb{C})$. Thus Φ extends uniquely to a map $\Phi : C(\sigma(T), \mathbb{C}) \to \mathcal{B}(H)$.

Properties $(*)$, (a), (b) hold by construction (and approximation by polynomials). Also, (c) is clearly true if f is a polynomial and by approximation for $f \in C(\sigma(T), \mathbb{C})$: if $Tx = \lambda x$ then for $f_n \in A$, $f_n \to f$,

$$f(T)x = \lim_{n \to \infty} f_n(T)x = \lim_{n \to \infty} f_n(\lambda)x = f(\lambda)x.$$

For (d), we have for any $f \in C(\sigma(T), \mathbb{C})$:

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

so $f(T)$ is normal (self-adjoint if f is real-valued). To show that $\sigma(f(T)) \subseteq f(\sigma(T))$ define for $\mu \notin f(\sigma(T))$ the continuous function

$$g(\lambda) = \frac{1}{\mu - f(\lambda)}, \quad \lambda \in \sigma(T)$$

$$\implies \text{id} = g(T)(\mu - f(T)) \implies \mu \notin \sigma(f(T)).$$

The direction $\sigma(f(T)) \subseteq f(\sigma(T))$ take $f_n \in A$ such that $f_n \to f$. By the previous lemma then $f_n(\sigma(T)) = \sigma(f_n(T))$. Thus for $\lambda \in \sigma(T)$ there are (x_n) with $\|x_n\| = 1$ such that

52

$f_n(T)x_n - f_n(\lambda)x_n \to 0.$

$$\implies (f(T) - f(\lambda))x_n = \underbrace{(f_n(T) - f_n(\lambda))}_{\to 0}x_n + \underbrace{(f(T) - f_n(T))}_{\to 0}x_n + \underbrace{(f_n(\lambda) - f(\lambda))}_{\to 0}x_n \to 0$$

$$\implies f(\lambda) \in \sigma(f(T)) \implies f(\sigma(T)) \subseteq \sigma(f(T)).$$

$\square$

Corollary. Let $T \geq 0$, i.e. $\sigma(T) \subset [0, \infty)$. Then there is $\sqrt{T} \in \mathcal{B}(H)$ such that $\sqrt{T}\sqrt{T} = T$.

Corollary. The resolvent satisfies

$$\|R_T(z)\| = \frac{1}{\operatorname{dist}(z, \sigma(T))}.$$