Linear Analysis *

Zhiyuan Bai

Compiled on May 4, 2023

This document serves as a set of revision materials for the Cambridge Mathematical Tripos Part II course *Linear Analysis* in Michaelmas 2021. However, despite its primary focus, readers should note that it is NOT a verbatim recall of the lectures, since the author might have made further amendments in the content. Therefore, there should always be provisions for errors and typos while this material is being used.

Contents

*Based on the lectures under the same name taught by Prof. C. Mouhot in Michaelmas 2021.

0 Introduction

WHat we want to study vector spaces of functions. Why? We get a lot of differential equations arising from e.g. Newtonian mechanics. Not all of them have a nice analytic solution that's easy to see, so we seek instead to understand the solutions from a structural standpoint.
Spaces of functions of course have a lot more attributes than the usual $\mathbb{R}^n$. They usually have infinite dimension, which requires a lot more analysis tools than in finite dimension. For example, not all linear forms are bounded anymore and the algebraic dual no longer guarantees to be isomorphic to the topological dual. The geometry of these spaces (topology, distance, norm, etc.) can also be very complicated, e.g. not all norms are equivalent anymore.
A key example of these spaces is the L^p spaces, which however is not going to be covered in full in this course (since it requires some measure theory), but will be done in Analysis of Functions next term.
We'll first cover some preliminary toolboxes (inequalities, some more analysis, etc.), and then introduce normed vector spaces (NVS). After interluding the course with a brief discussion of finite-dimensional vector spaces, we will go to the celebrated Hahn-Banach theorem ("how big can the dual be") and its variants. Afterwards, we will discuss the consequences of completeness in infinite dimension, e.g. Baire's category theorem, followed by a detailed study of the topology of $C(K)$, the space of continuous functions on a compact space. We'll then talk about "the" separable Hilbert spaces, namely ℓ^2, specifically results related to projections and duality. To finish the course, we will give a introduction to operators and spectral theory.

1 Young, Hölder and Minkowski's Inequalities

Proposition 1.1 (Young's Inequality for Products). *Let $p, q \in (1, \infty)$ be such that $1/p + 1/q = 1$, then for any $a, b \geq 0$, we have*

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

with equality iff $a^p = b^q$.

Proof. If $a = 0$ or $b = 0$ then the inequality is trivial. Assume henceforth that $a, b > 0$ and observe that $\log : (0, \infty) \to \mathbb{R}$ is strictly concave ($\log''(t) = -1/t^2 < 0$). Jensen's inequality then implies that $\forall A, B > 0, \lambda \in [0, 1]$, we have $\log(\lambda A + (1-\lambda)B) \geq \lambda \log A + (1-\lambda) \log B$ with equality iff $A = B$ or $\lambda \in \{0, 1\}$. Take $A = a^p, B = b^q$ and $\lambda = 1/p$ (so $1 - \lambda = 1/q$) gives the inequality. □

Proposition 1.2 (Hölder's Inequality for Vectors and Sequences). *Let $p, q \in (1, \infty)$ be such that $1/p + 1/q = 1$.*
(i) $\forall n \in \mathbb{N}, \forall x, y \in \mathbb{R}^n$,

$$\sum_{k=1}^{n} |x_k y_k| \leq \|x\|_p \|y\|_q, \|v\|_r = \left(\sum_{k=1}^{n} |v_k|^r\right)^{1/r}$$

(ii) Define $\ell^p = \{x \in \mathbb{R}^{\mathbb{N}} : \sum_k |x_k|^p < \infty\}$*, then* $\forall x \in \ell^p, y \in \ell^q$*, we have*

$$\sum_{k=1}^{\infty} |x_k y_k| \leq \|x\|_p \|y\|_q, \|v\|_r = \left(\sum_{k=1}^{\infty} |v_k|^r\right)^{1/r}$$

Proof. (i) implies (ii) by taking a limit. As for (i), clearly if either $\|x\|_p = 0$ or $\|y\|_q = 0$ then the inequality is obvious. Assume henceforth that $\|x\|_p, \|y\|_q > 0$, then we can normalise x, y to $\tilde{x} = x/\|x\|_p, \tilde{y} = y/\|y\|_q$. It then suffices to show that $\sum_{k=1}^n |\tilde{x}_k \tilde{y}_k| \leq 1$. Indeed, we know by Proposition 1.1 that $|\tilde{x}_k \tilde{y}_k| \leq |\tilde{x}_k|^p/p + |\tilde{y}_k|^q/q$. Summing all such terms together gives the result. □

Remark. If we have equality, then we must have the equality case for each $|\tilde{x}_k|, |\tilde{y}_k|$ in Proposition 1.1, i.e. $|\tilde{x}_k|^p = |\tilde{y}_k|^q$ for each k.

Proposition 1.3 (Minkowski's Inequality for Vectors and Sequences)**.** *Let* $p \in [1, \infty)$*, then:*
(i) $\forall x, y \in \mathbb{R}^n, \|x + y\|_p \leq \|x\|_p + \|y\|_p$.
(ii) $\forall x, y \in \ell^p, \|x + y\|_p \leq \|x\|_p + \|y\|_p$.

Proof. Again with a limit argument it suffices to show only (i). If $p = 1$, then the inequality follows from the ordinary triangle inequality in $\mathbb{R}$. For $p > 1$, we have

$$\begin{aligned} \|x + y\|_p^p = \sum_{k=1}^n |x_k + y_k|^p &= \sum_{k=1}^n |x_k + y_k||x_k + y_k|^{p-1} \\ &\leq \sum_{k=1}^n |x_k||x_k + y_k|^{p-1} + \sum_{k=1}^n |y_k||x_k + y_k|^{p-1} \\ &\leq \|x\|_p \|x + y\|_p^{p-1} + \|y\|_p \|x + y\|_p^{p-1} \end{aligned}$$

by Proposition 1.2 (with $q = p/(p-1)$). This implies the result. □

Remark. 1. For $p = 1$, the equality case happens exactly when $|x_k + y_k| = |x_k| + |y_k|$ for each k. For $p > 1$, the equality case happens exactly when $|\tilde{x}_k|^p = |\widetilde{x_k + y_k}|^p = |\tilde{y}_k|^p$ for each k.
2. For $p = 2$, there's a simpler proof based on the polynomial $P : \lambda \mapsto \|x + \lambda y\|^2$. Indeed, $0 \leq P(\lambda) = \|y\|^2 \lambda^2 + 2\langle x, y\rangle \lambda + \|x\|^2$ for all λ, so $\langle x, y\rangle^2 \leq \|x\|^2 \|y\|^2$.

2 Normed Vector Spaces

2.1 Normed and Topological Vector Spaces

Definition 2.1. A field $\mathbb{F}$ is a set with two binary operations, written as $+, \cdot$ (addition and multiplication) respectively, such that:
1. $(\mathbb{F}, +)$ is an abelian group with identity $0 \in \mathbb{F}$.
2. $(\mathbb{F}^\times, \cdot)$ (where $\mathbb{F}^\times = \mathbb{F} \setminus \{0\}$) is also an abelian group with identity $1 \in \mathbb{F}$.
3. $\forall a, b, c \in \mathbb{F}, a \cdot (b + c) = a \cdot b + a \cdot c$.

We of course often omit $\cdot$ and just write ab to denote $a \cdot b$.

Example 2.1. $\mathbb{Q}, \mathbb{R}, \mathbb{C}$.

Definition 2.2. A vector space V is over a field $\mathbb{F}$ is a set (whose elements are called vectors) with a binary operation $A : V \times V \to V$ (addition; also denoted as $+$) and an action $M : \mathbb{F} \times V \to V$ by $\mathbb{F}$ (scalar multiplication; notation sometimes suppressed) such that:
1. $(V, +)$ is an abelian group with identity $0 \in V$.
2. $\forall \lambda, \mu \in \mathbb{F}, v \in V, \lambda(\mu v) = (\lambda\mu)v$.
3. $\forall \lambda \in \mathbb{F}, v, w \in V, \lambda(v + w) = \lambda v + \lambda w$.

Note that intersections of vector subspaces (subset of vector space that is also a vector space under the same operations) are also vector subspaces.

Definition 2.3. Suppose V is a vector space over a field $\mathbb{F}$ and $S \subset V$ is a subset, then the span $\operatorname{Span} S$ of S in V is the smallest subspace of V containing S (or, alternatively, the intersection of all subspaces containing S). Indeed, it is the set of linear combinations of elements in S.

$$\operatorname{Span} S = \left\{ \sum_{i=1}^{N} \lambda_i s_i : \lambda_i \in \mathbb{F}, s_i \in S, N \in \mathbb{N} \right\}$$

Remark. It should be emphasised that linear combination means and will always mean finite linear combinations – who knows what will happen when you take a limit in V.

Definition 2.4. Let V be a vector space over a field $\mathbb{F}$ and $S \subset V$. We say that S is linearly independent if

$$\sum_{i=1}^{N} \lambda_i s_i = 0 \iff \forall i, \lambda_i = 0$$

S is a basis for V if it is linearly independent and $\operatorname{Span} S = V$. If V has a finite basis, we say that V is finite-dimensional. Otherwise, we say it is infinite-dimensional.

Remark. We will prove that every vector space has a basis using Zorn's lemma. Also, any two bases of a given vector space have the same cardinality.

These are all algebra so far, but the course has "analysis" in its name, so let's give it a topology.

Definition 2.5. A norm on a vector space V over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ is a map $\|\cdot\| : V \to \mathbb{R}_{\geq 0}$ such that:
1. (Positive definiteness) $\|v\| = 0 \iff v = 0$.
2. (Compatibility with scalar multiplication) $\forall \lambda \in \mathbb{F}, v \in V, \|\lambda v\| = |\lambda|\|v\|$.
3. (Compatibility with addition; Triangle inequality) $\forall v, w \in V, \|v + w\| \leq \|v\| + \|w\|$.

If we equip a norm $\|\cdot\|$ on a vector space V, we say $(V, \|\cdot\|)$ (or sometimes just V) is a normed vector space.

Unless otherwise specified, all vector spaces we come across in this course will be implicitly assumed to be either real or complex.
Two normed vector spaces are isomorphic if there is a linear isomorphism between them that is also an isometry. This is sometimes also known as an isometric isomorphism.

Example 2.2. On $V = \mathbb{R}^n$ over $\mathbb{R}$, we have the p-norms

$$\|v\|_p = \left(\sum_{i=1}^{n} |v_i|^p \right)^{1/p}$$

for $p \in [1, \infty)$. We also have the ∞-norm defined by $\|v\|_\infty = \sup_k |v_k|$.

Proposition 2.1. *If $(V, \|\cdot\|)$ is a normed vector space, then (V, d) is a metric space with $d(v, w) = \|v - w\|$.*

Proof. Just check the definitions. □

Definition 2.6. Given a set X, a topology on X is a collection $\mathcal{T} \subset 2^X$ of its subsets such that:
1. $\varnothing, X \in \mathcal{T}$.
2. $\forall A, B \in \mathcal{T}, A \cap B \in \mathcal{T}$.
3. $\forall (A_i)_{i \in I} \in \mathcal{T}, \bigcup_{i \in I} A_i \in \mathcal{T}$.

If $\mathcal{T}$ is a topology on X, we call the pair $(X, \mathcal{T})$ a topological space. Elements of $\mathcal{T}$ are called open sets of X.

It's a routine check that intersection of topologies is again a topology.

Definition 2.7. On a metric space (X, d), its metric topology is the smallest topology making all the open balls $\{x \in X : d(x, x_0) < r\}$ open, i.e. the intersection of all topologies containing these open balls.

So there is a natural topology on a normed vector space induced by the metric on it. Being able to do analysis doesn't mean that we have to throw away the algebra, so

Proposition 2.2. *Suppose $(V, \|\cdot\|)$ is a normed vector space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, then*
(i) A, M are continuous with respect to the topology on V induced by the norm and the standard topology on $\mathbb{F}$ (induced by $|\cdot|$).
(ii) For any $v_0 \in V$, the translation $T_{v_0} : V \to V, v \mapsto v + v_0$ is a homeomorphism; For any $\lambda_0 \in \mathbb{F}^\times$, the dilation $D_{\lambda_0} : V \to V, v \mapsto \lambda v$ is also a homeomorphism.

Proof. (i) For A, suppose $U \subset V$ is open and $(v_1, v_2) \in A^{-1}(U)$. Then $v_1 + v_2 \in U$. Since U is open, there is an open ball $B_V(v_1 + v_2, \epsilon)$ for some $\epsilon > 0$ contained in U. Triangle inequality implies that $B_V(v_1, \epsilon/2) + B_V(v_2, \epsilon/2) \subset B_V(v_1 + v_2, \epsilon)$, consequently $A(B_V(v_1, \epsilon/2) \times B_V(v_2, \epsilon/2)) \subset U$ which implies that $A^{-1}(U)$ is open, i.e. A is continuous.
For M, again suppose $U \subset V$ is open and $(\lambda, v) \in M^{-1}(U)$. Since U is open, there is $B_V(\lambda v, \epsilon) \subset U$ for some $\epsilon > 0$. WLOG $\epsilon < 1$, then it's clear that

$$M\left(B_{\mathbb{F}}\left(\lambda, \frac{\epsilon}{3\max\{1, \|v\|\}} \right) \times B_V\left(v, \frac{\epsilon}{3\max\{1, |\lambda|\}} \right) \right) \subset B(\lambda v, \epsilon) \subset U$$

which in turn means that M is continuous.
(ii) follows immediately from (i). □

Remark. One can do this with sequences as well (exercise!).

In order to better understand what is specific to the topology of normed vector spaces in this case, we want to characterise all the topologies on a vector space that can be induced by a norm. The preceding proposition means that for this to happen the topology must be compatible with the algebraic structure (i.e. the algebraic operations are continuous), so let's start there.

Definition 2.8. A topological vector space V is a vector space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ (equipped with their usual topology) which is also a T_1 (points are closed) topological space such that addition and scalar multiplication are continuous.

Two topological vector spaces are isomorphic if there is a linear isomorphism between them that is also a homeomorphism.

Remark. We'll later prove a even stronger separability axiom, namely that every topological vector space is T_2 (Hausdorff).

We now want to characterise topological vector spaces that are actually normed vector spaces by the "shape" of the neighbourhoods.

Definition 2.9. Given a (real or complex) vector space V, a subset $C \subset V$ is convex if $\forall v, w \in C, \forall \lambda \in [0,1], \lambda v + (1-\lambda)w \in C$.
A topological vector space is locally convex if every neighbourhood of 0 contains a convex neighbourhood of 0.

Equivalently, a topological vector space is locally convex if every neighbourhood of every $v \in V$ contains a convex neighbourhood of v.

Definition 2.10. Given a topological vector space V, a subset $B \subset V$ is bounded if for any open neighbourhood U of 0, there is some $t_0 > 0$ such that $\forall t > t_0, B \in tU$. V is said to be locally bounded if there is a bounded neighbourhood of 0.

Again, we can replace 0 by any vector $v \in V$.
Clearly, any normed vector space has a bounded convex neighbourhood of 0 given by an open ball around 0. What we are interested in is the converse.

Theorem 2.3 (Kolmogorov, 1934)**.** *Suppose V is a topological vector space that has a bounded convex neighbourhood C around 0, then there is a norm on V that induces the same topology.*

The proof is by construction (shocking!).

Proof. The first step is to construct some bounded convex neighbourhood $\tilde{C} \subset C$ around 0 that is balanced in the sense that $\lambda\tilde{C} \subset \tilde{C}$ for any $\lambda \in \mathbb{F}, |\lambda| \leq 1$. This is due to the continuity of scalar multiplication: $M^{-1}(C)$ is an open subset of $(0,0)$, so there is some $\epsilon > 0$ and an open neighbourhood U of 0 contained in C such that $B_{\mathbb{F}}(0,\epsilon) \times U \subset M^{-1}(C)$. We take $\tilde{C}$ to be the convex hull of $M(B_{\mathbb{F}}(0,\epsilon) \times U)$, which is open, contains 0, and is contained in C since C is convex and contains $M(B_{\mathbb{F}}(0,\epsilon) \times U)$ by definition. It is then immediate that $\tilde{C}$ is a bounded convex neighbourhood of 0 contained in C. It is also balanced. Suppose $|\lambda| \leq 1$, then $\lambda M(B_{\mathbb{F}}(0,\epsilon) \times U) \subset M(\lambda B_{\mathbb{F}}(0,\epsilon) \times U) \subset M(B_{\mathbb{F}}(0,\epsilon) \times U)$. Taking convex hulls on both sides shows the balancedness.
The second step is to consider the Minkowski functional (or Minkowski gauge) of $\tilde{C}$, defined as $\mu_{\tilde{C}}(v) = \inf\{t \geq 0 : v \in t\tilde{C}\}$ which is well-defined since $v/t \to 0$

as $t \to \infty$ via continuity of M.
We shall show that it is a norm. If $\mu_{\tilde{C}}(v) = 0$ but $v \neq 0$, then by the T_1 property of a topological vector space, there is an open set U containing 0 that does not contain v. Since $\tilde{C}$ is bounded, there is some $t_1 > 0$ such that $\tilde{C} \subset t_1 U$. Choose $t_2 \in (0, t_1^{-1})$, then $v \in t_2\tilde{C} \subset t_1^{-1}\tilde{C} \subset U$ by convexity of C, a contradiction. So $\mu_{\tilde{C}}(v) = 0$ iff $v = 0$.
The property $\mu_{\tilde{C}}(\lambda v) = |\lambda|\mu_{\tilde{C}}(v)$ follows from the balancedness of $\tilde{C}$. This is trivial for $\lambda = 0$. Now assume that $\lambda \neq 0$. For any $t > 0$ with $\lambda v \in t\tilde{C}$, $\lambda v/|\lambda| \in (t/|\lambda|)\tilde{C}$, hence $v \in (t/|\lambda|)\tilde{C}$ by balancedness. So $\mu_{\tilde{C}}(v) \leq t/|\lambda|$ and thus $\mu_{\tilde{C}}(v) \leq |\lambda|^{-1}\mu_{\tilde{C}}(\lambda v)$. Replace λ by λ^{-1} (and v by λv) shows the inequality from the other side. As for the triangle inequality, if $t_1, t_2 > 0$ have $v_1 \in t_1\tilde{C}, v_2 \in t_2\tilde{C}$, then $v_1 + v_2 \in t_1\tilde{C} \subset (t_1+t_2)((t_1/(t_1+t_2))\tilde{C} + (t_2/(t_1+t_2))\tilde{C}) \subset (t_1+t_2)\tilde{C}$ by the convexity of C. So $\mu_{\tilde{C}}(v_1+v_2) \leq t_1 + t_2$ which implies the result by taking infimum over t_1, t_2.
Lastly, we will show $\mu_{\tilde{C}}$ induces the topology on V. We first show that any open ball for $\mu_{\tilde{C}}$ is open. If $v \in B(v_0, \epsilon)$, then $B(v, \epsilon_1) \subset B(v_0, \epsilon)$ where $\epsilon_1 = \epsilon - \mu_{\tilde{C}}(v - v_0)$. $B(v, \epsilon_1)$ is a neighbourhood of v since it contains $v + (\epsilon_1/2)\tilde{C}$. Next, we show that any open $U \subset V$, WLOG around 0, contains $B(0, \epsilon_0)$ for some $\epsilon_0 > 0$. Choose $\epsilon_0 > 0$ such that $\tilde{C} \subset \epsilon_0^{-1}U$, then $U \supset \epsilon_0\tilde{C} \implies U \supset B(0, \epsilon_0)$. Combining these gives the result. □

Remark. The idea in the proof also allows one to show that any topological vector space is T_2 (a.k.a. Hausdorff). Indeed, suppose $v_1 \neq v_2$ in V, then $v_2 - v_1 \neq 0$, therefore there is some U open around 0 that does not contain $v_2 - v_1$. The continuity of A implies that there is some $\tilde{U}$ open around 0 such that $\tilde{U} + \tilde{U} \subset U$ (hence $\tilde{U} \subset U$ as well). The continuity of M implies that there is some balanced $\hat{U}$ open around 0 containing $\tilde{U}$, by the argument in start of the preceding proof. Then $U_1 = v_1 + \hat{U}, U_2 = v_2 + \hat{U}$, which are open and contain v_1, v_2 respectively, are disjoint. Indeed, suppose $w \in U_1 \cap U_2$, then $w = v_1 + w_1 = v_2 + w_2$ for some $w_1, w_2 \in \hat{U}$, therefore $v_2 - v_1 \in \hat{U} - \hat{U} \subset \tilde{U} + \tilde{U} \subset U$, contradiction.

Definition 2.11. Let V be a normed vector space. We say V is a Banach space if the metric induced by the norm is complete.

Example 2.3. 1. $\mathbb{R}^n, \mathbb{C}^n$ (over $\mathbb{R}, \mathbb{C}$ respectively) are Banach spaces when endowed with any norm (by using the completeness of $\mathbb{R}, \mathbb{C}$ on each component).
2. Given a topological space X, we define (for $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) $B_{\mathbb{F}}(X) = \{f : X \to \mathbb{F} \text{ bounded}\}, C_{\mathbb{F}}(X) = \{f : X \to \mathbb{F} \text{ continuous}\}, C_{\mathbb{F},b}(X) = C_{\mathbb{F}}(X) \cap B_{\mathbb{F}}(X)$. Of course, when X is compact, then $B_{\mathbb{F}}(X) \subset C_{\mathbb{F}}(X)$. Now all three of them are vector spaces over $\mathbb{F}$ with respect to the pointwise operations. Consider the ∞-norm $\|f\|_\infty = \|f\|_{\infty,X} = \sup_X |f|$. We claim that $\|\cdot\|_\infty$ makes $C_{\mathbb{F},b}(X)$ a Banach space. Indeed, $\|\cdot\|_\infty$ is well-defined by boundedness, and is a norm by simple verification. As for completeness, suppose $(f_k)_k \in C_{\mathbb{F},b}(X)$ is Cauchy, then $(f_k(x))_k$ is Cauchy in $\mathbb{F}$ for each k, so we can take a pointwise limit $f(x) = \lim_{k\to\infty} f_k(x)$ by the completeness of $\mathbb{F}$. Then clearly $f_k \to f$ in the ∞-norm, and f is continuous and bounded (by general properties of uniform convergence).
3. Given an open bounded nonempty $U \subset \mathbb{R}^n$ and $m \in \mathbb{N}$, we write $C^m(\bar{U})$ to denote the collection of functions $f : U \to \mathbb{R}$ such that f is m times differentiable in U and all $\partial^\alpha f$ for $|\alpha| < m$ are continuous and bounded on U. Then $\|f\| =$

$\sup_{|\alpha|\le m} \|\partial^\alpha f\|_{\infty,U}$ makes $C^m(\bar{U})$ a Banach space.
4. For $p \in [1, \infty)$, the p-norm

$$\|f\|_p = \left(\int_0^1 |f(x)|^p \, \mathrm{d}x \right)^{1/p}$$

is a norm on $C_{\mathbb{R}}([0,1])$, whose triangle inequality (for $p > 1$; the case $p = 1$ is trivial) follows from

$$\begin{aligned} \|f+g\|_p^p = \int_0^1 |f+g|^{p-1}|f+g| &\le \int_0^1 |f||f+g|^{p-1} + \int_0^1 |g||f+g|^{p-1} \\ &\le \|f\|_p \|f+g\|_p^{p-1} + \|g\|_p \|f+g\|_p^{p-1} \end{aligned}$$

by the integral analogue of Proposition 1.2, which we haven't proved but you can look up. The norm is, however, not complete: Just take

$$f_n(x) = \begin{cases} 0 & \text{for } 0 \le x \le 1/2 - 1/(4k) \\ 4k(x - 1/2 + 1/(4k)) & \text{for } 1/2 - 1/(4k) \le x \le 1/2 \\ 1 & \text{for } 1/2 \le x \le 1 \end{cases}$$

which is Cauchy for the p-norm for every $p \in [1, \infty)$ but does not converge to anything in $C_{\mathbb{R}}([0,1])$. However, we do have a way to complete it in a meaningful sense. The completion is the Banach space $L^p([0,1])$, which is the set of equivalent classes of measurable functions $[0,1] \to \infty$ with finite p-norms (in the sense of Lebesgue integral) with f equivalent to g iff $f = g$ almost everywhere.
5. Let X be the collection of sequences with entries in $\mathbb{F}$. For $p \in [1, \infty]$, define $\ell^p(\mathbb{F})$ as the set of sequences $x = (x_n)_n \in \mathbb{F}$ such that $\|x\|_p < \infty$ where $\|x\|_p = (\sum_n |x_n|^p)^{1/p}$ for $p < \infty$ and $\|x\|_\infty = \sup_k |x_k|$. Then $\ell^p(\mathbb{F})$ is a Banach space under the norm $\|\cdot\|_p$. Its triangle inequality for $p = \infty$ is trivial and that for $p < \infty$ is just Proposition 1.3. As for completeness, suppose $((x_k^n)_k)_n$ is Cauchy in ℓ^∞, then $(x_k^n)_k$ is Cauchy in $\mathbb{F}$ and it's clear that $(x_k^n)_n \to x_k$ for some x_k. Then $(x_k)_k$ is a limit of the Cauchy sequence in ℓ^∞. As for $p < \infty$, we consider the same $(x_k)_k$ but it takes some more effort to argue that it is indeed the limit.

Remark. There certainly exists topological vector spaces that are not normed vector spaces, e.g. $C_{\mathbb{R}}(U)$ where U is a nonempty open set in $\mathbb{R}^n$ with its topology induced by the basis consisting of the translations of the sets

$$U_n = \left\{ f \in C_{\mathbb{F}} \left(\bigcup_{m \ge 1} K_m \right) : \sup_{K_n} |f| < \frac{1}{n} \right\}$$

for each chain of compact sets $K_1 \subsetneq K_2 \subsetneq \cdots$,

2.2 Bounded Linear Maps and Duality

We have defined the structures of normed vector space and topological vector space. It's natural to ask what are the correct morphisms between them.

Definition 2.12. Let V, W be topological vector spaces and $T : V \to W$ is a linear map. T is bounded if it maps bounded sets to bounded sets.

Proposition 2.4. *Suppose V, W are topological vector spaces and V is locally bounded, then a linear map $T : V \to W$ is bounded iff it is continuous.*

Proof. If T is bounded, we shall show that T is continuous at 0. Let U_W be open around $0 \in W$, and U_V open and bounded around $0 \in V$ by local boundedness of V. By boundedness of T, $T(U_V)$ is bounded, so there is some $t > 0$ such that $T(U_V) \subset tU_W$. We then have $f^{-1}(U_W) \supset t^{-1}U_V$ and the latter is an open neighbourhood around 0. So T is continuous at 0. This implies that T is continuous everywhere: Let $w \in W$, U_W an open neighbourhood around w. If $T^{-1}(U_W) = \varnothing$ then $T^{-1}(U_W)$ is certainly open. Otherwise, choose $v \in T^{-1}(U_W)$. $-w + U_W$ is open around 0, so $T^{-1}(-w + U_W)$ is open around $0 \in V$. Hence $T^{-1}(U_W) \supset v + T^{-1}(-w + U_W)$ and the latter is an open neighbourhood around v.
Conversely, suppose T is continuous and B_V is bounded in V. For any open U_W around $0 \in W$, $T^{-1}(U_W)$ is open around $0 \in V$ by continuity. By boundedness of B_V, there is some $t > 0$ such that $B_V \subset tT^{-1}(U_W)$, so $T(B_V) \subset tU_W$. This means that T is bounded. □

Definition 2.13. Suppose V, W are normed vector spaces and $T : V \to W$ is a linear map. Then T is bounded (by the preceding proposition) and its operator norm is defined as

$$\|T\|_{\text{op}} = \inf\{t > 0 : T(B_V(0,1)) \subset B_W(0,t)\}$$

Remark. $\|\cdot\|_{\text{op}}$ can be equivalently defined as $\|T\|_{\text{op}} = \sup\{\|Tv\|_W : \|x\|_V \leq 1\} = \sup\{\|Tv\|_W : \|x\|_V < 1\} = \sup\{\|Tv\|_W : \|x\|_V = 1\}$.

Definition 2.14. Given V, W vector spaces, we denote by $L(V, W)$ the vector space of linear maps $V \to W$. Suppose V, W are topological vector spaces, we write $B(V, W)$ to denote the set of bounded linear maps $V \to W$.

Proposition 2.5. *When V, W are normed vector spaces, $B(V, W)$ is a normed vector space under $\|\cdot\|_{\text{op}}$.*

Proof. It's easy to check that $B(V, W)$ is indeed a vector space. $\|\cdot\|_{\text{op}}$ is clearly well-defined on $B(V, W)$. When $\|T\|_{\text{op}} = 0$, $T(B_V(0,1)) \subset B_W(0,t)$ for any $t > 0$, hence $T(B_V(0,1)) = \{0\}$. So for any $v \in V \setminus \{0\}$, $T(v) = 2\|v\|_V T(v/(2\|v\|_V)) = 0$, therefore $T = 0$. The other properties of a norm follow from the alternative definitions. □

Proposition 2.6. *Let V be a normed vector space and W a Banach space. Then $B(V, W)$ is also a Banach space under $\|\cdot\|_{\text{op}}$.*

Proof. Suppose $(T_n)_n$ is Cauchy in $B(V, W)$. For any $\epsilon > 0, \exists k_0 \in \mathbb{N}, \forall k_1, k_2 \geq k_0, \|T_{k_1} - T_{k_2}\|_{\text{op}} < \epsilon$. So

$$\sup_{k_1, k_2 \geq k_0} \|T_{k_1}(v) - T_{k_2}(v)\|_W \leq \|v\|_V \|T_{k_1} - T_{k_2}\|_{\text{op}} \to 0$$

as $k_0 \to \infty$. Hence for any $v \in V$, $(T_n(v))_n$ is Cauchy under W, so we can take a pointwise limit $T(v) = \lim_{n\to\infty} T_n(v)$. It is easy (but slightly tedious) to check that T is bounded, linear and $\|T_n - T\|_{\text{op}} \to 0$ as $n \to \infty$. □

Definition 2.15. Let V be an NVS over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. $L(V, \mathbb{F})$ is called the algebraic dual of V. $B(V, \mathbb{F})$ is called the topological dual, or simply the dual, of V, denoted V^* with norm $\| \cdot \|_* = \| \cdot \|_{V^*} = \| \cdot \|_{\text{op}}$.

Remark. Our preceding proposition means that V^* is always a Banach space, regardless of whether V is complete.

Definition 2.16. Let V, W be normed vector spaces and $T \in B(V, W)$. The adjoint $T^* : W^* \to V^*$ of T is defined by $T^*\psi = \psi \circ T$.

Proposition 2.7. *$T^* \in B(W^*, V^*)$ and we have $\|T^*\|_{\text{op}} \leq \|T\|_{\text{op}}$*

Remark. The equality, albeit true, is slightly harder to prove. We will show it later with the Hahn-Banach theorem.

Proof. Linearity is obvious. The rest follows from the inequality $|T^*\psi(v)|_{\mathbb{F}} \leq \|\psi\|_{W^*} \|T\|_{\text{op}} \|v\|_V$. □

Definition 2.17. Let V be a normed vector space. Its bidual V^{**} is the dual of V^*.

We immediately have a linear map $\Phi = \Phi_V : V \to V^{**}$ given by $\Phi(v)(\phi) = \phi(v)$, called the bidual embedding of V.

Proposition 2.8. *Given a normed vector space V, its bidual embedding Φ is an element of $B(V, V^{**})$ with $\|\Phi\|_{\text{op}} \leq 1$.*

Remark. Again, we will prove the equality after we get to harness the power of Hahn-Banach. It's also possible to prove that Φ is injective, however it is not always surjective. If it is indeed surjective, we say V is reflexive.

Proof. The linearity is again obvious. To check that it's well-defined, bounded and has operator norm at most 1, we use the inequality $|\phi(v)|_{\mathbb{F}} \leq \|\phi\|_{V^*} \|v\|_V$. □

Example 2.4. 1. For finite dimensional normed vector spaces V, W over $\mathbb{R}$, say with bases $\{v_1, \ldots, v_m\}, \{w_1, \ldots, w_n\}$. Define $v_i^* \in V^*$ via $v_i^* v_{i'} = \delta_{ii'}$. Similarly, we define $w_j^* \in W^*$ via $w_j^* w_{j'} = \delta_{jj'}$. Then $v_1^*, \ldots, v_m^*$ is a basis for V^* and $w_1^*, \ldots, w_n^*$ a basis for W^*. Under these bases, if $T : V \to W$ has matrix A, then T^* would have matrix $A^\top$.
2. Consider $\ell^2 = \ell^2(\mathbb{F})$ for $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Then there are linear endomorphisms, e.g. $(x_1, x_2, \ldots) \mapsto (0, x_1, x_2, \ldots)$, on ℓ^2 that is bounded, injective, but not surjective. There are also bounded linear endomorphisms that is surjective but not injective, e.g. $(x_1, x_2, x_3, \ldots) \mapsto (x_2, x_3, \ldots)$. Neither would've existed in finite dimension. To make things worse (well), we can even build maps that are not unbounded on this space, or infinite dimensional spaces in general – see example sheet.
The dual of ℓ^2, however, behaves quite well. In fact, we have $(\ell^2)^* \cong \ell^2$ as normed vector spaces. This will be an instance of the Reisz representation theorem (or indeed you can do it by hand, as in the next example).
3. In example sheet, you'll show that for $p \in (1, \infty)$ we have $(\ell^p)^* \cong \ell^q$ as normed vector spaces, where $p^{-1} + q^{-1} = 1$. Consequently, $\ell^p \cong (\ell^p)^{**}$ for every such p. In fact, ℓ^p is reflexive.
4. $(\ell^1)^* \cong \ell^\infty$ but $(\ell^\infty)^* \ncong \ell^1$.

3 Hahn-Banach Theorem

3.1 Insights from Finite Dimensional Spaces

Definition 3.1. Suppose V is a vector space. Two norms $\|\cdot\|_1, \|\cdot\|_2$ on V are said to be equivalent, written as $\|\cdot\|_1 \sim \|\cdot\|_2$ if there are $c, c' > 0$ such that $c\|v\|_1 \leq \|v\|_2 \leq c'\|v\|_1$ for all $v \in V$.

Remark. 1. This defines an equivalence relation on the norms on V.
2. Two norms are equivalent if and only if they induce the same topology on V.
3. When two norms are equivalent, the converging sequences, Cauchy sequences and bounded linear operators of the resulting spaces are the same.

Proposition 3.1. *(i) All norms on a finite-dimensional vector space are equivalent.*
(ii) Let $(V, \|\cdot\|_V)$ be a finite dimensional normed vector space, then it is complete and all of its closed bounded subsets are compact. Furthermore, every finite dimensional subspace of a (not necessarily finite dimensional) normed vector space is closed.
(iii) If $(V, \|\cdot\|_V), (W, \|\cdot\|_W)$ are normed vector spaces with V finite dimensional, then any linear map $T : V \to W$ is bounded.
(iv) A normed vector space has its closed unit ball compact if and only if it is finite dimensional.

Proof. (i) Identify (the algebraic structure of) our vector space with $V = \mathbb{F}^n$ with basis $\{e_i\}$. It suffices to show that any norm $\|\cdot\|$ on it is equivalent to the norm $\|v\|_\infty = \sup_i |v_i|$ where $v = \sum_i v_i e_i$ on V. Then

$$\|v\| = \left\|\sum_{i=1}^n v_i e_i\right\| \leq \sum_{i=1}^n |v_i|\|e_i\| \leq \|v\|_\infty \sum_{i=1}^n \|e_i\|$$

which in particular means that $\|\cdot\| : V\mathbb{R}_{\geq 0}$ is continuous with the topology on V induced by $\|\cdot\|_\infty$. Since we are in finite dimension, $S_{\|\cdot\|_\infty}(0,1) = \{v \in V : \|v\|_\infty = 1\}$ is compact, hence $\|\cdot\|$ attains its infimum on it. Let this infimum be C, then clearly $\|\cdot\| \geq C\|\cdot\|_\infty$ which means that $\|\cdot\| \sim \|\cdot\|_\infty$.
(ii) The first statement follows from the special case of $\mathbb{F}^n$ equipped with the standard Euclidean norm by (i). The second statement follows from completeness.
(iii) Suppose $T : V \to W$ is linear with V finite dimensional with basis $\{e_i\}$, then

$$\begin{aligned}\|Tv\|_W = \left\|T\left(\sum_{i=1}^n v_i e_i\right)\right\|_W &= \left\|\sum_{i=1}^n v_i T e_i\right\|_W \\ &\leq \sum_{i=1}^n \|Te_i\|_W |v_i| \leq \|v\|_\infty \sum_{i=1}^n \|Te_i\|_W\end{aligned}$$

which is bounded above by a constant multiple of $\|v\|_V$ by our result in (i).
(iv) Clearly closed unit ball is compact in finite dimensional normed vector spaces. Conversely, suppose $(V, \|\cdot\|_V)$ is a normed vector space with its closed unit ball compact. There is an open cover of it given by

$$\bar{B}(0,1) \subset \bigcup_{v \in \bar{B}(0,1)} B(v, 1/2)$$

By compactness, there are $v_1, \ldots, v_n \in \bar{B}(0,1)$ such that

$$\bar{B}(0,1) \subset \bigcup_{i=1}^{n} B(v_i, 1/2) = \bigcup_{i=1}^{n} (v_i + B(0,1/2)) \subset W + B(0,1/2) \subset W + \frac{1}{2}\bar{B}(0,1)$$

where $W = \text{Span}\{v_1, \ldots, v_n\}$. Iterating the argument and we get $\bar{B}(0,1) \subset W + 2^{-k}\bar{B}(0,1)$ for all k, so

$$\bar{B}(0,1) \subset \bigcap_{k\geq 1} (W + 2^{-k}\bar{B}(0,1)) \subset \bar{W} = W$$

Thus $V \subset W$ which implies that V is finite dimensional. □

Remark. An alternative proof of (iv) can be found in example sheet, where you'll show that the closed unit ball in an infinite dimensional normed vector space is not sequentially compact.

3.2 The Power of Choice

Zorn's lemma, a form of the axiom of choice, plays a key role in proving the Hahn-Banach theorem, which builds a family of bounded linear forms on a normed vector space.

Definition 3.2. A set S is partially ordered if there is a binary relation $\leq$ on S such that:
1. $\forall x \in S, x \leq x$.
2. $\forall x, y, z \in S$ with $x \leq y, y \leq z$, we have $x \leq z$.
3. $\forall x, y \in S$ with $x \leq y, y \leq x$, we have $x = y$.
We say $(S, \leq)$ is a partially ordered set. It is called totally ordered if $\forall x, y \in S$, either $x \leq y$ or $y \geq x$.

Definition 3.3. Let $(S, \leq)$ be a partially ordered set.
For a subset $S' \subset S$, $l \in S$ is called an upper bound of S' if $\forall x \in S', x \leq l$. It is called a least upper bound if $l \leq l'$ for any upper bound l' of S'.
A partially order set $(S, \leq)$ is said to have the least-upper-bound property if any totally ordered subset $S' \subset S$ has a least upper bound.

Definition 3.4. An element m in a partially ordered set $(S, \leq)$ is called maximal if $m \leq x \implies x = m$ for any $x \in S$.

Theorem 3.2 (Zorn's Lemma)**.** *Any nonempty partially ordered set with the least-upper-bound property has a maximal element.*

The proof is omitted, partially because it is equivalent to the axiom of choice and partially because the proof that it is equivalent to the axiom of choice can be found virtually everywhere.
One curious consequence of Zorn's lemma is the existence of a basis for any vector space.

Proposition 3.3. *Let $V \neq \{0\}$ be a vector space and let $S \subset V$ be a linearly independent subset. Then there is a basis B of V such that $S \subset B$. In particular, V has a basis.*

Proof. Let $\mathcal{S}$ be the set of linearly independent subsets of V containing S. It's nonempty since $S \in \mathcal{S}$. Also, $(\mathcal{S}, \subset)$ is a partially ordered set.
Suppose now that $\mathcal{O} \subset \mathcal{S}$ is totally ordered. Consider $\bar{S} = \bigcup_{S' \in \mathcal{O}} S'$. It clearly includes $\bar{S}$. Moreover, if $v_1, \ldots, v_n \in \bar{S}$ (with $v_i \in S_i' \in \mathcal{O}$), then the total order of $\mathcal{O}$ means that there is some $i_0 \in \{1, \ldots, n\}$ such that $S_i' \subset S_{i_0}'$ for all i. Then $v_1, \ldots, v_n \in S_{i_0}'$. Consequently, $v_1, \ldots, v_n$ are linearly independent. Thus $\bar{S}$ is linearly independent, which means that $\bar{S} \in \mathcal{O}$. It then follows easily that $\bar{S}$ is a least upper bound for $\mathcal{O}$, hence $\mathcal{S}$ has the least-upper-bound property.
Theorem 3.2 implies that there exists some maximal $B \in \mathcal{S}$. If $\operatorname{Span} B \neq V$, then we can choose $v_0 \in V \setminus \operatorname{Span} B$ and draw a contradiction to maximality by observing that B is strictly included in $B \cup \{v_0\} \in \mathcal{S}$. So $\operatorname{Span} B = V$. B is also linearly independent as $B \in \mathcal{S}$, therefore it is a basis for V. $\square$

Such existence argument using Theorem 3.2 are known as transfinite induction.

3.3 Hahn-Banach Theorem

Definition 3.5. For a real vector space V, we say $p : V \to \mathbb{R}$ is sublinear if $p(v + w) \leq p(v) + p(w)$ and $p(\lambda v) = \lambda p(v)$ for any $v, w \in V, \lambda > 0$.
For a vector space V over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, we say $p : V \to \mathbb{R}_{\geq 0}$ is symmetrically sublinear if $p(v + w) \leq p(v) + p(w)$ and $p(\lambda v) = |\lambda| p(v)$ for any $v, w \in V, \lambda \in \mathbb{F}$.

Theorem 3.4 (Hahn-Banach)**.** *Let V be a vector space over $\mathbb{R}$ and let $p : V \to \mathbb{R}$ be sublinear. Suppose W is a subspace of V and $f : W \to \mathbb{R}$ is a linear form with $f(w) \leq p(w)$ for all $w \in W$. Then there is some linear form $\tilde{f} : V \to \mathbb{R}$ extending f so that $\tilde{f}(v) \leq p(v)$ for all $v \in V$.*

Proof. We first show the case where $V = W \oplus \mathbb{R}v_0$ (the "codimension 1 case"). Any extension $\tilde{f}$ of f would be determined by its value at v_0. The condition $\tilde{f} \leq p$ is then just asking $\tilde{f}(v_0)$ to satisfy $p(w + av_0) \geq \tilde{f}(w + av_0) = f(w) + a\tilde{f}(v_0)$ for any $w \in W, a \in \mathbb{R}$.
The case $a = 0$ is guaranteed by hypothesis. For $a > 0$, this becomes $\tilde{f}(v_0) \leq p((w/a) + v_0) - f(w/a)$ for any $w \in W$; For $a < 0$, this translates to $\tilde{f}(v_0) \geq -p(-(w/a) - v_0) + f(-w/a)$, again for all $w \in W$.
So it suffices to show that $\sup_{w'' \in W}(-p(w'' - v_0) + f(w'')) \leq \inf_{w' \in W}(p(w' + v_0) - f(w'))$, i.e. $f(w') + f(w'') \leq p(w' + v_0) + p(w'' - v_0)$ for all $w', w'' \in W$, which is true by sublinearity.
How do we get from here to the general case? We will use transfinite induction (i.e. Theorem 3.2). Consider the collection $\mathcal{S}$ consisting of pairs $(\tilde{f}, \tilde{W})$ with $W \leq \tilde{W} \leq V$, $\tilde{f}$ a linear form on $\tilde{W}$ extending f, and $|\tilde{f}| \leq p$ on $\tilde{W}$. This is nonempty and we can define a partial order on it by saying $(f_1, W_1) \leq (f_2, W_2)$ iff $W_1 \leq W_2$ and $f_2|_{W_1} = f_1$. Let's verify the least-upper-bound property. Let $\mathcal{O} \subset \mathcal{S}$ be a totally ordered subset, then our candidate for a least upper bound for it would be $(\bar{f}, \bar{W})$ where $\bar{W} = \bigcup_{(f', W') \in \mathcal{O}} W'$ and $\bar{f}(v) = f'(v)$ for any $(f', W') \in \mathcal{O}$ with $v \in W'$. It is a well-defined element of $\mathcal{S}$ since $\mathcal{O}$ is totally ordered, and is clearly a least upper bound of $\mathcal{O}$.
The hypothesis in Theorem 3.2 is thus satisfied, so we conclude that $\mathcal{S}$ has a maximal element $(\tilde{f}, \tilde{W}) \in \mathcal{S}$. But the codimension 1 case means that necessarily $\tilde{W} = V$ which finishes the proof. $\square$

Theorem 3.5 (Symmetric Hahn-Banach)**.** *Let V be a vector space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and let $p : V \to \mathbb{R}_{\geq 0}$ be symmetrically sublinear. Suppose W is a subspace of V and $f : W \to \mathbb{F}$ is a linear form with $|f(w)| \leq p(w)$ for all $w \in W$. Then there is some linear form $\tilde{f} : V \to \mathbb{F}$ extending f so that $|\tilde{f}(v)| \leq p(v)$ for all $v \in V$.*

Proof. The case $\mathbb{F} = \mathbb{R}$ follows immediately from the preceding theorem.
As for $\mathbb{F} = \mathbb{C}$, observe that we can regard V, W as $\mathbb{R}$-vector spaces: Suppose $\{e_j\}_{j \in J}$ is a basis for V extending a basis $\{e_j\}_{j \in J_W}$ on W, then $V_0 = \operatorname{Span}_{\mathbb{R}}\{e_j\}_{j \in J}, W_0 = \operatorname{Span}_{\mathbb{R}}\{e_j\}_{j \in J_W}$ are $\mathbb{R}$-vector spaces and we can give V, W the $\mathbb{R}$-vector space structures of $V_0 \oplus iV_0$ and $W_0 \oplus iW_0$ respectively.
Now $g = \operatorname{Re} f$ is an $\mathbb{R}$-linear form on W with $|g| \leq p$, so by the theorem for $\mathbb{F} = \mathbb{R}$, there is some $\mathbb{R}$-linear $\tilde{g} : V \to \mathbb{R}$ extending g with $|\tilde{g}| \leq p$. Then $\tilde{f}(v) = \tilde{g}(v) - i\tilde{g}(iv)$ is a $\mathbb{C}$-linear form on V extending f. For any $v \in V$, choose $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ such that $\lambda f(v) \in \mathbb{R}$, then $|\tilde{f}(v)| = |\lambda \tilde{f}(v)| = |\tilde{f}(\lambda v)| = |\tilde{g}(\lambda v)| \leq p(\lambda v) = p(v)$ for any $v \in V$. □

Proposition 3.6. *(i) Let $(V, \|\cdot\|)$ be a normed vector space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Suppose W is a subspace of V and $f \in W^*$. Then there is some $\tilde{f} \in V^*$ extending f with $\|\tilde{f}\|_{V^*} = \|f\|_{W^*}$.*
(ii) If $V \neq \{0\}$ is a normed vector space, then $V^* \neq \{0\}$.
(iii) For any two distinct elements $v_1, v_2 \in V$ in a normed vector space V, there is some $f \in V^*$ with $f(v_1) \neq f(v_2)$.

Proof. (i) Applying the preceding theorem on $p : V \to \mathbb{R}_+$ given by $v \mapsto \|f\|_{W^*}\|v\|$ shows that f extends to $\tilde{f} : V \to \mathbb{F}$ with $|\tilde{f}(v)| \leq \|f\|_{W^*}\|v\|$. This inequality immediately shows that $\tilde{f}$ is bounded and $\|\tilde{f}\|_{V^*} \leq \|f\|_{W^*}$. But of course $\|\tilde{f}\|_{V^*} \geq \|f\|_{W^*}$ since $\tilde{f}$ extends f, so $\|\tilde{f}\|_{V^*} = \|f\|_{W^*}$.
(ii) Choose $v_0 \in V \setminus \{0\}$ and extend (using (i)) to V the linear form $f(\lambda v_0) = \lambda\|v_0\|$ on $\mathbb{F}v_0 \leq V$. Such an extension is sometimes known as a support functional of v_0.
(iii) Just use a support functional of $v_0 = v_1 - v_2$. □

Proposition 3.7. *Let V be a normed vector space, then the natural map $\Phi : V \to V^{**}$ given by $\Phi(v)(f) = f(v)$ is isometric. In particular, $\|\Phi\|_{\mathrm{op}} = 1$.*

Proof. We've already seen that $\|\Phi(v)\|_{V^{**}} \leq \|v\|_V$ for any $v \in V$. Conversely, for $v \in V \setminus \{0\}$, let $f_v \in V^*$ be a support functional of it, then $\|f_v\|_{V^*} = 1$ and $\Phi(v)(f_v) = f_v(v) = \|v\|_V$. This means that $\sup_{\|f\|_{V^*} \leq 1} |\Phi(v)(f)| \geq \|v\|_V$, so $\|\Phi(v)\|_{V^{**}} \geq \|v\|_V$, as desired. □

Proposition 3.8. *Given V, W normed vector spaces over the same field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and a bounded linear $T : V \to W$, we have $\|T^*\|_{\mathrm{op}} = \|T\|_{\mathrm{op}}$.*

Proof. We have seen the direction $\|T^*\|_{\mathrm{op}} \leq \|T\|_{\mathrm{op}}$. Conversely, we choose $v \in V, \|v\| = 1$ with $w = Tv \neq 0$. Let $g_w \in W^*$ be the support functional of w. Then

$$\begin{aligned}\sup_{\|g\|_{W^*} \leq 1} \|T^*g\|_{V^*} &\geq \|T^*g_w\|_{V^*} \geq (T^*g_w)(v) \\ &= g_w(Tv) = g_w(w) = \|w\|_W = \|Tv\|_W\end{aligned}$$

Taking supremum on the right hand side subject to $\|v\|_V = 1$ shows the inequality we desire. □

There is a more geometric form of Theorem 3.4, phrased in terms of separating hyperplanes.

Lemma 3.9. *Given a normed vector space V over $\mathbb{R}$.*
(i) For a linear form f on V and any $\alpha \in \mathbb{R}$, the hyperplane $H = f^{-1}(\{\alpha\})$ is closed iff f is continuous.
(ii) For any convex nonempty neighbourhood C around 0, the Minkowski gauge $p = \mu_C$ is sublinear and $p(v) \leq M\|v\|$ for some $M > 0$.

Proof. Exercise. □

Theorem 3.10 (Geometric Hahn-Banach)**.** *Let V be a normed vector space over $\mathbb{R}$.*
(i) Suppose $A, B \subset V$ are convex and nonempty and A is open. If $A \cap B = \varnothing$, then there is a closed hyperplane weakly separating A, B, i.e. there is some $f \in V^* \setminus \{0\}, \alpha \in \mathbb{R}$ with $\sup_A f \leq \alpha \leq \inf_B f$.
(ii) Suppose $A, B \subset V$ are convex and nonempty with A closed and B compact. If $A \cap B = \varnothing$, then there is a closed hyperplane strictly separating A, B in the sense that there exists $f \in V^* \setminus \{0\}, \alpha_1, \alpha_2 \in \mathbb{R}$ with $\sup_A f \leq \alpha_1 < \alpha_2 \leq \inf_B f$.

Proof. (i) $C_0 = A - B = \{a - b : a \in A, b \in B\}$ is open, convex, nonempty and does not contain 0 by hypothesis. Take $v_0 \in C_0$ and define $C = C_0 - v_0$. Then C is open, convex, nonempty, contains 0 and does not contain $-v_0$. Applying Theorem 3.4 to $p = \mu_C$ and $f : \mathbb{R}v_0 \to \mathbb{R}, \lambda v_0 \mapsto -\lambda$ (so $f \leq p$) shows that there is some $\tilde{f} : V \to \mathbb{R}$ extending f with $\tilde{f} \leq p$. As $p < 1$ on C, we have $\tilde{f}|_C < 1$, so $\tilde{f}|_{C_0} < 0$ which means that $\sup_A \tilde{f} \leq \inf_B \tilde{f}$. Note that $\tilde{f}$ is bounded since we can choose $B(0, \epsilon) \subset C$ and conclude that $\tilde{f}|_{B(0,\epsilon)} < 1 \implies \|\tilde{f}\|_{V^*} \leq \epsilon^{-1}$.
(ii) This time we consider $C_0 = B - A$. It is convex, nonempty, and closed: Suppose $(b_n - a_n)_n$ is a sequence in C_0 converging to $l \in V$. By compactness, we can assume that $b_n \to b \in B$ by passing to a subsequence. So $(a_n)_n$ has to converge to $a = b - l$, but then $a \in A$ as A is closed, therefore $l = b - a \in B - A$. C_0 also does not contain 0, so there is some $\epsilon > 0$ such that $B(0, \epsilon) \cap C_0 = \varnothing$. Applying (i) to $B(0, \epsilon)$ and C_0 gives the result. □

4 Baire Category Theorem

4.1 Rare and Meagre Sets

The moral of Hahn-Banach was to build a lot of bounded linear forms (or separating hyperplanes) and hence show that the dual space is "rather large" by exploiting the (symmetric) sublinearity of norms. Now, we are going to exploit completeness to deduce that Banach spaces are necessarily "big" in a certain sense. This is the content of Baire category theorem which will allow us to construct counter-intuitive objects and also formulate the local-global principle.

Definition 4.1. Suppose X is a topological space.
$B \subset X$ is said to be rare in X if $\bar{B}$ has empty interior. It is said to be meagre in X if it is a countable union of rare sets.
The space X is called meagre if it is meagre in itself.

It is useful to note that the closure of a rare set is still rare.
In the language of Baire, meagre sets are said to be of "first category" and non-meagre sets are of "second category".

Remark. $B \subset X$ being rare is equivalent to B being nowhere dense: If U is open, nonempty, and $U \cap B$ is dense in U, then $\bar{B} \supset \overline{B \cap U} \supset U$, so B cannot be rare. Conversely, if B is not rare, then there is some open $U \subset \bar{B}$ which means that $B \cap U$ is dense in U.

We want a criterion for a space to be non-meagre.

Proposition 4.1. *Let X be a topological space, then the followings are equivalent:*
(i) X is non-meagre.
(ii) For any countable collection $(C_n)_n$ of closed sets covering X, there is some n_0 such that C_{n_0} has nonempty interior.
(iii) For any countable collection $(O_n)_n$ of open dense sets in X, $\bigcap_{n \geq 1} O_n \neq \varnothing$.

Proof. (i) and (ii) are equivalent by definition. (ii) and (iii) are equivalent by taking complements. □

Definition 4.2. A topological space is called a Baire space if countable intersection of open dense sets are dense.

Equivalently, countable union of closed sets with empty interiors has empty interior.
It's clear that every Baire space is in particular non-meagre.

Theorem 4.2 (Baire)**.** *Any complete metric space is a Baire space.*

Proof. Let (X, d) be the complete metric space. Suppose $(O_n)_n$ is a countable collection of open dense set and let $U \subset X$ be open. We need to prove that $U \cap \bigcap_n O_n \neq \varnothing$.
Since O_1 is dense, $O_1 \cap U$ is nonempty, so we can choose $x_1 \in O_1 \cap U$. As $O_1 \cap U$ is open, there is some $r_1 > 0$ such that $B(x_1, r_1) \subset O_1 \cap U$. Replacing O_1 by O_2 and U by $B(x_1, r_1/2)$ gives us some $x_2 \in O_2 \cap B(x_1, r_1/2), r_2 > 0$ with $B(x_2, r_2) \subset O_2 \cap B(x_1, r_1/2)$.
For general n, once $x_1, \ldots, x_k, r_1, \ldots, r_k$ are constructed, $O_{k+1} \cap B(x_k, r_k/2)$ is open and nonempty, so we can choose $x_{k+1} \in O_{k+1} \cap B(x_k, r_k/2), r_{k+1} > 0$ with $B(x_{k+1}, r_{k+1}) \subset O_{k+1} \cap B(x_k, r_k/2)$.
The sequence $(x_n)_n$ thus constructed would be Cauchy, hence converges to some $x \in X$ which necessarily resides in $U \cap \bigcap_n O_n$ (noting that $\overline{B(x_k, r_k/2)} \subset B(x_k, r_k)$ for all k). □

Definition 4.3. A topological space X is called normal if for any pair of disjoint nonempty closed sets $C_1, C_2 \subset X$, there exists disjoint open sets $U_1, U_2 \subset X$ with $C_1 \subset U_1, C_2 \subset U_2$.

Theorem 4.3. *Any compact Hausdorff topological space is normal and is a Baire space.*

Proof. Let X be the topological space.
For normality, pick C_1, C_2 closed in X. For any $x \in C_1, y \in C_2$, we can pick $U^1_{x,y}, U^2_{x,y} \subset X$ disjoint and open such that $x \in U^1_{x,y}, y \in U^2_{x,y}$. Fix $y \in C_2$, C_1 is covered by $\{U^1_{x,y} : x \in C_1\}$. As C_1 is closed in a Hausdorff space, it is necessarily compact, so there exists $x_1, \ldots, x_m \in C_1$ such that C_1 is covered by $\{U^1_{x_i,y} : 1 \leq i \leq m\}$. Denote by $\mathcal{U}^1_y$ the union of this open cover, which in particular is open and contains C_1. Let $\mathcal{U}^2_y = \bigcap_{i=1}^m U^2_{x_i,y}$ which is open and

contains y. As C_2 is closed hence compact, we can choose $y_1, \dots, y_n \in C_2$ such that $C_2 \subset \bigcup_{j=1}^n \mathcal{U}_{y_j}^2 = \mathcal{U}_2$. Let $\mathcal{U}_1 = \bigcap_{j=1}^n \mathcal{U}_{y_j}^1$, then $\mathcal{U}_1$ and $\mathcal{U}_2$ are open, disjoint, and $C_i \subset \mathcal{U}_i$.
To show that it is a Baire space, we pick a countable collection $(O_n)_n$ of open dense sets. Fix U open in X, we shall show that $U \cap \bigcap_n O_n \neq \varnothing$. We again use induction: Choose $x_1 \in O_1 \cap U$. By normality, there exists U_1, U_1' with $x_1 \in U_1$ and $X \setminus (U \cap O_1) \subset U_1'$. In particular, $x_1 \in \bar{U}_1 \subset O_1 \cap U$. We can continue this by replacing O_1 by O_2 and U by U_1, etc., and construct a sequence $(x_n)_n$. Since X is compact and Hausdorff $(x_n)_n$ has an accumulation point which will reside in $U \cap \bigcap_n O_n$. □

Remark. By the end of the proof, what we did is essentially that we replaced the sequence convergence in the proof of Theorem 4.2 by the fact that a nonincreasing sequence $(\bar{U}_k)_{k\geq 1}$ of nonempty compact sets has nonempty intersection.

Example 4.1. If $\mathbb{Q}$ were complete, then it has to be a Baire space. However, singletons in $\mathbb{Q}$ are closed and has empty interior, which gives a contradiction as $\mathbb{Q} = \bigcup_{q\in\mathbb{Q}} \{q\}$.

Corollary 4.4. *There exists $f : [0,1] \to \mathbb{R}$ continuous and nowhere differentiable.*

Proof. Since the space $(C([0,1]), \|\cdot\|_\infty)$ is a complete metric space hence a Baire space, it suffices to show that the subset $\mathcal{D}$ of continuous functions differentiable at at least one point is meagre. Consider $A_n = \{f \in C([0,1]) : \exists x \in [0,1], \forall y \in [0,1] \cap [x - 1/n, x + 1/n], |f(x) - f(y)| \leq n|x - y|\}$. By the general theory on uniform convergence, it's easy to see that each A_n is closed. Each (A_n) also have empty interior: Suppose there is an open ball $B(f_0, \epsilon) \subset A_n$, then there is some piecewise linear $f_1 \in B(f_0, \epsilon/2)$. But then $f_1 + g_\delta \in B(f_0, \epsilon)$ whenever $\delta < \epsilon/2$ where g_δ is a "sawtooth" function bounded by 0 and δ (with slopes $\pm\delta^{-1}$ and lengths δ^2). As $\delta \to 0$, however, $f_1 + g_\delta$ would leave A_n, contradiction.
But it's clear that $\mathcal{D} \subset \bigcup_{n\geq 1} A_n$, so $\mathcal{D}$ is meagre. □

Is there a relation between being meagre and being null in some (sensible) measure on the Borel σ-algebra on the space? Sadly, even in $\mathbb{R}$ (under the Lebesgue measure) there exists some disturbing examples. The set $\mathcal{D} = \bigcap_{n\geq 1} D_n$ where $D_n = \bigcup_{k\geq 1}(q_k - 2^{-n-k}, q_k + 2^{-n-k})$ with $(q_k)_{k\geq 1}$ a enumeration of the rationals has Lebesgue measure 0 but is not meagre.

4.2 Uniform Boundedness Principle

Theorem 4.5 (Uniform Boundedness Principle)**.** *Suppose V, W are Banach spaces.*
(i) Let $(T_i)_{i\in I}$ be an arbitrary collection of bounded linear operators from V to W. Suppose $(T_i)_{i\in I}$ is locally bounded in the sense that for each $v \in V$, $\|T_i v\|_W$ is bounded over $i \in I$. Then $\sup_{i\in I} \|T_i\|_{\text{op}} < \infty$.
(ii) Suppose $(T_n)_{n\geq 1} : V \to W$ is a sequence of bounded linear maps converging pointwise to some linear $T : V \to W$, then T is bounded with $\|T\|_{\text{op}} \leq \liminf_n \|T_n\|_{\text{op}}$.
(iii) $B \subset V$ is bounded iff $f(B)$ is bounded in $\mathbb{R}$ for any $f \in V^*$.
(iv) $B' \subset V^*$ is bounded iff $\Phi(v)(B')$ is bounded in $\mathbb{R}$ for any $v \in V$.

Proof. (i) Consider $C_n = \{v \in V : \sup_{i\in I} \|T_i v\|_W \le n\}$ for $n \ge 1$. Each C_n is clearly closed. The local boundedness condition implies that V is covered by $\{C_n\}$. Theorem 4.2 then tells us that there is some $n \ge 1$ such that C_n has nonempty interior. That is, there is some $v_0 \in V$ and $\epsilon > 0$ such that $\forall i \in I, v \in B(v_0, \epsilon), \|T_i v\|_V \le n$. We are actually done: Any $v \in B(0, \epsilon)$ would have $\|T_i v\|_W \le n + \|T_i v_0\|_W$ for any $i \in I$. So

$$\sup_{i\in I} \|T_i\|_{\text{op}} \le \frac{1}{\epsilon}\left(n_0 + \sup_{i\in I} \|T_i v_0\|_W\right) < \infty$$

(ii) Take the sequence to be the collection of the operators, then it is locally bounded by pointwise convergence, so (i) gives $\sup_n \|T_n\|_{\text{op}} < \infty$ which implies the boundedness of T. As for the bound, given any $\epsilon > 0$ we take unit $v_\epsilon \in V$ so that $\|T\|_{\text{op}} \le \epsilon + \|Tv_\epsilon\|_W$. By pointwise convergence, there is some $N \ge 1$ such that $\|T\|_{\text{op}} \le 2\epsilon + \|T_n v_\epsilon\|_W$ whenever $n \ge N$. This implies that $\|T\|_{\text{op}} \le 2\epsilon + \liminf_n \|T_n\|_{\text{op}}$. Taking $\epsilon \to 0$ gives the result.
(iii) The "only if" direction is trivial. For the "if" direction, suppose $f(B)$ is bounded for any $f \in V^*$. Consider collection of bidual embeddings $\{\Phi(v)\}_{v\in B}$. This collection is locally bounded since given $f \in V^*$ we have $\sup_{v\in B} |\Phi(v)(f)| = \sup_{v\in B} |f(v)| < \infty$. Applying (i) shows that $\sup_{v\in B} \|\Phi(v)\|_{V^*} < \infty$. But we already know that Φ is isometric, so $\sup_{v\in B} \|v\| < \infty$, i.e. B is bounded.
(iv) Again the "only if" direction is clear. Conversely, suppose $\Phi(v)(B')$ is bounded for all $v \in V$. We apply (i) to B' by viewing it as a collection of bounded linear forms. The local boundedness is because any $v \in V$ would have $\sup_{f\in B'} |f(v)| = \sup_{f\in B'} |\Phi(v)(f)| < \infty$. Then we have $\sup_{f\in B'} \|f\|_{V^*} < \infty$, i.e. B' is bounded. □

Theorem 4.6 (Banach-Schauder (1929))**.** *Let V, W be Banach spaces.*
(i) (Open mapping theorem) Any surjective bounded linear map $T : V \to W$ is open (i.e. takes open sets to open sets).
(ii) (Inverse mapping theorem) Any bijective bounded linear map $T : V \to W$ has bounded linear inverse.
(iii) (Closed graph theorem) A linear $T : V \to W$ is bounded iff the graph of it, i.e. $\{(v, T(v)) : v \in V\} \subset V \times W$ is closed.

Remark. The topology on $V \times W$ induced by the product topology is the same as the topology on the vector space $V \times W$ (sometimes also known as $V \oplus W$) induced by the norm $\|(v, w)\|_{V\times W} = \|v\|_V + \|w\|_W$.

Proof. (i) It suffices to show that there is some $\epsilon > 0$ with $B(0, \epsilon) \subset T(B(0, 1))$. As T is surjective,

$$W \subset \bigcup_{n\ge 0} T(B(0, n)) = \bigcup_{n\ge 0} nT(B(0, 1)) \subset \bigcup_{n\ge 0} n\overline{T(B(0, 1))}$$

Since W is a Banach space, there is some n such that $n\overline{T(B(0, 1))}$ has nonempty interior by Theorem 4.2, hence $\overline{T(B(0, 1))}$ has nonempty interior. So there is some $w_0 \in W, \epsilon > 0$ such that $\overline{T(B(0, 1))} \supset w_0 + B(0, 2\epsilon)$. Observe that $T(B(0, 1))$ is balanced and convex, so

$$\overline{T(B(0, 1))} \supset \frac{1}{2}(w_0 + B(0, 2\epsilon)) + \frac{1}{2}(-w_0 + B(0, 2\epsilon)) \supset B(0, 2\epsilon)$$

We claim that $T(B(0,1)) \supset B(0,\epsilon)$. Take $w = w_1 \in B(0,\epsilon)$, then $w_1 \in (1/2)B(0,2\epsilon) \subset (1/2)\overline{T(B(0,1))} = \overline{T(B(0,1/2))}$. So there exists some $v_1 \in B(0,1/2)$ with $\|w_1 - Tv_1\|_W < \epsilon/2$. Then we construct $w_2 = w_1 - Tv_1 \in B(0,\epsilon/2) \subset \overline{T(B(0,1/4))}$ and $v_2 \in B(0,1/4)$ with $\|w_2 - Tv_2\|_W < \epsilon/4$, and so on. So we get sequences $(v_i)_i \in V, (w_i)_i \in W$ such that $\|w_k\|_W < \epsilon/2^{k-1}, \|v_k\|_V < 1/2^k$. By the completeness of V, since $\sum_k \|v_k\|_V < 1 < \infty$, there is some $v \in V$ with $v = \sum_k v_k$. Moreover, $\|v\|_V \leq \sum_k \|v_k\|_V < 1$ and $Tv = w$ which is what we need.
(ii) Immediate from (i).
(iii) The "only if" part is clear. For the "if" direction, observe that the graph G is in fact a closed subspace of $V \times W$ viewed as a Banach space. Consider the bounded linear map $\pi : G \to V$ by projection onto the first coordinate. Then π is bijective, so by (ii) we know that π^{-1} is bounded. That is, for any $v \in V$ we have $\|v\|_V + \|Tv\|_W \leq C\|v\|_V$ for some constant C. This gives boundedness. □

Remark. The closed graph theorem is often useful in proofs of continuity of a certain linear operator between Banach spaces: Doing it directly for $T : V \to W$ would require one to show that any $v_n \to v$ has $Tv_n \to Tv$. With the help of the theorem, we can just show that if both $v_n \to v$ and $Tv_n \to w$ for some $w \in W$, then $Tv = w$.

5 Topology of $C(K)$

Our main object of interest now is the space $C(K) = \{f : K \to \mathbb{F} \text{ continuous}\}$ equipped with the norm $\|\cdot\|_\infty$ where K is often taken as a compact Hausdorff space (for the norm to be defined). This appears very frequently in contexts of functional analysis and PDEs.

5.1 Tietze Extension Theorem

Definition 5.1. A topological space X is T_0 if for any distinct $x, y \in X$, either there is an open U with $x \in U, y \notin U$ or there is an open V with $y \in V, x \notin V$.
It is T_1 if for any distinct $x, y \in X$, there are open sets U, V such that $x \in U, y \in V, x \notin V, y \notin U$.
It is T_2 (Hausdorff) if for any distinct $x, y \in X$, there are disjoint open sets U, V with $x \in U, y \in V$.
It is normal if for any disjoint closed $C_1, C_2 \subset X$, there are disjoint open sets U, V with $C_1 \subset U, C_2 \subset V$.

Example 5.1. 1. Any metric space is normal.
2. Any compact Hausdorff space is normal.

Lemma 5.1 (Urysohn Lemma)**.** *A topological space X is normal iff for any disjoint closed nonempty $C_1, C_2 \subset X$, there is a continuous $f : X \to [0,1]$ with $f|_{C_1} \equiv 0, f|_{C_2} \equiv 1$.*

Proof. The "if" part is clear.
For the "only if" part, given any U_0, U_1 open with $\varnothing \neq U_0 \subset \bar{U}_0 \subset U_1 \subsetneq X$, we can find $U_{1/2}$ open such that $U_0 \subset \bar{U}_0 \subset U_{1/2} \subset \bar{U}_{1/2} \subset U_1$ by normality of X (used on the disjoint closed $\bar{U}_0, U_1^c$). We want to extend this to an inductive construction. Let $D_n = \{k/2^n : 0 \leq k \leq 2^n\}$. Our intended induction statement

is the following: Given $\varnothing \neq U_0 \subset \bar{U}_0 \subset U_1 \subsetneq X$, we can construct open sets $(U_r)_{r\in D_n}$ so that $\bar{U}_{r_1} \subset U_{r_2}$ for all $r_1, r_2 \in D_n$ with $r_1 < r_2$. The base case is immediate. The induction process is analogous to what we did at the beginning when we construct $U_{1/2}$. Let $r \in D_{n+1} \setminus D_n$, then there is some odd $k = 2k_0+1$ such that $r = k/2^{n+1}$. Then we construct U_r with $\bar{U}_{k_0/2^n} \subset U_r \subset \bar{U}_r \subset U_{(k_0+1)/2^n}$ by applying normality on $\bar{U}_{k_0/2^n}$ and $U^c_{(k_0+1)/2^n}$ which are disjoint and closed by the induction hypothesis. There are only finitely such r, so we can fill them this one-by-one this way.
Hence we can construct U_r for all $r \in D = \bigcup_n D_n$ with $\bar{U}_{r_1} \subset U_{r_2}$ for all $r_1, r_2 \in D$ with $r_1 < r_2$. We will use them to construct f. Given disjoint closed nonempty C_1, C_2, there is some U_0 such that $C_1 \subset U_0 \subset \bar{U}_0 \subset U_1 = C_2^c \subsetneq X$ by normality. Let $(U_r)_{r\in D}$ be as above and construct

$$f(x) = \begin{cases} 0 & \text{if } x \in C_1 \\ \inf\{r \in D : x \in U_r\} & \text{if } x \notin C_1 \cup C_2 \\ 1 & \text{if } x \in C_2 \end{cases}$$

To see f is continuous, it suffices to show that $f^{-1}((a,1])$ and $f^{-1}([0,b))$ are open for any $a \in [0,1)$ and $b \in (0,1]$. Suppose $a \in [0,1)$ and $x \in f^{-1}((a,1])$ (i.e. $f(x) > a$). Since D is dense in $[0,1]$, there is some $r, r' \in D$ such that $f(x) > r' > r > a$. This means that $x \notin U_{r'} \supset \bar{U}_r$, so $x \in (\bar{U}_r)^c$ which is open. Also, if $y \notin \bar{U}_r \supset U_r$, then $f(y) > a$, i.e. $y \in f^{-1}((a,1])$. So in fact $(\bar{U}_r)^c \subset f^{-1}((a,1])$. Hence $f^{-1}((a,1])$ is open. A similar argument would work for $f^{-1}([0,b))$. ☐

Corollary 5.2. *Suppose X is normal and T_1, then $C(X)$ separates points of X (i.e. for any $x, y \in X$ distinct, there is some continuous $f : X \to \mathbb{F}$ with $f(x) \neq f(y)$).*

Theorem 5.3 (Tietze Extension Theorem)**.** *Let X be a normal topological space and $C \subset X$ closed and nonempty. Suppose $f : C \to \mathbb{C}$ (resp. $\mathbb{R}$) is continuous and bounded, then there is some continuous $\tilde{f} : X \to \mathbb{C}$ (resp. $\mathbb{R}$) extending f with $\sup_X \max\{|\operatorname{Re}(\tilde{f})|, |\operatorname{Im}(\tilde{f})|\} = \sup_C \max\{|\operatorname{Re}(f)|, |\operatorname{Im}(f)|\}$.*

Proof. By rescaling and translation, it suffices to show that whenever the image of f is contained in $[0,1]$, we can always extend it to $\tilde{f}$ on X whose image is also contained in $[0,1]$. Lemma 5.1 tells us that there is some continuous $g_1 : X \to [0,1/3]$ continuous such that $g_1|_{f^{-1}([0,1/3])} = 0, g_1|_{f^{-1}([2/3,1])} = 1/3$ so that $f_2 = f_1 - g_1|_C$ has image in $[0,2/3]$ (where $f_1 = f$). Inductively, given $f_k : C \to [0,(2/3)^{k-1}]$ continuous, we apply Lemma 5.1 to conclude the existence of a continuous $g_k : X \to [0,(1/3)(2/3)^{k-1}]$ such that $g_k|_{C_1} = 0, g_k|_{C_2} = (1/3)(2/3)^{k-1}$ where

$$C_1 = f_k^{-1}\left(\left[0, \frac{1}{3}\left(\frac{2}{3}\right)^{k-1}\right]\right), C_2 = f_k^{-1}\left(\left[\frac{2}{3}\left(\frac{2}{3}\right)^{k-1}, \left(\frac{2}{3}\right)^{k-1}\right]\right)$$

which allows us to construct $f_{k+1} = f_k - g_k$ with image in $[0,(2/3)^k]$.
Consequently,

$$\sup_C \left| f - \sum_{k=1}^n g_k \right| = \sup_C |f_{n+1}| \leq \left(\frac{2}{3}\right)^n \to 0$$

Also, $\sum_n g_n$ converges since we have the uniform size estimate

$$\sum_{n=1}^{\infty} \sup_X |g_n| \leq \sum_{n=1}^{\infty} \frac{1}{3} \left(\frac{2}{3}\right)^{k-1} = 1$$

The same estimate also tells us that its limit $\tilde{f}$ has $\sup_X |\tilde{f}| \leq 1$. Clearly $\tilde{f} \geq 0$, so the proof is completed. □

5.2 Arzelà-Ascoli Theorem

Recall that metric spaces with all continuous functions to $\mathbb{R}$ bounded are compact. This inspires us to study compactness in general in terms of some kind of boundedness.

Definition 5.2. Let X be a metric space. A subset $Y \subset X$ is totally bounded if for all $\epsilon > 0$ there is some finite collection $N = \{x_1, \dots, x_n\} \subset X$ so that $Y \subset \bigcup_i B(x_i, \epsilon)$. Such an N is called an ϵ-net of Y.

Remark. 1. Total boundedness implies boundedness.
2. If $Z \subset Y \subset X$ and Y is totally bounded, then Z too is totally bounded.
3. The definition does not change if we require $N \subset Y$ instead.

Proposition 5.4. *Let X be a complete metric space and $Y \subset X$. Then Y is relatively compact (i.e. $\bar{Y}$ is compact) iff Y is totally bounded.*

Proof. We shall show that both statements are equivalent to every sequence in Y having a Cauchy subsequence.
$\bar{Y}$ is compact iff any sequence in Y has a Cauchy subsequence:
Indeed, suppose we have this, then for any $(z_n)_n \in \bar{Y}$, there is a sequence $(y_n)_n \in Y$ such that $d(z_n, y_n) \leq 1/n$. By passing to a subsequence, we can assume that $(y_n)_n$ is Cauchy, hence by completeness converges to some $z \in \bar{Y}$ which has to be the limit of $(z_n)_n$. Hence $\bar{Y}$ is sequentially compact, therefore compact. The other direction is clear.
Y is totally bounded iff any sequence in Y has a Cauchy subsequence:
If Y is not totally bounded, then there is some ϵ so that there is no ϵ-net. With this, one can inductively build a sequence $(y_n)_n \in Y$ such that $d(y_n, y_m) > \epsilon$ for any $m \neq n$. $(y_n)_n$ then cannot have a Cauchy subsequence.
Conversely, suppose Y is totally bounded and $(y_n)_n \in Y$ is a sequence. Take $\epsilon_k = 1/k$. There is an ϵ_1-net $\{x_i^1\}_i \in X$ such that $Y \subset \bigcup_i B(x_i^1, \epsilon_1)$. As $\{x_i^1\}_i$ is a finite collection, there is some $i = i_1$ such that $B(x_{i_1}^1, \epsilon_1)$ contains infinitely many terms of the sequence $(y_n)_n$. Now $Y \cap B(x_{i_1}^1, \epsilon_1)$ is also totally bounded, so there is an ϵ_2-net $\{x_i^2\}_i \in X$ such that $Y \cap B(x_{i_1}^1, \epsilon_1) \subset \bigcup_i B(x_i^2, \epsilon_2)$. By the same argument, we find i_2 such that $B(x_{i_2}^2, \epsilon_2)$ contains infinitely many terms of the sequence $(y_n)_n$. Inductively, we can find $(x_{i_k}^k)_k$ with $Y \cap B(x_{i_k}^k, \epsilon_k)$ contains infinitely many terms of $(y_n)_n$. By passing to a subsequence, we can make $y_n \in Y \cap B(x_{i_k}^k, \epsilon_k)$ for all $n \geq k \geq 1$. Then $(y_n)_n$ is Cauchy. □

Definition 5.3. Let K be a compact Hausdorff space, then a subset $\mathcal{F} \subset C(K)$ is equibounded at $x \in K$ if $\sup_{f \in \mathcal{F}} |f(x)| < \infty$. We say it is equibounded on K if it is equibounded at every $x \in K$.
$\mathcal{F}$ is called equicontinuous at $x \in K$ if for all $\epsilon > 0$, there is some U open in K and $x \in U$ such that $\sup_{y \in U} \sup_{f \in \mathcal{F}} |f(x) - f(y)| < \epsilon$. It is equicontinuous on K if it is equicontinuous at every $x \in K$.

Remark. We say $\mathcal{F}$ is uniformly equibounded on K if $\sup_{x\in K}\sup_{f\in\mathcal{F}}|f(x)| < \infty$. One can also define uniform equicontinuity if K has a metric structure, which is not a notion we care about at the moment.

Example 5.2. Any finite collection $\mathcal{F} \subset C(K)$ is equibounded and equicontinuous on K.

Theorem 5.5 (Arzelà-Ascoli)**.** *Suppose K is compact and Hausdorff. A collection $\mathcal{F} \subset C(K)$ is relatively compact iff it is equibounded and equicontinuous on K.*

Proof. Recall that $C(K)$ is a Banach space, hence complete as a metric space, so we can replace "relatively compact" by "totally bounded" by the preceding proposition.
If $\mathcal{F}$ is totally bounded, then it is bounded, thus equibounded on K. For any $\epsilon > 0$, there is an $\epsilon/3$-net such that there exists $f_1, \ldots, f_n \in C(K)$ such that $\mathcal{F} \subset \bigcup_i B(f_i, \epsilon/3)$. Since each f_i is continuous, there are open $U_i \ni x$ such that $f_i(U_i) \subset B(f_i(x), \epsilon/3)$. Take $U = \bigcap_i U_i$, then $\sup_{y\in U}\sup_{f\in\mathcal{F}}|f(x) - f(y)| < \epsilon$, so $\mathcal{F}$ is equicontinuous.
Suppose now that $\mathcal{F}$ is equibounded and equicontinuous. Fix $\epsilon > 0$. We shall construct a ϵ-net of $\mathcal{F}$. For all $x \in K$, there is some $U_x \ni x$ open such that every $f \in \mathcal{F}$ has $f(U_x) \subset B(f(x), \epsilon/3)$ by equicontinuity. By compactness of K, there is a finite collection $\{x_1, \ldots, x_n\} \subset K$ with $K = \bigcup_{i=1}^n U_{x_i}$. The set $A = \{(f(x_1), \ldots, f(x_n)) : f \in \mathcal{F}\} \subset \mathbb{F}^n$ is bounded (taking the norm on $\mathbb{F}^n$ to be $\|\cdot\|_\infty$) due to equiboundedness, hence is totally bounded by Bolzano-Weierstrass. This means that there is a collection $N = \{f_1, \ldots, f_m\} \subset \mathcal{F}$ such that

$$A \subset \bigcup_{j=1}^m B\left((f_j(x_1), \ldots, f_j(x_n)), \frac{\epsilon}{3}\right)$$

N is in fact an ϵ-net of $\mathcal{F}$, so we are done. □

Remark. If K has the structure of a (compact) metric space, one can give a sequential proof of the theorem: First show that K is separable, i.e. contains a countable dense subset $S \subset K$. Then, show that if $\mathcal{F} \subset C(K)$ is equibounded and equicontinuous and $(f_n)_{n\geq 1} \in \mathcal{F}$ then one can use a diagonal argument to construct a subsequence of $(f_n)_{n\geq 1}$ converging pointwise (say to some function $f : K \to \mathbb{F}$) on S. One can use equicontinuity to deduce that f is continuous on S, hence can be extended continuously to the entirety of K.

Recall a cool theorem on the theory of ODEs.

Theorem 5.6 (Picard-Lindelöf)**.** *Suppose we have a continuous function $\phi : \mathbb{F}\times\mathbb{F} \to \mathbb{R}^n$ locally Lipschitz in the second variable, then the initial value problem $f'(t) = \phi(t, f(t))$ subject to $f(t_0) = y_0$ has a unique solution on $[t_0 - \epsilon, t_0 + \epsilon]$ for some $\epsilon > 0$.*

The modern proof of this is via the Banach fixed point theorem (aka contraction mapping theorem), i.e. by arguing that the map $\Phi : C([t_0 - \epsilon, t_0 + \epsilon]) \to C([t_0 - \epsilon, t_0 + \epsilon])$ given by

$$\Phi(f)(t) = y_0 + \int_{t_0}^{t} \phi(s, f(s))\, \mathrm{d}s$$

with some cleverly chosen ϵ.

Example 5.3. 1. If $\phi(t,y) = y$, then there is in fact a unique global solution to the problem given by $f(t) = y_0 e^{t-t_0}$.
2. If $\phi(t,y) = y^2, t_0 = 0, y_0 > 0$, then we get a unique solution on $(-\infty, y_0^{-1})$ given by $f(t) = 1/(y_0^{-1} - t)$.
3. $\phi(t,y) = \sqrt{|y|}$ is, alas, not Lipschitz at $y = 0$. Indeed, under the conditions $t_0 = y_0 = 0$, we have two distinct solutions given by $f \equiv 0$ and

$$f(t) = \begin{cases} t^2/4 & \text{if } t > 0 \\ -t^2/4 & \text{if } t \leq 0 \end{cases}$$

However, we still have existence of solution!

Theorem 5.7 (Cauchy-Peano)**.** *Suppose we are in the setting of Picard-Lindelöf but without the assumption of ϕ being locally Lipschitz in the second variable, then there still exists a solution to the initial value problem which is in general not unique.*

Sketch of Proof. Take the sequence with $f_0 \equiv y_0$ and

$$f_{k+1}(t) = y_0 + \int_{t_0}^{t} \phi(s, f_k(s))\,\mathrm{d}s$$

One can show that for small enough ϵ, $(f_k)_{k\geq 0}$ is equibounded and equicontinuous on $[t_0 - \epsilon, t_0 + \epsilon]$. Picking a converging subsequence by Theorem 5.5 finishes the proof. □

5.3 Stone-Weierstrass Theorem

Definition 5.4. V is an algebra over a field $\mathbb{F}$ if V is a vector space over $\mathbb{F}$ with an additional product operation $P : V \times V \to V$ (sometimes written as $P(v_1, v_2) = v_1 \times v_2$) such that:
(i) $\forall w, v_1, v_2 \in V, (v_1 + v_2) \times w = v_1 \times w + v_2 \times w, w \times (v_1 + v_2) = w \times v_1 + w \times v_2$.
(ii) $\forall \lambda_1, \lambda_2 \in \mathbb{F}, v_1, v_2 \in V, (\lambda_1 v_1) \times (\lambda_2 v_2) = (\lambda_1 \lambda_2)(v_1 \times v_2)$.
V is a commutative algebra if $\forall v, w \in V, P(v,w) = P(w,v)$. It is a unitary algebra if there is some $1 \in V$ with $P(1,v) = P(v,1) = v$.
We say V is a normed algebra if it is a NVS and $\forall v_1, v_2 \in V, \|v_1 \times v_2\|_V \leq \|v_1\|\|v_2\|$ (in particular, P is bilinear and continuous). It is a Banach algebra if it is also complete.

Example 5.4. 1. Suppose X is a compact topological space, then $C_{\mathbb{R}}(X)$ is a unitary commutative Banach algebra under $P(f_1, f_2) = f_1 f_2$ and $\|\cdot\|_\infty$.
2. For a vector space V, $L(V,V)$ is a unitary algebra for $P(T_1, T_2) = T_1 \circ T_2$.
3. If V is a NVS, then $B(V,V)$ is a unitary normed algebra for the same operation $P(T_1, T_2) = T_1 \circ T_2$.
4. If V is a Banach space, then $B(V,V)$ is a unitary Banach algebra, again for $P(T_1, T_2) = T_1 \circ T_2$.

Theorem 5.8 (Stone-Weierstrass)**.** *Let K be a compact Hausdorff space and suppose $\mathcal{A} \subset C_{\mathbb{R}}(K)$ is a subalgebra that separates points (in the sense that for all $x, y \in K$ distinct, there is some $f \in \mathcal{A}$ with $f(x) \neq f(y)$). Then $\bar{\mathcal{A}}$ either equals $C_{\mathbb{R}}(K)$ or has the form $\{f \in C_{\mathbb{R}}(K) : f(x_0) = 0\}$ for some $x_0 \in K$.*

Proof. First, we shall show that if $\mathcal{L} \subset C_{\mathbb{R}}(K)$ is closed under min, max, then any $g \in C_{\mathbb{R}}(K)$ with $\forall x, y \in K, \forall \epsilon > 0, \exists f \in \mathcal{L}, |f(x)-g(x)| < \epsilon, |f(y)-g(y)| < \epsilon$ has $g \in \bar{\mathcal{L}}$. Indeed, let $g \in C_{\mathbb{R}}(K)$ be such a function and pick $\epsilon > 0$. For $x, y \in K$, let $f_{x,y} \in \mathcal{L}$ be such that $|f_{x,y}(x) - g(x)| < \epsilon, |f_{x,y}(y) - g(y)| < \epsilon$. $f_{x,y} - g$ is continuous for all x, y. Choose $U_{x,y}$ open around x and $V_{x,y}$ open around y such that $\sup_{U_{x,y}} |f_{x,y} - g| \leq \epsilon, \sup_{V_{x,y}} |f_{x,y} - g| \leq \epsilon$. We have $K = \bigcup_{y \in K} V_{x,y}$, so by compactness there is some $y_1, \ldots, y_n \in K$ such that $K = \bigcup_{i=1}^{n} V_{x,y_i}$. $\tilde{U}_x = \bigcap_{i=1}^{n} U_{x,y_i}$ is open around x. Consider $f_x = \min\{f_{x,y_1}, \ldots, f_{x,y_n}\}$, then $f_x \leq g + \epsilon$ on K and $f_x \geq g - \epsilon$ on $\tilde{U}_x$. Again by compactness we can choose $x_1, \ldots, x_m$ such that $K = \bigcup_{i=1}^{m} \tilde{U}_{x_i}$. Then $f = \max\{f_{x_1}, \ldots, f_{x_m}\}$ has $g - \epsilon \leq f \leq g + \epsilon$ and $f \in \mathcal{L}$, as desired.
Next, we show that if $\mathcal{A}$ is a closed subalgebra of $C_{\mathbb{R}}(K)$, then it is closed under min, max. Indeed, it suffices to show that $\mathcal{A}$ is closed under taking absolute value. Take $f \in \mathcal{A}$ with WLOG $|f| \leq 1$. For $\epsilon > 0$, consider $\phi_\epsilon(r) = \sqrt{\epsilon^2 + r}$. Clearly $|\phi_\epsilon(r^2) - r| \leq \epsilon$ for $r \in [0,1]$. Also, for fixed ϵ, ϕ_ϵ has a Taylor series around $1/2$ that uniformly converges to ϕ_ϵ on $[0,1]$, i.e.

$$\phi_\epsilon(r) = G_N^\epsilon(r) + R_N^\epsilon(r), G_N^\epsilon(r) = \sum_{k=0}^{N} a_k^\epsilon \left(r - \frac{1}{2}\right)^k$$

with $|R_N^\epsilon|_\infty \to 0$ as $N \to \infty$. Note that $G_N^\epsilon(0) \to \phi_\epsilon(0) = \epsilon$ as $N \to \infty$, so we have the desired approximation

$$|f| = (G_N^\epsilon(f^2) - G_N^\epsilon(0)) + G_N^\epsilon(0) + R_N^\epsilon(f^2) + (|f| - \phi_\epsilon(f^2))$$

We proceed to conclude the theorem. Suppose every $x \in K$ admits some $f \in \bar{\mathcal{A}}$ with $f(x) \neq 0$. Then for any $x, y \in K, x \neq y$, there are some $f_x, f_y, f_{x,y} \in \bar{\mathcal{A}}$ such that $f_x(x) \neq 0, f_y(y) \neq 0, f_{x,y}(x) \neq f_{x,y}(y)$. Choose $\alpha, \beta \in \mathbb{R}$ with $\tilde{f} = f_x + \alpha f_y + \beta f_{x,y} \in \bar{\mathcal{A}}$ and $\tilde{f}(x) \neq 0, \tilde{f}(y) \neq 0, \tilde{f}(x) \neq \tilde{f}(y)$. So $(\tilde{f}(x), \tilde{f}(y))$ and $(\tilde{f}(x)^2, \tilde{f}(y)^2)$ are linearly independent over $\mathbb{R}$. This then means that $\bar{\mathcal{A}} = C_{\mathbb{R}}(K)$ by earlier discussion.
Otherwise, there is some $x_0 \in K$ such that $\bar{\mathcal{A}} \subset \{f \in C_{\mathbb{R}}(K) : f(x_0) = 0\}$. Then by the previous case we know that $\overline{\mathcal{A} \oplus \mathbb{R}} = C_{\mathbb{R}}(K)$. But then we must have $\bar{\mathcal{A}} = \{f \in C_{\mathbb{R}}(K) : f(x_0) = 0\}$, as desired. ☐

Theorem 5.9 (Stone-Weierstrass, Complex Version). *Suppose K is compact and Hausdorff and $\mathcal{A}$ is a subalgebra of $C_{\mathbb{C}}(K)$ that separates points and is closed under conjugation, then either $\bar{\mathcal{A}} = C_{\mathbb{C}}(K)$ or there is some $x_0 \in K$ with $\bar{A} = \{f \in C_{\mathbb{C}}(K) : f(x_0) = 0\}$.*

Proof. Since $\mathcal{A}$ is closed under conjugation, any $f \in \mathcal{A}$ has $\operatorname{Re} f, \operatorname{Im} f \in \mathcal{A}$. $\mathcal{A}_{\mathbb{R}} = \{\operatorname{Re} f : f \in \mathcal{A}\}$ is then a subalgebra of $\mathcal{A}$ that separates points, so the preceding theorem applies and gives the result. ☐

Example 5.5. 1. The set of real (resp. complex) polynomials in $[0,1]$ is dense in $(C_{\mathbb{R}}([0,1]), \|\cdot\|_\infty)$ (resp. $(C_{\mathbb{C}}([0,1]), \|\cdot\|_\infty)$). One can prove this particular case alternatively by "kernel analysis" (example sheet). There is another proof by Bernstein that involves the construction of Bernstein polynomials

$$p_n(x) = \sum_{k=1}^{n} b_{k,n}(x) f(1/n), b_{k,n}(x) = \binom{n}{k} x^k (1-x)^{n-k}$$

which approximate f as $n \to \infty$.
2. The set of linear combinations of positive powers of the (real) exponential function is dense in $C_{\mathbb{R}}([0,1])$.
3. $\mathrm{Span}_{\mathbb{C}}\{x \mapsto e^{ikx} : k \in \mathbb{Z}\}$ is dense in $C_{\mathbb{C}}(S^1)$ (with the identification $S^1 \cong \mathbb{R}/(2\pi\mathbb{Z})$). Consequently, any $f \in C_{\mathbb{C}}(S^1)$ has

$$\int_0^{2\pi} f(x)e^{ikx}\,\mathrm{d}x \to 0$$

as $|k| \to \infty$. This is known as the Riemann-Lebesgue lemma.
4. $\mathrm{Span}_{\mathbb{F}}\{(x_1, x_2) \mapsto f_1(x_1)f_2(x_2) : f_1, f_2 \in C_{\mathbb{F}}([0,1])\}$ is dense in $C_{\mathbb{F}}([0,1]^2)$ (for $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$). This is can be used to prove Fubini's theorem for continuous functions.

6 Hilbert Spaces

The idea of Hilbert spaces is to generalise Euclidean geometry to infinite dimensional spaces with nice enough structures.

6.1 Euclidean Spaces and Orthogonality

Definition 6.1. Let V be a vector space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. An inner product on V is a map $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{F}$ such that:
(i) $\forall v_1, v_2 \in V, \langle v_1, v_2\rangle = \overline{\langle v_2, v_1\rangle}$.
(ii) $\forall v_1, v_2, w \in V, \lambda_1, \lambda_2 \in \mathbb{F}, \langle \lambda_1 v_1 + \lambda_2 v_2, w\rangle = \lambda_1\langle v_1, w\rangle + \lambda_2\langle v_2, w\rangle$.
(iii) $\forall v \in V, \langle v, v\rangle \in \mathbb{R}_{\geq 0}$ and $\langle v, v\rangle = 0 \iff v = 0$.
When V is equipped with such an inner product, we say V is an inner product space.

Remark. In the real case, an inner product is a positive definite symmetric bilinear form; In the complex case, an inner product is a positive definite Hermitian sesquilinear form.

Proposition 6.1. *(i) (Cauchy-Schwartz Inequality) Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space, then $|\langle v_1, v_2\rangle|^2 \leq \langle v_1, v_1\rangle\langle v_2, v_2\rangle$ for all $v_1, v_2 \in V$, with equality iff $v_1 = \lambda v_2$ for some $\lambda \in \mathbb{F}$. Consequently, $\|v\| = \sqrt{\langle v, v\rangle}$ is a norm on V. An inner product space normed as such is called a Euclidean space.*
(ii) (Polarisation Identity) Let $(V, \|\cdot\|, \langle\cdot,\cdot\rangle)$ be a Euclidean space, then

$$\begin{cases} \langle v, w\rangle = (\|v+w\|^2 - \|v-w\|^2)/4 & \text{if } \mathbb{F} = \mathbb{R} \\ \langle v, w\rangle = (\|v+w\|^2 - \|v-w\|^2 + i\|v+iw\|^2 - i\|v-iw\|^2)/4 & \text{if } \mathbb{F} = \mathbb{C} \end{cases}$$

(iii) (Jordan-von Neumann Theorem) A normed vector space comes from a Euclidean space iff its norm satisfies the parallelogram law $\|v+w\|^2 + \|v-w\|^2 = 2\|v\|^2 + 2\|w\|^2$.

Proof. (i) This is obvious when $\langle v_1, v_2\rangle = 0$. Otherwise, given $v_1, v_2 \in V$, there is some $\alpha \in \mathbb{F}$ with $|\alpha| = 1$ such that $\langle \alpha v_1, v_2\rangle \in \mathbb{R}_{\geq 0}$. WLOG replace v_1 by αv_1. Then $0 \leq \langle v_1 + tv_2, v_1 + tv_2\rangle$ is a real quadratic polynomial in $\mathbb{R}$ that is nonnegative, hence has nonpositive discriminant. This is just the inequality we wanted.

(ii) Simple expansion.
(iii) The parallelogram law is clearly satisfied by any Euclidean space. Conversely, suppose the parallelogram law is true, then one can check (via a sequence of annoying calculations and exploiting continuity) that the form given by the polarisation identity is indeed an inner product inducing the original norm. □

Definition 6.2. Let V be a Euclidean space.
Vectors $v, w \in V$ are orthogonal, denoted $v \perp w$ if $\langle v, w \rangle = 0$. A subset $S \in V$ is orthogonal if $v \perp w$ for any distinct $v, w \in S$. It is orthonormal if in addition $\langle v, v \rangle = 1$ for any $v \in S$.
The orthogonal complement of $S \subset V$ is $S^{\perp} = \{v \in V : \forall w \in S, v \perp w\}$.

Proposition 6.2. *(i) $S^{\perp}$ is a subspace of V for any $S \subset V$, and $S^{\perp} = (\overline{\operatorname{Span} S})^{\perp}$.*
(ii) There exists a maximal orthonormal family $\tilde{S}$ (i.e. not contained in any strictly bigger orthonormal sets in V) extending any given orthonormal family $S \subset V$.

Proof. (i) $S^{\perp}$ is a subspace simply because of the linearity (rep. sesquilinearity) of inner products. Again by (sesqui)linearity we have $S^{\perp} = (\operatorname{Span} S)^{\perp}$. And $(\operatorname{Span} S)^{\perp} = (\overline{\operatorname{Span} S})^{\perp}$ by continuity of inner products.
(ii) Cast Theorem 3.2 on $\{S' \subset V$ orthonormal and contains $S\}$. □

We might not just be satisfied with mere existence of orthonormal families.

Proposition 6.3 (Gram-Schmidt Orthonormalisation)**.** *For a Euclidean space V and a countable linearly independent family $(v_n)_{n=1}^{N}$ for $N \in \mathbb{N} \cup \{\infty\}$, there is an orthonormal family $(e_n)_{n=1}^{N}$ given exclusively in terms of $(v_n)_{n=1}^{N}$ such that* $\operatorname{Span}((e_n)_{n=1}^{k}) = \operatorname{Span}((v_n)_{n=1}^{k})$ *for all $k < N$.*

Proof.
$$e_1 = \frac{v_1}{\|v_1\|}, e_{k+1} = \frac{v_{k+1} - \sum_{n=1}^{k} \langle v_{k+1}, e_n \rangle e_n}{\left\| v_{k+1} - \sum_{n=1}^{k} \langle v_{k+1}, e_n \rangle e_n \right\|}$$
gives the desired family. □

Theorem 6.4. *Let V be a Euclidean space.*
(i) Suppose $v_1, v_2 \in V$ are orthogonal, then $\|v_1 + v_2\|^2 = \|v_1\|^2 + \|v_2\|^2$.
(ii) Suppose $\{e_1, \ldots, e_N\} \subset V$ is a finite orthonormal family and $v \in V$, then
$$\|v\|^2 = \left\| v - \sum_{i=1}^{N} \langle v, e_i \rangle e_i \right\|^2 + \sum_{n=1}^{N} |\langle v, e_n \rangle|^2$$
(iii) (Bessel's Inequality) Suppose $\{e_1, e_2, \ldots\} \subset V$ is a countably infinite orthonormal family and $v \in V$, then
$$\sum_{k=1}^{\infty} |\langle v, e_n \rangle| \leq \|v\|^2$$
with equality iff $\sum_{k=1}^{N} \langle v, e_n \rangle e_n \to v$ as $N \to \infty$.

Proof. (i) is obvious. (ii) follows from (i) and

$$\left(v - \sum_{n=1}^{N} \langle v, e_n \rangle e_n\right) \perp \sum_{n=1}^{N} \langle v, e_n \rangle e_n$$

(iii) follows from (ii). □

Example 6.1. 1. ℓ^2 has the structure of a Euclidean space with $\langle x, y \rangle = \sum_i x_i \bar{y}_i$. It is complete under the induced norm and $(e^j)_j$ where $e^j_i = \delta_{ij}$ is a maximal orthonormal family in ℓ^2.
2. On $C_{\mathbb{F}}([0,1])$, we have the inner product

$$\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)}\,\mathrm{d}x$$

but it is not complete under the induced norm. However, we can in fact complete it into the space $L^2([0,1])$.

Theorem 5.8 often comes in handy in the contexts of Euclidean spaces of continuous functions.

Proposition 6.5. *Consider the inner product*

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x)\,\mathrm{d}x$$

on the space $C_{\mathbb{C}}(S^1)$. $e_k = (x \mapsto e^{ix})$ is an orthonormal family in this space. Write $\hat{f}(k) = \langle f, e_k \rangle$ and $S_n(f) = \sum_{k \in \mathbb{Z}} \hat{f}(k) e_k$, then:
(i) $\|S_N(f) - f\|_2 \to 0$ where $\|\cdot\|_2$ is the induced norm.
(ii) $\sum_{k \in \mathbb{Z}} |\hat{f}(k)|^2 = \langle f, f \rangle$.

Proof. Theorem 5.8 gives (i). (ii) follows from Theorem 6.4. □

6.2 Completeness

It's well-known that the space $(C([0,1]), \|\cdot\|_2)$ has a completion given by $L^2([0,1])$. Let's generalise such kind of construction.

Theorem 6.6 (Completion)**.** *Given a normed vector space V, there is a complete normed vector space $\bar{V}$ with a linear isometry $\Phi : V \to \bar{V}$ such that $\overline{\Phi(V)} = \bar{V}$. Such a completion is unique up to isometric isomorphism.*

One can show the existence of completion from the level of metric spaces. But we've got a norm here, so we can make things much easier.

Proof. Recall that the bidual embedding $\Phi : V \to V^{**}$ is isometric. Consider $\bar{V} = \overline{\Phi(V)}$. This is closed in the complete space V^{**}, hence is also complete.
For uniqueness, suppose $(\bar{V}_1, \|\cdot\|_1), (\bar{V}_1, \|\cdot\|_2)$ are completions of V with embeddings Φ_1, Φ_2 respectively. Then $\Phi_0 : \Phi_2 \circ \Phi_1^{-1} : \Phi_1(V) \to \Phi_2(V)$ is a linear isometry. There's a unique linear isometric extension $\tilde{\Phi}_0 : \bar{V}_1 \to \bar{V}_2$ of it. By hypothesis its image is dense in $\bar{V}_2$. If we can show that the image is closed, then we are essentially done. Suppose $(y_n)_n \in \tilde{\Phi}_0(\bar{V}_1)$ with $y_n \to y \in \bar{V}_2$. Choose $(x_n)_n \in \bar{V}_1$ with $\tilde{\Phi}_0(x_n) = y_n$, then $(x_n)_n$ is Cauchy as Φ is an isometry, hence $x_n \to x$ for some $x \in \bar{V}_1$ by completeness of $\bar{V}_1$. Then $\tilde{\Phi}_0(x) = y$, as desired. □

Definition 6.3. A complete Euclidean space is called a Hilbert space.

Let H be a Hilbert basis of infinite dimension. Then H would have to be "too big" to have a countable basis by Theorem 4.2. This inspires us to consider an alternative notion of basis.

Theorem 6.7. *Let H be a separable Hilbert space of infinite dimension, then:*
(i) There exists a countable orthonormal family $(e_n)_n$ (called a "Hilbertian basis") so that $\overline{\mathrm{Span}\{e_n : n \geq 1\}} = H$.
(ii) Let $(e_n)_n$ be a Hilbertian basis in H, then for any $v, w \in H$, if we write $v_n = \langle v, e_n \rangle, w_n = \langle w, e_n \rangle$, then we have

$$\sum_{n\geq 1} v_n e_n \to v, \sum_{n\geq 1} w_n e_n \to w, \sum_{n\geq 1} v_n \bar{w}_m \to \langle v, w \rangle$$

(iii) $\Phi : H \to \ell^2$ given by $v \mapsto (v_n)_{n\geq 1}$ (as in (ii)) is a linear isometric bijection.

Proof. (i) Take a countable dense set S, produce a countable linearly independent set whose span contains S, and cast Gram-Schmidt.
(ii) Let $s_n = \sum_{k=1}^n v_k e_k$. $(s_n)_n$ is Cauchy as for any $n > m$,

$$\|s_n - s_m\|^2 = \left\| \sum_{k=m+1}^{n} v_k e_k \right\|^2 = \sum_{k=m+1}^{n} |v_k|^2 \to 0$$

as $m \to \infty$ since $(v_k)_k \in \ell^2$ by Theorem 6.4.
Completeness shows that $s_n \to s$ for some $s \in H$. For all n, we have $(v - s_n) \perp \mathrm{Span}\{e_1, \ldots, e_n\}$, so $(v - s) \perp \mathrm{Span}\{e_n : n \geq 1\}$ which implies that $(v - s) \perp \overline{\mathrm{Span}\{e_n : n \geq 1\}} = H$, i.e. $v - s = 0$.
As for the inner product, observe that

$$\left\langle \sum_{k=1}^{n} v_k e_k, \sum_{k=1}^{n} w_k e_k \right\rangle = \sum_{k=1}^{n} v_k \bar{w}_k$$

which is absolutely convergent (combining Proposition 1.2 and Theorem 6.4), therefore convergent as H is complete. We are then done by the continuity of inner product.
(iii) Clear from (ii). □

6.3 Projection, Decomposition and Representation

Our goal now is to exploit the convexity of unit ball in a Hilbert space to build some nice orthogonal projections. We can further show that this projection we'll construct is 1-Lipschitz by the uniform strict convexity of the unit ball.

Proposition 6.8 (Projection to a Convex Set)**.** *Let V be a Euclidean space and suppose $C \subset V$ is nonempty, convex and complete. Then:*
(i) For every $v \in V$, there is a unique $P_C(v) \in C$ such that

$$\|v - P_C(v)\| = d(v, C) = \inf_{z\in C} \|v - z\|$$

(ii) $P_C(v)$ is characterised by the condition

$$\forall z \in C, \mathrm{Re}\langle z - P_C(v), v - P_C(v) \rangle \leq 0$$

(iii) $P_C : V \to C$ *is* 1*-Lipschitz, i.e.* $\|P_C(v_1) - P_C(v_2)\| \le \|v_1 - v_2\|$ *for all* $v_1, v_2 \in V$*.*

Proof. (i) Existence is clear when $v \in C$ (in which case $P_C(v) = v$). Suppose $v \notin C$. Let $w_n \in C$ be such that $\|v - w_n\|^2 \le d(v, C)^2 + n^{-1}$. As $\|w_m - w_n\|^2 + \|w_m + w_n - 2v\|^2 = 2\|w_n - v\|^2 + 2\|w_m - v\|^2$, we have

$$\begin{aligned}\frac{\|w_m - w_n\|^2}{2} &= \|w_n - v\|^2 + \|w_m - v\|^2 - 2\left\|\frac{w_m + w_n}{2} - v\right\|^2 \\ &\le 2d(v,C)^2 + \frac{1}{m} + \frac{1}{n} - 2d(v,C)^2 \le \frac{1}{m} + \frac{1}{n}\end{aligned}$$

So $(w_n)_n$ is Cauchy and we can take $P_C(v)$ as its limit.
As for uniqueness, suppose $w_1, w_2 \in C$ both has $\|v - w_1\| = \|v - w_2\| = d(v, C)$, then we have

$$\frac{\|w_1 - w_2\|^2}{2} = \|w_1 - v\|^2 + \|w_2 - v\|^2 - 2\left\|\frac{w_1 + w_2}{2} - v\right\|^2 \le 0$$

Thus $w_1 = w_2$.
(ii) Suppose the condition holds, then for any $z \in C$ we have

$$\begin{aligned}\|z - v\|^2 &= \|(z - w) + (w - v)\|^2 \\ &= \|z - w\|^2 + \|v - w\|^2 + 2\operatorname{Re}\langle z - w, w - v\rangle \ge \|z - w\|^2 + \|v - w\|^2\end{aligned}$$

So $z = P_C(v)$. Conversely, for any $z \in C, \lambda \in (0, 1]$ we have $\|(\lambda z + (1 - \lambda)P_C(v)) - v\|^2 \ge \|v - P_C(v)\|$. Expanding gives $2\operatorname{Re}\langle P_C(v) - v, z - P_C(v)\rangle + \lambda\|z - P_C(v)\|^2 \ge 0$. Sending $\lambda \to 0$ gives the result.
(iii) Apply (ii) to $v = v_1, z = P_C(v_2)$ and $v = v_2, z = P_C(v_1)$ gives

$$\begin{cases}\operatorname{Re}\langle P_C(v_2) - P_C(v_1), v_1 - P_C(v_1)\rangle \le 0 \\ \operatorname{Re}\langle P_C(v_1) - P_C(v_2), v_2 - P_C(v_2)\rangle \le 0\end{cases}$$

Adding them together and rearranging gives $\|P_C(v_1) - P_C(v_2)\|^2 \le \operatorname{Re}\langle P_C(v_1) - P_C(v_2), v_1 - v_2\rangle$ which implies the result by Proposition 6.1. □

Example 6.2. 1. For any Euclidean space V, any finite dimensional subspace $C \ne \varnothing$ is convex and complete.
2. Closed subspaces of Hilbert space are convex and complete.
3. In the case where H is a Hilbert space and $C = \bar{B}(0, 1)$, the projection is given by

$$P_C(v) = \begin{cases} v & \text{if } \|v\| \le 1 \\ v/\|v\| & \text{if } \|v\| > 1\end{cases}$$

Theorem 6.9. *Suppose* V *is Euclidean and* $W \le V$ *is a complete subspace, then we can decompose every* $v \in V$ *into* $v = w + w^\perp$ *where* $w = P_W(v)$ *and* $w^\perp = v - P_W(v) \in W^\perp$ *(in particular* $V = W \oplus W^\perp$*). Furthermore,* P_W *is linear, idempotent, identity on* W *and zero on* $W^\perp$*.*

Proof. The only thing requiring a proof is $v - P_W(v) \in W^\perp$. In the real case, by (ii) in the preceding proposition we have, for any $z \in W$, $\langle \pm z, v - P_W(v)\rangle \le 0$, hence $\langle z, v - P_W(v)\rangle = 0$ which is exactly what's needed. In the complex case, we can get (again for any $z \in W$) $\pm\operatorname{Re}\langle z, v - P_W(v)\rangle \le 0, \pm\operatorname{Im}\langle z, v - P_W(v)\rangle \le 0$, giving the same result. □

Remark. 1. If $W = \text{Span}\{e_1, \ldots, e_n\}$ with $\{e_1, \ldots, e_n\}$ orthonormal, then essentially

$$P_W(v) = \sum_{k=1}^{n} \langle v, e_k \rangle e_k$$

2. If $V = H$ is a Hilbert space and $W = \overline{\text{Span}\{e_n : n \geq 1\}}$ with $\{e_1, \ldots, e_n\}$ orthonormal, then

$$P_W(v) = \sum_{n=1}^{\infty} \langle v, e_n \rangle e_n$$

Consequently, $S \subset H$ has $\overline{\text{Span}(S)} = H$ iff $S^\perp = \{0\}$.

Theorem 6.10 (Riesz-Fréchet Representation Theorem)**.** *For any Hilbert space H, $\phi : H \to H^*, v \mapsto \phi_v$ with $\phi_v(w) = \langle w, v \rangle$ is a bijective (sesqui)linear isometry.*

We can of course define such a ϕ in any Euclidean space, but surjectivity depends on completeness.

Proof. (Sesqui)linearity is clear. ϕ is isometric (hence injective) due to Proposition 6.1. To check surjectivity, suppose $f \in H^* \setminus \{0\}$. Then $W = \ker f$ is a closed subspace of H. Write $H = W \oplus W^\perp$. Take $v_0 \in W^\perp$ such that $f(v_0) \neq 0$, then essentially $W^\perp = \mathbb{F}v_0$. Choose $\alpha \in \mathbb{F}$ such that $f(\alpha v_0) = \|\alpha v_0\|_H^2$, then $f = \phi_{\alpha v_0}$ on both W and $W^\perp$, therefore on H. □

6.4 Spectrum and Resolvent

The goal is to generalise the theory of eigenvalues and eigenvectors to Hilbert spaces. Eventually, we will prove the celebrated spectral theorem, which generalises orthonormal diagonalisation of symmetric matrices you see in finite dimension.
From now on, we take $\mathbb{F} = \mathbb{C}$ to avoid problems caused by the non-algebraic closedness of $\mathbb{R}$.

Definition 6.4. Let H be a Hilbert space and $T \in B(H, H) = B(H)$. The resolvent of T is $\rho(T) = \{\lambda \in \mathbb{C} : T - \lambda \text{ bijective with bounded inverse}\}$. The spectrum of T is $\sigma(T) = \mathbb{C} \setminus \rho(T)$. The resolvent map of T is $R_T : \rho(T) \to B(H), \lambda \mapsto (T - \lambda)^{-1}$.

Remark. By Theorem 4.6, $T - \lambda$ automatically has a bounded inverse if it is bijective. However, $T - \lambda$ can be injective but not surjective, so do NOT interpret $\sigma(T)$ as the "set of eigenvalues" in the usual sense.

Proposition 6.11. *Let H be a Hilbert space and $T \in B(H)$. Then $\rho(T)$ is open and R_T is analytic. Moreover, $\sigma(T) \neq \varnothing$ and $\sigma(T) \subset \{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|_{\text{op}}\}$.*

Proof. Observe that whenever $U \in B(H)$ has $\|U\|_{\text{op}} < 1$, then $1 - U$ is invertible with inverse $(1 - U)^{-1} = \sum_{n \geq 0} U^n$ (the series converges due to completeness). Suppose $\lambda_0 \in \rho(T)$, then $T - \lambda = (T - \lambda_0) - (\lambda - \lambda_0) = (T - \lambda_0)(1 - R_T(\lambda_0)(\lambda - \lambda_0))$. Take $U = R_T(\lambda_0)(\lambda - \lambda_0)$, then $\|U\|_{\text{op}} < 1$ whenever $\lambda \in$

$B_{\mathbb{C}}(\lambda_0, \|R_T(\lambda_0)\|^{-1})$, so $\rho(T)$ is open. Moreover, for $\lambda \in B_{\mathbb{C}}(\lambda_0, \|R_T(\lambda_0)\|^{-1})$ we have

$$\begin{aligned} R_T(\lambda) &= (T-\lambda)^{-1} = (1-U)^{-1} R_T(\lambda_0) \\ &= \sum_{n\geq 0} U^n R_T(\lambda_0) = \sum_{n\geq 0} R_T(\lambda_0)^{n+1}(\lambda-\lambda_0)^n \end{aligned}$$

So R_T is analytic around λ_0 with $R_T'(\lambda_0) = R_T(\lambda_0)^2$.
If $|\lambda| > \|T\|_{\mathrm{op}}$, then $T - \lambda = -\lambda(1 - \lambda^{-1}T)$. But $\|\lambda^{-1}T\|_{\mathrm{op}} < 1$, so $T - \lambda$ is invertible and we have $\|R_T(\lambda)\| \leq (|\lambda| - \|T\|_{\mathrm{op}})^{-1}$. This in particular shows that $\sigma(T) \subset \{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|_{\mathrm{op}}\}$.
To see $\sigma(T) \neq \varnothing$, suppose $\sigma(T) = \varnothing$, then R_T is an analytic function defined over all of $\mathbb{C}$. For any $v \in H, \phi \in H^*$, we can consider the entire function $F_{v,\phi}(\lambda) = \phi(R_T(\lambda)v)$ which is bounded since for $|\lambda| > \|T\|_{\mathrm{op}}$ we have $|F(v,\phi)(\lambda)| \leq \|v\|\|\phi\|/(|\lambda| - \|T\|_{\mathrm{op}})$. Hence it is constant by Liouville's theorem, which then forces R_T to be constant, contradiction. □

Definition 6.5. The point spectrum $\sigma_{\mathrm{p}}(T)$ is the set of eigenvalues $\sigma_{\mathrm{p}}(T) = \{\lambda \in \mathbb{C} : T - \lambda \text{ not injective}\}$. The continuous spectrum $\sigma_{\mathrm{c}}(T)$ is $\sigma_{\mathrm{c}}(T) = \{\lambda \in \mathbb{C} : T - \lambda \text{ injective}, (T-\lambda)(H) \subsetneq H \text{ dense}\}$. The residual spectrum $\sigma_{\mathrm{r}}(T)$ is the rest, i.e. $\sigma_{\mathrm{r}}(T) = \{\lambda \in \mathbb{C} : T - \lambda \text{ injective}, (T-\lambda)(H) \text{ not dense in } H\}$.

Remark. Observe that if T is bounded below (i.e. there is some $C > 0$ with $\|Tv\| \geq C\|v\|$ for all $v \in H$), then $T(H)$ would have to be closed. So T is invertible iff it is bounded below and $T(H)$ is dense in H.
A consequence of this is that $T - \lambda$ is not bounded below whenever $\lambda \in \sigma_{\mathrm{c}}(T)$, which is a nice way to think about the continuous spectrum.
It also inspires to introduce the notion of approximate eigenvalues: We say $\lambda \in \mathbb{C}$ is an approximate eigenvalue of T, written $\lambda \in \sigma_{\mathrm{ap}}(T)$ if there is a sequence $(v_n)_n \in H$ of unit vectors such that $(T-\lambda)(v_n) \to 0$ as $n \to \infty$. Clearly $\sigma_{\mathrm{p}}(T) \subset \sigma_{\mathrm{ap}}(T)$.

Example 6.3. 1. In $\mathbb{C}^d$ equipped with the Euclidean norm, $\sigma(T) = \sigma_{\mathrm{p}}(T)$ since injectivity is equivalent to invertibility. We also have $|\sigma(T)| \leq d$.
2. Take $H = \ell^2$ with the usual norm $\|x\| = (\sum_n |x_n|^2)^{1/2}$, then the left shift operator $T_{\mathrm{left}}((x_n)_{n\geq 1}) = (x_{n+1})_{n\geq 1}$ has unit norm and $\sigma(T_{\mathrm{left}}) = \bar{B}(0,1) = B(0,1) \sqcup S^1 = \sigma_{\mathrm{p}}(T_{\mathrm{left}}) \sqcup \sigma_{\mathrm{c}}(T_{\mathrm{left}})$.
3. Again take $H = \ell^2$, but this time we consider instead the right shift operator $T_{\mathrm{right}}((x_n)_{n\geq 1}) = (x_{n-1})_{n\geq 1}$ (with $x_0 = 0$), which again has unit norm and $\sigma(T_{\mathrm{right}}) = \bar{B}(0,1) = B(0,1) \sqcup S^1 = \sigma_{\mathrm{r}}(T_{\mathrm{right}}) \sqcup \sigma_{\mathrm{c}}(T_{\mathrm{right}})$.
4. For any $K \subset \mathbb{C}$ compact, there is some $T \in B(\ell^2)$ with $\sigma(T) = K$.

Definition 6.6. For all $T \in B(H)$, there is a unique $T^* \in B(H)$ (called the adjoint of T) with $\forall v, w \in V, \langle Tv, w\rangle = \langle v, T^*w\rangle$.
We say T is normal if $TT^* = T^*T$, self-adjoint if $T = T^*$ and unitary if $TT^* = \mathrm{id}_H$.

By definition $(T^*)^* = T$.

Example 6.4. 1. In finite dimension, we can just write down $A^* = \bar{A}^\top$ after expressing T as a matrix A in terms of some basis.
2. $T_{\mathrm{right}} = T_{\mathrm{left}}^*$.

Remark. The uniqueness of adjoints is obvious. As for existence, we can take $T^* = \phi^{-1} \circ T \circ \phi$ where $\phi : H \to H^*$ is as in Theorem 6.10.

Proposition 6.12. *(i) If T is either unitary or self-adjoint, then it's normal.*
(ii) $\ker T = (\operatorname{Im} T^*)^\perp$. *Moreover, if T is normal, then* $\langle Tv, Tv\rangle = \langle T^*Tv, v\rangle = \langle TT^*v, v\rangle = \langle T^*v, T^*v\rangle$, *i.e.* $\ker T = \ker T^*$.
(iii) If T is normal, then eigenvectors of distinct eigenvalues are orthogonal.
(iv) If T is self-adjoint, then $\langle Tv, v\rangle \in \mathbb{R}$ for all $v \in H$ and $\sigma(T) \subset \mathbb{R}$.

Definition 6.7. $T \in B(H)$ is compact if $T(\bar{B}(0,1))$ is relatively compact.

Proposition 6.13. *(i) T is compact iff T is a ($\|\cdot\|_{\text{op}}$-)limit of finite rank operators (i.e. bounded linear maps with finite dimensional image).*
(ii) T is compact iff T^* is compact.
(iii) If T is compact and $\lambda \in \sigma(T) \setminus \{0\}$, then $\lambda \in \sigma_{\text{p}}(T)$ with $\dim\ker(T-\lambda) < \infty, \operatorname{codim}(\operatorname{Im}(T-\lambda)) < \infty$.

Proof. (i) The "if" part is clear. Conversely, if T is compact, then $T(\bar{B}_H(0,1))$ is totally bounded. So for any $\epsilon > 0$, there are $v_1, \dots, v_n \in H$ such that $T(\bar{B}_H(0,1)) \subset \bigcup_{j=1}^n \bar{B}_H(v_d, \epsilon/2)$. Take $T_n = P_{\text{Span}\{v_1,\dots,v_n\}}T$, then $\|T_n - T\|_{\text{op}} < \epsilon$.
(ii) Example sheet.
(iii) If $\lambda \neq 0$ and $\ker(T-\lambda) \neq \{0\}$, then $\ker(T-\lambda)$ has compact closed unit ball by compactness of T, hence finite dimensional. By (ii), $\ker(T-\lambda)^*$ too is finite dimensional, so $\operatorname{Im}(T-\lambda) = (\ker(T-\lambda)^*)^\perp$ has finite codimension.
To finish, we shall show that any $\lambda \in \sigma(T) \setminus \{0\}$ is an eigenvalue. It suffices to show that if $\lambda \neq 0, \ker(T-\lambda) = \{0\}$, then $T - \lambda$ is invertible.
$T - \lambda$ is bounded below by example sheet. Choose finite rank operators $(T_n)_n$ converging to T, then $(T_n - \lambda)$ is uniformly bounded below. As T_n has finite rank and $\lambda \neq 0$, $T_n - \lambda$ would have to be surjective. Indeed, consider $U_n = T_n P_{(\ker T_n)^\perp}(T_n - \lambda) : (\ker T_n)^\perp \to \operatorname{Im} T_n$ is injective, hence bijective. Therefore $\dim(\operatorname{Im}(T_n - \lambda) \cap (\ker T_n)^\perp) = \dim(\operatorname{Im}(T_n)) = \dim(\ker T_n)^\perp$. So $\operatorname{Im}(T_n - \lambda) \supset (\ker T_n)^\perp$, but also $\operatorname{Im}(T_n - \lambda) \supset \ker T_n$, so $\operatorname{Im}(T_n - \lambda) \supset H$, i.e. $T_n - \lambda$ is surjective.
Now, for any unit $w \in W$, choose $v_n \in H$ such that $U_n v_n = w$. (v_n) is bounded as $T_n - \lambda$ is uniformly bounded below. Compactness of T then shows that $(Tv_n)_n$ converges after passing to a subsequence. We also have

$$v_n = \frac{1}{\lambda}(Tv_n - w - (T - T_n)v_n)$$

which has to converge, say to some $v \in H$. But then $(T-\lambda)v = w$ is surjective, as desired. □

Example 6.5. 1. Suppose H is a separable Hilbert space and let $(e_n)_n$ be a Hilbertian basis of H. We define the Hilbert-Schmidt norm on H as $\|T\|_{\text{HS}} = \sqrt{\sum_n \|Te_n\|^2}$ which is clearly independent of the choice of $(e_n)_n$. Also, we always have $\|\cdot\|_{\text{HS}} \geq \|\cdot\|_{\text{op}}$.
We say $T \in B(H)$ is a Hilbert-Schmidt operator if $\|T\|_{HS} < \infty$. Not every bounded operator is Hilbert-Schmidt. In fact, the set of Hilbert-Schmidt operators is not closed in $(B(H), \|\cdot\|_{\text{op}})$. Indeed, any Hilbert-Schmidt operator must be compact (exercise).

2. Again let H be a separable Hilbert space with Hilbertian basis $(e_n)_n$. Suppose we have a diagonal linear operator T, i.e. there are complex numbers $(\lambda_n)_n$ such that $Te_n = \lambda_n e_n$, then T is bounded iff (λ_n) is bounded, compact iff $(\lambda_n) \to 0$, Hilbert-Schmidt if $(\lambda_n)_n \in \ell^2$, and finite rank iff $(\lambda_n)_n$ is eventually zero.

Theorem 6.14 (Spectral Theorem)**.** *If H is a Hilbert space and $T \in B(H)$ is a compact self-adjoint operator, then $\sigma(T) \setminus \{0\} \subset \sigma_{\mathrm{p}}(T) \subset \mathbb{R}$. Furthermore, $\sigma_{\mathrm{p}}(T)$ is at most countable and its only possible accumulation point is 0.*
Let $\sigma_{\mathrm{p}}(T) \setminus \{0\} = (\lambda_n)_n$, then $E_{\lambda_n} = \ker(T - \lambda_n)$ has finite dimension for each n and

$$H = \ker T \oplus^{\perp} \bigoplus_{n}^{\perp} E_{\lambda_n}, T = \sum_{n \geq 1} \lambda_n P_{E_{\lambda_n}}$$

Proof. The idea to inductively exhaust the eigenvalues of T in an order of non-increasing modulus.
The modulus of any eigenvalue of an operator U would have to be bounded by $\|U\|_{\mathrm{op}}$. Wouldn't it be nice if, under certain conditions on U, we can actually find an eigenvalue attaining the bound (hence allowing us to execute the idea)? In fact, if $U \in B(H)$ is compact and self-adjoint, then at least one of $\pm\|U\|_{\mathrm{op}}$ is an eigenvalue.
We have $\|U\|_{\mathrm{op}} = \sup_{\|v\| \leq 1} \|Uv\| = \sup_{\|v\| \leq 1, \|w\| \leq 1} |\langle Uv, w\rangle|$. The last expression actually equals $\sup_{\|v\| \leq 1} |\langle Uv, v\rangle|$. Indeed, for any $\|v\| \leq 1, \|w\| \leq 1$, we choose $\alpha \in S^1$ such that $\alpha\langle Uv, w\rangle \in \mathbb{R}_{\geq 0}$. Then

$$\begin{aligned}
|\langle Uv, w\rangle| &= |\operatorname{Re}\langle Uv, \alpha w\rangle| = \frac{1}{4}|\langle U(v + \alpha w), v + \alpha w\rangle - \langle U(v - \alpha w), v - \alpha w\rangle| \\
&\leq \left(\sup_{\|z\| \leq 1} |\langle Uz, z\rangle|\right)\left(\frac{\|v + \alpha w\|^2 + \|v - \alpha w\|^2}{4}\right) \\
&= \left(\sup_{\|z\| \leq 1} |\langle Uz, z\rangle|\right)\left(\frac{\|v\|^2 + \|\alpha w\|^2}{2}\right) \leq \sup_{\|z\| \leq 1} |\langle Uz, z\rangle|
\end{aligned}$$

So we can find a unit length sequence $(v_n)_n$ such that $|\langle Uv_n, v_n\rangle| \to \|U\|_{\mathrm{op}}$. But U is self-adjoint, so $\langle Uv_n, v_n\rangle$ is always real and therefore $\langle Uv_n, v_n\rangle \to \lambda$ for some $\lambda \in \{\pm\|U\|_{\mathrm{op}}\}$. Then $\|(U - \lambda)v_n\|^2 = \|Uv_n\|^2 + \lambda^2\|v_n\|^2 - 2\lambda\langle Uv_n, v_n\rangle \leq 2\|U\|_{\mathrm{op}}^2 - 2\lambda\langle Uv_n, v_n\rangle \to 0$, so $\lambda \in \sigma(U)$. As $\lambda \neq 0$, we have $\lambda \in \sigma_{\mathrm{p}}(U)$ as U is compact.
Using this, we shall construct a (possibly finite) real sequence $(\lambda_n)_n$ exhausting $\sigma_{\mathrm{p}}(T) \setminus \{0\}$ such that $|\lambda_n|$ is nonincreasing, $|\lambda_{n+2}| < |\lambda_n|$ and $\lambda_n \to 0$ if the sequence is infinite. The above shows that T has an eigenvalue λ_1 that is $\pm\|T\|_{\mathrm{op}}$. Now assuming that $\lambda_1, \ldots, \lambda_n$ have been constructed, we consider $V_n = \left(\bigoplus_{k \leq n}^{\perp} E_{\lambda_k}\right)^{\perp}$ which is stable under T. $T|_{V_n}$ is then a compact self-adjoint operator on T_n. If it is zero, then we stop the process; Otherwise, it has an eigenvalue $\pm\|T|_{V_n}\|_{\mathrm{op}}$ which we shall assign as λ_{n+1}.
If $(\lambda_n)_n$ is infinite but does not converge to zero, then there is a subsequence $(\lambda_{n_k})_k$ such that there is some $\delta > 0$ with $|\lambda_{n_k}| > \delta$. Choose unit length sequence $(v_k)_k$ with $Tv_k = \lambda_{n_k} v_k$, then $\|Tv_{n_k} - Tv_{n_m}\| \geq \sqrt{2}\delta$ whenever $m \neq k$, contradicting compactness of T.

Let $V_\infty = \left(\bigoplus_n^\perp E_{\lambda_n}\right)^\perp$. $T|_{V_\infty}$ has no nonzero eigenvalue by construction. But its operator norm is an eigenvalue (up to a sign), so we must have $T|_{V_\infty} = 0$. Hence $H = \ker T \oplus^\perp \bigoplus_n^\perp E_{\lambda_n}$. Also, $\left\| T - \sum_{k=1}^n \lambda_k P_{E_{\lambda_k}} \right\| \leq |\lambda_n|$, so $T = \sum_n \lambda_n P_{E_{\lambda_n}}$ □

Example 6.6. For nonnegative $q \in C([a,b])$, the operator $Tf = -f'' + qf$ has a bounded right inverse T (i.e. putting in homogeneous boundary conditions) on $(C([a,b]), \|\cdot\|_2)$ that happens to be compact and self-adjoint. Applying the theorem then recovers the Sturm-Liouville theory you've seen in Methods.