Functional Analysis

Krish Nigam *

March, 2024

Functional analysis brings together the ideas of continuity and linear algebra; many consider this subject the study of *infinite dimensional vector spaces*. As such, we will often look at spaces of functions. The ideas formulated here are an important requirement for further study in many areas of mathematical analysis, including PDEs, stochastic analysis and quantum mechanics.

In this course, we assume familiarity with the foundations of linear algebra and analysis, including metric spaces and topological spaces. This documents constitutes my notes taken from lectures by *Dr Pierre-François Rodriguez* at *Imperial College London*.

Contents

1 Introduction **2**
- 1.1 Motivation 2
- 1.2 Non-examples 2
- 1.3 Banach Spaces and Examples 3

2 Linear Spaces **6**
- 2.1 Separability 6
- 2.2 Hilbert Spaces 7
- 2.3 Finite vs. Infinite Dimensions 9
- 2.4 Compactness 10

3 Linear Operators **13**
- 3.1 Duality 15
- 3.2 Dual Operators 17
- 3.3 Hahn-Banach and Applications 17

4 Baire Category and UBP **21**
- 4.1 Baire Category 21
- 4.2 Uniform Boundedness Principle 22
- 4.3 Open Mapping Theorem 25
- 4.4 Closed Graph Theorem 26

5 Spectral Theory **29**
- 5.1 Weak vs. Strong Topologies 29
- 5.2 Compact Operators 32
- 5.3 Spectrum and Resolvent 34
- 5.4 Spectral Theory in Hilbert Spaces 36
- 5.5 Spectral Theorem for Compact Self-Adjoint Operators 38

*Email: kn221@ic.ac.uk

§1 Introduction

§1.1 Motivation

Broadly speaking, in this course we are interested in solving linear equations of the form $\boxed{\text{Ax} = \text{y}}$, where $x \in X$, $y \in Y$ and X, Y are vector spaces (which we may also refer to as **linear spaces**). We have $A : X \to Y$ with A linear, *i.e.* $A(\alpha x + \beta y) = \alpha A(x) + \beta A(y)$. We want to find x such that the above equation holds, with $y \in Y$ and A given.

For X, Y finite dimensional, we have traditional linear algebra, which I hope we're familiar with. However, we are concerned with X, Y being infinite dimensional, with topological properties such as completeness, compactness *etc.* which come about from a metric or norm. It will become clear later on as to why certain topological properties like completeness are nice for us to work with; after all **complete** normed vector spaces are given a special name - *Banach spaces.*

Example 1.1 Consider $f \in C_0^\infty(\mathbb{R}^n)$ where

$$C_0^\infty(\mathbb{R}^n) = \{f : \mathbb{R}^n \to \mathbb{K} : f \text{ differentiable } \infty \text{ often}, \ \operatorname{supp}(f) \text{ bounded}\}$$

and recall that $\operatorname{supp}(f) = \overline{\{x \in \mathbb{R}^n : f(x) \neq 0\}}$. Now, we may be interested in solving *Poisson's equation* $-\Delta u = f$. Here 'A' is $-\Delta$, and the question becomes we want to find suitable function spaces X, Y to find u.

Note that in the above example an adequate choice of space is necessary to find a solution. In this course, X, Y are almost always Banach spaces (or even Hilbert spaces).

§1.2 Non-examples

We will begin by presenting some non-examples, in order to demonstrate where things can go wrong. We adopt the structure (X, d) a metric space, $d : X \times X \to \mathbb{R}_+$.

1. Take $X = \mathbb{Q}$, $d(x, y) = |x - y|$, and consider the sequence $x_1 = 1$, $x_{n+1} = \frac{2(1+x_n)}{2+x_n}$ so $(x_n) \subset \mathbb{Q}$. Clearly, $\forall n x_n < x_{n+1}$ and $1 \leq x_n < 2$. By Bolzano-Weierstrass $\lim\limits_{n\to\infty}$ exists, but solving gives $x_n \to \sqrt{2} \notin \mathbb{Q}$, even though (x_n) is still Cauchy in $\mathbb{Q}$.

   This emphasises the importance of requiring the completion of $\mathbb{Q}$ in this case, which is $\mathbb{R}$, *i.e.* complete spaces are nice.

2. A more relevant example - take $X = C([0, 1])$, $d(f, g) = \int_0^1 |f - g| \, dx$, $f, g \in X$ and we define a sequence $(f_n) \subset X$ with

$$f_n(x) = \begin{cases} 0, & 0 \leq x \leq \frac{1}{2} \\ nx - \frac{n}{2}, & \frac{1}{2} < x < \frac{1}{2} + \frac{1}{n} \\ 1, & x \geq \frac{1}{2} + \frac{1}{n} \end{cases}$$

   We can check that $(f_n) \subset X$ is Cauchy and $\nexists f \in X$ such that $\lim\limits_{n\to\infty} d(f_n, f) = 0$. This is an example of a sequence of functions with no converging solution in the space. Note imagining a picture of the above function is good for intuition.

Remark 1.2 (Fixes) We present potential fixes to the problems demonstrated in non-example 2) above.

- We can change the notion of convergence to uniform convergence, in which case f_n as above does not 'converge'.
- We can X to L^1-space, so $X = L^1([0,1])$ and then f_n converges to the indicator function at $\frac{1}{2}$. Note that the L^1-space is complete as we will soon show.

§1.3 Banach Spaces and Examples

Definition 1.3 (Banach Space) — A **Banach space** is a normed space $(X, \|\cdot\|)$ which is complete w.r.t $\|\cdot\|$. Note that the norm induces a metric $d(x,y) = \|x-y\|$.

For Banach spaces, there are many key examples to be familiar with when working through this course. First, take a measure space $(X, \mathcal{A}, \mu)$, then the 'L^p-space', $L^p(\mu) = L^p(X, \mathcal{A}, \mu) = \{f : X \to \overline{\mathbb{R}}/\ \ : \|f\|_{L^p} < \infty\}$, where

$$\|f\|_{L^p} = \begin{cases} (\int |f|^p \,\mathrm{d}\mu)^{\frac{1}{p}}, & p < \infty \\ \mathrm{esssup}|f|, & p = \infty \end{cases}$$

If two functions are equivalent μ almost-everywhere, they have the same L^p norm. The choice of measure space is important; in traditional linear algebra, we typically have $X = \{1, 2, \dots, n\}$, $\mathcal{A} = 2^X$, $\mu(\{k\}) = 1\ \forall k \in X$ (counting measure).

Then, every function $f : X \to \mathbb{R}$ is **simple**, *i.e.* f is of the form $f(x) = \sum_{k=1}^n f(k)\mathbb{1}_{\{k\}}(x)$. This motivates the following norm $\|\cdot\|_p$.

$$\implies \|f\|_p^p = \int |f|^p \,\mathrm{d}\mu = \sum_{k=1}^n |f(k)|^p \int \mathbb{1}_{\{k\}} \,\mathrm{d}\mu = \sum_{k=1}^n |f(k)|^p$$

So, the finite dimensional space $L^p(\{1, \dots, n\}, \mu) \cong \mathbb{R}^n$, endowed with the norm given above $\|f\|_p = (\sum_{k=1}^n |f(k)|^p)^{\frac{1}{p}}$.

This space is finite dimensional, but we want to go to infinite dimensions! So, let's send $n \to \infty$. This gives rise to the "little-ℓ-p" space.

Example 1.4 (ℓ^p, "little-ℓ-p") We adopt the same format as demonstrated above, with $X = \mathbb{N} = \{1, 2, 3, \dots\}$. Now, every $f : X \to \mathbb{R}$ is not always simple, since it is of the form $f(x) = \sum_{k=1}^\infty f(k)\mathbb{1}_{\{k\}}(x)$.

Since f is not a finite sum, we 'approximate' f by f_n (which is the same sum for f, just with n instead of ∞) and use monotone convergence to get $\|f\|_{\ell^p}^p = \sum_{k=1}^\infty |f(k)|^p$.

An element of ℓ^p, $f : X \to \mathbb{R}$ $f = (f(1), f(2), \dots) \equiv (f_1, f_2, \dots) \equiv (f_k)_{k \geq 1}$ is just a **sequence**!

$$\therefore \boxed{\ell^p = \{\text{all } \mathbb{R}\text{-valued sequences } f = (f_k) : \sum_{k=1}^\infty |f(k)|^p < \infty\}}$$

Remark 1.5 Note the case $p = 1$, for ℓ^1, is just the space of all absolutely convergent series. Furthermore, if $p = \infty$ the norm instead becomes $\|f\|_\infty = \sup_n |x_n|$.

However, this example is only for the countably infinite case. Now let's consider the uncountable case.

Example 1.6 Again, adopting a similar convention to above, take $X = \mathbb{R}^n$ for some $n \in \mathbb{N}$, $\mathcal{A}$ the Borel σ-algebra, μ the Lebesgue measure. Then,

$$\|f\|_{L^p} = \left(\int |f|^p \, \mathrm{d}\mu\right)^{\frac{1}{p}}$$

where $1 \leq p \leq \infty$ and for the $p = \infty$ case we use the esssup, as demonstrated earlier in this section.

More generally, if $X \subset \mathbb{R}^n$ is open or closed, then $L^p(X, \mu)$ is the set of all measurable functions $f : X \to \mathbb{R}$ with finite L^p norm (as given above).

Theorem 1.7 Let $(X, \mathcal{A}, \mu)$ be any measure space. Then,

i $\|f\|_{L^p}$ defines a norm $\forall p \in [1, \infty]$. (The triangle inequality, also known as *Minkowski's inequality*, is $\|f + g\|_p \leq \|f\|_p + \|g\|_p \; \forall f, g \in L^p(\mu)$.

ii Hölder's inequality, if $\frac{1}{p} + \frac{1}{q} = 1$, $p, q \in [1, \infty]$, $\forall f \in L^p(\mu), \forall g \in L^q(\mu)$, then $fg \in L^1(\mu)$ and $\underline{\|fg\|_1 \leq \|f\|_p \|g\|_q}$.

iii $L^p(\mu)$ is complete.

Remark 1.8 i) and iii) together imply that $L^p(\mu)$ is Banach.

§1.3.1 Further examples of Banach spaces

- $C([a, b]) = \{f : [a, b] \to \mathbb{R} \text{ continuous}\}$ with $\|f\|_\infty = \sup_{[a,b]} |f(x)|$.
- $C^r([a, b]) = \{f : [a, b] \to \mathbb{R} \; r\text{−times cont. diff'able}\}$ with $\|f\|_{r,\infty} = \sup_{[a,b], 0 \leq k \leq r} |f^{(k)}(x)|$.
- Sobolev spaces (not specifically discussed in this course but very relevant for PDEs).

Proposition 1.9 $(C([0, 1]), \|\cdot\|_\infty)$ is a Banach space.

Proof. The general strategy to prove completeness of some $(X, \|\cdot\|)$ is as follows: given a Cauchy sequence $(f_n) \subset X$, find a candidate limit f. Show $\lim_{n\to\infty} \|f_n - f\| = 0$ and show $f \in X$.

Let $(f_n) \subset C([0, 1])$ be Cauchy $\implies \forall \varepsilon > 0, \exists N(\varepsilon)$ s.t. $\forall n, m \geq N$, $\|f_n - f_m\| < \varepsilon$. Consider $|f_n - f_m|$, $(f_n(x)) \subset \mathbb{R}$ is a Cauchy sequence in $\mathbb{R}$. By completeness of $\mathbb{R}$, so

$(f_n(x))$ converges in $\mathbb{R}$ to $f(x)$, *i.e.* $\lim_{n\to\infty} |f_n(x) - f(x)| = 0 \forall x \in [0,1]$.

Going back to $C([0,1])$, $\forall x, \forall n, m \geq N$: $|f_n(x) - f_m(x)| < \varepsilon$. But, $|f_n - f_m| \leq \|f_n - f_m\|_\infty \forall f$, so $\|f(x) - f_m(x)\|_\infty < \varepsilon$.

We also need to argue that $f \in C([0,1])$. For any $n \in \mathbb{N}$ and $\forall x, y \in [0,1]$, $|f(x) - f(y)| < |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)|$.

Now apply the '$\frac{\varepsilon}{3}$ argument'. Since $f_n \to f$, pick n s.t. $\|f_n - f\|_\infty < \frac{\varepsilon}{3}$. So, $|f_n(x) - f(x)| < \frac{\varepsilon}{3}$, $|f_n(y) - f(y)| < \frac{\varepsilon}{3}$. As n fixed, by continuity of $f_n \in C([0,1])$, pick δ such that $|f_n(x) - f_n(y)| < \frac{\varepsilon}{3}$ whenever $|x - y| < \delta$. Then $f \in C([0,1])$. $\square$

Remark 1.10 The above is a well-known result; the **uniform** convergence limit of a sequence of continuous is also continuous.

Note that pointwise limits are not always continuous, *e.g.* $f_n(x) = x^n$ for $x \in [0,1]$. Although if we take the space to be $L^1[0,1]$ instead of $C[0,1]$, then this converges to an indicator function - L^1 is a 'bigger' space and $C[0,1] \subset L^1[0,1]$.

For proofs, the following inequalities can be quite useful.

Lemma 1.11 (Triangle-like inequality 1) for $|x+y|^p$ with $p \geq 1$, we have

$$|x+y|^p \leq (|x| + |y|)^p \leq 2^p \max\{|x|, |y|\}^p \leq 2^p(|x|^p + |y|^p)$$

Or, for a better bound, we can use the **convexity** of $x \mapsto |x|^p$. Note: f convex if $f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}$.

Lemma 1.12 (Triangle-like inequality 2)

$$\frac{|x+y|^p}{2^p} \leq \frac{1}{2}\left(|x|^p + |y|^p\right) \implies |x+y|^p \leq 2^{p-1}\left(|x|^p + |y|^p\right)$$

§2 Linear Spaces

Let $(V, \bigoplus, \bigodot)$ be a linear space (or *vector space*) over $(\mathbb{K}, +, \cdot)$.

Definition 2.1 (Linear Metric Space) — Suppose the linear space, as given above, has $\mathbb{K}$, V equipped with the metrics $\rho_{\mathbb{K}}$, ρ_V respectively. Then, this linear space is called a **metric linear space** if $\bigoplus$ and $\bigodot$ are both continuous.

§2.1 Separability

In the following discussion, we take (V, ρ) with V a linear space and ρ a metric, so (V, ρ) is a metric space (we are not assuming any additional structure so it could *e.g.* normed, Banach, Hilbert,... *etc.*).

Recall the definition of ρ-open balls $B_\rho(x, \varepsilon) = \{y \in V : \rho(x, y) < \varepsilon\}$, $x \in V$, $\varepsilon > 0$.

Definition 2.2 (Dense) — A subset $D \subset V$ is called **dense** if $B_\rho(x, \varepsilon) \cap D \neq \forall x \in V$, $\varepsilon > 0$, (*i.e.* any open ball of any radius will contain a point of D).

Definition 2.3 (Separable) — V is **separable** if it has a <u>countable</u>, <u>dense</u> subset.

Proposition 2.4 ℓ^p, for $p \in [1, \infty)$, is separable (where ℓ^p is space of all sequences with finite p-norm).

Note that here ℓ^p is actually (ℓ^p, ρ), where ρ is the metric induced by $\|\cdot\|_p$.

Proof. Our goal is to find a countable, dense subset. Define $D := \bigcup_{n\geq 1} D_n$ where $D_k = \{(x_n) : x_n \in \mathbb{Q} \forall n \text{ and } x_n = 0 \forall n \geq k\}$ (all rational sequences which eventually equal zero). Notice $D_k \cong \mathbb{Q}^k$, which is countable so D is the union of countable sets which is countable.

Therefore we just need to show D is dense. Let $x = (x_n) \in \ell^p$. Pick $\varepsilon > 0$, first find $\tilde{x} = (\tilde{x}_n)$ with values in $\mathbb{Q}$ and $|x_n - \tilde{x}_n \leq \frac{\varepsilon}{2} \cdot 2^{-\frac{n}{p}}$.

$$\implies \|x - \tilde{x}\|^p = \sum_{n=1}^{\infty} |x_n - \tilde{x}_n|^p \leq \left(\frac{\varepsilon}{2}\right)^p \sum_{n=1}^{\infty} 2^{-n} \leq \left(\frac{\varepsilon}{2}\right)^p \tag{1}$$

Then, $\tilde{x} \in \overline{B(x, \frac{\varepsilon}{2}}$ and $\tilde{x} \in \ell^p$ (since $\|\tilde{x}\|_p \leq \|x\|_p + \|x - \tilde{x}\|_p < \infty$). Since $\tilde{x} \in \ell^p$, $\|\tilde{x}\|_p = \sum_{n\geq 1} |\tilde{x}_n|^p < \infty$, hence can find $k = k(x)$ s.t. $\sum_{n>k} ||^p < (\frac{\varepsilon}{2})^p$.

Define $y = (\tilde{x}_1, \ldots, \tilde{x}_k, 0, \ldots) \in D$, $\|\tilde{x} - y\|_p < \frac{\varepsilon}{2}$. This together with 1 implies $y \in B(x, \varepsilon)$. □

Proposition 2.5 $L^p(\mathbb{R}^n)$, for $p \in [1, \infty)$, is separable.

The definition of a basis is less clear in infinite dimensions than in finite dimensions, hence we demonstrate two specific examples of infinite dimensional bases below; *Schauder* basis and *Hamel* basis.

Definition 2.6 — Take a normed vector space $(X, \|\cdot\|)$. A **Schauder basis** of X is a sequence (e_n), $e_n \in X$ s.t. $\forall x \in X$, $\exists (x_n)$ with

$$\lim_{n\to\infty} \|x - \sum_{k=1}^{n} x_k e_k\| = 0$$

Remark 2.7 In contrast, a **Hamel basis** is a sequence (e_n), $e_n \in X$ s.t. $\forall x$, $\exists n, x_1, \ldots, x_n \in \mathbb{R}$ s.t. $x = \sum_{k=1}^{n} x_k e_k$ (*i.e.* we can write each x as a finite linear combination).

Clearly, Hamel basis $\implies$ Schauder basis.

Remark 2.8 Note that there exist Banach spaces (even separable spaces) which do not have a Schauder basis. There is emphasis on separability since if a Banach space has a Schauder basis then it is separable.

Lemma 2.9 If normed vector space $(X, \|\cdot\|)$ has a Schauder basis, then $(X, \|\cdot\|)$ is separable.

Proof. Define $D := \{\sum_{i=1}^{n} q_i e_i : n \in \mathbb{N}, q_i \in \mathbb{Q}\}$. This is clearly countable since Schauder basis is countable, and we are taking finite linear combinations.

Now we must show it's dense (*i.e.* we must show that we can approximate any point in X by a point in D to arbitrary precision).

Fix $x \in X$, $\varepsilon > 0$. We can find $n = n(x, \varepsilon)$ and $x_1, \ldots, x_n \in \mathbb{R}$ s.t. $\|\sum_{i=1}^{n} x_i e_i - x\| < \frac{\varepsilon}{2}$ (by definition of Schauder basis). Choose $q_k \in \mathbb{Q}$ for $k = 1, \ldots, n$ s.t. $|x_k - q_k| < \frac{\varepsilon}{2n \sum_{i=1}^{n} \|e_i\|}$ (since $\mathbb{Q}$ is dense in $\mathbb{R}$).

$$\implies \|\sum_{k=1}^{n} x_k e_k - \sum_{k=1}^{n} q_k e_k\| \le \sum_{k=1}^{n} |x_k - q_k| \|e_k\| < \frac{\varepsilon}{2}$$

□

Example 2.10 ℓ^p is separable for $p \in [1, \infty)$ with Schauder basis $e_n = (0, \ldots, 0, 1, 0, \ldots) \in \ell^p$, where the 1 is in the n-th position.

For L^p, the Schauder basis is made up of so-called '*Haar functions*'.

§2.2 Hilbert Spaces

Let H be a vector space over $\mathbb{R}$.

Definition 2.11 (Inner Product) — A **bilinear map** (*i.e* left and right linear) $H \times H \to \mathbb{R}$, $(x, y) \mapsto \langle x, y\rangle$, which is

- symmetric: $\langle x, y\rangle = \langle y, x\rangle \ \forall x, y \in H$
- and positive definite: $\langle x, x\rangle \geq 0$ with equality iff $x = 0$

is called an **inner product** over H.

Remark 2.12 $(H, \langle\cdot,\cdot\rangle)$ is an inner product space. For example, $H = \mathbb{R}$, $\langle x, y\rangle = xy$ $\forall x, y \in H$.

Theorem 2.13 (Cauchy-Scwharz) Let $(H, \langle\cdot,\cdot\rangle)$ be an inner product space with $\|x\| = \sqrt{\langle x, x\rangle}$. Then,

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\| \ \forall x, y \in H$$

Notice that Cauchy-Schwarz is true for all inner product spaces, ad it doesn't require any additional structure such as completeness. However, completeness can be extremely important as we saw with Banach spaces, and adding completeness to inner product spaces gives rise to the idea of *Hilbert* spaces.

Definition 2.14 (Hilbert Space) — A **Hilbert space** is an inner product space $(H, \langle\cdot,\cdot\rangle)$ which is complete with respect to the norm $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$.

Example 2.15 An example of a Hilbert space is $L^2(\mu)$ for the Lebesgue measure μ on $\mathbb{R}$ and $\langle f, g\rangle = \int fg \, \mathrm{d}\mu$.

As you can see, here the norm induced by the inner product above is just the L^2-norm.

Theorem 2.16 For Hilbert space H, $K \subset H$ subspace which is closed and convex (*i.e.* $\forall x, y \in K$, $t \in [0, 1]$ then $tx + (1-t)y \in K$), then $\forall y \in H$, there exists a **unique** $x_0 \in K$ s.t.

$$\delta := \inf_{x \in K} \|x - y\| = \|x_0 - y\|$$

Proof. We can assume $y = 0$ (else replace K by $K - y$). If $\delta = 0$, $x_0 = y$ and we are good.

So let $\delta > 0$, then $\exists (x_n) \subset K$ s.t. $\lim_{n\to\infty} \|x_n\| = \delta$. We want to show (x_n) is Cauchy, since then this would imply $\exists x_0 \in H$ s.t. $\|x_n - x_0\| \to 0$ and since K is closed, $x_0 \in K$ and we would be done.

Let $\varepsilon > 0$, $\exists N$ s.t. $\forall n \geq N$ $\|x_n\|^2 < \delta^2 + \frac{\varepsilon^2}{4}$. For all $n, m \geq N$,

$$\|x_n - x_m\|^2 = \underbrace{2\left(\|x_n\|^2 + \|x_m\|^2\right)}_{<4\delta^2+\varepsilon^2} - \underbrace{\|x_n + x_m\|^2}_{\leq -4\delta^2} < \varepsilon^2$$

Then $x_n, x_m \in K \implies \frac{x_n+x_m}{2} \in K$ by convexity of K, so $\|\frac{x_n+x_m}{2}\| \geq \delta$. $\square$

Lemma 2.17 (Parallelogram Law) For Hilbert space H,

$$2\|x\|^2 + 2\|y\|^2 = \|x+y\|^2 + \|x-y\|^2 \quad \forall x, y \in H$$

Definition 2.18 (Orthogonal) — For $x, y, \in H$, x is orthogonal to y, written $x \perp y$, if $\langle x, y\rangle = 0$.

Definition 2.19 (Orthogonal Complement) — For $S \subset H$, define its **orthogonal complement** $S^\perp = \{y \in H : \langle x, y\rangle = 0 \ \forall x \in S\}$.

Remark 2.20 You can show that $S^\perp$, as above, is closed using that $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{R}$ is continuous.

Corollary 2.21 For H Hilbert, if there exists $E \subset H$ closed subspace. Then $H = E \bigoplus E^\perp$ (*i.e.* $\forall x \in H, \exists e \in E, e' \in E^\perp$ s.t. $x = e + e'$), and $E \cap E^\perp = \{0\}$.

Proof. If $x \in E \cap E^\perp$, then $\langle x, x\rangle = 0 \implies x = 0$.

$\forall x \in H$, define $K = E + x$ which is closed and convex. Then by theorem 2.16, $\exists x_0 \in E$ s.t. $\|x - x + 0\| \leq \|x - \eta\| \ \forall \eta \in E$ (x_0 unique).

Consider map $\psi : t \in [0,1] \mapsto \frac{1}{2}\|(x - x_0) + t\eta\|^2$ has $t = 0$ as its minimum.

$$\begin{aligned} 0 = \psi'(t)\big|_{x=0} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{t^2}{2}\|\eta\|^2 + t\langle x - x_0, \eta\rangle\right)\bigg|_{t=0} &= t\|\eta\|^2 + \langle x - x_0, \eta\rangle\big|_{t=0} \\ &= \langle x - x_0, \eta\rangle \end{aligned}$$

This implies $x - x_0 \perp E$ and $x = (x - x_0) + x$. $\square$

§2.3 Finite vs. Infinite Dimensions

Take X to be a linear space (vector space).

Definition 2.22 (Equivalent Norms) — Two norms are **equivalent**, $\|\cdot\|_1$, $\|\cdot\|_2$, if $\exists C \in [1, \infty)$ s.t. $C^{-1}\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1$.

Proposition 2.23 If $\dim X < \infty$ (*i.e* $X \cong \mathbb{R}^n$), any two norms on X are equivalent.

Proof. (sketch) Write $x = \sum_{j=1}^n x_j e_j$, as some basis. Assume WLOG $\sum_{j=1} n|x_j| = 1$. The claim is: $\forall x \in X$,

$$c \leq \frac{\|x\|}{\sum_{j=1}^n |x_j|} \leq C$$

One can obtain this C by bounding $\|x\|$ with the maximum of finitely many terms by the triangle inequality. $\square$

Remark 2.24 Note that proposition 2.23 fails if $\dim X = \infty$.

Example 2.25 $X = C([0,1])$ with L^1-norm and L^∞-norm. Consider $f_n = t^n$, $t \in [0,1]$.

$$\|f_n\|_\infty = 1, \;\; \|f_n\|_1 = \int_0^1 |f_n(t)| \,\mathrm{d}t = \frac{1}{n+1} \to 0.$$

Proposition 2.26 If $(X, \|\cdot\|)$ is a normed space, consider a subspace $Y \subset X$ with $\dim Y < \infty$. Then, $(Y, \|\cdot\|)$ is complete. (here Y takes the induced norm)

Remark 2.27

1. If $\dim X < \infty$, then choose $Y = X$.
2. The above proposition fails if $\dim Y = \infty$, *e.g.* take $C([0,2]) = Y \subset X = L^1([0,2])$. Then take $f_n(t) = t^n$ when $t \in [0,1]$, and $f_n(t) = 1$ when $t \in [1,2]$. Then $(f_n) \subset Y \subset X$, $f_n \to f = \mathbb{1}_{[1,2]} \in X$, but clearly $f \notin Y$.

Corollary 2.28 For $(X, \|\cdot\|)$ normed space, $Y \subset X$ subspace with $\dim Y < \infty$. Then $Y \subset X$ is closed.

Proof. Let $(x_n) \subset Y$ be convergent, *i.e.* $\exists x \in X$, $\|x_n - x\| \to 0$. Need to show: $x \in Y$.

(x_n) is Cauchy ($\|x_n - x_m\| \leq \|x_n - x\| + \|x_m - x\|$). By proposition 2.26, $\|x_n - y\| \to 0$ for $y \in Y$, and we must have $y = x$ by uniqueness of limit. $\square$

§2.4 Compactness

Unlike in finite dimensions, in infinite dimensions, the idea of compactness as we know it with open covers is not easily as well defined. Hence, in this course we adopt the equivalent notion of **sequential compactness** instead as the definition.

Definition 2.29 (Compact) — For a metric space (X, ρ), a set $K \subset X$ is (sequentially) **compact** if every sequence $(x_n) \subset K$ has a ρ-convergent subsequence with its limit in K.

In finite dimensions, one may remember the well-known *Heine-Borel* theorem, which states that K compact $\iff$ K closed and bounded.

In infinite dimensions, this is not completely true however. The forward implication, K compact $\implies$ K closed and bounded, still remains true when $\dim X = \infty$.

However, the reverse implication fails if $\dim X = \infty$. For example take $K = \{e_n = (0, 0, \ldots, 1, 0, \ldots\} \in \ell^1$ (with the 1 in the n-th position). Then, $\|e_n\| = 1 \implies$

bounded, and use $\|e_n - e_m\| = 2\mathbb{1}_{n\neq m}$ to show K closed. So any convergent sequence in K is an eventually zero sequence. But, if $(x_n) := e_n$ then (x_n) has no convergent subsequence.

Example 2.30 As another example of "closed and bounded $\implies$ compact" failing in infinite dimensions, take $(C([0,1]), \|\cdot\|_\infty)$, $\overline{B_1} = \{f \in C : \|f\|_\infty \leq 1\}$. Then $\overline{B_1}$ is closed, bounded - but not compact.

Consider $f_n(t) = \sin(2^m \pi t)$ for $t \in [0,1]$ (it helps to draw out f_1 and f_2). Then, $\|f_n - f_m\|_\infty \geq 1 \ \forall n \neq m \implies (f_n)$ has no convergent subsequence.

N/B: the closed unit ball is **never** compact in infinite dimensions.

Theorem 2.31 Let $(X, \|\cdot\|)$ be a normed space. Then the following are equivalent,

1. $\dim X < \infty$ (finite dimensional).
2. $\overline{B_1} = \{x \in X : \|f\| \leq 1\}$ is compact.

For 1) $\implies$ 2), this is just Heine-Borel. For 2) $\implies$ 1), this uses the following lemma 2.32.

Lemma 2.32 (Riesz's Lemma) For $(X, \|\cdot\|)$ a normed space, $Y \subset X$ closed subspace with $Y \neq X$. Then for all $\varepsilon \in (0,1)$, $\exists x \in X \backslash Y$ s.t.

a) $\|x\| = 1$

b) $d(x, Y) := \inf_{y\in Y}\{\|x - y\|\} > 1 - \varepsilon$ (so it can be made arbitrarily close to 1).

Proof. (of theorem 2.31) We shall use lemma 2.32 to prove theorem 2.31, however the proof of the lemma itself is given after this proof for the theorem. We can use the lemma to show the case 2) $\implies$ 1) by considering the contrapositive, *i.e.* not 1) $\implies$ not 2). So assume $\dim X = \infty$.

We claim: $\exists (x_n) \subset \overline{B_1}$ s.t. $\|x_n - x_m\| \geq \frac{1}{2} \ \forall n \neq m$ (which implies that $\overline{B_1}$ is not sequentially compact).

Let (y_n) be a sequence of linearly independent vectors (if we couldn't have this for some n, then $\dim X$ would be finite).

$Y_n = \operatorname{span}\{y_1, \dots, y_n\} \subset X$ is closed. Pick $x_1 = \frac{y_1}{\|y_1\|} (\in \overline{B_1})$. For $n \geq 2$, suppose $x_1, \dots, x_{n-1}$ are given.

Apply lemma 2.32 with $X = Y_n$, $Y = Y_{n-1}$, $\varepsilon = \frac{1}{2}$.

Clearly, $(x_n) \subset \overline{B_1}$, and we have from the lemma $\forall m > n$,

$$\|x_m - x_n\| \geq d(x_m, Y_n) \geq \underbrace{d(x_m, Y_m - 1)}_{\text{since } Y_n \subset Y_{m-1}} > \frac{1}{2}$$

□

Proof. (of lemma 2.32) Pick $x^* \in X \backslash Y$. Since Y is closed, $d(x^*, Y) > 0$. This is an infimum, so we can find $y^* \in Y$ s.t. $d(x^*, Y) \leq \|x^* - y^*\| < \frac{d(x^*,Y)}{1-\varepsilon}$.

Set $x = \frac{x^* - y^*}{\|x^* - y^*\|}$. Now, a) is clear and for b), we have $\forall y \in Y$,

$$\|x - y\| = \left\| \frac{x^* - \overbrace{y^* - \|x^* - y^*\| y}^{\in Y}}{\|x^* - y^*\|} \right\| \geq \frac{d(x^*, Y)}{\|x^* - y^*\|} > 1 - \varepsilon$$

□

Example 2.33 (Application of lemma 2.32) If $\dim X < \infty$, and $X = \mathbb{R}^3$, $Y \cong \mathbb{R}^2$, so there is a vector which has distance 1 from the plane Y.

N/B: we can take $\varepsilon = 0$ in lemma 2.32 iff X is 'reflexive', which we shall see later in the course.

§3 Linear Operators

In this section, we take $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ to be normed spaces, and $A : X \to Y$ is **linear**.

Definition 3.1 (Bounded) — $A : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ is **bounded** if $\exists C \in (0, \infty)$ s.t. $\|Ax\|_Y \le C\|x\|_X$ $\forall x \in X$.

If A is bounded, then $\|A\| := \sup_{\|x\|\le 1} \|Ax\|_Y$ is the best possible 'C', and $\|A\|$ is called the **operator norm** (which is a norm on $\mathcal{L}(X, Y)$ - this is defined later on in this section).

Note that in the sense of linear operators, boundedness is the same as continuity (boundedness $\iff$ continuity), as demonstrated in the following theorem.

Theorem 3.2 The following are equivalent,

i) A is continuous in $x_0 \in X$.

ii) A is continuous in every $x \in X$.

iii) A is Lipschitz continuous ($\exists L : \|Ax - Ay\|_Y \le L\|x - y\|_X \forall x, y \in X$).

iv) A is bounded.

Proof. iv) $\implies$ iii) by linearity of A, $\forall x_1 \ne x_2 \in X$,

$$\|Ax_1 - Ax_2\|_Y = \|A(x_1 - x_2)\|_y = \|x_1 - x_2\|_X \left(\|A\left(\frac{x_1 - x_2}{\|x_1 - x_2\|}\right)\|_Y\right)$$
$$\le \|A\| \cdot \|x_1 - x_2\|_x \quad \text{so take } L = \|A\|.$$

Further, iii) $\implies$ ii) $\implies$ i) is clear. Now, there's just left to show i) $\implies$ iv).

Assume $\|A\| = \infty$, so we can find $(x_n) \subset X$ with $\|x_n\| \le 1$, and $0 < \|Ax_n\|_Y \to \infty$ as $n \to \infty$.

Set $z_n := \frac{x_n}{\|Ax_n\|_Y}$, then $\|z_n\|_X \to 0$ as $n \to \infty$, but

$$\|A(x_0 + z_n) - Ax_0\|_Y = \|Az_n\|_Y = 1 \not\to 0 \text{ as } n \to \infty$$

□

Corollary 3.3 If $\dim X < \infty$, with $A : X \to Y$ linear. Then A is continuous

Proof. $\|x\|_* := \|x\|_X + \|Ax\|_Y$ defines a norm on X. By proposition 2.23, $\exists C > 0$ s.t. $\|x\|_* \le C\|x\|_X$ $\forall x \in X$. But since $\|Ax\|_Y \le \|x\|_X$, A is bounded, hence continuous by theorem 3.2. □

Example 3.4 Take $X = Y = C([0,1])$ with $\|\cdot\|_X = \|\cdot\|_1$, $\|\cdot\|_y = \|\cdot\|_\infty$ and $A = \text{id}$. Then A is not continuous; we can show that it is not bounded. Take,

$$f_n(t) = \begin{cases} 2n^2 t & 0 \leq t \leq \frac{1}{2n} \\ 2n - 2n^2 t & \frac{1}{2n} < t \leq \frac{1}{n} \end{cases}$$

Essentially, this looks like a 'tent function', which attains value of n at its peak, for $t \in [0, \frac{1}{n}]$ and 0 everywhere else.

Then $\|f_n\|_1 = 1$, $f_n \in X\ \forall n$, but $\|A\| \geq \sup_n \|A \cdot f_n\|_Y = \sup_n \|f_n\|_\infty = \sup_n n = \infty$.

In fact, unboundedness is rather common, so care is needed.

Example 3.5 A more classical example is $X = C^1([0,1])$, $Y = C([0,1])$, $A : X \to Y$ with "$A = \frac{\mathrm{d}}{\mathrm{d}x}$". This is well-defined since if $f \in X$, $Af \in Y$.

Take $\|\cdot\|_Y = \|\cdot\|_\infty$ and $\|\cdot\|_x = \|\cdot\|_Y$, then A is unbounded. Indeed, take $f_n(t) = \sin(nt)$, or even $f_n(t) = t^n$. Then it is easy to check that $\|f_n\|_X = 1$ but $\|Af_n\|_Y = n \to \infty$.

Remark 3.6 If instead one sets $\|\cdot\|_X = \|\cdot\|_{C^1} = \|f\|_\infty + \|f'\|_\infty$ then the above f_n's are of no use (and in fact A is bounded!).

Now let's define the following important space

$$\mathcal{L}(X,Y) = \{A : X \to Y \quad : \quad A \text{ linear and continuous}\}$$

$\mathcal{L}(X,Y)$ is a normed vector space with the following norm:

$$\|A\|_{\mathcal{L}(X,Y)} = \|A\| = \sup_{\|x\|_X \leq 1} \|Ax\|_Y = \sup_{x \neq 0} \frac{\|Ax\|_y}{\|x\|_X}$$

and one has the following useful inequality, $\forall x \in X$, $\|Ax\|_Y \leq \|A\| \cdot \|x\|_X$.

Remark 3.7 By $\mathcal{L}(X,Y)$, we really mean $\mathcal{L}\left((X, \|\cdot\|_X), (Y, \|\cdot\|_Y)\right)$.

N/B: If $X = Y$ and $\|\cdot\|_X = \|\cdot\|_Y$, one sets $\mathcal{L}(X,X) = X$

Theorem 3.8 If $(Y, \|\cdot\|_Y)$ is Banach, then so is $\left(\mathcal{L}(X,Y), \|\cdot\|_{\mathcal{L}(X,Y)}\right)$.

Corollary 3.9 Take $A : X \to Y$ to be continuous, and $K \subset X$ compact. Then $A(K) = \{Ax : x \in K\}$ and $A(K) \subset Y$ is compact.

Proof. Fix some $(y_n) \subset A(K)$, then we need to find a convergent subsequence (y_{n_k}). By definition of $A(K)$, $y_n = Ax_n$ for some $x_n \in K$, so $(x_n) \subset K$ has a convergent subsequence by compactness of K - call it (x_{n_k}).

We claim: $y_{n_k} = Ax_{n_k}$ is convergent. Indeed let $\|x_{n_k} - x\|_X \to 0$ as $k \to \infty$. By (Lipschitz) continuity,

$$\|y_{n_k} - Ax\|_Y = \|Ax_{n_k} - Ax\|_Y \leq L\|x_{n_k} - x\|_X \to 0 \quad \text{as } k \to \infty$$

so $(y_{n_k}) \subset Y$ converges and the limit is Ax. □

§3.1 Duality

Take 2 normed spaces $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$. Recall,

$$\mathcal{L}(X,Y) = \{A : X \to Y \quad : \quad A \text{ linear and bounded}\}$$

and $\mathcal{L}(X,Y)$ is Banach if Y is with norm $\|A\| = \|A\|_{\mathcal{L}(X,Y)} = \sup_{\|x\|\leq 1} \|Ax\|_Y = \sup_{\|x\|=1} \|Ax\|_Y$.

An important special case is take $Y = \mathbb{R}$ (which is finite dimensional so we can take any norm on it).

Definition 3.10 (Dual Space) — The **dual space** of X is X^*, with $X^* = \mathcal{L}(X, \mathbb{R}$ (the space of all bounded, linear operators from X to $\mathbb{R}$).

Remark 3.11 X^* is always Banach (as soon as X is normed, even though X may not be Banach itself). So going to the dual space gives us completeness for free. The elements of X^* are called bounded, linear functionals on X.

§3.1.1 Duality in Hilbert spaces

Take an inner product space $(H, \langle\cdot,\cdot\rangle)$ over $\mathbb{R}$. Consider the map $\forall y \in H$,

$$\Lambda_y : H \to \mathbb{R} \quad , \quad x \mapsto \langle y, x\rangle$$

This is a linear map, since the inner product is linear.

Lemma 3.12 i) $\Lambda_y \in H^*$ (the dual space of H).

ii) the map $\Lambda : H \to H^*$ is a **linear isometry** (*i.e.* $\|\Lambda(y)\| = \|y\|$, or the input and output have the same norm).

Proof. For i), linearity is clear since the inner product is linear. For boundedness of Λ_y, $\forall x \in H$, $\|x\|_H \leq 1$,

$$\underbrace{|\Lambda_y(x)|}_{\text{Euc. norm on } \mathbb{R}} = |\langle y, x\rangle| \leq \underbrace{\|y\|_H \cdot \|x\|_H}_{\text{Cauchy-Schwarz}} \leq \|y\|_H (< \infty)$$

For ii) Λ is linear so just need to check isometry. Check $\|\Lambda(y)\| = \|\Lambda_y\|_* \overset{?}{=} \|y\|_H$.

We know from i) $\|\Lambda_y\|_* \leq \|y\|_H$ so take $x = \frac{y}{\|y\|_H}$ in Λ_y to find

$$|\Lambda_y(y)| = \|y\|_H \left(\implies \|\Lambda_y\| \geq \|y\|_H \right)$$

□

Theorem 3.13 (Riesz's Representation Theorem) For every $l \in H^*$, $\exists!\ y \in H$ s.t. $l = \Lambda_y$.

Notice how this theorem states that Λ_y produces an isomorphism between H and H^*! For example, $(\ell^2)^* \cong \ell^2$.

Proof. (sketch) At the moment this is just a sketch of the proof, when I get some more time I may come back later and fill in the explicit details.

First show uniqueness of l. Then show existence, for which one can assume WLOG $\|l\|_* = 1$ (since if $l(\cdot) \equiv 0$, then choose $y = 0$).

Consider a sequence $(y_n) \subset H$ s.t. $l(y_n) \to 1 (= \|l\|_*)$, and $\|y_n\| = 1\ \forall n$.

Claim 1: (y_n) is Cauchy.

Claim 2: $l(\cdot) = \Lambda_y(\cdot)$. And then we are done. □

§3.1.2 Duality of Banach Spaces

Note that ideas from Hilbert spaces in the previous few results do not follow over to Banach spaces (but the converse is true).

Theorem 3.14 $\forall p \in (1, \infty)$, $(\ell^p)^* \cong \ell^q$, where $\frac{1}{p} + \frac{1}{q} = 1$.

For the proof, we take $y = \ell^q$, define $\Lambda_y : \ell^p \to \mathbb{R}$, with $x \mapsto \sum_{n\geq 1} x_n y_n$. We need the following lemma to make sure this sum doesn't explode, and along with the fact that Λ is surjective (which one would need to show explicitly), then one can conclude theorem 3.14.

Lemma 3.15

i) $\Lambda_y \in (\ell^p)^*$.

ii) $\Lambda : \ell^q \to (\ell^p)^*$, $y \mapsto \Lambda_y$ is a linear isometry.

> **Remark 3.16** In fact, theorem 3.14 extends to $p = 1$, and it is also true that for any measure space $(X, \mathcal{A}, \mu)$, $(L^p(\mu))^* \cong L^q(\mu)$.

Consider the space $c_0 = \{(x_n) : \lim_{n\to\infty} x_n = 0\} \subset \ell^\infty$, then $(c_0, \|\cdot\|_\infty)$ is Banach.

We claim: $(c_0)^* \cong \ell^1$.

To show this one needs to show the following:

1. $\Lambda_y : c_0 \to \mathbb{R}$, $x \mapsto \Lambda_y(x) = \sum_{n\geq 1} x_n y_n$.
2. The map $\ell^1 \to c_0^*$, $y \mapsto \Lambda_y$ is a linear isometry.

§3.2 Dual Operators

Consider normed spaces $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ over $\mathbb{R}$. $A : X \to Y$ is bounded, linear (so $A \in \mathcal{L}(X, Y)$), and the duals are X^*, Y^*.

Definition 3.17 (Dual Operator) — The **dual operator** of A is the linear operator $A^* : Y^* \to X^*$ defined by

$$(X^* \ni)\ A^* y^* := y_0^* A \quad : \quad X \to \mathbb{R} \quad \forall y^* \in Y^*$$

Notation: Often for a linear map $l \in X^*$, instead of writing $l(x)$, we write $\langle l, x\rangle$.

Then the above definition is equivalent to: $\langle A^* y^*, x\rangle = \langle y^*, Ax\rangle \quad \forall x \in X, \forall y \in Y^*$.

Later, we will see A^* is in fact bounded and $\|A^*\| = \|A\|$ (using Hahn-Banach). In finite dimensions, $A \in \mathbb{R}^{m\times n}$ induces a linear operator $L_A : \mathbb{R}^n \to \mathbb{R}^m$, $L_A x = Ax$. If $A^T \in \mathbb{R}^{m\times n}$ is the transpose of A, and $i_k : \mathbb{R}^k \to \left(\mathbb{R}^k\right)^*$ is the canonical isomorphism (by *Riesz's* representation thm), then

$$(L_A)^* \circ i_m = i_n \circ L_{A^T} \quad : \quad \mathbb{R}^m \to (\mathbb{R}^n)^*$$

$\implies$ dual operators in finite dimensions are nothing but transposes.

- More generally, if H Hilbert (or $\mathbb{R}$ for instance), $A : H \to H$, $A : \mathcal{L}(H)$ then with $I : H \to H^*$, the canonical isomorphism (by *Riesz's* representation thm), the operator
$$\tilde{A}^* := i^{-1} \circ A^* \circ i \quad : \quad H \to H$$
is called the **adjoint** operator of A (and one writes $A^* \equiv \tilde{A}^*$).

Hence, $\langle \tilde{A}^* y, x\rangle = \langle y, Ax\rangle$ $\forall x, y \in H$, where $\langle \cdot, \cdot\rangle$ is the inner product on H. If $A^* = A$ then $\langle Ay, x\rangle = \langle y, Ax\rangle$ and A is called **self-adjoint**.

§3.3 Hahn-Banach and Applications

Now we will go back to the case of X simply as a linear space (vector space) over $\mathbb{R}$.

Definition 3.18 (Sublinear Map) — The map $p : X \to \mathbb{R}$ is called **sublinear** if

i) $p(\alpha x) = \alpha p(x)$, $\quad \forall \alpha \geq 0$, $x \in X$.

ii) $p(x + y) \leq p(x) + p(y)$, $\quad \forall x, y \in X$.

Notice that linear $\implies$ sublinear. In the definition the map p could be $p \in X^*$, or $p(x) :=$ any norm on X.

Theorem 3.19 (Hahn-Banach) Let $M \subset X$ be a linear subspace. Let $p : X \to \mathbb{R}$ be sublinear, $f : M \to \mathbb{R}$ be linear, with

$$f(x) \leq p(x)\ \forall x \in M$$

Then, there exists a linear map $F : X \to \mathbb{R}$ with $F|_H = f$ and $F(x) \leq p(x)\ \forall x \in X$.

Remark 3.20 (Geometrical Intuition) Take $X = \mathbb{R}^n$, $0 \in M \subset X$ open, convex and fix a $x_1 \notin M$. Then I can find a linear map $f : X \to \mathbb{R}$ s.t. f is "separating x_1 from M".

$$i.e. \begin{cases} f(x) < a & \forall x \in M \\ f(x_1) \geq a & \text{for some } a \neq 0 \end{cases}$$

We can normalise so WLOG take $a = 1$. We introduce $p(x) := \inf\{r > 0 : \frac{x}{r} \in M\}$ (*e.g.* if $M = \{x \in X : \|x\| \leq 1\}$ in a normed space $(X, \|\cdot\|)$, then $p(x) = \|x\|$).

We can check that p is sublinear (by convexity) and $p(x) < 1 \iff x \in M$. To show this, if $p(x) < 1$ then $\exists \varepsilon > 0$ s.t. $\frac{x}{1-\varepsilon} \in M$ so

$$x = (1-\varepsilon)\underbrace{\frac{x}{1-\varepsilon}}_{\in M} + \underbrace{\varepsilon \cdot 0}_{\in M} \in M \qquad \text{by convexity}$$

For the reverse direction, $x \in M \implies \frac{x}{1-\varepsilon} \in M$ for some $\varepsilon > 0$ by openness of M. This implies $p(x) \leq 1 - \varepsilon < 1$.

N/B: to find f with the desired separation property, we need to ensure: $f(x_1) = 1$ **and** $f(x) \leq p(x)$ $\forall x \in X$.

We omit the proof of *Hahn-Banach* (H-B) from this document (for one, it is rather long to write up), although there are plenty of proofs of it which can be found online.

There is an analogous complex version of H-B for X over $\mathbb{C}$, $p : X \to \mathbb{R}$ is called $\mathbb{C}$-sublinear if i) $p(\alpha x) = |\alpha| p(x)$ $\forall x \in X$, $\alpha \in \mathbb{C}$ and ii) holds. But in this course we focus on spaces X over $\mathbb{R}$

§3.3.1 Applications of Hahn-Banach

We take X to be a normed space $(X, \|\cdot\|)$. We present some further results which come as a consequence of H-B, although we do not give proofs for all of the results.

Corollary 3.21 (Extending Linear Functionals) Let $M \subset X$ be a linear space (which inherits $\|\cdot\|$ from X), $f \in M^*$. Then $\exists F \in X^*$ s.t. $F|_m = f$ and $\|F\|_{X^*} = \|f\|_{M^*}$.

Proof. Define $p : X \to \mathbb{R}$ via: $p(x) = \|x\|_X \cdot \|f\|_{M^*}$ (this is sublinear) and $\forall x \in M$, $f(x) \leq |f(x) = \|x\|_X \cdot \frac{|f(x)|}{\|x\|_X} \leq \|x\|_X \cdot \|f\|_{M^*} = p(x)$. Then apply H-B. □

For $x^* \in X^*$, recall the notation $x^*(x) = \langle x^*, x \rangle$, $x \in X$.

Theorem 3.22 $\forall x \in X$, $\exists x^* \in X^*$ s.t. $\langle x^*, x \rangle = \|x\|_X^2 = \|x^*\|_{X^*}^2$.

Proof. Let $M := \text{span}\{x\}$. Define $f(tx) = t\|x\|_X^2$ $\forall t \in \mathbb{R}$, then $f : M \to \mathbb{R}$ is linear and $\|f\|_{M^*} = \sup_{\|tx\|_X \leq 1} |f(tx)| = \|x\|_X$, so $f \in M^*$.

Apply corollary 3.21 to extend f to $x^* := F \in X^*$, with

$$\|x^*\|_{X^*} = \|f\|_{M^*} = \|x\|_X \text{ and } \langle x^*, x\rangle = \underbrace{f(x)}_{x^*|_M = f} = \|x\|_X^2$$

$\square$

Remark 3.23 This gives the dual characterisation of the norm, which we shall see later.

We can use H-B to 'separate' all sorts of things! For instance, separating points using dual elements as the following proposition demonstrates.

Proposition 3.24 $\forall x, y \in X$ with $x \neq y$, $\exists l \in X^*$ s.t. $l(x) \neq l(y)$.

Proof. Choose a $l \in X^*$ according to theorem 3.22 with $y - x$ in place of x. Then, $l(x-y) = l(x) - l(y) = \|y - x\|_X^2 > 0$. $\square$

Further, we can separate points from closed subspaces, as in the following theorem.

Theorem 3.25 Let $M \subset X$ be a linear, closed space, and assume $x_0 \notin M$ s.t.

$$d = \operatorname{dist}(x_0, M) := \inf_{x \in M} \|x_0 - x\|_M > 0$$

Then $\exists l \in X^*$ with $l|_M = 0$ and $\|l\|_{X^*} = 1$, $l(x_0) = d$.

This theorem has a lot of mileage. For instance, you get,

- If we apply theorem 3.25 with $M = \{0\}$, $x_0 = \frac{x}{\|x\|_X}$, we recover theorem 3.22 with $x^* := \|x\|_X l$.
- We can use theorem 3.25 to prove the following new theorem.

Theorem 3.26 If X^* is separable, then X is separable.

In particular, this shows $(\ell^\infty)^* \ncong \ell^1$, since ℓ^1 is separable but ℓ^∞ is not separable.

From theorem 3.22, we can obtain a **dual characterisation of the norm**, as given in the following corollary.

Corollary 3.27 i) $\forall x \in X$, $\|x\|_X = \sup_{\|x^*\|_{X^*} \leq 1} |\langle x^*, x\rangle| \quad (= \sup_{x^* \in X^*} |\langle x^*, x\rangle|)$

ii) $\forall x^* \in X^*$, $\|x^*\|_{X^*} = \sup_{\|x\|_X \leq 1} |\langle x^*, x\rangle| \quad (= \sup_{x \in X} |\langle x^*, x\rangle|)$

N/B: the sup in i) is always achieved.

Proof. For $x = 0$, the RHS of i) is $= 0$, so let $x \neq 0$. Consider "$\geq$", by homogeneity, we can assume $\|x\|_X = 1$ if $x^* \in X^*$ s.t. $\|x^*\|_{X^*} \leq 1$, then

$$|\langle x^*, x\rangle| \leq \|x^*\|_{X^*}\|x\|_X \leq \|x\|_X$$

Now consider "$\leq$". By theorem 3.22, $\exists x^* \in X^*$ s.t. $|\langle x^*, x\rangle| = 1 = \|x\|_X$, and so the sup is achieved.

Item ii) is just the consequence of the definition of $\|\cdot\|_{X^*}$. □

Theorem 3.28 Let X, Y be normed and $A \in \mathcal{L}(X, Y)$. The dual operator $A^* : Y^* \to X^*$ is bounded and $\|A^*\|_{\mathcal{L}^*(Y^*,X^*)} = \|A\|_{\mathcal{L}(X,Y)}$.

Proof.

$$\begin{aligned}\|A^*\| &= \overbrace{\sup_{\|y^*\|_{Y^*}=1} \|A^*y^*\|_{X^*}}^{\text{defn of } \|\cdot\|} = \overbrace{\sup_{\|y^*\|_{Y^*}=1} \sup_{\|x\|_X=1} |\langle A^*y^*, x\rangle|}^{\text{defn of } \|\cdot\|_{X^*}} \\ &= \underbrace{\sup_{\|x\|_X=1} \sup_{\|y^*\|_{Y^*}=1} |\langle y^*, Ax\rangle|}_{\text{defn of } A^*} = \sup_{\|x\|_X=1} \|Ax\|_Y = \|A\|\end{aligned}$$

□

§4 Baire Category and UBP

§4.1 Baire Category

The motivation for *Baire* category arises from Baire's original question:

> *Let* $f_n : [0,1] \to \mathbb{R}$, $n \geq 1$ *be continuous* $\forall n$. *Assume* $f(x) := \lim_{n\to\infty} f_n(x)$ *exists* $\forall x \in [0,1]$. *Then, does* f *have at least one point of continuity?*

What we seek here, in this section, is a topological characterisation of the 'size' of sets.

Recall for a metric space (X, d) with $A \subset X$, the following definitions of the interior and closure of A.

Definition 4.1 (Interior of A) — We can define the interior of $A \subset X$ as follows

$$\operatorname{int}(A) = \bigcup_{G \subset A,\ G \text{ open}} G$$

Definition 4.2 (Closure of A) — We can define the closure of $A \subset X$ as follows

$$\overline{A} = \bigcap_{U \supset A,\ U \text{ closed}} U$$

Lemma 4.3 If (X, d) is a complete metric space, $X \neq \emptyset$, and if you can write $X = \bigcup_{k=1}^{\infty} A_k$ with A_k closed (*i.e.* $A_k = \overline{A_k}$) then $\exists k$ s.t. $\operatorname{int}(A_k) \neq \emptyset$.

Proof. Assume for contradiction, not true, then $X = \bigcup_{k=1}^{\infty} A_k$ with A_k closed and $\operatorname{int}(A_k) = \emptyset$ $\forall k \geq 1$.

Pick $x_1 \in X \backslash A_1$. Since $X \backslash A_1 = X \cup A^C$ is open, then one can find $0 < r_1 < {}^{\prime\prime -1}$ s.t. $B(x_1, r_1) \subset X \backslash A_1$.

Repeat to find $x_2 \in B(x_1, \frac{r_1}{2}) \backslash A_2$ and $0 < r_k < 2^{-2}$ s.t. $B(x_2, r_2) \subset X \backslash A_2$.

We claim: $\forall k \geq 1$, $\exists x_k \in X \backslash A_k$ and $0 < r_k < 2^{-k}$ s.t. (by induction over k)

$$B\left(x_{k+1}, r_{k+1}\right) \subset B\left(x_k, \frac{r_k}{2}\right) \subset B(x_k, r_k) \subset X \backslash A_k$$

The sequence (x_k) is Cauchy (since $d(x_m, x_k) \leq \frac{r_k}{2} < 2^{-(k+1)}$ $\forall m \geq k \geq 1$). By completeness, $\exists x_* \in X$, $d(x, x_*) \to 0$ if true then, $x_* \in X \backslash \bigcup_{k=1}^{\infty} A_k = \emptyset$, which is a contradiction. $\square$

Definition 4.4 (Meager and Fat Sets) — For a metric space (X, d)

i) $A \subset X$ is called **meager** (or of 1^{st} Baire category, *i.e.* $\operatorname{cat}(A) = 1$) if you can write $A = \bigcup_{k=1}^{\infty} A_k$ with nowhere dense sets A_k. (A_k is nowhere dense if $\operatorname{int}(\overline{A_k}) = \emptyset$)

ii) $A \subset X$ is called **fat** (or of 2^{nd} Baire category, *i.e.* $\operatorname{cat}(A) = 2$) if it is not meager.

We can rewrite lemma 4.3 with the above language, which becomes the following **Baire category** theorem.

Theorem 4.5 (Baire Category) Let (X, d) be a complete metric space, $X \neq \emptyset$, then $\operatorname{cat}(X) = 2$.

Remark 4.6 *Baire category* (BC) shows the existence of fat sets (*e.g.* $\mathbb{R}, \mathbb{C}, \dots$ etc.). Here are some useful properties of BC.

- Note that $\mathbb{Q} \subset \mathbb{R}$ is meager.
- If A is meager and $A' \subset A$, then A' is meager.
- 'Meager-ness' is closed under countable unions, so if A_k meager, then $\bigcup_{k=1}^{\infty} A_k$ is meager.

What about the set $\mathbb{R} \backslash \mathbb{Q}$? You can show this is fat with the following corollary of BC.

Corollary 4.7 Let (X, d) be complete, $X \neq \emptyset$, and if $A \subset X$ with $\operatorname{cat}(A) = 1$, then $\operatorname{cat}(X \backslash A) = 2$ and $X \backslash A$ is dense.

The proof of this arises from the fact that $U \subset X$ open, dense $\iff$ $A = X \backslash U$ closed, nowhere dense.

Corollary 4.8 If $\emptyset \neq U \subset X$ is open, then $\operatorname{cat}(U) = 2$.

Proof. If $\operatorname{cat}(U) = 1$, then this implies $X \backslash U$ is dense, so we can write $X = \overline{X \backslash U} = X \backslash U$, *i.e.* $U = \emptyset$, which is a contradiction. □

What about topological vs measure-theoretic notions of size? Take $X = \mathbb{R}$ and λ the Lebesgue measure, then does $\lambda(A) = 0 \implies A$ meager? Or, does $A \subset \mathbb{R}$ meager $\implies \lambda(A) = 0$? The answer is no and no!

Example 4.9 Take $\mathbb{Q} = \{q_1, q_2, \dots\}$ and define $U_j = \bigcup_{k \geq 1} \left(q_k - \frac{1}{2^{j+k+1}}, q_k + \frac{1}{2^{j+k+1}} \right)$, then $\lambda(U_j) \leq \sum_{k \geq 1} 2 \cdot \frac{1}{2^{j+k+1}} = 2^{-j}$.

U_j is open and $\overline{U_j} \supset \overline{\mathbb{Q}} = \mathbb{R}$ so $\overline{U_j} = \mathbb{R}$, *i.e.* U_j is dense. Then by corollary 4.7 $A := X \backslash U_j$ is nowhere dense and so $A := \bigcup_{j \geq 1} A_j$ is meager.

Hence, $X \backslash A = \bigcap_{j \geq 1} U_j$ is fat and $\lambda(U) = \lim_{n \to \infty} \lambda(U_j) = 0 \implies \lambda(A) = \infty$.

§4.2 Uniform Boundedness Principle

Baire's category theorem leads to the uniform boundedness principle.

Theorem 4.10 Let (X, d) be complete and $(f_\lambda)_{\lambda\in\Lambda}$ be a family of continuous functions $f_\lambda : X \to \mathbb{R}$. If (f_λ) is bounded pointwise, *i.e.* (note that the bound below can depend on x)

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

then $\exists B \subset X$ open ball s.t. $\sup_{\lambda\in\Lambda, x\in B} |f_\lambda(x)| < \infty$ (this is uniform in $x \in B$), *i.e.* (f_λ) is **uniformly bounded** on B.

Remark 4.11 (f_λ) need not be linear.

Proof. For $k \geq 1$, consider the closed set $A_k = \{x \in X : \forall\lambda \in \Lambda : |f_\lambda(x)| \leq k\} = \bigcap_{\lambda\in\Lambda}\{|f_\lambda| \leq k\}$ (which is closed as f_λ continuous).

Clearly $\bigcup_{k=1}^{\infty} A_k = X$ and since X is complete, by lemma 4.3 $\exists k_0$ s.t. $\operatorname{int}(A_{k_0}) \neq \emptyset$. Then we can pick $B \subset A_{k_0}$. $\square$

Incorporating the linear structure, we can get the following useful result known as the Uniform Boundedness Principle (which is also called Banach-Steinhaus theorem).

Corollary 4.12 (Banach-Steinhaus) If X, Y are normed spaces, and X is complete with $(A_\lambda)_{\lambda\in\Lambda} \subset \mathcal{L}(X, Y)$. If $(A_\lambda)_{\lambda\in\Lambda}$ are bounded pointwise, *i.e.*

$$\sup_{\lambda\in\Lambda} \|A_\lambda x\|_Y < \infty \ \forall x \in X$$

then (A_λ) is bounded uniformly, *i.e.* $\sup_{\lambda\in\Lambda} \|A_\lambda\|_{\mathcal{L}(X,Y)} < \infty$.

Proof. For $\lambda \in \Lambda$, we define the continuous map $f_\lambda : X \to \mathbb{R}$ by $f_\lambda(x) = \|A_\lambda x\|_Y$. By assumption on A_λ, theorem 4.10 applies and yields $B = B_r(x_0) \subset X$ s.t. $\sup_{\lambda\in\Lambda, x\in B} |f_\lambda(x)| < \infty$. This gives for all $\|x\|_X < 1$ and $\lambda \in \Lambda$:

$$\|A_\lambda x\|_Y = \frac{1}{r}\|A_\lambda(x_0 + rx) - A_\lambda(x_0)\|_Y \leq \frac{1}{r}\|A_\lambda(x_0 + rx)\|_Y + \frac{1}{r}\|A_\lambda(x_0)\|_Y \leq M$$

and so $\|A_\lambda\|_{\mathcal{L}(X,Y)} \leq M$. $\square$

§4.2.1 Applications of UBP

Theorem 4.13 Let X, Y be normed spaces and let X be complete. let $A_k \in \mathcal{L}(x, Y)$ and let (A_k) converge pointwise to $A : X \to Y$, (*i.e* $\forall x \in X$, $\|A_k x - Ax\|_Y \to 0$ as $k \to \infty$).

Then A is linear and continuous, *i.e.* $A \in \mathcal{L}(X, Y)$, and

$$\|A\|_{\mathcal{L}(X,Y)} \leq \lim_{k\to\infty} \inf_k \|A_k\|_{\mathcal{L}(X,Y)} < \infty$$

Proof. $(A_k x) \subset Y$ is convergent and hence bounded, so corollary 4.12 (BS) applies and yields $\sup_k \|A_k\|_{\mathcal{L}(X,Y)} < \infty$; this shows $\lim_{k\to\infty} \inf_k \|A_k\|_{\mathcal{L}(X,Y)} < \infty$.

Pick a subsequence k_j s.t. $\|A_{k_j} \to_{j\to\infty} \lim_{k\to\infty} \inf_k \|A_k\| := M$. (A_{k_j}) converges pointwise, and we can check that A is linear and

$$\forall x \in X,\ \ \|Ax\|_Y = \lim_{j\to\infty} \|A_{k_j}x\|_Y \leq \lim_{j\to\infty} \|A_{k_j}\|\|x\|_X \leq M\|x\|_X$$

$\square$

Remark 4.14 The completeness of X is important for Banach-Steinhaus 4.12 (as for Baire).

Example 4.15 Take $X = C^0([0,1])$ with $\|\cdot\|_X = \|\cdot\|_1$ the L^1-norm, $\int_0^1 |f(x)|\,\mathrm{d}x$, and let

$$A_k f = k\int_{1-\frac{1}{k}}^1 f(t)\,\mathrm{d}t,\ \ k \geq 1.$$

Clearly, $|A_k f| \leq k \cdot \|f\|_1$ so $A_k : X \to \mathbb{R}$ is continuous with $\|A_k\|_{\mathcal{L}(X,\mathbb{R})} \leq k\ \forall k$.

Moreover, $\forall f \in X$, $A_k f \to_{k\to\infty} Af := f(1)$. But $A : X \to \mathbb{R}$ is **not** continuous, *e.g.* take $f_n(t) = t^n$, which $f_n \to_{L^1} 0$ but

$$Af_n = f_n(1) = 1 \not\to 0 (= A0) \text{ as } n \to \infty$$

Of course, $(C([0,1]), \|\cdot\|_1)$ is **not** complete, so this is not a contradiction.

N/B: if we instead took $(C([0,1]), \|\cdot\|_\infty)$ then $|A_k f| \leq \|f\|_\infty$ and $Af = f(1)$ **is** continuous because it is bounded by $\|A\|_{\mathcal{L}(X,\mathbb{R}} \leq 1$ as $\|Af\|_\infty \leq \|f\|_\infty$.

§4.2.2 Baire's Original Problem

We can go back to Baire's original question from the beginning of this section, which is as follows: for $f_n : [0,1] \to \mathbb{R}$ continuous $\forall n \geq 1$ and pointwise convergent (*i.e.* $f(x) = \lim_{n\to\infty} f_n(x)$ exists $\forall x \in [0,1]$). Then, does f have point(s) of continuity?

Theorem 4.16 (Baire) Let (X,d) be complete and (f_n) be a sequence of continuous functions $f_n : X \to \mathbb{R}$, and for each $x \in X$ the pointwise limit

$$\lim_{n\to\infty} f_n(x) := f(x) \in \mathbb{R}$$

exists. Then the set

$$R := \{x \in \mathbb{R} : f \text{ is continuous at } x\}$$

is dense.

Proof. (sketch) For $\varepsilon > 0$, $n \geq 1$ set $P_{n,\varepsilon} = \{x \in \mathbb{R} : |f_n(x) - f(x)| \leq \varepsilon\}$; $R_\varepsilon = \bigcup_{n=1}^\infty \operatorname{int}(P_{n,\varepsilon})$ $(\subset R_{\varepsilon'}\ \forall \varepsilon \leq \varepsilon')$.

Claim 1: $\bigcap_{n=1}^\infty R_{\frac{1}{n}} = R$.

Claim 2: R_ε is open and dense $\forall \varepsilon > 0$. $\square$

Remark 4.17 As a corollary, $\mathbb{1}_{\mathbb{Q}}$ is nowhere continuous ; hence $\nexists f_n : \mathbb{R} \to \mathbb{R}$ continuous with $f_n \to \mathbb{1}_{\mathbb{Q}}$ pointwise.

§4.3 Open Mapping Theorem

In this section we take X, Y to be normed spaces and $A : X \to Y$ to be a linear operator. Recall the definition of an open ball in X

$$B_X(x,r) = \{y \in X : \|y - x\|_X < r\}, \;\; x \in X, r > 0,$$

Notation: We have sets $A, B \subset X$ linear over $\mathbb{K}$, $\lambda \in \mathbb{K}$; $A + B = \{a + b : a \in A, b \in B\}$, $\lambda A = \{\lambda a : a \in A\}$.

Definition 4.18 (Open Mapping) — A is **open** if $A(U) = \{Ax : x \in U\}$ $(\subset Y)$ is open $\forall U$ $(\subset X)$ open.

Remark 4.19

1. A continuous means $A^{-1}(V)$ $(\subset X)$ is open $\forall V$ $(\subset Y)$ open.
2. A continuous need not be open *e.g.* $Ax := 0 \in Y$.

Theorem 4.20 (Open Mapping Theorem) Let X, Y be Banach spaces, and $A \in \mathcal{L}(X,Y)$. Then:

i) if A is surjective, A is open.

ii) if A is bijective, then $A^{-1} \in \mathcal{L}(Y,X)$. (inverse operator theorem)

Remark 4.21 Statement ii) is important in applications. If $A \in \mathcal{L}(X,Y)$ is bijective, then $A^{-1} : Y \to X$ is linear and A^{-1} is also bounded (or equivalently continuous).

The proof for open mapping theorem is not given here but it can be easily found online; it makes use of Baire category and the completeness of X and Y. I intend to add in the details of the proof in the future though when I get some more time for it.

Example 4.22 Let $X = Y$ with norms $\|\cdot\|_1$ and $\|\cdot\|_2$ and assume $\exists C > 0$ s.t.

$$\|x\|_2 \le C\|x\|_1 \;\; \forall x \in X$$

If X is complete w.r.t both $\|\cdot\|_1$ and $\|\cdot\|_2$, then $A = \mathrm{id}$: $(X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ is open by the open mapping theorem (indeed this applies since A is bounded by the above: $\|x\|_2 \le C\|x\|_1$). Since A is bijective, ii) gives that $A^{-1} = \mathrm{id} : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$ is bounded, *i.e.*

$$\exists C : \;\; \|A^{-1}x\|_1 = \|x\|_1 \le C\|x\|_2$$

so $\|\cdot\|_1$ and $\|\cdot\|_2$ are actually equivalent.

Example 4.23 Consider $X = C^0([0,1])$ with $\|\cdot\|_1 = \|\cdot\|_\infty$, $\|\cdot\|_2 = \|\cdot\|_{L^1}$. Then, $A = \mathrm{id}\colon (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ is continuous, since

$$\|Af\|_2 = \|f\|_2 = \int_0^1 |f(t)\,\mathrm{d}t \leq \|f\|_\infty = \|f\|_1$$

but A is not open. Otherwise by example 4.22 $\|\cdot\|_1$ and $\|\cdot\|_2$ would be equivalent. If we take $f_n(t)$ to be the tent map which peaks at n for $t \in [0, \frac{1}{n}]$ (and 0 everywhere else), then $\|f_n\|_2 = 1$, $\|f_n\|_1 = n \to \infty$. This shows Y in theorem 4.20 needs to be complete.

Example 4.24 This example shows that completeness of X is also a requirement. Take

$$X = Y = \{(x_n) \in \ell^\infty : \ \exists N : \ x_n = 0 \ \forall n \geq N\} \subset \ell^\infty$$

with $\|\cdot\|_X = \|\cdot\|_Y = \|\cdot\|_\infty$. This is a linear normed space. Although it is not complete. Define

$$A : X \to X, \ \ Ax = \left(x_1, \frac{x_2}{2}, \frac{x_3}{3, \dots}\right) \text{ if } x = (x_1, x_2, \dots)$$

Then A is linear and bijective with $A^{-1} : X \to X$, $A^{-1}x = (x_1, 2x_2, 3x_3, \dots)$ and A is bounded:

$$\|Ax\|_\infty = \sup_{n\geq 1} \frac{|x_n|}{n} \leq \sup_{n \geq 1} |x_n| = \|x\|_\infty$$

so $\|A\| \leq 1$. but A^{-1} is unbounded. Pick $x^{(n)} = (\overbrace{1, \dots, 1}^{n}, 0, \dots)$ then $\|x^{(n)}\|_\infty = 1$ but $\|A^{-1}x^{(n)}\| = n$.

Hence, $A^{-1} \notin \mathcal{L}(X)$ and X cannot be complete, otherwise by ii) in theorem 4.20 A^{-1} would be bounded.

§4.4 Closed Graph Theorem

Again, we take X, Y to be normed, linear spaces. Often, the operator A is not defined on all of A but instead on a '**domain**' $D(A)$. So we assume that:

- $D(A) \subset X$ is a linear subspace, on which the linear operator $A : D(A)\ (\subset X) \to Y$ is defined.

Definition 4.25 (Graph) — The **graph** of A (really of $(A, D(A))$) is the linear space

$$\Gamma_A = \{(x, Ax) : x \in D(A)\} \subset X \times Y$$

We endow $X \times Y$ with the norm $\|(x, y\|_{X\times Y} = \|x\|_x + \|y\|_Y$, $x \in X, y \in Y$ (note this norm preserves the product topology).

Definition 4.26 (Closed Mapping) — A is called **closed** if Γ_A is closed in $(X \times Y, \|\cdot\|_{X\times Y})$.

Proposition 4.27 Let $A \in \mathcal{L}(X, Y)$ with $D(A) = X$. Then A is closed.

Proof. Let $(x_k, y_k) \subset \Gamma_A$ with $\|(x_k, y_k) - (x, y)\|_{X\times Y} \to_{k\to\infty} 0$ for some $(x, y) \in X \times Y$. We need to show: $(x, y) \in \Gamma_A$, *i.e.* $y = Ax$.

We know $y_k = Ax_k$ and $\|x_k - x\|_X \to 0$, $\|Ax_k - y\| \to 0$. But $\forall k \geq 1$

$$\|y - Ax\|_Y \leq \underbrace{\|y - Ax_k\|_Y}_{\to 0} + \underbrace{\|Ax_k - Ax\|_Y}_{\leq \|A\| \|x_k - x\|_X \to 0}$$

$\square$

Theorem 4.28 (Closed Graph Theorem) Let X, Y be Banach spaces and $A : X \to Y$ be linear. Then, the following are equivalent

i) $A \in \mathcal{L}(X, Y)$

ii) A is closed

Proof. For i) $\implies$ ii), see proposition 4.27. So we need to show ii) $\implies$ i). If X, Y are complete, then so is $(X \times Y, \|\cdot\|_{X\times Y})$.

A closed means Γ_A is closed in $(X \times Y, \|\cdot\|_{X\times Y})$, and so $(\Gamma_A, \|\cdot\|_{X\times Y})$ is Banach. Consider

$$\begin{aligned} \pi_X : \Gamma_A &\to X \\ (x, Ax) &\mapsto x \end{aligned} \qquad \begin{aligned} \pi_Y : \Gamma_A &\to Y \\ (x, Ax) &\mapsto Ax \end{aligned}$$

π_X, π_Y are continuous (with $\|\pi_X\|$, $\|\pi_Y\| \leq 1$). π_X is injective, and surjective, so by ii) in the open mapping theorem 4.20, $\pi_X^{-1} \in \mathcal{L}(X, \Gamma_A)$ and so

$$A = \pi_Y \circ \pi_X^{-1} \in \mathcal{L}(X, Y)$$

$\square$

Remark 4.29 Notice how statement ii) is simpler than statement i), but they are in fact equivalent. We can unpack statements i) and ii) from 4.28 a little further as follows.

i) says that A is continuous, *i.e* if $(x_n) \subset X$, $x \in X$, $\|x_n - x\|_X \to 0 \implies \|Ax_n - Ax\|_Y \to 0$, so there are **two** things to check: (Ax_n) converges and the limit is in Ax.

ii) says that A is closed, *i.e.* both $\|x_n - x\|_X \to 0$ and $\|Ax_n - y\|_Y \to 0$ implies that $Ax = y$, so conversely here there is only **one** condition to check.

A running example to have in mind for this section is $Y = X = C^0([0, 1])$ with $\|\cdot\|_X = \|\cdot\|_\infty$ and $A = \frac{\mathrm{d}}{\mathrm{d}t}$, with $D(A) =^{e.g.} C^1([0, 1]) \subset X$.

Then, $(D(A), \|\cdot\|_\infty)$ is **not** Banach and $A : D(A) \to C^0([0,1])$ is therefore an example of an operator which is closed, but not continuous (see theorem 4.28).

For non-continuous, take $f_n(t) = t^n \in D(A)$, $Af_n = nf_{n-1}$ so $\|f_n\|_\infty = 1$, $\|Af_n\|_\infty = n\|f_{n-1}\|_\infty = n$, so

$$\sup_{f\in D(A), \|f\|_\infty \leq 1} \|Af\|_\infty = \infty$$

For closed, if $(f_n, f_n') \to (f, g)$ (in $D(A) \times C^0([0,1])$) then $\|f_n - f\|_\infty \to 0$, $\|f_n' - g\|_\infty \to 0$, but

$$\forall t \in (0,1], \quad \overbrace{f_n(t)}^{\to f(t)} = \underbrace{\int_0^t f_n'(\tilde{t})\,\mathrm{d}\tilde{t}}_{\to \int_0^t g(\tilde{t})\,\mathrm{d}\tilde{t}} + \overbrace{f_n(0)}^{\to f(0)}$$

by dominated convergence theorem. So, $f' = g$ by fundamental theorem of calculus, *i.e.* $(f, g) = (f, f') \in \Gamma_A$.

Using closed graph theorem, we can now replace continuity by closedness in ii) of open mapping theorem.

§4.4.1 Closable Operators

Let X, Y be normed and $A : D(A) \subset X \to Y$ be a linear operator. This leads to the well-defined notion of a graph on A, Γ_A.

Definition 4.30 (Extension of Linear Operator) — The map $B : D(B) \subset X \to Y$ is called an **extension** of A if $D(B) \supset D(A)$ and $B|_{D(A)} = A$.

Definition 4.31 (Closable) — A is called **closable** if it has an extension B with graph $\Gamma_B = \overline{\Gamma_A}$. We write $B = \overline{A}$ (the closure of A).

N/B: the closure of A is itself a closed operator.

Lemma 4.32 The following are equivalent

i) A is closable.

ii) $\forall (x_k, Ax_k) \subset \Gamma_A$ with $x_k \to 0$, $Ax_k \to y$, one has $y = 0$.

§5 Spectral Theory

In general, spectral theory is an extension of the theory of certain matrices being diagonalisable to a much broader class of linear operators in a variety of mathematical spaces being 'diagonalisable'. More on this later in this section.

§5.1 Weak vs. Strong Topologies

Take $(X\|\cdot\|_X)$ be a normed space, and consider the dual X^* over $\mathbb{R}$.

Definition 5.1 (Weak Convergence) — A sequence $(x_n) \subset X$ **converges weakly** to $x \in X$, if $\forall l \in X^*$

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

We write $x_n \to^w x$.

Definition 5.2 (Strong Convergence) — A sequence $(x_n) \subset X$ converges **strongly** (or **in norm**) to x if

$$\lim_{n\to\infty} \|x_n - x\| = 0$$

We write $x_n \to x$.

Remark 5.3 $x_n \to x$ implies $x_n \to^w x$, since $|l(x_n) - l(x)| \leq \|l\|_* \|x_n - x\|$.

However, the converse is false. Let $x_n = (0, \dots, 0, \overbrace{1}^{n\text{th}}, 0, \dots) \in \ell^2$, $\|x_n - x_m\| = \sqrt{2}$ $(n \neq m)$. So, x_n does not converge (strongly) but $x_n \to^w 0$.

By Riesz's representation theorem, $l(\cdot) = \langle y, \cdot \rangle_{\ell^2}$ for some $y \in \ell^2$. Hence, with $y = (y^n)_n$

$$l(x_n) = \langle y, x_n \rangle = y^n \leq \sqrt{\sum_{k\geq n} |y^k|^2} \to 0 \quad (\text{since } \|y\|_2 < \infty)$$

The above follows from the fact that $(y_n) \in \ell^p$ implies $y_n \to 0$.

Proposition 5.4 If $x_n \to^w x$ and $x_n \to^w y$, then $x = y$.

Proof. Assume not. By proposition 3.24, $\exists l \in X^*$ with $l(x) \neq l(y)$ if $x \neq y$. With this l,

$$l(x) =^{x_n \to^w x} \lim_{n\to\infty} l(x_n) =^{x_n \to^w y} l(y)$$

which is a contradiction. □

Proposition 5.5 If $x_n \to^w x$, then $\sup_n \|x_n\|_X < \infty$.

Proof. Consider $A_n \in \mathcal{L}(X^*, \mathbb{R})$ $(=X^{**}$: bidual) with $A_n(l) := l(x_n)$, $l \in X^*$.

$x_n \to^w x$ implies $\sup_n |A_n(l)| < \infty \ \forall l \in X^*$ and X^* is complete so by Banach-Steinhaus 4.12, $\sup_n \|A_n\|_{\mathcal{L}(X^*,\mathbb{R})} < \infty$.

But $\|A_n\|_{\mathcal{L}(X^*,\mathbb{R})} = \sup_{l\in X^*:\|l\|\le 1} |l(x_n)| = \|x_n\|_X$, by the dual characterisation of the norm in corollary 3.22. □

This naturally leads to the following definition.

Definition 5.6 (Bidual) — $X^{**} := (X^*)^* \ (= \mathcal{L}(X^*, \mathbb{R}))$ is called the **bidual** of X.

X embeds canonically into X^{**} via the embedding

$$i : X \to X^{**}, \ \ i(x)(x^*) := x^*(x) = \langle x^*, x\rangle \ \ \forall x \in X \forall x^* \in X^*$$

Here, $i : X \to X^{**}$ assigns each $x \in X$ the linear functional $i \in X^{**}$, whose value at x^* is obtained by evaluating i at x.

Remark 5.7 i is a linear isometry, and one has $\forall x \in X \ \|x\|_X = \sup_{l\in X^*:\|l\|\le 1} |l(x)| = \|i(x)\|_{**}$.

Definition 5.8 (Reflexive) — X is **reflexive** if i, as above, is surjective. Or, equivalently, X is reflexive if $X \cong X^{**}$.

Example 5.9 i) If $\dim X < \infty$, then X is reflexive.

ii) If H is Hilbert, then H is reflexive.

iii) $L^p(\mu)$, $1 < p < \infty$ is reflexive.

iv) L^1, L^∞ are, in general, not reflexive.

Proposition 5.10 If X is reflexive, then X^* is reflexive.

Recall that we showed unit balls in ∞ dimensions are never (sequentially) compact. Weak convergence allows us to restore a (weak) version of this, known as Banach-Algaoglu - see below. For reflexive spaces, that's the whole story. Since X may not be reflexive, one must consider an even weaker topology.

§5.1.1 Banach-Alaoglu

Take $(X, \|\cdot\|_X)$ to be a normed space, X^* to be the dual, X^{**} to be the bidual, $i : X \to X^{**}$ to be an isometry.

Definition 5.11 (Weak* Convergent) — $(l_n) \subset X^*$ is **weak*-convergent** to $l \in X^*$ if

$$\lim_{n\to\infty} l_n(x) = l(x) \quad \forall x \in X$$

We write $l_n \to^{w^*} l$. In fact, this is just **pointwise** convergence in X.

Remark 5.12 We now have 3 notions of convergence on the dual X^*:

i) norm/strong convergence $\|l_n(x) - l\|_* \to 0$.

ii) weak convergence $l_n \to^w l$, *i.e.* $\forall z \in X^{**}$, $\lim_n z(l_n) = z(l)$.

iii) weak* convergence $l_n \to^{w^*} l$, which is equivalent to asking ii) but for $z \in i(X)$ only (where $i(X)$ is the image of the canonical embedding).

If X is reflexive then ii) $\iff$ iii) (*e.g.* when X is Hilbert). But, in general, i) $\implies$ ii) $\implies$ iii).

Theorem 5.13 (Banach-Alaoglu) Let X be separable. If $(l_n) \subset X^*$ is bounded (in X^*, *i.e* $\sup_{n\geq 1} \|l_n\| < \infty$) there exists $l \in X^*$ and a subsequence $\Lambda \subset \mathbb{N}$ s.t. $l_n \to^{w^*} l$ ($n \to \infty$, $n \in \Lambda$).

Proof. Let $(x_j) \subset X$ be dense. Using boundedness, pick a subsequence $\mathbb{N} \supset \Lambda_1 \supset \Lambda_2 \supset \cdots \supset \Lambda_j \supset \Lambda_{j+1}$ (inductively) such that, for all $j \in \mathbb{N}$, $l_n(x_j) \to a_j \in \mathbb{R}$.

$\Lambda :=$ diagonal sequence of $(\Lambda_j)_j$ so $\forall j$ $l_n(x_j) \to a_j$ $(n \to \infty, n \in \Lambda)$. $l(x_j) := a_j$, extend it linearly on $M = \text{span}\{x_j : j \in \mathbb{N}\}$ and for all $x \in M$

$$\underbrace{|l(x)|}_{\text{finite lin. comb. of } x_j\text{s}} = \lim_{k\to\infty, k\in\Lambda} |l_k(x)| \leq \underbrace{\sup_k \|l_k\|_*}_{\leq C} \|x\|_X$$

so $l \in M^*$, hence it can be extended to $l \in X^*$ using corollary 3.21.

We now show $l_n \to^{w^*} l$ ($n \to \infty$, $n \in \Lambda$). Let $x \in X$, and pick $I \subset \mathbb{N}$ s.t $x_j \to x$ ($j \to \infty$, $j \in J$). For such j and $n \geq 1$

$$\begin{aligned} |l_n(x) - l(x)| &\leq |l_n(x - x_j)| + |l(x - x_j)| + |l_n(x_j) - l(x_j)| \\ &\leq \left(\sup_n \|l_k\|_* + \|l\|_*\right) \|x - x_j\|_X + |l_n(x_j) - l(x_j)| \end{aligned}$$

Letting first $n \to \infty$ yields $\lim_{n\to\infty} |l_n(x) - l(x)| \leq C\|x - x_j\|_X$, $j \in J$. Now let $j \to \infty$ and we are done. □

If X is reflexive, the separability condition above in 5.13 can be removed.

Corollary 5.14 Let H be Hilbert. If $(x_n) \subset H$ is bounded ($\sup_n \|x_n\|_H < \infty$), then (x_n) has a weakly convergent subsequence.

Note that we cannot replace weak convergence by strong convergence in the above corollary, unless H is finite dimensional.

Example 5.15 $X = L^1[0,1]$ is separable, $X^* \cong L^\infty$. If $(f_n) \subset L^\infty$ is bounded, *i.e.* $\sup_n \|f_n\|_\infty < \infty$, theorem 5.13 yields a subsequence $(n_k)_k \subset \mathbb{N}$ and $f \in L^\infty$ s.t.

$$\lim_{k\to\infty} \int f_{n_k} g \,\mathrm{d}x = \int f g \,\mathrm{d}x \;\; \forall g \in L^1[0,1]$$

Example 5.16 $X = L^\infty[0,1]$ is not separable (and also not reflexive). The following example shows that the conclusions of the theorem fail in this case. For $0 \le \varepsilon \le 1$ consider

$$T_\varepsilon : L^\infty \to \mathbb{R} \quad T_\varepsilon f = \frac{1}{\varepsilon}\int_0^\varepsilon f \, \mathrm{d}x, f \in L^\infty$$

Then $\|T_\varepsilon\|_{(L^\infty)^*} \le 1$, *i.e* $T_\varepsilon \in (L^\infty)^*$.

One can then show $\{T_\varepsilon : 0 \le \varepsilon \le 1\}$ is not weak*-sequentially compact.

Remark 5.17 If instead we considered $X = C^0([0,1]) \subset L^\infty$, for the above example, which is a separable, closed subspace then theorem 5.13 applies to $T_\varepsilon|_X$. Indeed one can see that

$$T_\varepsilon f \to^{\varepsilon \to 0} f(0), \;\; i.e\; T_\varepsilon \to^{w^*} \delta_0(\text{ Dirac functional at } 0)$$

§5.2 Compact Operators

Compact operators form a very important class of bounded operators. Roughly speaking, these are the closest thing to a matrix in infinite dimensions (which we shall see in spectral theory later on).

Definition 5.18 (Compact Operator) — Let X, Y be normed spaces and $T : X \to Y$ be linear. T is **compact** is for all bounded $B \subset X$ (*i.e.* $\sup\{\|x\|_X : x \in B\} < \infty$) $\overline{T(B)}$ is (sequentially) compact, where $T(B) = \{Tx : x \in B\} \subset Y$.

Lemma 5.19 Let X, Y be Banach spaces. Then the following are equivalent.

i) T is compact.

ii) $\overline{T(B_X(0,1))} \subset Y$ is compact (here, $B_X(0,1)$ is the unit ball of radius 1 around the origin in X).

iii) $\forall (x_n) \subset X$ bounded, (Tx_n) has a Cauchy subsequence.

Remark 5.20 Note that if X, Y are simply just normed, then the lemma 5.19 is still true, except that you replace "Cauchy" by "convergent".

Example 5.21 Take $T = \mathrm{id} : X \to X$ is compact iff $\dim X < \infty$. For $\dim X = \infty$, recall $B = \overline{B_X(0,1)}$ is not compact.

An operator T has **finite rank** if $\dim(\mathrm{im}(T)) < \infty$. If $T \in \mathcal{L}(X,Y)$ has finite rank then T is compact. Using iii) in 5.19, let $(x_n) \subset X$ be bounded. Then $\|Tx_n\| \le \|T\|\|x_n\| \le C$ so $(Tx_n) \subset \mathrm{im}(T)$ is bounded. Since $\mathrm{im}(T)$ is finite-dimensional, we can choose a convergent subsequence.

Remark 5.22 If $\dim X < \infty$, T is compact.

Theorem 5.23 Let X, Y be Banach spaces. If $T_n : X \to Y$ is a sequence of compact operators and for some $T \in \mathcal{L}(X, Y)$

$$\|T_n - T\|_{\mathcal{L}(X,Y)} \to 0$$

Then T is compact.

Remark 5.24 This means that the space $\Big(\{T \in \mathcal{L}(X,Y) : T \text{ compact}\}, \|\cdot\|_{\mathcal{L}(X,Y)}\Big) \subset \Big(\mathcal{L}(X,Y), \|\cdot\|_{\mathcal{L}(X,Y)}\Big)$ is closed, *i.e.* a Banach space.

Consider the operator $T_\lambda : \ell^p \to \ell^p$ with $T_\lambda x := (\lambda_m x_n)_{n\in\mathbb{N}}$ for $x = (x_n)_{n\in\mathbb{N}}$, where $1 \le p \le \infty$, $\lambda = (\lambda_n)_{n\in\mathbb{N}}$ and $\sup_n |\lambda_n| < \infty$.

Then T_λ is well-defined and

$$T_\lambda \text{ is compact} \iff \lim_{k\to\infty} \lambda_k = 0$$

For " $\Longleftarrow$ " use theorem 5.23 and show $\|T - T_n\| \to 0$, where $T_n : \ell^p \to \ell^p$ where $x \mapsto T_n x = (\lambda_0 x_0, \dots \lambda_n x_n, 0, 0, \dots)$.

For " $\Longrightarrow$ ", if $\Lambda \subset \mathbb{N}$ is s.t. $|\lambda_n| \ge \delta$, $n \in \Lambda$ for some $\delta > 0$, then the sequence e_n, $n \in \Lambda$, is bounded but $T_\lambda e_n$, $n \in \Lambda$, has no Cauchy subsequence since for all $n, m, \in \Lambda$ (with $n \ne m$) $\quad \|T_\lambda e_n - T_\lambda e_m\|_p \ge \delta 2^{\frac{1}{p}}$.

Example 5.25 (Hilbert-Schmidt Integral Operator) Let $X = L^2[0,1]$ and take $a = a(x,y) \in C^0[0,1]^2$. Now we define the operator $A : X \to X$

$$Af(x) = \int_0^1 a(x,y) f(y)\,\mathrm{d}y, \quad f \in L^2[0,1]$$

Then A is well-defined, bounded since

$$\|Af\|_2^2 = \int_0^1 |Af(x)|^2\,\mathrm{d}x \overbrace{\le}^{\text{Cauchy-Schwarz}} \underbrace{\int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y |a(x,y)|^2}_{\le C} \|f\|_2^2$$

Now, we claim: A is compact. Let $(f_) \subset X$, $\|f_n\|_2 \le M$. We can check that (Af_n) is continuous, and $\sup_n \|Af_n\|_\infty < \infty$. Moreover,

$$\begin{aligned} |Af_n(x) - Af_n(y)| &\le \int_0^1 \underbrace{|a(x,z) - a(y,z)|}_{<\varepsilon \text{ if } |x-y|<\delta} |f_n(z)|\,\mathrm{d}z \\ &\le \varepsilon \|f_n\|_2 \le M\varepsilon \text{ if } |x-y| < \delta \end{aligned}$$

so (Af_n) is **equicontinuous**. By *Arzelà-Ascoli* theorem, (Af_n) has a subsequence which converges in $\|\cdot\|_{L^\infty[0,1]}$, hence in $\|\cdot\|_{L^2[0,1]}$.

§5.3 Spectrum and Resolvent

Let $(X, \|\cdot\|_X)$ be a Banach space over $\mathbb{C}$, and $A : D_A \subset X \to X$ a linear operator

Definition 5.26 (Resolvent Set) — The **resolvent set** of A is

$$\rho(A) = \{\lambda \in \mathbb{C} : (\lambda \mathrm{id} - A) : D_A \to X \text{ is bijective with } (\lambda - A)^{-1} \in \mathcal{L}(X)\}$$

Definition 5.27 (Spectrum) — $\sigma(A) = \mathbb{C} \backslash \rho(A)$ is called the spectrum of A.

Definition 5.28 (Resolvent) — The **resolvent** of A is the map $R : \rho(A) \to \mathcal{L}(X)$

$$\rho(A) \in \lambda \mapsto R_\lambda = (\lambda - A)^{-1} \in \mathcal{L}(X)$$

Example 5.29 Let $X = \mathbb{C}^n$, $A \in \mathcal{L}(X)$ (*i.e.* $D_A = X$), $\lambda \in \mathbb{C}$.

$(\lambda - A)$ invertible $\iff$ $p(\lambda) := \det(\lambda - A) \neq 0$. Since $p(\cdot) = 0$ has at least one, and at most n solutions, we obtain $\sigma(A) \neq \emptyset$, and $\sigma(A)$ contains at most n points. Hence, $\rho(A) \neq \emptyset$ and $\rho(A) \subset \mathbb{C}$ is dense.

Lemma 5.30 If $z_0 \in \rho(A)$, then

$$D := \{z \in \mathbb{C} : |z - z_0| < \frac{1}{\|R_{z_0}\|_{\mathcal{L}(X)}}\} \subset \rho(A)$$

hence $\rho(A)$ is open (and $\sigma(A)$ is closed).

Proof. $z - A = (z - z_0) + (z_0 - A) = (1 + (z - z_0)R_{z_0})\,(z_0 - A)$.

If $z \in D$ then $1 + (z - z_0)R_{z_0}$ is invertible with

$$(1 + (z - z_0)R_{z_0})^{-1} = \sum_{n \geq 0} (z_0 - z)^n R_{z_0}^n$$

Hence also

$$R_z = (z - A)^{-1} = R_{z_0}\,(1 + (z - z_0)R_{z_0})^{-1} \in \mathcal{L}(X)$$

□

We will revisit the operator $T = T_\lambda : \ell^p \to \ell^p$ from previously, where $T_\lambda x = (\lambda_n x_n)_{n \in \mathbb{N}}$. Here, the spectrum of T is $\sigma(T) = \overline{\{\lambda_k : k \in \mathbb{N}\}}$. To show this, we consider each inclusion.

For "$\supset$", if $x = e_k$ then $Tx = \lambda_k x$ so $\lambda_k = T$ is not injective, so $\{\lambda_k : k \in \mathbb{N}\} \subset \sigma(T)$, hence by lemma 5.30 above $\overline{\{\lambda_k : k \in \mathbb{N}\}} \subset \sigma(T)$.

For "$\subset$", if $\mu \notin \overline{\{\lambda_k : k \in \mathbb{N}\}}$, then $\exists \delta > 0$ s.t. $|\mu - \lambda_k| > \delta\ \forall k \in \mathbb{N}$. Let $x \in \ell^2$ and define y as $y := (\mu - T)x = ((\mu - \lambda_k)x_k)_k$.

So $x_k = (\mu - \lambda_k)^{-1} y_k$ and $\|x\|_{\ell^2} \le \delta^{-1}\|y\|_{\ell^2}$. This implies $(\mu - T)^{-1} \in \mathcal{L}(\ell^2)$ is a bounded operator. This operator is doing the mapping $z \in \ell^2 \mapsto \left(\frac{z_1}{\mu-\lambda_1}, \frac{z_2}{\mu-\lambda_2}, \dots\right) \in \ell^2$. Hence, $\mu \in \rho(T)$.

Remark 5.31 In infinite dimensions, $(\lambda - A)$ is not invertible $\iff$ $(\lambda - A)$ is not injective. One may wonder if "lack of injectivity" is the only reason for $\lambda \notin \sigma(A)$.

We now introduce a few definitions. Take a linear operator A to be closed with a spectrum $\sigma(A)$.

Definition 5.32 (Point Spectrum) — Then the **point spectrum** of A $\sigma_p(A)$ is defined as

$$\sigma_p(A) := \{\lambda \in \mathbb{C} : \lambda - A \text{ is not injective}\}$$

Definition 5.33 (Eigenvalue) — The elements $\lambda \in \sigma_p(A)$ (in the point spectrum of A) are called **eigenvalues** of A.

Definition 5.34 (Eigenvector) — If you have $x \neq 0$ with $Ax = \lambda x$ (and of course you do have such an x since you are not injective), this x is called an **eigenvector** of A.

Definition 5.35 (Eigenspace) — The **eigenspace** of A is given by $\ker(\lambda - A) = \{x \in D_A : Ax = \lambda x\} \neq \{0\}$.

The above notions should be somewhat familiar already from finite dimensions; we are just extending these so they work in infinite dimesnions as well.

Definition 5.36 (Continuous Spectrum) — The **continuous spectrum** of A $\sigma_c(A)$ is defined by the set

$$\sigma_c(A) := \{\lambda \in \mathbb{C}\backslash\rho(A) : (\lambda - A) \text{ injective}, \operatorname{im}(\lambda - A) \text{ dense}\}$$

Clearly, the point spectrum of A and continuous spectrum of A are disjoint. But, we also have the residual spectrum of A, which is whatever is left over from the other two.

Definition 5.37 (Residual Spectrum) — The **residual spectrum** of A $\sigma_r(A)$ is defined as $\sigma_r(A) := \sigma(A)\backslash\left(\sigma_p(A) \cup \sigma_c(A)\right)$.

If we have $X = \mathbb{C}^n$, then the spectrum is the pure point spectrum, *i.e* $\sigma(\cdot) = \sigma_p(\cdot)$. Howver, below is a more interesting example.

Example 5.38 (Shift Operator) Consider the shift operator $S : \ell^2 \to \ell^2$ with $S(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$. Then $0 \in \sigma(S)$, and indeed S is not invertible (since not surjective *e.g.* $\forall y \in \ell^2$, $y_1 \neq 0$, $y \notin \operatorname{im}(S)$). But, $0 \notin \sigma_p(S)$ so S is injective.

In fact, $Sx = \lambda x \implies 0 = \lambda x_1$, $x_1 = \lambda x_2$, $\dots \implies x_k = 0$ for all k, so $\sigma_p(S) = \emptyset$.

In fact, $\sigma(S) = \overline{D} = \{z \in \mathbb{C} : |z| \le 1\}$, $\sigma_r(S) = D$, and $\sigma_c(D) = \delta D$.

§5.4 Spectral Theory in Hilbert Spaces

From now on, up until basically the end of the course, we're going to work in Hilbert spaces.

Let $(H, \langle \cdot, \cdot \rangle)$ be a Hilbert space over $\mathbb{C}$, and $A : D_A \subset H \to H$ linear operator (D_A is a linear, dense subdomain that is not necessarily equal to H), with adjoint $A^* : D_{A^*} \subset H \to H$. Recall that A^* is characterised by

$$\forall x \in D_A, y \in D_{A^*} \quad \langle A^* y, x \rangle = \langle y, Ax \rangle$$

and $D_{A^*} = \{y \in H : l_y : D_A \to \mathbb{C}, x \mapsto \langle y, Ax \rangle \text{ is continuous}\}$.

Notation: we write $A \subset B$, read "B is extension of A", if $D_A \subset D_B$ and $B|_{D_A} = A$. In finite dimensions, we would usually talk about symmetric matrices, but in infinite dimensions we have to be a little more careful here. There is something called symmetric and something called self-adjoint (which are usually squashed together into one in finite dimensions).

Definition 5.39 (Symmetric) — A, as given above, is **symmetric** if $A \subset A^*$, *i.e.* $D_A \subset D_{A^*}$ and $\langle Ax, y \rangle = \langle x, Ay \rangle$ $\forall x, y \in D_A$.

Definition 5.40 (Self-Adjoint) — A, as given above, is **self-adjoint** if $A = A^*$, *i.e* A is symmetric with $D_{A^*} = D_A$.

What can we say about the spectrum of A, $\sigma(A)$ for such A?

Lemma 5.41 If A is symmetric, then its pure point spectrum is real, *i.e.* $\sigma_p(A) \subset \mathbb{R}$.

Proof. Note that in our inner product above $\langle \cdot, \cdot \rangle$ for H, we take the left input to be linear and the right input to be anti-linear (so if we pull out a scalar from the second entry, it gets conjugated).

Let $\lambda \in \sigma_p(A)$ with eigenvalue $0 \neq x \in \ker(\lambda - A)$. Then

$$\lambda \|x\|_H^2 = \langle Ax, x \rangle =^{\text{symm.}} \langle x, Ax \rangle = \overline{\langle Ax, x \rangle} = \overline{\lambda} \|x\|_H^2$$

So $\lambda = \overline{\lambda} \in \mathbb{R}$ □

Is this true for all the spectrum of A (*i.e.* if we replace σ_p by σ)? The answer is **no**. So, symmetric operators in infinite dimensions can have imaginary spectrum, for instance, as the following example demonstrates.

Example 5.42 Take $H = L^2(0,1)$, with $\langle f, g\rangle = \int_0^1 f\overline{g}\,\mathrm{d}t$. We define "$A := i\frac{\mathrm{d}}{\mathrm{d}t}$".

Furthermore, $f \in H$ is said to have a **weak derivative** f' if $f' := v$ for some $v \in H$ and

$$\int_0^1 fg'\,\mathrm{d}t = -\int_0^1 vg\,\mathrm{d}t \quad \forall g \in C_c^\infty(0,1)$$

Consider $A_\infty = i\frac{\mathrm{d}}{\mathrm{d}t} : C_C^\infty(0,1) \subset H \to H$, and extension A_3 with

$$H^1 := \{f \in H : f \text{ has a weak derivative } f'\}$$
$$D_{A_3} = \{f \in H^1 : f(0) = 0 = f(1)\} \text{ (known as Dirichlet boundary conditions)}$$

We can show that A_3 is symmetric, $\sigma_p(A_3) = \emptyset$, $\rho(A_3)\emptyset$, $\sigma(A_3) = \mathbb{C}$.

Note that for $g \in H$, the general solution of: $\lambda f - if' = g$ is

$$f(t) = ae^{it} + i\int_0^t e^{i\lambda(s-t)}g(s)\,\mathrm{d}s \text{ for some } a \in \mathbb{C}$$

To show $\sigma_p(A) = \emptyset$, if $A_3 f = if' = \lambda f$ for some $\lambda \in \mathbb{C}$, then by the solution above, $f(t) = ae^{i\lambda t}$, $f \in D_{A_3}$ and $f(0) = 0 \implies a = 0 \implies f \equiv 0$.

For $\sigma(A_3) = \mathbb{C}$, notice that $\lambda - A_3$ is never surjective (so $\lambda \notin \rho(A_3)$). Consider $g(s) := e^{i\lambda s}$, from the solution above, you can obtain $a = 0$ from the boundary conditions. And so, $f(t) = ie^{-i\lambda t}t$ ($f(1) \neq 0$, so $f \notin D_{A_3}$ so not surjective).

This shows that you can absolutely have a symmetric operator which develops a non-real spectrum. But, if A is self-adjoint, its whole spectrum is real - as we shall see below.

Lemma 5.43 Let $A \subset A^*$ (i.e. A is symmetric). Then,

$$\forall z \in \mathbb{C}\ \forall u \in D_A, \quad \|(z-a)u\|_H \geq |\mathrm{Im}(z)|\|u\|_H$$

(so for $z \notin \mathbb{R} \implies (z - A)$ is injective, *i.e.* $z \notin \sigma_p(A)$).

Proof. For $u \in D_A$, $\langle u, Au\rangle = \langle Au, u\rangle$ (by symmetric) $= \overline{\langle u, Au\rangle} \in \mathbb{R}$. Hence,

$$\begin{aligned}|\mathrm{Im}(z)|\|u\|_H^2 &= |\mathrm{Im}(\langle u, (z-A)u\rangle)| \\ &\leq |\langle u, (z-A)u\rangle| \leq \|u\|_H\|(z-A)u\|_H\end{aligned}$$

□

Proposition 5.44 If $A = A^*$ (*i.e.* A is self-adjoint), then A is closed.

It is not difficult to show the proposition above and it follows directly from the definitions.

Proposition 5.45 If $A = A^*$ (*i.e.* A is self-adjoint), then $\sigma(A) \subset \mathbb{R}$.

Proof. Let $z \in \mathbb{C}\backslash\mathbb{R}$, then we want to show $z \in \rho(A)$ *i.e* $z - A : D_A \to H$ is bijective, with $(z-A)^{-1} \in \mathcal{L}(H)$.

We will show $z - A$ is surjective (since for injectivity apply Lemma 5.43 and $\|(z-A)^{-1}\|_{\mathcal{L}(H)} \leq \frac{1}{|\mathrm{Im}(z)|} < \infty$).

To show this, we first argue $M := \mathrm{im}(z - A) \subset H$ is closed.

Let $v_k = (z-A)u_k \to v$ (in norm) as $k \to \infty$. By lemma 5.43, $\|u_k - u_l\|_H \leq \frac{1}{|\mathrm{Im}(z)|}\|v_k - v_l\|_H \to 0$ (as $k, l \to \infty$).

Hence (u_k) is Cauchy and $u_k \to u$ for some $u \in H$. But since $A = A^*$, it has a closed graph by proposition 5.44, so $v = (z-A)u$, *i.e.* we have shown M, as above, is closed.

Now we show this implies the surjectivity of $z - A$. Assume $M \neq H$. Pick $v \in M^\perp\backslash\{0\}$, then $\forall u \in D_A$

$$\langle v, (z-A)u\rangle = 0 \text{ or } \langle v, Au\rangle = \overline{z}\langle v, u\rangle$$

Hence, $D_A \ni u \mapsto \langle v, Au\rangle$ is continuous, $v \in D_{A^*} = D_A$ and $Av = A^*v = \overline{z}v$.

But by lemma 5.43,

$$\underbrace{|\mathrm{Im}(z)|}_{\neq 0}\|v\|_H \leq \|(\overline{z} - A)v\|_H = 0$$

yielding $v = 0$, which is a contradiction. □

§5.5 Spectral Theorem for Compact Self-Adjoint Operators

As is customary in this section, we take H to be Hilbert over $\mathbb{C}$ with an inner product $\langle\cdot,\cdot\rangle$ and $\|x\|_H^2 = \langle x, x\rangle$.

The following is an extension of the familiar result from linear algebra concerning the diagonalisation of symmetric matrices.

Theorem 5.46 (Riesz-Schauder) Let $T : H \to H$ be a linear operator which is compact and self-adjoint. Then

i) $\sigma_p(T) \subset \mathbb{R}$

ii) $\sigma_p(T)$ contains at most countably many eigenvalues $\lambda_k \in \mathbb{R}\backslash\{0\}$, which accumulate at most at $\lambda = 0$ (although of course they may not accumulate at all).

iii) Let λ_ks be counted with multiplicity (*i.e.* if for instance the eigenspace of λ_1 was two dimensional then you write two copies of λ_1 *etc.*). One can choose eigenvectors e_k (corresponding to λ_k) s.t. $e_k \perp e_l$ $\forall k \neq l$ (*i.e* they are pairwise orthogonal) and one has

$$\forall x \in H \quad Tx = \sum_k \lambda_k e_k\langle x, e_k\rangle$$

Remark 5.47 To clarify point ii) in the statement of Riesz-Schauder: when we have a sequence, we can talk about accumulation points which are points whose neighbourhood you visit infinitely often, and here the only such case is zero.

Point iii) is really just the spectral decomposition of the operator T.

Example 5.48 Consider $T_\lambda : \ell^2 \to \ell^2$ (as seen before with $T_\lambda x := (\lambda_m x_n)_{n\in\mathbb{N}}$). We have previously seen T_λ is compact $\iff \lim_{k\to\infty} \lambda_k = 0$.

And also T_λ is self-adjoint $\iff \lambda_k \in \mathbb{R}$, which implies $\sigma_p(T_\lambda) = \overline{\{\lambda_k : k \in \mathbb{R}\}}$.

Note that compactness of T implies that $T \in \mathcal{L}(H)$. In the lemma below we drop compactness and we just assume we have a bounded operator (that doesn't necessarily have to be compact).

Lemma 5.49 let $T \in \mathcal{L}(H)$ be self-adjoint. If $\lambda_1 \neq \lambda_2$, with $\lambda_1, \lambda_2 \in \sigma_p(T)$ and with eigenvectors $e_i \neq 0$ for $i = 1, 2$ (*i.e.* $\lambda_i e_i = Te_i$), then

$$\langle e_1, e_2 \rangle = 0$$

Proof. We have

$$\lambda_1 \langle e_1, e_2 \rangle = \langle \lambda e_1, e_2 \rangle = \langle Te_1, e_2 \rangle \overbrace{=}^{\text{self-adjoint}} \langle e_1, Te_2 \rangle$$
$$= \langle e_1, \lambda_2 e_2 \rangle \overbrace{=}^{\overline{\lambda_2}=\lambda_2} \lambda_2 \langle e, e_2 \rangle$$

Since $\lambda_1 \neq \lambda_2$, $\langle e_1, e_2 \rangle = 0$. ☐

We can also formulate lemma 5.49 in a slightly diferent way, taking $T : H \to H$ to be compact and self-adjoint (which we can do since this is a stronger condition than given in the lemma).

If we define for $\lambda \in \sigma_p(T)\backslash\{0\}$, $X_\lambda = \ker(\lambda - T)$ $(\neq \{0\})$, then by lemma 5.49

$$X_\lambda \perp X_{\lambda'} \quad \forall \lambda \neq \lambda' \text{ with } \lambda, \lambda' \in \sigma_p(T)\backslash\{0\}$$

Lemma 5.50 Let T be compact, self-adjoint and $\lambda \in \sigma_p(T)\backslash\{0\}$. Then,

i) $\dim(X_\lambda) < \infty$

ii) $\forall r > 0$, $\sigma_p(T)\backslash B_r(0)$ is finite.

We won't prove this lemma, although this lemma can be used to prove statement ii) in Riesz-Schauder theorem 5.46. By lemma 5.50

$$A_n = \sigma_p(T) \cup \{z : \frac{1}{n+1} \leq |z| < \frac{1}{n}\} \quad (\subset \sigma_p(T)\backslash B_{\frac{1}{n+1}}(0))$$

is finite and $\sigma_p(T)\backslash\{0\} = \bigcup_n A_n$ is thus countable. This also implies that $\sigma_p(T)\backslash\{0\}$ has no accumulation point.

Remark 5.51 Lemma 5.50 gives important structural information on the spectrum. In particular, ii) implies that $\sigma_p(T)\backslash\{0\}$ is countable, with no accumulation point, and by i) each eigenvalue has finite multiplicity.

Remark 5.52 From statement iii) in Riesz-Schauder theorem 5.46, we actually have that H admits the orthogonal decomposition

$$H = \ker(T) \oplus \overline{\underbrace{\bigoplus_{\lambda \in \sigma_p(T)\backslash\{0\}}}_{\text{countably many}} X_\lambda}$$

where $X_\lambda = \ker(\lambda - T)$, which are finite dimensional.

Lemma 5.53 If $A \in \mathcal{L}(H)$ is self-adjoint, then

$$\|A^n\|_{\mathcal{L}(H)} = \|A\|^n_{\mathcal{L}(H)}$$

Proof. We proceed by induction. The $n = 1$ case is trivial. Assume it is true for n and consider $n+1$. For $x \in H$, $\|\| \leq 1$,

$$\begin{aligned}\|A^n x\| &= \langle A^n x, A^n x\rangle = \langle A^{n+1}x, A^{n-1}x\rangle \\ &\leq \|A^{n+1}x\|\|A^{n-1}x\| \leq \|A^{n+1}\|\|A^{n-1}\|\end{aligned}$$

By the inductive hypothesis,

$$\begin{aligned}\|A\|^{2n} = \|A^n\|^2 = \sup_{\|x\|\leq 1} \|A^n x\|^2 \leq \|A^{n+1}\| \underbrace{\|A^{n-1}\|}_{\|A\|^{n-1}} \\ \implies \|A^{n+1}\| \geq \|A\|^{n+1}\end{aligned}$$

Now, the "$\leq$" case is very similar (and applying self-multiplicativity of $\|\cdot\|$). □

This motivates the following definition of the spectral radius of an operator A.

Definition 5.54 (Spectral Radius) — The spectral radius r_A of any operator A is defined as

$$r_A := \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}}_{\mathcal{L}(H)}$$

Lemma 5.55 If A is self-adjoint, then

$$\left(\lim_{n\to\infty} \|A^n\|^{\frac{1}{n}}_{\mathcal{L}(H)} =\right) \; r_A = \|A\|_{\mathcal{L}(H)}$$

Theorem 5.56 If $A \in \mathcal{L}(H)$, then $r_A = \sup_{z\in\sigma(A)} |z|$.

Proofs for the above are not given here, although can be easily found online. More details about spectral theory and functional analysis in general can be read in textbooks and other, more thorough, lecture notes online. But for this course, we end the discussion here.