Functional Analysis

2022

This set of notes is based on the lecture notes of Dr. Pierre-François Rodriguez and compiled by Rongkai Zhang, Tim Wang and Ziang Yan.

The `theoremstyle.tex` is adapted from the Imperial-Probability-Theory notes, which is subject to Creative Common Attribution 4.0 International (CC BY 4.0) license.

This is **NOT** an official note.

- The content in gray boxes like this are either content from the original notes or included to facilitate understanding of the material.

Contents

1 Introduction **4**

2 Separability **9**

3 Hilbert Space **11**

4 Finite vs. Infinite Dimensional Spaces **16**

4.1 Compactness 17

5 Linear Operators **20**

6 Duality **23**

6.1 Duality in Hilbert Spaces 23

6.2 Duality in Banach Spaces 26

7 Hahn-Banach Theorem **31**

7.1 Applications of Hahn-Banach (H-B) 35

8 Baire Category and UBP **39**

8.1 Baire Category 39

8.2 Uniform boundedness principle (UBP) 42

8.3 Baire's Original Problem 44

9 Open mapping theorem **47**

10 Closed Graph Theorem **51**

11 Weak vs. Strong topologies **54**

12 Compact operators **60**

13 Spectral Theory **64**

13.1 Spectral Theory in Hilbert space 66

13.2 Spectral theorem for compact self-adjoint operators 70

1 Introduction

What is the course about?

Roughly: solving linear systems of the form

$$Ax = y$$

where $A : X \to Y$ is linear, $y \in Y$ given, find solutions $x \in X$. X and Y are linear spaces.

Remark 1.1 When X and Y are finite dimensional spaces, this is linear algebra.

∞-dim brings into play additional structure, then it's functional analysis. (completeness, compactness, metric, norm)

Example: $f \in C_0^\infty(\mathbb{R}^n)$, solve

$$\underbrace{-\Delta}_{A} u = f \in \mathbb{R}^n$$

where

$$C_0^\infty(\mathbb{R}^n) = \{f : \mathbb{R}^n \to \mathbb{R} : f \text{ infinitely differentiable }, supp(f) = \{x : f(x) \neq 0\} \text{ compact }\}$$

What space? Can take (in PDE) $X = Y = C_\infty(\mathbb{R}^n)$: function spaces. Adequate choice of space to find solutions necessary! con vary, metric/normed linear space, locally convex topological space...

In this course: linear space is (almost always): Banach space or Hilbert space.

Example 1.2 — Running example, L^P. cf. MATH50006 notes §2.6 . $(X, \mathcal{A}, \mu)$:measure space.

$$L^P(\mu) \equiv l^P(X, \mathcal{A}, \mu) = \{f : X \to \mathbb{K}/\sim : \|f\|_{L^P} < \infty\}, \quad p \in [1, \infty]$$

Norm of $f \in L^p$:

$$\|f\|_p \equiv \|f\|_{L^p} = \begin{cases} \left(\int |f|^p \mathrm{d}\mu\right)^{1/p} & p < \infty \\ \operatorname*{ess\,sup}_X |f| & p = \infty \end{cases}$$

Choices of measure space:

<u>Example 0</u>:

$X = \{1, 2, ..., n\}$, $\mathcal{A} = 2^X$, $\mu(\{k\}) = 1, \forall k = 1, ..., n$, counting measure on X, extended to measure by additivity.

Every function $f : X \to \mathbb{R}$ is simple

$$f(x) = \sum_{k=1}^{n} f(k) 1_{\{k\}}(x) \qquad 1_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

Thus

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

So $L^p(\{1, ..., n\}, \mu \cong \mathbb{R}^n$ endowed with norm

$$\|p\|_p = \left(\sum_{k=1}^{n} |f(k)|^p \right)^{1/p} \qquad f \in \mathbb{R}^n$$

which are finite dimensional vector spaces over $\mathbb{R}$ (linear algebra).

Example 1:

Same with $X = \mathbb{N} = \{1, 2, 3, ...\} \rightsquigarrow \ell^p$, "Little-l-p",

i.e. $\mu(\{k\}) = 1 \, \forall k \in \mathbb{N}$, extended to measure by σ-additivity, i.e.

$$\mu(A) = \sum_{k \in A} \mu(\{k\}) = |A| \quad A \subset \mathbb{N}$$

Now every $f : X \to \mathbb{R}$ of form

$$f(x) = \sum_{k=1}^{\infty} f(k) 1_{\{k\}}(x)$$

Approximated by $f_n := \sum_{k=1}^{n} f(k) 1_{\{k\}}(x)$ and use of monotone convergence theorem to get

$$\|f\|_p^p = \sum_{k=1}^{\infty} |f(k)|^p$$

An element $f : X \to \mathbb{R}$, $f = (f(1), f(2), ...) \equiv (f_1, f_2, ..._{\equiv} (f_k)_k$ is a sequence!

$$\ell^p = \{\text{all real-valued sequences } f = (f_k)_k \text{ s.t. } \sum_{k=1}^{\infty} |f_k|^p < \infty\}$$

Specially, when $p = 1$, the space ℓ^1 is the set of all absolutely convergent series!.

Remark 1.3 We can add weights to the measures in examples above, the resulting space is denoted $\ell^p(\eta)$, where we define $\mu(\{j\}) = \eta_j, \eta_j \geq 0, \forall j \in \mathbb{N}$, thus the norm is

$$\|f\|_{\ell^p(\eta)} = \sum_{j=1}^{\infty} |f_j|^p \eta_j$$

Example 2

$X = \mathbb{R}^n$ for some $n \in \mathbb{N}$, $\mathcal{A}$ = Borel σ-algebra, μ =Lebesgue meausre, we denote the space as $L^P(\mathbb{R}^n)$, which has norm:

$$\|f\|_p = \left(\int |f|^p \mathrm{d}x\right)^{\frac{1}{p}}$$

where we are integrating w.r.t the Lebesgue measure.

More generally, $x \subset \mathbb{R}^n$ open/closed $\rightsquigarrow L^P(X)$, e.g. $n = 1$, $X = [0, 1]$.

Theorem 1.4 Let $(X, \mathcal{A}, \mu)$ be any measure space. Then:

i) $\|f\|_p$ defines a norm $\forall p \in [1, \infty]$, triangular inequality (Minkowski ineq.) holds: $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

ii) Holders inequality holds: if $\frac{1}{p} + \frac{1}{q} = 1$ with $p, q \in [1, \infty]$, then $\forall f \in L^p(\mu)$, $\forall g \in L^q(\mu)$, then $f \cdot g \in L^1(\mu)$ and $\|fg\|_{L^1(\mu)} \leq \|f\|_{L^1(\mu)} \|g\|_{L^1(\mu)}$

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

Definition 1.5 — Banach Space. A normed linear space $(X, \|\cdot\|$ which is complete w.r.t the induced metric is called **Banach space**.

Explanation:

- $\|\cdot\|$ is a norm. See definition 3 in notes.
- the induced metric is $d(x, y) = \|x - y\|$. It is a metric (proposition 2 in the notes)
- X complete w.r.t. d: every Cauchy sequence converges.
- Cauchy sequence: $(f_n)_n \subset X$ s.t. $\forall \varepsilon > 0$, $\exists N \, \forall m, n \geq N$, $d(f_m, f_n) < \varepsilon$

Remark 1.6 Theorem asserts $L^P(\mu)$ is a Banach space. Apply in the case of 1 to get immediately all of theorem 2, exercise 10-12 proposition 7 example 12-14 and much more !

Furthur examples

3. $C([a, b]) = \{f : [a, b] \to \mathbb{R}, \text{continuous}\}$ with $\|f\|_\infty = \sup_{x \in [a,b]} |f(x)|$

4. $C^r([a, b])$, similar to above but f is set to be r-times continuously differentiable. In particular

$C = C[(a,b)] = C^0[(a,b)]$. $\|f\|_{r,\infty} = \sup_{x\in[a,b],1\leq k\leq r} |f^{(k)}(x)|$

5. Sobolev space, for solving PDE(not this course).

Proposition 1.7 $(C[0,1], \|\cdot\|_\infty)$ is complete.

General strategy to show completeness of $(X, \|\cdot\|)$: For a given Cauchy sequence $(f_n) \subset X$

1. find candidate limit f
2. show $\|f_n - f\| \xrightarrow{n\to\infty} 0$
3. show $f \in X$

Proof.

STEP I Let $(f_n) \subset C$ be Cauchy sequence, i.e. $\forall \varepsilon > 0$, $\exists N$, $\forall m, n \geq N$: $\|f_m - f_n\|_\infty < \varepsilon$.

But $\|f_n - f_m\|_\infty \geq |f_n(x) - f_m(x)|$, $\forall x \in [0,1]$ so $(f_n(x))$ is Cauchy sequence in $\mathbb{R}$ $\forall x \in [0,1]$. Since $\mathbb{R}$ is complete, it has a limit. Call it $f(x) = \lim_{n\to\infty} f_n(x)$.

STEP II $|f_n(x) - f_m(x)| < \varepsilon \, \forall n, m \geq N$, $\forall x \in [0,1]$, which implies $\lim_{n\to\infty} |f_n(x) - f_m(x)| < \varepsilon$, we also have that $lim_{n\to\infty}|f_n(x) - f_m(x)| = |f(x) - f_m(x)|$ by continuity of f.

STEP III To show $f \in C$, need to argue

$$\forall x,\, \varepsilon > 0,\, \exists \delta : |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon \tag{C1}$$

Now we use $\varepsilon/3$ argument. Write for any $n \in \mathbb{N}$ and $x, y \in [0,1]$,

$$|f(x) - f(y)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \tag{C2}$$

First, using STEP II, pick n s.t. $\|f_n - f\|_\infty < \frac{\varepsilon}{3}$, whence

$$|f(x) - f_n(x)| < \frac{\varepsilon}{3},\ |f(y) - f_n(y)| < \frac{\varepsilon}{3},\ \forall x, y \in [0,1] \tag{C3}$$

Then, with n now fixed, using continuity of f_n, pick $\delta > 0$ s.t.

$$|f_n(x) - f_n(y)| < \frac{\varepsilon}{3} \text{ whenever } |x - y| < \delta \tag{C4}$$

Substitute Equation $(C4)$,Equation $(C3)$ into Equation $(C2)$ to get Equation $(C1)$.

■

Remark 1.8

- eq. $(C3)$ is well-known from analysis I-II. $\lim_{n\to\infty} \|f_n - f\|_\infty = 0$ is precisely the <u>uniform convergence</u> of (f_n) towards f. so 3 asserts that "uniform limit of a sequence of continuous functions is again continuous".

- $f_n(x) = x^n$ is not a Cauchy sequence in $(c, \|\cdot\|\infty$, yet $f_n(x) \xrightarrow{n\to\infty} f(x) = 1_{\{1\}}(x)$.

- if instead consider $f_n(x) = x^n$ to be elements of $L^1([0,1])$:

$$\int_0^1 f_n \mathrm{d}x = \left.\frac{x^{n+1}}{n+1}\right|_0^1 = \frac{1}{n+1} < \infty$$

  Then $(f_n) \subset L^1([0,1])$ a Cauchy sequence(Ex.), and converges by completeness of $L^p(\mu)$. The limit in $L^1([0,1])$ is $f = 0$:

$$\|f_n\|_{L^1([0,1])} = \|f_n - 0\|_{L^1([0,1])} = (n+1)^{-1} \xrightarrow{n\to\infty} 0$$

2 Separability

In this section, we will be working with metric spaces (V, ρ).

Definition 2.1 — Separable. A metric space (V, ρ) is **separable** if $\exists D \subset V$ countable, such that $B_\rho(x, \varepsilon) \cap D \neq \varnothing, \forall x \in V, \forall \varepsilon > 0$ (*i.e.* D is dense in V)

Here $B_\rho(x, \varepsilon) = \{y \in V : \rho(y, x) < \varepsilon\}$ is an open ball of radius ε centred at x. For convenience, ρ is often dropped in the notation, writes $B(x, \varepsilon)$

Proposition 2.2 ℓ^p space is separable for $p \in [1, \infty)$

Here ℓ^p actually denotes (ℓ^p, ρ), where ρ is the metric induced by the p-norm $\|\cdot\|_p$, *i.e.* $\rho(x, y) = \|x - y\|_p$

Proof. Consider $D = \cup_{n \geq 1} D_n$, where

$$D_n = \{x = (x_n) : x_n \in \mathbb{Q}, \forall n \in \mathbb{N}, \text{ and } x_k = 0, \forall k > n\}$$

Clearly $D_n \subset \ell^p$, hence $D \subset \ell^p$ and $D_n \cong \mathbb{Q}^n$ is countable, hence D is also countable.

Claim: D is dense in ℓ^p.

Let $x = (x_n) \in \ell^p$ and $\varepsilon > 0$. First we build a $\|\cdot\|_p$ −close sequence $\widetilde{x} = (\widetilde{x}_n)$ with values in $\mathbb{Q}$. Since $\mathbb{Q} \subset \mathbb{R}$ is dense, we find for every n a number $\widetilde{x}_n \in \mathbb{Q}$ s.t.

$$|x_n - \widetilde{x}_n| \leq \left(\frac{\varepsilon}{2}\right)(2^{-\frac{n}{p}})$$

This implies

$$\|x - \widetilde{x}\|_p^p = \sum_{n=1}^{\infty} |x_n - \widetilde{x}_n|^p \leq \left(\frac{\varepsilon}{2}\right)^p \sum_{n=1}^{\infty} 2^{-n} = \left(\frac{\varepsilon}{2}\right)^p \tag{2.1}$$

Note that this also implies $\widetilde{x} \in \ell^p$, since $\|\widetilde{x}\|_p \leq \underbrace{\|\widetilde{x} - x\|_p}_{<\infty} + \underbrace{\|x\|_p}_{<\infty}$.

Since $\widetilde{x} \in \ell^p$, we have $\sum_n |\widetilde{x}_n|^p < \infty$ hence we can pick an n s.t.

$$\sum_{k \geq n} |\widetilde{x}_n|^p < \left(\frac{\varepsilon}{2}\right)^p \tag{2.2}$$

Now define $y = (\widetilde{x}_1, \ldots, \widetilde{x}_n, 0, 0, \ldots,)$. Clearly, $y \in D_n$. Moreover, Equation (2.2) asserts that $\|\widetilde{x} - y\| < \frac{\varepsilon}{2}$. Combining with Equation (2.1) and using the triangle inequality yields $\|x - y\|_p < \varepsilon$, *i.e.* $x \in B(y, \varepsilon)$. ∎

Proposition 2.3 $L^p(\mathbb{R}^n), p \in [1, \infty)$ is separable

Proof. (Sketch) Consider

$$C = \{Q \text{ dyadic cube}, i.e.\ Q = x + [0, 2^{-l}) \text{ for some } x \in 2^{-l}\mathbb{Z}^n (\subset \mathbb{R}^n) \text{ and } l \in \mathbb{N} \cup \{0\}\}$$

Define

$$D = \left\{ g = \sum_{k=1}^{n} a_k \mathbf{1}_{Q_k} : n \in \mathbb{N}, a_k \in \mathbb{Q}, Q_k \in C \right\}$$

Claim: D is dense in $L^p(\mathbb{R}^n), p \in [1, \infty)$

Let $f \in L^p(\mathbb{R}^n)$. Assume $f \geq 0$ (else split into $f = f^+ - f^-$)

Step 1 By approximation of simple functions, we can find $\widetilde{g}$ simple, s.t. $0 \leq \widetilde{g} \leq f$ and

$$\|f - \widetilde{g}\| < \frac{\varepsilon}{3}$$

with $\widetilde{g} = \sum_{k=1}^{m} a_k \mathbf{1}_{A_k}$ for suitable $A_k \in \mathcal{B}(\mathbb{R}^n)$.

Step 2 We can find a sequence of simple functions $\hat{g}$ with coefficients $a_l \to a_k$ where $a_l \in \mathbb{Q}$ such that

$$\|\hat{g} - \widetilde{g}\| < \frac{\varepsilon}{3}$$

Step 3 For each A_k, we can find O_k open s.t.

$$\lambda(O_k \setminus A_k) < \frac{\varepsilon}{6} 2^{-k}$$

And we can approximate O_k using dyadic cubes with precision $\frac{\varepsilon}{6} 2^{-k}$ ∎

It is crucial that $\lambda(A_k) < \infty$, as

$$\forall k \in \mathbb{N} : |a_k|^p \lambda(A_k) \leq \|\widetilde{g}\|_p \leq \|f\|_p < \infty$$

see also MATH50006 proof of (4.13).

Definition 2.4 — Schauder basis. Let $(X, \|\cdot\|)$ be a normed linear space. A **Schauder basis** of X is a sequence of linearly independent $(e_i)_{i \in \mathbb{N}}, e_i \in X$, such that $\forall x \in X$, there is a *unique* sequence $(a_n)_{n \in \mathbb{N}}, a_n \in \mathbb{R}$ with

$$\left\| x - \sum_{i=1}^{n} a_i e_i \right\| \xrightarrow{n \to \infty} 0$$

Proposition 2.5 — Schauder implies separability. If $(X, \|\cdot\|)$ has a Schauder basis, then it is separable.

Proof. Define the set $D \subset X$ as,

$$D = \left\{ \sum_{i=1}^{n} q_i e_i : q_i \in \mathbb{Q} \right\}$$

where (e_i) is a Schauder basis. (if X is over $\mathbb{C}$, then use $q \in \mathbb{Q} + i\mathbb{Q}$)

Then by definition, one can find n and x_i such that

$$\left\| x - \sum_{i=1}^{n} x_i e_i \right\| \leq \frac{\varepsilon}{2} \tag{2.3}$$

Choose $q_i \in \mathbb{Q}$ such that $|q_i - x_i| < \frac{\varepsilon}{2n \sum_{i=1}^{n} \|e_i\|}$, we have

$$\left\| \sum_{i=1}^{n} x_i e_i - \sum_{i=1}^{n} q_i e_i \right\| \leq \sum_{i=1}^{n} |x_i - q_i| \, \|e_i\| \leq \frac{\varepsilon}{2} \tag{2.4}$$

Using triangle inequality and Equation (2.3), Equation (2.4) above, we see

$$\left\| x - \sum_{i=1}^{n} q_i e_i \right\| < \frac{\varepsilon}{2}$$

■

Remark 2.6 The converse of Proposition 2.5 is not true, Per Enflo constructed a counter example that is Banach in this paper.

Example 2.7 A Schauder basis of $\ell^p, p \in [1, \infty)$ is $e_n = (0, \ldots, 0, 1, 0, \ldots, 0, \ldots), n \in \mathbb{N}$ (the n^{th} entry is 0). Take $x = (x_n) \in \ell^p$

$$\left\| x - \sum_{i=1}^{n} x_i e_i \right\|_p^p = \sum_{i=n+1}^{\infty} |x_i|^p \overset{n \to \infty}{\longrightarrow} 0$$

since $\|x\|_p < \infty$.

3 Hilbert Space

In this section we work with linear space H over $\mathbb{K} = \mathbb{R}$. For convenience, we do not study Hilbert space over $\mathbb{C}$ in this section.

Definition 3.1 — Inner Product. Let H be a vector space over $\mathbb{R}$. An **inner product** is a bilinear map (i.e. linear in both argument): $H \times H \to \mathbb{R}$, $(x, y) \mapsto \langle x, y\rangle$ satisfying:

- Symmetric: $\langle x, y\rangle = \langle y, x\rangle, \forall x, y \in H$
- Positive definite: $\langle x, x\rangle \geq 0$,$\langle x, x\rangle = 0$ iff $x = 0$

Definition 3.2 — Inner Product Space. $(H, \langle\cdot,\cdot\rangle)$, a vector space equipped with an inner product is called an inner product space.

Theorem 3.3 If $\langle\cdot,\cdot\rangle$ is an inner product on X, define $\|x\| \stackrel{\text{def}}{=} \sqrt{\langle x, x\rangle}$.

i) (Cauchy-Schwarz) $\forall x, y \in X$,

$$|\langle x, y\rangle| \leq \|x\| \, \|y\|$$

ii) $\|x\|$ is a norm

Proof. i) If $x = 0$ or $y = 0$, the inequality holds. Else, let $\xi = \frac{x}{\|x\|}, \eta = \frac{y}{\|y\|}$, so $\|\xi\| = \|\eta\| = 1$. Hence

$$0 \leq \|\eta - \langle \xi, \eta\rangle \xi\|^2 = \|\eta\|^2 - |\langle \xi, \eta\rangle|^2 = 1 - |\langle \xi, \eta\rangle|^2$$

so $|\langle \xi, \eta\rangle| \leq 1$

ii) Positivity and homogeneity follows from definition of $\langle\cdot,\cdot\rangle$; and triangle inequality follows from i)

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \leq (\|x\| + \|y\|)^2$$

■

Definition 3.4 — Hilbert space. An inner product space $(H, \langle\cdot,\cdot\rangle)$ which is *complete* w.r.t. the metric induced by $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$ is called a **Hilbert space**

Example 3.5 — L^2−spaces. The space $L^2(\mu)$ for all measures μ is a Hilbert space with inner product $\langle f, g\rangle = \int f g d\mu$ and $\langle f, f\rangle = \|f\|_2^2$

Example 3.6 — l^2−spaces. The sequence space $\ell^2 = \left\{\{x_k\}_{k\in\mathbb{N}} : \sum_{k=1}^{\infty} |x_k|^2 < \infty\right\}$ is a Hilbert space with inner product defined by $\langle x, y\rangle = \sum_{k=1}^{\infty} x_k y_k$

Theorem 3.7 — Nearest Point Property.

Let H be a Hilbert space, $K \subset H$ be a closed, convex subset, then $\forall y \in H$ there exists a **unique** $x_0 \in K$ such that

$$\delta \stackrel{\text{def}}{=} \inf_{x \in K} \|x - y\| = \|x_0 - y\|$$

Proof. By considering the set $K - y = \{x - y : x \in K\}$ (still closed and convex), we can assume $y = 0$.

Existence:

By definition of δ, $\exists (x_n)_{n\in\mathbb{N}}$, $x_n \in K$ such that $\lim_{n\to\infty} \|x_n\| = \delta$. We show that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Let $\varepsilon > 0$. Pick $N \in \mathbb{N}$ such that

$$\|x_n\|^2 < \delta^2 + \frac{\varepsilon^2}{4} \qquad \forall n \geq N$$

K being convex implies that $\frac{x_n + x_n}{2} \in K, \forall n, m \in \mathbb{N}$, which implies by definition of δ, $\|x_n + x_m\| \geq 2\delta$.

It follows that for all $n, m \geq N$,

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

where we have used the Parallelogram law (Proposition 3.9).

By completeness, $\exists x_0$ s.t. $x_k \to x_0$ as $k \to \infty$. Since K is closed, the limit $x_0 \in K$ and $\|x_0\| = \delta$ by continuity of the norm $\|\cdot\|$.

Uniqueness:

Take $x_0, x_1 \in K$ with $\|x_0\| = \|x_1\| = \delta$ and assume $x_0 \neq x_1$, then $\frac{1}{2}(x_0 + x_1) \in K$ by convexity and so $\|x_0 + x_1\| \geq 2\delta$. By the Parallelogram law,

$$\|x_0 - x_1\|^2 = 2(\|x_0\|^2 + \|x_1\|^2) - \|x_0 + x_1\|^2 \leq 4\delta^2 - 4\delta^2 = 0$$

So $x_0 = x_1$, a contradiction. ■

Remark 3.8 A good example of K convex is $K \subset H$ a subspace.

Proposition 3.9 — Parallelogram law.

Let $x, y \in H$, then

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

Proof. Then

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \overline{\langle x,y\rangle} + \langle y,y\rangle \\
&= \|x\|^2 + 2Re(\langle x,y\rangle) + \|y\|^2
\end{aligned} \tag{3.1}$$

Similarly,

$$\begin{aligned}
\|x-y\|^2 &= \langle x-y, x-y\rangle \\
&= \langle x,x\rangle - \langle x,y\rangle - \langle y,x\rangle + \langle y,y\rangle \\
&= \langle x,x\rangle - \langle x,y\rangle - \overline{\langle x,y\rangle} + \langle y,y\rangle \\
&= \|x\|^2 - 2Re(\langle x,y\rangle) + \|y\|^2
\end{aligned} \tag{3.2}$$

Adding up, we obtain the identity. ■

If H is Hilbert space, a lot of geometric intuition from linear algebra prevails. For instance, call $x \perp y$ if $\langle x, y\rangle = 0$.

Then if $K \subset H$ is a closed subspace, $y \in H$, then Theorem 3.7 applies with x_0 being a nearest point in K to y and

$$z = (y - x_0) \perp K$$

in other words, $z \perp x, \forall x \in K$.

To see this, assume for the sake of contradiction, $\exists x \in K : \langle z, x\rangle \neq 0$, then

$$\left\| z - \frac{\langle z,x\rangle}{\|x\|^2} x \right\|^2 = \underbrace{\|z\|^2}_{=\delta} - \frac{|\langle z,x\rangle|^2}{\|x\|^2} < \delta$$

which violates the minimality of x_0.

More generally, one has the following definition.

Definition 3.10 — Orthogonal Complement.

For $S \subset H$, H a Hilbert space, we define the **orthogonal complement**

$$S^\perp \stackrel{\text{def.}}{=} \{y \in H : \langle x, y\rangle = 0, \forall x \in S\}$$

We can also check that $S^\perp$ is closed.

Corollary 3.11 — Orthogonal Decomposition.

Let H be a Hilbert space and $E \subset H$ be a closed subspace. Then

$$H = E \oplus E^{\perp}$$

($i.e.$ $E \cap E^{\perp} = \{0\}$ and $H = E + E^{\perp}$, that is $\forall x \in H, x = e + e^{\perp}$ for some $e \in E, e^{\perp} \in E^{\perp}$)

Proof. If $x \in E \cap E^{\perp}$, then $\langle x, x\rangle = 0$, so $\|x\| = 0$, $x = 0$.

For all $x \in H$, the subspace $K \overset{\text{def.}}{=} x + E$ is closed and convex. Thus, by Theorem 3.7, $\exists x_0 \in E$, s.t. $\|x - x_0\| \leq \|x - \eta\|, \forall \eta \in E$.

We show that every $x \in H$ can be written as $x = x_0 + (x - x_0)$.

We have, $\forall \eta \in E$:

$$t = 0 \text{ is a minimum of } t \in [0,1] \mapsto \frac{1}{2}\|x - x_0 + t\eta\|^2$$

which is a quadratic function of t, therefore

$$0 = \frac{d}{dt}\frac{1}{2}\|x - x_0 + t\eta\|^2\Big|_{t=0} = t\|\eta\|^2 + \langle x - x_0, \eta\rangle\Big|_{t=0} = \langle x - x_0, \eta\rangle$$

$i.e.$ $(x - x_0) \in E^{\perp}$ and $x = x_0 + (x - x_0) \in E + E^{\perp}$ ■

4 Finite vs. Infinite Dimensional Spaces

In this section, let X be a linear space.

Definition 4.1 — Equivalence of Norm.

Let X be a linear space. Two norms $\|\cdot\|_a$, $\|\cdot\|_b$ on X are **equivalent** if $\exists C \in [1, \infty)$, such that:

$$\forall x \in X : \frac{1}{C} \|x\|_1 \leq \|x\|_2 \leq C \|x\|_1$$

Proposition 4.2 — Norms are equivalent in finite dim.

If $\dim X < \infty$ (*i.e.* $X \cong \mathbb{K}^n$ for some $n \geq 1$), any two norms on X are equivalent.

Proof. cf. Page 44 Notes, Theorem 9. ■

However, this is not true in infinite dimensional spaces.

Example 4.3 Take $X = C[0,1]$, with $\|\cdot\|_1$ and $\|\cdot\|_\infty$

$$f_n(t) = t^n, \qquad n \geq 1,\ t \in [0,1]$$

then

$$\|f_n\|_1 = \int_0^1 t^n dt = \frac{1}{n+1} \overset{n\to\infty}{\longrightarrow} 0$$

but $\|f_n\|_\infty = 1$.

Proposition 4.4 If $(X, \|\cdot\|)$ is normed, $Y \subset X$ a finite dimensional subspace $\dim Y < \infty$, then $(Y, \|\cdot\|)$ is complete.

Proof. cf. Page 45 Notes, Theorem 10. ■

Remark 4.5 In particular, if $\dim X < \infty$, then one can choose $Y = X$.

Example 4.6 Proposition 4.4 fails if $\dim Y = \infty$. Consider $C[0,2] = Y \subset X = L^1[0,2]$, $(f_n) \subset Y$ with

$$f_n(t) = \begin{cases} t^n, & 0 \leq t < 1 \\ 1, & t \geq 1 \end{cases}$$

(*i.e.* view $(Y, \|\cdot\|_1)$ as a subspace of $(X, \|\cdot\|_1)$).

Then (f_n) is Cauchy in L^1 with $f_n \xrightarrow{L^1} f$ where

$$f(t) = \begin{cases} 0, & t < 1 \\ 1, & t \geq 1 \end{cases}$$

but clearly $y \notin Y$. So $(Y, \|\cdot\|_1)$ is not complete.

As a consequence of Proposition 4.4, one also gets

Corollary 4.7 If $(X, \|\cdot\|)$ is normed, $Y \subset X$ a finite dimensional subspace $\dim Y < \infty$, then $(Y, \|\cdot\|)$ is closed.

Proof. Let $(x_n) \subset Y$ be convergent, $\|x_n - x\| \xrightarrow{n\to\infty} 0$ for some $x \in X$. Need to show $x \in Y$. Since (x_n) is convergent, it is Cauchy, but $x_n \in Y$, $\forall n$, so (x_n) is Cauchy in Y. By Proposition 4.4 $(Y, \|\cdot\|)$ is complete, hence (y_n) converges to a point $y \in Y$, thus Y is closed. ■

4.1 Compactness

In the following, we will let (X, ρ) be a metric space.

Definition 4.8 — Compact.

A set $K \subset X$ is **(sequentially) compact** if every sequence in $(x_n) \subset K$ has a convergent subsequence with limit in K.

Definition 4.9 — Closed.

A set $K \subset X$ is **closed** if for $(x_n) \subset K$, $x_n \xrightarrow{\rho} x \in X$.

Remark 4.10 From Year 2 Analysis, if $\dim X < \infty$,

$$\text{K compact} \iff \text{K closed and bounded}$$

where K is bounded if there exists $R > 0$, $\forall x, y \in K$, $\rho(x, y) \leq R$ ($\rho(\cdot, \cdot)$ is the metric on K).

Remark 4.11 "$\implies$" remains true even if $\dim X = \infty$: let K be a compact set.

Proposition 4.12 K compact $\implies$ K closed and bounded

Proof. Let $K \subset X$ be compact.

Closed:

Let $(x_n) \subset K$, $x_n \xrightarrow{\rho} x \in X$.

We would like to show $x \in K$. By compactness, $\exists (x_{n_k})_k \subset K$ with $\rho(x_{n_k}, \tilde{x}) \xrightarrow{k\to\infty} 0$ for some $\tilde{x} \in K$. But we know $\rho(x_{n_k}, x) \xrightarrow{k\to\infty} 0$ [subsequence of a convergent sequence has the same limit]. So $\tilde{x} = x$.

Bounded:

Assume that K is not bounded. Fix $x_0 \in K$. By definition, $\forall n \geq 1$, we can find x_k s.t.

$$\rho(x_k, x_0) \geq n \tag{*}$$

By compactness, (x_n) has a convergent subsequence (x_k) with limit $x \in K$. But then

$$\rho(x_{n_k}, x_0) \leq \underbrace{\rho(x_{n_k}, x)}_{\xrightarrow{k\to\infty} 0} + \rho(x, x_0) \leq C$$

for some $C > 0$, violating (*) for large k. ■

The " $\Longleftarrow$ " breaks down if $\dim X = \infty$.

Example 4.13 Take the following set in ℓ^1:

$$K = \{e_n = (0, \ldots, 0, 1, 0, \ldots), n \in \mathbb{N}\}$$

Then $\|e_n\|_{\ell^1} = 1$, so K is closed and bounded($\|e_n - e_m\| = 2 \times \mathbf{1}_{n \neq m}$, so any convergent sequence $(x_n) \subset K$ is eventually constant *i.e. it equals e_k for some k*). But the bounded sequence $(x_n) := (e_n)$ has no convergent subsequences, since no subsequence is Cauchy. So K is **not** compact.

Another (illustrative) example is the following

Example 4.14 Consider the set $\overline{B_1} \subset C[0,1]$,

$$\overline{B_1} = \{f \in C[0,1] : ||f||_\infty \leq 1\}$$

Then $\overline{B_1}$ is closed and bounded, but **not** compact. To see this, consider,

$$f_n(t) = \sin(2^n \pi t), \quad 0 \leq t \leq 1$$

Then $||f_n - f_m||_\infty \geq 1$ for all $n \neq m$, so it has no convergent subsequences.

Compactness of the unit ball characterises in fact finite dimensional spaces.

Theorem 4.15 — Characterisation of finite dim. spaces.

In a normed space $(X, \|\cdot\|)$, the following statements are equivalent:

i) $\dim X < \infty$

ii) The unit ball $\overline{B_1}$ is compact

Proof. **i)** $\implies$ **ii)** $\overline{B_1}$ is closed and bounded, which implies compactness since X has finite dimension.

For **ii)** $\implies$ **i)**, one uses

Lemma 4.16 — Riesz.

Let $Y \subset X$, $Y \neq X$ be a *proper* closed subspace of $(X, \|\cdot\|)$. Then for all $\varepsilon \in (0,1)$, $\exists x \in X \setminus Y$ such that,

i) $\|x\| = 1$

ii) $d(x, Y) \overset{\text{def.}}{=} \inf_{y \in Y} \|x - y\| > 1 - \varepsilon$

Proof. Pick any $x^* \in X \setminus Y$. Since Y is closed, $d := d(x^*, Y) > 0$. By the definition of $d(x, Y)$, we can thus find $y^* \in Y$ s.t.

$$d \leq \|x^* - y^*\| < \frac{d}{1 - \varepsilon}$$

Set $x = \frac{x^* - y^*}{\|x^* - y^*\|}$, then i) is satisfied and for all $y \in Y$ one has

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

■

Returning to **ii)** $\implies$ **i)**. We show the contrapositive.

Assume $\dim X = \infty$, let (y_n) be a sequence of linearly independent vectors, define

$$Y_n = \operatorname{span}\{y_k : 1 \leq k \leq n\}$$

note $\dim Y_n < \infty$, so by Corollary 4.7, Y_n is closed.

Pick $x_1 = \frac{y_1}{\|y_1\|}$ and for all $n \geq 2$ using Lemma 4.16 with $X = Y_n$, $Y = Y_{n-1}$, and $\varepsilon = \frac{1}{2}$, we can choose $x_n \in Y_n \setminus Y_{n-1}$ with

$$\|x_n\| = 1 \qquad \text{and} \qquad d(x_n, Y_{n-1}) > \frac{1}{2}$$

Then $\forall m > n$, one has:

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

So (x_n) has no convergent subsequences. Clearly $(x_n) \subset \overline{B_1}$. So $\overline{B_1}$ is not compact. ■

5 Linear Operators

Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed spaces, $A : X \to Y$ linear.

Definition 5.1 — Bounded Operator.

A linear operator $A : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ is bounded if $\exists C \in (0, \infty)$, such that

$$\|Ax\|_Y \leq C \|x\|_X \qquad \forall x \in X$$

If A is bounded,

$$\|A\| \overset{\text{def.}}{=} \sup_{\|x\|_X \leq 1} \|Ax\|_Y \, (< \infty)$$

is the best possible C.

$\|A\|$ is called **operator norm** (it is a norm on $\mathcal{L}(X, Y)$).

For linear operators, boundedness is the same as continuity.

Theorem 5.2 The following are equivalent:

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

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

iii) A is Lipschitz continuous $(\exists L > 0 : \|Ax - Ay\|_Y \leq L \|x - y\|_X, \forall x, y \in X)$

iv) A is bounded

Proof. **iv)** $\implies$ **iii)**

By linearity of A, $\forall x_1 \neq x_2 \in X$,

$$\begin{aligned} \|Ax_1 - Ax_2\|_Y = \|A(x_1 - x_2)\|_Y &= \|x_1 - x_2\|_X \left\| A\left(\frac{x_1 - x_2}{\|x_1 - x_2\|_X}\right) \right\|_Y \\ &\leq \|A\| \|x_1 - x_2\|_X \qquad \text{so can take } L = \|A\| \end{aligned}$$

iii) $\implies$ **ii)** $\implies$ **i)** is clear. To show **i)** $\implies$ **iv)**:

Assume $\|A\| = \infty$, so can find $(x_n) \subset X$ with

$$\|x_n\|_X \leq 1 \qquad 0 < \|Ax_n\|_Y \to \infty \text{ as } n \to \infty$$

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

$$\|A(x_0 + z_n) - Ax_0\|_Y = \|Az_n\|_Y = 1$$

which does not converge to 0 as $n \to \infty$. ∎

Corollary 5.3 If $\dim X < \infty$, $A : X \to Y$ linear. Then A is continuous.

Proof. $\|x\|_* \stackrel{\text{def.}}{=} \|x\|_X + \|Ax\|_Y$ defines a norm on X (check this).

By Proposition 4.2, $\exists C > 0$ 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 5.2. ∎

Example 5.4 Let $X = Y = C[0,1]$, $\|\cdot\|_X = \|\cdot\|_1, \|\cdot\|_Y = \|\cdot\|_\infty$ and $A = id$, then A is not continuous: we show it is not bounded.

Take

$$f_n(x) = \begin{cases} n^2 x & x \in [0, \frac{1}{n}] \\ -n^2 x + 2n & x \in (\frac{1}{n}, \frac{2}{n}] \\ 0 & x \in (\frac{2}{n}, 1] \end{cases} \quad \text{then} \quad \|f_n\|_1 = 1 \quad f_n \in X, \forall n \in \mathbb{N}$$

But we have

$$\|A\| \underbrace{\ge}_{\|f_n\|_1 = 1} \sup_n \|Af_n\|_Y = \sup_n \underbrace{\|f_n\|_\infty}_{n} = \infty$$

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

A classical example is:

Example 5.5 Let $X = C^1[0,1]$, $Y = C[0,1]$,

$$A : X \to Y$$
$$f \mapsto f'$$

by "$A = \frac{d}{dx}$", which is well-defined (*i.e.* if $f \in X$ then $Af \in Y$)

Take $\|\cdot\|_Y = \|\cdot\|_\infty$ and endow X with $\|\cdot\|_X = \|\cdot\|_Y$. Then A is unbounded. Indeed, take

$$f_n(t) = \sin(nt)$$

(or $f_n(t) = t^n$) then $\|f_n\|_X = 1$ but $\|Af_n\|_Y = n \to \infty$.

Note: 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**.(see Definition 5.1)

Now set

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

(really we are setting $\mathcal{L}((X, \|\cdot\|_X), (Y, \|\cdot\|_Y)))$ is a normed linear space with norm (check!)

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

and one has the useful inequality:

$$\forall x \in X : \qquad \|Ax\|_Y \le \|A\| \, \|x\|_X$$

If $X = Y$ **and** $\|\cdot\|_X = \|\cdot\|_Y$ one sets $\mathcal{L}(X,X) = X$.

Theorem 5.6 If $(Y, \|\cdot\|_Y)$ is Banach, then so is $(\mathcal{L}(X,Y), \|\cdot\|_{\mathcal{L}(X,Y)})$.

Proof. Notes Page 53, Theorem 17. ■

Corollary 5.7 If $A : X \to Y$ is continuous and $K \subset X$ is compact, then

$$A(K) = \{Ax : x \in K\} \subset Y$$

is compact

Proof. Fix $(y_n) \subset A(K)$, we need to find a convergent subsequence (y_{n_k}).

By definition of $A(K)$,

$$y_n = Ax_n \qquad \text{for some } x_n \in K$$

So $(x_n) \subset K$, has a convergent subsequence (x_{n_k}) by compactness of K.

Claim: $(y_{n_k})(= Ax_{n_k})$ is convergent.

Indeed, let $\|x_{n_k} - x\|_X \to 0$, $k \to \infty$. By continuity:

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

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

6 Duality

Recall that the space of all bounded linear operators is defined as

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

$\mathcal{L}(X,Y)$ is Banach if Y is Banach and it has norm

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

An Important special case is

$$X^* \overset{\text{def}}{=} \mathcal{L}(X,\mathbb{R})$$

which is the *dual space* of X.

Definition 6.1 — Dual Spaces.

The space of all *continuous* linear operators $\mathcal{L}(X,\mathbb{R})$ is called the **dual space** of X and is denoted as X^*

Remark 6.2 X^* is always Banach (even though X may not be). We often abbreviate $\|\cdot\|_* = \|\cdot\|_{X^*}$. The elements of X^* are called (bounded, linear) **functionals**.

Dual spaces play a central role in functional analysis. They are easiest to grasp in the following contexts.

6.1 Duality in Hilbert Spaces

In this section, let $(H, \langle\cdot,\cdot\rangle_H)$ be a Hilbert space over $\mathbb{R}$. For $y \in H$, we define the map

$$\Lambda_y : X \to \mathbb{R}, \qquad x \mapsto \langle y, x\rangle_H$$

We note that this is an injective map from H to its dual H^* and we will show that this is in fact a bijective isometry.

Lemma 6.3 — Mapping to dual space.

i) $\Lambda_y \in H^*$

ii) The map $\Lambda : H \to H^*$ is a linear isometry with $||\Lambda_y||_* = ||y||$

Proof. i) We need to check the linearity and boundedness of the operator Λ_y^*. The former follows from the linearity of the inner product and the latter is proved by applying Cauchy-Schwarz

$$||\Lambda_y||_* = \sup_{x \in H, ||x|| \leq 1} |\langle y, x \rangle_H| \leq \|y\|_H$$

which implies $\Lambda_y \in H^*$

ii) Choose $x = \frac{y}{\|y\|_H}$ to attain the equality in the equation above , whence we have $\|\Lambda_y\|_* = \|y\|$. ■

Theorem 6.4 — Riesz Representation.

For every $\ell \in H^*$, there is a unique $y \in H$, such that $\ell = \Lambda_y$

Proof. We show the existence and uniqueness of such a linear operator.

- **(Existence)** If $\ell(x) \equiv 0$, then take $y = 0$. Otherwise, assume $\|\ell\|_* = 1$ (as we can replace $\ell(\cdot)$ by $\frac{\ell(\cdot)}{\|\ell\|_*}$). By the definition of $\|\cdot\|_*$, there is a sequence of $(y_n)_{n\in\mathbb{N}} \subset H$ with

$$\ell(y_n) \to \|\ell\|_*, \qquad \|y_n\| = 1, \forall n \in \mathbb{N}$$

We will show that the limit of this sequence is the desired y.

Note that for negative $\ell(y_n)$, we may multiply by -1 using linearity

Claim 1: The sequence $(y_n)_{n\in\mathbb{N}}$ is Cauchy

Apply the parallelogram identity to $x = \frac{y_n}{2}$ and $y = \frac{y_m}{2}$, so we have

$$\forall n, m \geq 1, \qquad \left\| \frac{y_n - y_m}{2} \right\|^2 = 1 - \left\| \frac{y_n + y_m}{2} \right\|^2$$

Using linearity and boundedeness of ℓ,

$$\frac{1}{2} l(y_n) + l(y_m) = l(\frac{y_n + y_m}{2}) \leq \|l\|_* \left\| \frac{y_n + y_m}{2} \right\|$$

The LHS of the equation above converges to 1 by assumption on $(y_n)_{n\in\mathbb{N}}$, which implies and $(y_n)_{n\in\mathbb{N}}$ is Cauchy. Since H is complete, there is a unique y, such that $y_n \to y$

Claim 2: $\ell = \Lambda_y$

Since $\text{span}\{y\}$ is closed, we can consider the orthogonal decomposition $H = \text{span}\{y\} \oplus (\text{span}\{y\})^\perp$. It suffices to show:

(1) $\ell(y) = \Lambda_y(y), \forall y \in \text{span}\{y\}$

(2) $\ell(x) = \Lambda_y(x), \forall x \in (\text{span}\{y\})^\perp$

To show (1), assume wlog $\|y\| = 1$, we note by continuity of ℓ

$$\ell(y) = \lim_{n\to\infty} |\ell(y_n)| = \|\ell\|_* = 1$$

and

$$\|y\|_H^2 = \langle y, y\rangle_H = \Lambda_y(y) = 1$$

So $\Lambda_y(y) = \ell(y)$.

To show (2), we need to argue that

$$\ell(x) = 0, \qquad \forall x \in (\text{span}\{y\})^\perp$$

Now take $y_a = \frac{y+ax}{\sqrt{1+a^2}}$ and $\|y_a\| = 1$, where $y \in \text{span}\{y\}, a \in \mathbb{R}$ and $x \in (\text{span}\{y\})^\perp$. By definition of the norm $\|\cdot\|_*$ and (1),

$$\ell(y_a) \leq |\ell(y_a)| \leq 1 = \ell(y)$$

So $\ell(y_a)$ has a global maximum at $a = 0$ $(y_0 = y)$. Therefore,

$$0 = \frac{d}{da}\ell(y_a)\Big|_{a=0} = \frac{d}{da}\frac{1}{\sqrt{1+a^2}}(\ell(y) + a\ell(x)) = \ell(x)$$

So $\ell(x) = \Lambda_y(x), \forall x \in (\text{span}\{y\})^\perp$.

- **(Uniqueness)** If $\ell = \Lambda_y = \Lambda_z$ for some $y, z \in H$, then

$$\forall x \in H \qquad \Lambda_y(x) = \langle y, x\rangle_H = \langle z, x\rangle_H = \Lambda_z(x)$$

Pick $x = y - z$, then $\langle y, z - x\rangle_H = \langle z, z - x\rangle_H \implies \|y - z\|_H^2 = 0$. Hence $y = z$.

■

Corollary 6.5 All Hilbert spaces H are isomorphic to their duals H^*.

Example 6.6 Some of the examples are:

- $(l^2)^* \cong l^2$

- $(L^2(\mu))^* \cong L^2(\mu)$

6.2 Duality in Banach Spaces

Theorem 6.7 — Conjugates are duals.

For all $p \in (1, \infty)$, $(\ell^p)^* \cong \ell^q$, where $\frac{1}{p} + \frac{1}{q} = 1$

Proof. For $y \in \ell^q$ define

$$\begin{aligned} \Lambda_y : \ell^p &\to \mathbb{R} \\ x \mapsto \Lambda_y(x) &= \sum_{n \in \mathbb{N}} y_n x_n \end{aligned}$$

Lemma 6.8

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

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

i) By Hölder's inequality,

$$|\Lambda_y(x)| \leq \sum_{n=1}^{\infty} |y_n x_n| \leq \|y\|_q \|x\|_p$$

in particular Λ_y is well-defined (i.e. maps to $\mathbb{R}$). The inequality implies $\|\Lambda_y\|_* \leq \|y\|_q$. In fact, one has equality. Let $x = (x_n)_n$ with

$$x_n = \text{sign}(y_n)|y_n|^{q-1}$$

where the sign function is

$$\text{sign}(t) = \begin{cases} 1 & t \geq 0 \\ -1 & t < 0 \end{cases}$$

with $\text{sign}(t)t = |t|, \forall t \in \mathbb{R}$. Then $x \in \ell^p$ as $|x_n|^p = |y_n|^{p(q-1)} = |y_n|^q$ so

$$\|x\|_p = \left(\sum_{n \in \mathbb{N}} |y_n|^q \right)^{\frac{1}{p}} = \|y\|_q^{\frac{q}{p}} = \|y\|_q^{q-1}.$$

and

$$|\Lambda_y(x)| = |\sum_{n \in \mathbb{N}} x_n y_n| = \sum_{n \in \mathbb{N}} |y_n|^q = \|y\|_q^q = \|y\|_q \|x\|_p$$

which implies

$$\|\Lambda_y\|_* \geq \frac{|\Lambda_y(x)|}{\|x\|_p} = \frac{\Lambda_y(x)}{\|x\|_p} = \|y\|_q$$

from which ii) follows.

To complete the proof of the theorem, we have to show the following analogue of Riesz representation theorem (which applies only when $p = q = 2$).

Lemma 6.9 $\Lambda : \ell^q \to (\ell^p)^*$ given by $y \mapsto \Lambda_y$ is surjective (onto).

Let $e_n = (0, \dots, 0, 1, 0, \dots) \in \ell^p$ and define $y_n = \ell(e_n) \in \mathbb{R}$ for $\ell \in (\ell^p)^*$

Claim:

i) $y = (y_n)_{n\in\mathbb{N}} \in \ell^q$

ii) For every $\ell \in (\ell^p)^*$, $\ell = \Lambda_y$ for some $y \in \ell^q$

Consider the "truncated y": $y^{(n)} = (y_1, \dots, y_n, 0, \dots) = \sum_{i=1}^n y_i e_i \in \ell^q$ and let

$$x^{(n)} = \sum_{i=1}^n |y_i|^{q-1} \operatorname{sign}(y_i) e_i \in \ell^p$$

Then as before: $\left\|x^{(n)}\right\|_p = \left\|y^{(n)}\right\|_q^{q-1}$ with

$$\ell(x^{(n)}) = \sum_{i=1}^n |y_i|^{q-1} \operatorname{sign}(y_i) \ell(e_i) = \sum_{i=1}^n |y_i|^q = \left\|y^{(n)}\right\|_q^q$$

where $y_i = \ell(e_i)$. Hence

$$\left(\sum_{i=1}^n |y_i|^q\right)^{\frac{1}{q}} = \left\|y^{(n)}\right\|_q = \frac{\ell(x^{(n)})}{\left\|y^{(n)}\right\|_q^{q-1}} = \frac{\ell(x^{(n)})}{\left\|x^{(n)}\right\|_p} \leq \|\ell\|_* < \infty$$

and letting $n \to \infty$, Claim (i) follows.

For ii), let $x \in \ell^p$ and $\varepsilon > 0$. Since e_n's form a Schauder basis, we know that

$$\left\|x^{(n)} - x\right\|_p \overset{n\to\infty}{\to} 0$$

(since by definition of Schauder basis, we have unique rep. of $x = \sum_{n\in\mathbb{N}} x_n e_n$)

By choosing n large, we can ensure that

$$|\ell(x) - \ell(x^{(n)})| < \frac{\varepsilon}{2} \qquad |\Lambda_y(x) - \Lambda_y(x^{(n)})| < \frac{\varepsilon}{2}$$

using continuity of $\ell(\cdot)$ and $\Lambda_y(\cdot)$, where the latter is a consequence of Lemma 6.8.

But writing

$$|\ell(x) - \Lambda_y(x)| \leq |\ell(x) - \ell(x^{(n)})| + |\ell(x^{(n)}) - \Lambda_y(x^{(n)})| + |\Lambda_y(x^{(n)}) - \Lambda_y(x)| \tag{6.1}$$

and observing that $\ell(x^{(n)}) = \sum_{i=1}^n x_i \ell(e_i) = \sum_{i=1}^n x_i y_i = \Lambda_y(x^{(n)})$, it follows that

$$|\ell(x) - \Lambda_y(x)| \leq \varepsilon \quad \text{and} \quad \ell = \Lambda_{\mathrm{y}}$$

by letting $\varepsilon \downarrow 0$ ■

Remark 6.10 1) The proof extends to $p = 1$, so $(\ell^1)^* = \ell^\infty$. In fact, for $(X, \mathcal{A}, \mu)$, one has

$$L^p(\mu)^* \cong L^q(\mu) \qquad \forall p \in [1, \infty)$$

2) For $p = \infty$, one can still define

$$\Lambda : \ell^1 \to (\ell^\infty)^* \qquad y \mapsto \Lambda_y$$

as before and check that Λ is a linear isometry between Banach spaces.

However, it is **not** surjective.

To see this, consider

$$c_0 = \{(x_n) : \lim_{n \to \infty} = 0\} \subset \ell^\infty$$

Claim: $(c_0)^* \cong \ell^1$

Proof. (Sketch) We show the following statements are true:

1) $\Lambda_y : c_0 \to \mathbb{R}, x \mapsto \Lambda_y(x)$ is well-defined and bounded for any $y \in \ell^1$.

$$|\Lambda_y(x)| \leq \|x\|_\infty \|y\|_1$$

2) The map $\ell^1 \to c_0^*, y \mapsto \Lambda_y$ is a linear isometry. To check $\|\Lambda_y\|_* \geq \|y\|_1$, use $x = \sum_{i=1}^n \mathrm{sign(y_i)e_i}$, as before, where $n \geq 1$.

Clearly, $x \in c_0$ and

$$\Lambda_y(x) = \sum_{i=1}^n \mathrm{sign(y_i)}\Lambda_{\mathrm{y}}(\mathrm{e_i}) = \sum_{\mathrm{i=1}}^{\mathrm{n}} |\mathrm{y_i}|$$

so for $y \neq 0$

$$\|\Lambda_y\|_* \geq \frac{|\Lambda_y(x)|}{\|x\|_\infty} = |\Lambda_y(x)| = \Lambda_y(x) = \sum_{i=1}^{n} |y_i|$$

and letting $n \to \infty$ gives the result.

3) $y \mapsto \Lambda_y$ is onto: similar as before (exercise),

■

Definition 6.11 — Dual Operators.
$(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ are normed spaces over $\mathbb{R}$, and $A : X \to Y$ is a bounded linear operator. Then the dual operator

$$A^* : Y^* \to X^*$$

is defined by

$$A^* y^* \stackrel{\text{def}}{=} y^* \circ A \qquad \forall y^* \in Y^*$$

where $A^* y^* : X \to \mathbb{R}$ is a linear functional in X^*.

A note on the notation: For $\ell \in X^*$, instead of $\ell(x)$, we write $\langle \ell, x \rangle$. Then the above is equivalent to

$$\langle A^* y^*, x \rangle = \langle y^*, Ax \rangle \qquad \forall x \in X, y^* \in Y^*$$

Later we will show that if A is a bounded linear operator, then A^* is also bounded and $\|A^*\| = \|A\|$ using Hahn-Banach; if X, Y are Hilbert spaces, then we can use the Riesz representation theorem instead.

Example 6.12 1) $A \in \mathbb{R}^{m \times n}$ a real matrix, which induces linear map $L_A : \mathbb{R}^n \to \mathbb{R}^m, L_A x = Ax$. If A^T is the transpose of A and $i_k : \mathbb{R}^k \to (\mathbb{R}^k)^*$ is the canonical isomorphisim, then

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

2) More generally, if H is a Hilbert space, $A : H \to H$ is a bounded linear operator with $\Lambda : H \to H^*$ as the canonical isomorphism, the operator

$$\tilde{A^*} \stackrel{\text{def}}{=} \Lambda^{-1} \circ A^* \circ \Lambda : H \to H$$

is called the adjoint of A (and one writes A^* for $\tilde{A}^*$ with abuse of notation). Thus,

$$\left\langle \tilde{A}^* y, x \right\rangle = \langle y, Ax \rangle, \qquad x, y \in H$$

where $\langle \cdot, \cdot \rangle$ is the inner product on H. If $A = \tilde{A}^*$, then A is **self-adjoint**.

7 Hahn-Banach Theorem

Definition 7.1 — Sublinear functional.

Let X be a vector space. $p : X \to \mathbb{R}$ is called **sublinear** if the following holds

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

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

Example 7.2 Any linear functional is also sublinear. Also, $p(x) = \|x\|$ on X is sublinear.

Theorem 7.3 — Hahn-Banach.

Let $M \subset X$ be a linear subspace, $p : X \to \mathbb{R}$ is sublinear, and $f : M \to \mathbb{R}$ is linear with

$$f(x) \leq p(x) \qquad \forall x \in M \tag{$*$}$$

Then, there exists a linear map $F : X \to \mathbb{R}$ with $F|_M = f$ and

$$F(x) \leq p(x) \qquad \forall x \in X$$

Remark 7.4 — Geometric Intuitions.

Take $x = \mathbb{R}^n$, $0 \in M \subset X$ an open convex subset, if $x_1 \notin< M$, then one can find a $f : X \to \mathbb{R}$ linear "separating" x_1 from M:

i.e.

$$\begin{cases} f(x) < a & , x \in M \\ f(x) \geq a & x \notin M \end{cases}$$

for some $a \neq 0$. One can assume $a = 1$ by replacing f by $\frac{f}{a}$.

To find f, we introduce

$$p(x) \stackrel{\text{def.}}{=} \inf\{r > 0 : \frac{x}{r} \in M\}$$

Minkowski functional, cf. PS6

(e.g. $p(x) = \|x\|$ if $M = \{x : \|x\| < 1\}$ in a normed space $(X, \|\cdot\|)$).

One can check (using continuity) that p is sublinear and

$$p(x) < 1 \iff x \in M$$

" $\implies$ ": if $p(x) < 1$, then $\exists \varepsilon > 0$ s.t. $\frac{x}{1-\varepsilon} \in M$, hence

$$x = (1-\varepsilon)\frac{x}{1-\varepsilon} + \varepsilon \cdot 0 \in M$$

by convexity of M.

" $\impliedby$ ": $x \in M \overset{M \text{ open}}{\implies} \frac{x}{1-\varepsilon} \in M$ which implies that $p(x) \leq 1 - \varepsilon$.

To find f with above properties, it is thus enough to ensure that $f(x_1) = 1$ and $f(x) \leq p(x)$, for all $x \in X$.

Proof. If $M = X$, take $F = f$. Else, choose $x_1 \notin M$ and set

$$M_1 = \{x + tx_1 : x \in M, t \in \mathbb{R}\}$$

($M \subset X$ linear). Extend f to M_1, such that $(*)$ holds on M_1, then 'repeat'

Step (I)

For all $x, y \in M$:

$$f(x) + f(y) \overset{\text{f linear}}{=} f(x+y) \overset{(*)}{\leq} p(x+y) \leq p(x - x_1) + p(x_1 + y)$$

hence

$$\forall x, y \in M : \qquad f(x) - p(x - x_1) \leq p(x_1 + y) - f(y) \tag{$**$}$$

Take supremum over x, with y fixed, we get

$$\alpha \overset{\text{def.}}{=} \sup_{x \in M} (f(x) - p(x - x_1)) < \infty$$

Define $f_1 : M \to \mathbb{R}$ by

$$f_1(x + tx_1) = f(x) + t\alpha \qquad , x \in M, t \in \mathbb{R}$$

Lemma 7.5

1) f_1 is linear

2) $f_1 |_M = f$

3) $f_1(x) \leq p(x), \forall x \in M_1$, (*i.e.* $(*)$ for (f_1, M_1) instead of (f, M))

The first two properties are easy to check.

Proof of 3):

By definition of α, $\forall x \in M : f(x) - \alpha \leq f(x) - (f(x) - p(x - x_1)) = p(x - x_1)$. On the other hand, taking supremum over x in $(**)$ yields, for all $y \in M$

$$f(y) + \alpha \leq p(y + x_1)$$

Overall, we have

$$\forall x \in M : \qquad f_1(x \pm x_1) = f(x) \pm \alpha \leq p(x \pm x_1)$$

Apply with $t^{-1}x (\in M)$ for $t > 0$ in place of x and multiply both sides by t to find

$$\forall x \in M, t > 0 : \qquad f_1(x \pm t x_1) \leq t p(t^{-1} x \pm x_1) = p(x \pm t x_1)$$

■

NB: If X has a countable basis $\{e_1, i \geq 1\}$ [e.g. $\ell^p, p \in [1, \infty)$], then we can take $x_1 = e_1$ and proceed by induction. For the general case, we use:

Step (II)

Proof. (continued:)

Lemma 7.6 — Zorn's lemma.

Let $(P, \leq), P \neq \emptyset$ is a nonempty partially ordered set and every totally ordered subset has an upper bound, then P has a maximal element.

Definition 7.7 — Partial Order.

A **partial order** on set X, is a binary relation, written generically $\leq$, satisfying following property.

- transitivity: if $a \leq b$ and $b \leq c$ then $a \leq c$
- reflexivity: $a \leq a$
- anti-symmetry: if $a \leq b$ and $b \leq a$ then $a = b$

If we also have that for any a and b, either $a \leq b$ or $b \leq a$, then we say $\leq$ is a total order.

Definition 7.8 — Upper bound.

Let X be a set partially ordered by $\leq$ and $Y \subset X$, we say an element $x \in X$ is an **upper bound** of Y if $y \leq x$ for all $y \in Y$

Definition 7.9 — Maximal element.

Let X be a set partially ordered by $\leq$ and $Y \subset X$. say $x \in X$ is a **maximal element** of X if $x \leq m$ implies $m = x$.

Take

$$P = \{(N, g) : N \subset X \text{ linear subspace}, g : N \to \mathbb{R} \text{ linear}, \ g\,|_N = f, \ g\,|_N \leq p\}$$

and define

$$(N, g) \leq (O, h) \overset{\text{def.}}{\Longleftrightarrow} N \subset O, h\,|_N = g$$

Then $(P, \leq)$ is partially ordered, $(M, f) \in P$ so $P \neq \emptyset$. Assume $(N_i, g_i)_{i \in I}$ is a totally ordered subset. Set $N = \cup_{i \in I} N_i$ and for $x \in N$,

$$g(x) = g_i(x) \qquad \text{if } x \in N_i$$

Then $(N, g) \in P$. Indeed $N \subset X$ is linear, and g is well-defined and linear with $g \leq p$ on N.

- (Well-defined) If $x \in N_i \cap N_k$, $N_i \subset N_k$, then $g_k\,|_{N_i} = g_i$, hence $g_i(x) = g_k(x)$
- (Linear) If $x, y \in N$, then $x \in N_i, y \in N_k$, for some $i, k \in I$ and $N_i \subset N_k$ (or vice versa), so $x, y \in N_k$ and
$$g(x + y) = g_k(x + y) = g_k(x) + g_k(y) = g(x) + g(y)$$
- (Bounded by $p(x)$) Similarly, one can check $g \leq p$ on N (exercise)

(N, g) is an upper bound for $(N_i, g_i)_{i \in I}$, since $N_i \subset N$ and $g\,|_{N_i} = g_i$ by definition

So Lemma 7.6 applies and yields that $(P, \leq)$ has a maximal element $(N, g) \in P$. Set $F = g$. By definition of P, all properties required of F hold and $N = X$. For, otherwise, one can apply **Step(I)** to (N, g) and find $(N_1, g_1) \in P$ with $(N, g) < (N_1, g_1)$, which violates the maximality of (N, g). ■

Remark 7.10 There is a version of Hahn-Banach theorem for X over $\mathbb{C}$ and $p : X \to \mathbb{R}$ is called $\mathbb{C}$-sublinear if

i) $p(\alpha x) = |\alpha| p(x), \forall x \in X, \alpha \in \mathbb{C}$

ii) $p(x + y) \leq p(x) + p(y), \ \forall x, y \in X$, same as in over $\mathbb{R}$

Then Theorem 7.3 holds for X over $\mathbb{C}$, if p is $\mathbb{C}$-sublinear and $(*)$ replaced by

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

The conclusion is the same with $|F| \leq p$ on X instead.

Proof. Apply Hahn-Banach to $f_1 = \operatorname{Re}(f)$ (linear!) and $f_2 = \operatorname{Im}(f)$. ∎

From now on assume X is over $\mathbb{R}$.

7.1 Applications of Hahn-Banach (H-B)

Let $(X, \|\cdot\|_X)$ be a normed vector space, we have the following corollaries.

Corollary 7.11 Let $M \subset X$ be a linear subspace, f a bounded linear functional on M. Then $\exists F \in X^*$ with

$$F|_M = f \text{ and } \|F\|_{X^*} = \|f\|_{M^*}$$

Proof. Define $p : X \to \mathbb{R}$ via

$$p(x) = \|x\|_X \|f\|_{M^*}$$

Note that p is sublinear and $\forall x \in M$ and,

$$f(x) \leq |f(x)| = \|x\|_X \frac{|f(x)|}{\|x\|_X} \leq \|x\|_X \|f\|_{M^*} = p(x)$$

Now apply Hahn-Banach to obtain $F : X \to \mathbb{R}$, with

$$\|F(x)\|_X \leq \|x\|_X \|f\|_{M^*} \implies \|F\|_{X^*} \leq \|f\|_{X^*}$$

and the other direction of the inequality follows as $F|_M = f$

∎

Theorem 7.12 Let X be a normed linear space. $\forall x \in X$, $\exists x^* \in X^*$ s.t.

$$\langle x^*, x\rangle \equiv x^*(x) = \|x\|_X^2 = \|x^*\|_{X^*}^2$$

Proof. Let $M = \operatorname{span}(x)$. Define $f : M \to \mathbb{R}$

$$f(tx) = t\|x\|_X^2 \qquad \forall t \in \mathbb{R}$$

Then f is linear, and

$$\|f\|_{M^*} = \sup_{\|tx\|_X \le 1} |f(tx)| = \|x\|_X$$

Then we apply Corollary 7.11 to extend f to $x^* = F \in X^*$, with $\|x^*\|_{X^*} = \|f\|_M = \|x\|_X$ and $\langle x^*|_M, x\rangle = f(x) = \|x\|_X^2$ ■

Remark 7.13 Theorem 7.12 gives dual characterisation of the norm later.

When the space is a Hilbert space, this theorem becomes Riesz representation theorem. (without changing notation! That's why bracket is a good notation here) In short, this theorem says that you can always find a linear functional such that for its value for a chosen element is precisely the norm of this element.

Using Hahn-Banach, one can "separate" all sorts of things. Two examples:

1) Separating points

Proposition 7.14 $\forall x, y \in X$, $x \neq y$, there exists $\ell \in X^*$, such that $\ell(x) \neq \ell(y)$

Proof. Choose $\ell \in X^*$ according to Corollary 7.11 with $y - x$ in place of x. Then

$$\ell(x - y) = \ell(x) - \ell(y) = \|y - x\|_X^2 > 0$$

Thus, $\ell(x) \neq \ell(y)$. ■

2) Separating points from closed subspaces (Urysohn-type result)

Theorem 7.15 $M \subset X$ linear, closed. Assume $x_0 \in X \setminus M$, such that

$$d = \mathrm{dist(x_0, M)} = \inf_{\mathrm{x \in M}} \|\mathrm{x_0 - x}\|_{\mathrm{M}} > 0$$

Then $\exists \ell \in X^*$ with $\ell|_M = 0$ and

$$\|\ell\|_{X^*} = 1, \ \ell(x_0) = d$$

Proof. (Sketch) Let $M_0 = \{x + tx_0 : x \in M\}$. Define a linear functional $f : M_0 \to \mathbb{R}$,

$$f(x + tx_0) = td$$

Then $f\,|_M = 0$ and $f(x_0) = d$. Check $\|f\|_{M_0^*} = 1$ (exercise).

Now apply Corollary 7.11 to obtain extension $\ell \overset{\text{def.}}{=} F \in X^*$ with $\|\ell\|_{X^*} = 1$. ■

Remark 7.16 In the proof above, we utilized the fact M is a linear subspace to construct subspace M_0. For a similar result but with convex, closed subset, see Problem Sheet 6.

The previous Theorem 7.15 has a lot of mileage. For instance, one get

3) Alternative for Theorem 7.12

Proposition 7.17 Let X be a normed linear space. $\forall x \in X$, $\exists x^* \in X^*$ s.t.

$$\langle x^*, x\rangle \equiv x^*(x) = \|x\|_X = \|x^*\|_{X^*}$$

Proof. (Sketch) Apply Theorem 7.15 with $M = \{0\}$, $x_0 = \frac{x}{\|x\|_X}$ to recover Theorem 7.12 with $x^*(x) \overset{\text{def.}}{=} \|x\|_X$. ■

4) X^* separable $\implies$ X separable

Proof. See notes Page 82, which uses Theorem 7.15. ■

In particular, this gives

Corollary 7.18 $(\ell^\infty)^* \ncong \ell^1$

Proof. We know ℓ^1 is separable (use the set $\{\sum_{i=1}^n x_i e_i : x_i \in \mathbb{Q}, n \in \mathbb{N}\}$), so if $(\ell^\infty)^* \cong \ell^1$, then $(\ell^\infty)^*$ is separable and so is ℓ^∞ by the theorem above. But ℓ^∞ is not separable. ■

From Theorem 7.12, one gets a dual characterisation of the norm:

Corollary 7.19

i) $\forall x \in X$: $\|x\|_X = \sup\limits_{\|x^*\|_{X^*} \le 1} |\langle x^*, x\rangle|$

ii) $\forall x^* \in X^*$: $\|x^*\|_{X^*} = \sup\limits_{\|x\|_X \le 1} |\langle x^*, x\rangle|$

The supremum in i) is always achieved.

Proof. For $x = 0$, the RHS of i) is 0 by linearity. Let $x \neq 0$, we show two directions of inequality.

- "$\ge$": By homogeneity, we can assume $\|x\|_X = 1$. If $x^* \in X^*$ satisfies $\|x^*\|_{X^*} \le 1$, then

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

- "$\le$": By Theorem 7.12, $\exists x^* \in X^*$, such that $|\langle x^*, x\rangle| = \|x\|_X^2 = 1$. So the supremum is achieved.

For ii), note that this is the definition of operator norm. ■

Another consequence of Theorem 7.12 is (cf. notes Theorem 23 on Page 67 for the special case $X = Y$ Hilbert, in which case the use of Hahn-Banach can be substituted by the Riesz representation theorem.)

Theorem 7.20 Let X, Y be normed linear spaces 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.

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

where in the last step we used the "$\leq$" direction in the proof of Corollary 7.19 holds and the supremum over y^* is attained by Theorem 7.12. ∎

8 Baire Category and UBP

8.1 Baire Category

We seek a topological characterisation of the ”size” of the sets. First recall some relevant topological terms.

Definition 8.1 — Closure and interior. Let (X, d) be a metric space, $A \subset X$. The **interior** of A, written as $\text{int}(A)$ or sometimes A^o:

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

The closure of A, written as $\bar{A}$ or $cl(A)$:

$$\bar{A} = \bigcap_{U \supset A\, closed} U$$

Definition 8.2 — nowhere dense. A set A is **nowhere dense** if $\text{int}\bar{A} = \emptyset$

Proposition 8.3 $G \subset X$ is open and dense in X if and only if $X \backslash G$ is closed and nowhere dense

Proof. This is left as an exercise. ■

We now present a key lemma.

Lemma 8.4 Let (X, d) be a complete metric space over $\mathbb{R}$, $X \neq \emptyset$. If $X = \bigcup_{k=1}^{\infty} A_k$, where each A_k is closed (i.e. $\bar{A}_k = A_k$), then at least one of the A_k contains an open ball. ($\exists k \in \mathbb{N}$ such that $\text{int}A_k \neq \emptyset$)

Proof. Assume for the sake of contradiction, that

$$\text{int}(A_k) = \emptyset, \qquad \forall k \in \mathbb{N} \tag{$\star$}$$

We pick $x_1 \in X \setminus A_1$, by Proposition 8.3.

Since $X \setminus A_1 = X \cap A_1^c$ is open, we can find $r_1 \in (0, 2^{-1})$ such that

$$B(x_1, r_1) \subset X \setminus A_1$$

We can repeat the construction above and find $x_2 \in B(x_1, 2^{-1}r_1) \setminus A_2$ and $r_2 \in (0, 2^{-2}r_1)$, such that

$$B(x_2, r_2) \subset X \setminus A_2$$

Claim: $\forall n \in \mathbb{N}$, $\exists x_n \in X$ and $0 < r_n < 2^{-n} r_1$ such that,

$$B(x_{n+1}, r_{n+1}) \subset B(x_n, r_n/2) \subset B(x_n, r_n) \subset X \setminus A_n$$

We prove this claim by induction on n.

The base case $n = 1$ is shown above.

For $n = k$, we have by the inductive hypothesis, $B(x_k, r_k) \subset X \setminus A_k$, which is open and dense, so we can choose $x_{k+1} \in B(x_k, r_k) \setminus A_k$ and $B(x_{k+1}, r_{k+1}) \subset B(x_k, r_k) \setminus A_k$.

The sequence $(x_n)_{n \geq 1}$ is Cauchy. By (1)

$$d(x_m, x_n) \overset{x_m \in B(x_n, r_n/2)}{\leq} r_n/2 < 2^{-(k+1)} r_1, \qquad \forall m \geq n \geq 1$$

By completeness, $\exists x_* \in X$, such that $d(x_n, x_*) \to 0$ as $n \to \infty$.

Claim: $x_* \in B(x_n, r_n),\ \forall k \geq 1$

If the claim is true then, $x_* \in X \setminus A_k\, \forall k \in \mathbb{N}$ i,e,

$$x_* \in \bigcap_{k \geq 1} (X \setminus A_k) = X \setminus \bigcup_{k \geq 1} A_k = \emptyset$$

which is a contradiction. To get claim, note that for all $b \geq 0$:

$$d(x_*, x_n) \leq \underbrace{d(x_*, x_{n+1})}_{\to 0\ n \to \infty} + \underbrace{d(x_{n+1}, x_n)}_{\leq r_n/2}$$

so by choosing n large enough, the sum is less than r_n ■

The lemma motivates the following.

Definition 8.5 Let (X, d) be a metric space.

i) $A \subset X$ is called **meager** (or of the 1$^{\text{st}}$ Baire Category) if $A = \cup_{k=1}^{\infty} A_k$ with nowhere dense sets A_k; denoted $\text{cat}(A) = 1$. (A_k is nowhere dense if $\text{int}(\overline{A_k}) = \emptyset$)

ii) $A \subset X$ is called **fat** (or of the 2$^{\text{nd}}$ Baire Category) if it is not meager; denoted $\text{cat}(A) = 2$.

In this language, Lemma 8.4 becomes

Theorem 8.6 — Baire category theorem.

Let (X, d) be a complete, non-empty metric space. Then $\text{cat}(X) = 2$.

Note that the "completeness" assumption depends and by conclusion is purely topological; thus, the theorem extends to (X, d) not complete if one can find a metric $\tilde{d}$ such that $(X, \tilde{d})$ is complete and d induces the same topology as d.

Remark 8.7 Every subset of a meager set A is meager (in particular, Theorem 8.6 implies the **existence** of fat sets, examples are $(\mathbb{R}, |\cdot|), (C, \|\cdot\|_\infty)$)

Remark 8.8 1) $\mathbb{Q} = \bigcup_{x\in\mathbb{Q}}\{x\} \subset \mathbb{R}(= X)$ is meager. What about $\mathbb{R} \setminus \mathbb{Q}$? (this is fat. Why?)

2) Under the assumption of Theorem 8.6, if $A \subset X$

$$\text{cat}(A) = 1 \implies \text{cat}(X \setminus A) = 2 \text{ and } X \setminus A \text{ is dense}$$

use (exercise): $U \subset X$ open and dense $\iff$ $A = X \backslash U$ closed and nowhere dense, hence:

$$\text{Lemma 8.4} \iff U_k \subset X, k \geq 1, \text{open, dense. Then } U = \cap_{k=1}^{\infty} U_k \text{ dense}$$

3) If $\emptyset \neq U \subset X$ open $\implies$ $\text{cat}(U) = 2$ (under the same assumption in 2))

Proof. If $\text{cat}(U) = 1$, by 2), $X \setminus U$ is dense, *i.e.*

$$X = \overline{X \setminus U} \overset{X\setminus U \text{ closed}}{=} X \setminus U$$

in other words, $U = \emptyset$, which is also closed, a contradiction. ∎

4) (Topological vs. measure-theoretic size) $X = \mathbb{R}$ with λ the Lebesgue measure. Does $\lambda(A) = 0 \implies A$ meager? Does $A \subset \mathbb{R}$ meager $\implies \lambda(A) = 0$?

The answer to both questions are NO.

Take $\mathbb{Q} = \{q_1, q_2, \ldots\}$, for $j \geq 1$:

$$U_j = \cup_{k\geq 1}(q_k - \frac{1}{2^{j+k+1}}, q_k + \frac{1}{2^{j+k+1}})$$

which is decreasing in j

$$\lambda(U_j) \leq \sum_{k\geq 1} 2 \cdot \frac{1}{2^{j+k+1}} = 2^{-j}$$

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. By 2) (see exercise), $A_j \overset{\text{def.}}{=} X \setminus U_j$ is nowhere dense, $A \overset{\text{def.}}{=} \cup_{j=1}^{\infty} A_j$ is meager and hence $U = X \setminus A = \cap_{j=1}^{\infty} U_j$ is fat. But

$$\lambda(U) = \lim_{j\to\infty} \lambda(U_j) = 0 \qquad \lambda(A) = \infty$$

8.2 Uniform boundedness principle (UBP)

Baire's category theorem leads to UBP. The first instance of this is not a 'linear' property at all.

Theorem 8.9 Let (X, d) be a complete metric space and $(f_\lambda)_{\lambda\in\Lambda}$ be a family of continuous functions $f_\lambda : X \to \mathbb{R}$. If $(f_\lambda)_{\lambda\in\Lambda}$ is bounded pointwise, *i.e.*

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

then $\exists B \subset X$ an open ball s.t.

$$\sup_{\lambda\in\Lambda, x\in B} |f_\lambda(x)| < \infty$$

i.e. $(f_\lambda)_{\lambda\in\Lambda}$ uniformly bounded on B.

Remark 8.10 The $(f_\lambda)_{\lambda\in\Lambda}$ need NOT to be linear.

Proof. For $k \geq 1$, consider the closed set

$$A_k = \left\{ x \in X : \forall \lambda \in \Lambda : |f_\lambda(x)| \leq k = \bigcap_{\lambda\in\Lambda} \underbrace{\{|f_\lambda| \leq k\}}_{\text{closed as } f_\lambda \text{ continuous}} \right\}$$

Clearly $\cup_{k=1}^{\infty} A_k = X$. Since X is complete, by Lemma 8.4, $\exists k_0 \in \mathbb{N}$, s.t. $\text{int}(A_{k_0}) \neq \emptyset$. Pick $B \subset A_{k_0}$. ∎

Incorporating the linear structure, this leads to the following.

Corollary 8.11 — Banach-Steinhaus.

Let X, Y be normed vector spaces and X is complete.

Let $(A_\lambda)_{\lambda\in\Lambda} \subset \mathcal{L}(X, Y)$. If $(A_\lambda)_{\lambda\in\Lambda}$ are bounded pointwise *i.e.*

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

then $(A_\lambda)_{\lambda\in\Lambda}$ is bounded uniformly, *i.e.*

$$\sup_{\lambda\in\Lambda} \|A_\lambda\|_{\mathcal{L}(X,Y)} < \infty$$

Proof. For $\lambda \in \Lambda$ define the continuous (check this!) map $f_\lambda : X \to \mathbb{R}$ by

$$f_\lambda(x) = \|A_\lambda x\|_Y, \qquad x \in X$$

By assumption on A_λ, Theorem 8.9 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 $\lambda \in \Lambda$:

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

■

An application of Corollary 8.11 is the following.

Proposition 8.12 With X, Y normed and X **complete**, let $A_k \in \mathcal{L}(X,Y)$, (A_k) converges pointwise to $A : X \to Y$, *i.e.*

$$\forall x \in X: \qquad \|A_k x - Ax\|_Y \overset{k\to\infty}{\longrightarrow} 0$$

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

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

Proof. $(A_k x) \subset Y$ is convergent hence bounded, so Corollary 8.11 applies and yields $\sup_k \|A_k\|_{\mathcal{L}(X,Y)} < \infty$. This shows $\liminf_{k\to\infty} \|A_k\|_{\mathcal{L}(X,Y)}$ is finite hence well-defined. Pick subsequence k_j s.t.

$$\|A_{k_j}\| \overset{j\to\infty}{\longrightarrow} \liminf_k \|A_k\|_{\mathcal{L}(X,Y)} \overset{\text{def.}}{=} M$$

$(A_{k_j})_j$ converges pointwise, A is linear (check!) and $\forall x \in X$:

$$\|Ax\|_Y = \lim_j \|A_{k_j} x\|_Y \leq \lim_j \|A_{k_j}\|_{\mathcal{L}(X,Y)} \|x\|_X \leq M \|x\|_X$$

■

Remark 8.13 **Completeness** of X is important (as for Baire)

Take $X = C(= C^0[0,1])$, $\|\cdot\|_X = \|\cdot\|_1$. Let

$$A_k f = k \int_{1-1/k}^{1} f(t)dt \qquad k \geq 1$$

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

$$\forall f \in X : \qquad A_k f \overset{k\to\infty}{\longrightarrow} Af \overset{\text{def.}}{=} f(1)$$

But $A : X \to \mathbb{R}$ is not continuous: take for examples, $f_n(t) = t^n$ then $\|f_n\| = \frac{1}{n+1}$ so $f_n \overset{L^1}{\longrightarrow} 0$ but

$$Af_n = f_n(1) = 1 \nrightarrow 0$$

Of course $(X, \|\cdot\|_1)$ is not complete (why?), so there is no contradiction.

Remark 8.14 If instead we consider $(C, \|\cdot\|_\infty)$, then

$$|A_k f| \leq \|f\|_\infty$$

and $Af = f(1)$ is continuous as it's bounded $\|A\|_{\mathcal{L}(X,Y)} \leq 1$ as $\|Af\|_\infty \leq \|f\|_\infty$

The following section is not examinable.

8.3 Baire's Original Problem

Let $(f_n)_{n\geq 1}$ be a sequence of continuous functions $f_n : [0,1] \to \mathbb{R}$ and converges point-wise, *i.e.*

$$f(x) = \lim_{n\to\infty} f_n(x) \text{ exists } \forall x \in [0,1]$$

Question: Does f have points of continuity?

The answer is given by the following theorem.

Theorem 8.15 — Baire.

Let (X, d) be a complete metric space and (f_n) a sequence of continuous functions $f_n : X \to$

$\mathbb{R}$, and for each $x \in X$, the point-wise limit

$$\lim_{n\to\infty} f_n(x) \overset{\text{def.}}{=} f(x) \in \mathbb{R}$$

exists. Then

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

is **dense** in X.

Proof. For $\varepsilon > 0, n \geq 1$ set

$$P_{n,\varepsilon} = \{x \in \mathbb{R} : |f_n(x) - f(x)| \leq \varepsilon\}$$

and

$$R_\varepsilon = \bigcup_{n=1}^{\infty} \operatorname{int}(P_{n,\varepsilon})$$

We have the following two claims.

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

Proof. This can be checked by writing out explicitly what $x \in \bigcap_{n=1}^{\infty} R_{\frac{1}{n}} = \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \operatorname{int}(P_{m,\frac{1}{n}})$ means, namely $\forall n \in \mathbb{N}, \exists M \in \mathbb{N}, \forall m \geq M, |f_m(x) - f(x)| < \frac{1}{n}$, as the sets $\operatorname{int}(P_{m,\frac{1}{n}})$ are increasing in n. By the usual '3ε' argument, for any $\varepsilon > 0$, we can find $|x - y| < \delta$, so that $|f_n(x) - f_n(y)| < \frac{\varepsilon}{3}$ and we choose n accordingly to let $\frac{1}{n} < \frac{\varepsilon}{3}$ such that

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

which shows the continuity. ∎

Claim 2: R_ε is open and dense for all $\varepsilon > 0$.

Proof. R_ε is open, as it is a countable union of open sets.

For density, consider $F_{n,\varepsilon} = \bigcap_{k\geq 1} \{x : |f_n(x) - f_{n+k}(x)| \leq \varepsilon\}$, which is closed.

Since $f_{n+k}(x) \overset{k\to\infty}{\longrightarrow} f(x)$, we have $F_{n,\varepsilon} \subseteq P_{n,\varepsilon}$. Also, since $f_n(x) \overset{n\to\infty}{\longrightarrow} f(x)$, we have

$$\bigcup_{n=1}^{\infty} P_{n,\varepsilon} = X \qquad \forall \varepsilon > 0$$

Let $A_{n,\varepsilon} = \partial F_{n,\varepsilon}$, then $A_{n,\varepsilon}$ is closed and

$$\operatorname{int}(A_{n,\varepsilon}) = \operatorname{int}(\bar{A}_{n,\varepsilon} \setminus \operatorname{int}(A_{n,\varepsilon})) = \emptyset$$

i.e. $A_{n,\varepsilon}$ is nowhere dense. Hence the set

$$A_\varepsilon = \bigcup_{n\geq 1} A_{n,\varepsilon}$$

is meager, also we have

$$\begin{aligned} A_\varepsilon &= \bigcup_{n\geq 1} A_{n,\varepsilon} \\ &\supseteq \left(\bigcup_{n\geq 1} F_{n,\varepsilon}\right) \setminus \left(\bigcup_{n\geq 1} \operatorname{int}(F_{n,\varepsilon})\right) \\ &\supseteq X \setminus \left(\bigcup_{n\geq 1} \operatorname{int}(P_{n,\varepsilon})\right) \\ &= X \setminus R_\varepsilon \end{aligned}$$

In other words, $R_\varepsilon \supseteq X \setminus A_\varepsilon$, which is dense by (Baire Category Thm. Remark 8.8).

■

By Claim 2, $R_{\frac{1}{n}}$ is open and dense, so by Claim 1 R is dense, as the countable intersection of open and dense sets is dense (Baire Category Thm.). ■

Remark 8.16 As a corollary, the Dirichlet function $\mathbf{1}_\mathbb{Q}$ is nowhere continuous; hence there does not exist any continuous function $f_n : \mathbb{R} \to \mathbb{R}$ with $f_n \to \mathbf{1}_\mathbb{Q}$ point-wise.

9 Open mapping theorem

Definition 9.1 — Open Ball.

An open ball in normed linear space X with radius $r > 0$ centered at $x \in X$ is

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

Also, when $x = 0$ we write

$$B_X(0, r) \equiv B_x(r)$$

Definition 9.2 — Open map.

Let X, Y be linear spaces. $A : X \to Y$ is **open** if $A(U) \subset Y$ is open whenever $U \subset X$ is open.

Remark 9.3

- A being continuous means $A^{-1}(V) \subset X$ open $\forall V \subset Y$ open.
- A being continuous need not be open. e.g. $Ax \overset{\text{def.}}{=} 0 \in Y$

Theorem 9.4 — Open Mapping Theorem.

Let X, Y be Banach, $A \in \mathcal{L}(X, Y)$. Then:

i) if A is surjective, A is open.

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

Remark 9.5

ii) is important in application. If $A \in \mathcal{L}(X, Y)$ is bijective then $A^{-1} : X \to Y$ linear is easy (why?). The point is A^{-1} is also bounded, or equivalently continuous.

The main step of the proof is the following:

Lemma 9.6

Let A be surjective and bounded as in i), then $\exists r > 0$ s.t. $B_Y(r) \subset \overline{A(B_X(1))}$

Proof. Since A is surjective,

$$Y = \bigcup_{k=1}^{\infty} A(B_X(k))$$

Since Y is complete, by Baire Category theorem, $\exists k_0$ s.t.

$$\mathrm{int}(\overline{A(B_X(k_0))}) \neq \emptyset$$

So by surjectivity of A, one can find $y_0 = Ax_0 \in Y$, $r_0 > 0$ s.t.

$$\underbrace{B_Y(y_0, r_0)}_{=Ax_0 + B_Y(r_0)} \subset \overline{A(B_X(k_0))}$$

By linearity of A,

$$\begin{aligned} B_Y(r_0) &\subset \overline{A(B_X(k_0))} - Ax_0 \\ &= \overline{A(B_X(k_0) - x_0)} \\ &\subset \overline{A(B_X(k_0 + M))} \\ &= (k_0 + M)\overline{A(B_X(1))} \end{aligned}$$

Where $M \overset{\text{def.}}{=} \|x_0\|_X$. So pick $r = \frac{r_0}{k_0 + M}$. ■

Proof of Theorem 9.4:

Proof. i) Pick r as in Lemma 9.6.

$\boxed{\textbf{Claim:} \quad B_Y(r/2) \subset A(B_X(1)))}$

If claim holds, then for $U \subset X$ open, pick $x_0 \in U$, $s > 0$ small so that $B_X(x_0, s) \subset U$. Letting $y_0 \overset{\text{def.}}{=} {=}Ax_0$, get

$$B_Y(y_0, \frac{rs}{2}) = y_0 + sB_Y(\frac{r}{2}) \overset{claim}{\subset} Ax_0 + sA(B_X(1)) \overset{linearity}{=} A(B_X(x_0, s)) \subset A(U)$$

which proves i). To see i) $\implies$ ii), it's enough to show that $B = A^{-1} : Y \to X$ is continuous; but for any $U \subset X$ open, $B^{-1}(U) = (A^{-1})^{-1}(U) = A(U)$ which is open by i). ■

Proof of claim:

Proof. Fix $y \in B_Y(r/2)$. Need to show: $y = Ax$ for some $x \in X$ with $\|x\|_X < 1$.

We construct a sequence $(x_k) \subset X$ with

$$\sum_{k=1}^{\infty} \|x_k\|_X < 1 \text{ and } \sum_{k=1}^{\infty} Ax_k \xrightarrow{\|\cdot\|_Y} y,\ n \to \infty$$

By completeness of X, $\sum_{k=1}^{\infty} x_k \overset{\text{def.}}{=} x$ exists, $x \in B_X(1)$ and by continuity of A,

$$Ax = \sum_{k=1}^{\infty} Ax_k = y$$

By Lemma 9.6 above with $\tilde{r} = \frac{r}{2}$,

$$\forall s > 0, \qquad B_Y(s\tilde{r}) \subset \overline{A(B_X(s/2)} \tag{$*$}$$

Let $s = 1$. Pick $x_1 \in B_X(1/2)$, such that $\|Ax_1 - y\| < \tilde{r}/2$.

Now set $y_1 = y - Ax_1$, where ($y_1 \in B_Y(\tilde{r}/2)$). Iterate.

Assume that for some $k \geq 1$ have $x_1, ...x_k, y_1, ...y_k$ s.t.

$$\forall 1 \leq \tilde{k} \leq k : \qquad \|x_{\tilde{k}}\|_X < 2^{-k} \qquad y_{\tilde{k}} = y_{\tilde{k}-1} - Ax_{\tilde{k}} \in B_Y(2^{-\tilde{k}}\tilde{r})$$

Then using ($*$) with $s = 2^{-(k+1)}$ find $x_{k+1} \in B_X(2^{-(k+1)})$ such that

$$y_{k+1} \overset{\text{def.}}{=} y_k - Ax_{k+1} \in B_Y(2^{-(k+1)}\tilde{r})$$

This yields $\sum_{k=1}^{\infty} \|x_k\|_X < 1$ and

$$y - \sum_{k=1}^{n} Ax_k = y_1 - \sum_{k=2}^{n} Ax_k = ... = y_n \to 0 \, (n \to \infty)$$

■

1) Equivalence of norm

Example 9.7 — Equivalence of Norm.

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

$$\|x\|_2 \leq C \|x\|_1, \forall x \in X \tag{9.1}$$

If X is complete with respect to both $\|\cdot\|_1$ and $\|\cdot\|_2$, then $A = id : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ is open by Theorem 9.4 (indeed thm applies, as A is bounded by Equation (9.1)).

Since A is bijective, ii) gives that $A^{-1} = id : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$ is bounded, i.e.

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

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

2) Y needs to be complete in Theorem 9.4

Example 9.8 — Completeness of Y.

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

$$\|Af\|_2 = \|f\|_2 = \int_0^1 |f(t)|dt \leq \|f\|_\infty = \|f\|_1$$

but not open. Else by 1), $\|\cdot\|_1$ and $\|\cdot\|_2$ would be equivalent. However consider counter example:

$$f_n(x) = \begin{cases} 2n^2 x & x \in [0, \frac{1}{2n}] \\ -2n^2 x + 2n & x \in (\frac{1}{2n}, \frac{1}{n}] \\ 0 & x \in (\frac{1}{n}, 1] \end{cases} \quad \text{satisfy} \quad \|f_n\|_2 = 1, \|f_n\|_1 = n \to \infty$$

This shows Y needs to be complete in theorem.

3) X needs to be complete in Theorem 9.4

Example 9.9 — Completeness of X.

This example shows completeness of X is also required. Take

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

with norm $\|\cdot\|_X = \|\cdot\|_Y = \|\cdot\|_\infty$. This is a linear normed space. It's not complete (Exercise: show directly $\overline{X} = c_0$). Another way: Define $A : X \to X$,

$$Ax = (x_1, \frac{x_2}{2}, \frac{x_3}{3}, \underbrace{......}_{0\,\text{eventually}}) \quad \text{if } x = (x_1, x_2......)$$

Then A is linear, bijective with

$$A^{-1} : X \to X, \ \ A^{-1}x = (x_1, 2x_2, 3x_3 \underbrace{......}_{0\,\text{eventually}})$$

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 $x^{(n)} = (\overbrace{1, 1, \ldots, 1, 1}^{n}, 0, ...)$ then $\left\|x^{(n)}\right\|_\infty = 1$ but $\left\|A^{-1}x^{(n)}\right\| = n$. Hence $A^{-1} \notin \mathcal{L}(X)$ and X cannot be complete, else by Theorem 9.4 ii),

A^{-1} would be bounded.

10 Closed Graph Theorem

Consider X, Y normed spaces. Often an operator A not defined on all of A but on a "domain" $D(A)$. So we assume that $D(A) \subset X$ is a linear subspace on which $A : D(A)(\subset X) \to Y$, linear is defined.

Example 10.1 — running example.

$Y = X = C = C^0[0,1]$ with $\|\cdot\|_X = \|\cdot\|_\infty$ and

$$A = \frac{d}{dt}$$

with $D(A) \overset{eg}{=} C^1[0,1] \subset X$ or subspaces thereof.

Prime example of (unbounded) operator with dense domain $D(A)$: indeed $\overline{C^1[0,1]} = C$ using e.g. Weierstrass Approximation Theorem (Polynomials are already $\|\cdot\|_\infty$-dense in C)

Definition 10.2 — Graph of an operator.

Let X, Y be normed space, $A : D(A)(\subset X) \to Y$. 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 endowed $X \times Y$ with the norm $\|(x,y)\|_{X\times Y} = \|x\|_X + \|y\|_Y$, for all $x \in X$, $y \in Y$.

Definition 10.3 — Closed Operator. A is called **closed** if Γ_A is closed in $(X \times Y, \|\cdot\|_{X\times Y})$

Example 10.4 Let $A \in \mathcal{L}(X,Y)$ with $D(A) = X$. Then A is closed.

Proof. Let $(x_k, y_k)_k \subset \Gamma_A$ with $\|(x_k, y_k) - (x,y)\|_{X\times Y} \xrightarrow{k\to\infty} 0$ for some $(x,y) \in X \times Y$

Need to show: $(x,y) \in \Gamma_A$ i.e. $y = Ax$.

We know $y_k = Ax_k$ and $\|x_k - x\|_X \xrightarrow{k\to\infty} 0$, $\|Ax_k - y\|_Y \xrightarrow{k\to\infty} 0$ But $\forall k \geq 1$,

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

Thus

$$\lim_{k\to\infty} \|y - Ax\|_Y \leq \lim_{k\to\infty} \left(\|y - Ax\|_Y + \|A\| \, \|x_k - x\|_X\right) = 0$$

■

Theorem 10.5 — Closed Graph.

Let X, Y be Banach $A : X \to Y$ linear. The following are equivalent:

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

ii) A is closed

Proof. i) $\implies$ ii): see example

ii) $\implies$ i): If X, Y complete, then so is $(X \times Y, \|\cdot\|_{X \times Y})$ (exercise). A closed means Γ_A is closed in $(X \times Y, \|\cdot\|_{X \times Y})$, so $(\Gamma_A, \|\cdot\|_{X \times Y})$ is complete. Consider:

$$\begin{aligned} \Pi_X :\ \Gamma_A &\to X & \qquad \Pi_Y : \Gamma_A &\to Y \\ (x, Ax) &\mapsto x & (x, Ax) &\mapsto Ax \end{aligned} \tag{10.1}$$

Π_X, Π_Y are continuous with $\|\Pi_X\|, \|\Pi_Y\| \leq 1$, Π_X is injective, and surjective. By the Open Mapping Theorem 9.4 ii), ${\Pi_X}^{-1} \in \mathcal{L}(X, \Gamma_A)$ and so

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

■

Remark 10.6 ii) is simpler than i), but equivalent.

i) says A is continuous, i.e. if $(x_n) \subset X$, $x \in X$

$$\|x_n - x\|_X \overset{n\to\infty}{\longrightarrow} 0 \implies \|Ax_n - Ax\|_Y \overset{n\to\infty}{\longrightarrow} 0$$

This contains two things to check: (Ax_n) converges and limit is Ax.

ii) says A is closed, i.e.

$$\begin{cases} \|x_n - x\|_X \overset{n\to\infty}{\longrightarrow} 0 \\ \|Ax_n - y\|_Y \overset{n\to\infty}{\longrightarrow} 0 \end{cases} \implies Ax = y \tag{10.2}$$

Which is only one condition to check.

Example 10.7 — running example continues.

$(D(A), \|\cdot\|_\infty)$ with $D(A) = C^1[0, 1]$ is NOT Banach, and $A : D(A) \to C$ is an example of an operator which is:

Claim:

i) closed, but

ii) not continuous

For ii), 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

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

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

$$\forall t \in (0,1], \quad \underbrace{f_n(t)}_{\xrightarrow{n\to\infty} f(t)} = \underbrace{\int_0^t f_n'(x)dx}_{\xrightarrow{DCT} \int_0^t g(x)dx} + f_n(0)$$

so $f' = g$ by fundamental theorem of calculus(FTC), i.e. $(f, g) = (f, f') \in \Gamma_A$.

The second convergence uses dominated convergence theorem

Corollary 10.8 — Continuous Inverse.

X, Y Banach, $A : D(A) \subset X \to Y$ linear, closed and bijective. Then $\exists B = A^{-1} \in \mathcal{L}(Y, X)$ with $AB = id_Y$ and $BA = id_{D(A)}$.

Proof. exercise(Hint: similar to the Closed Graph Theorem 10.5, consider $\Pi_Y : \Gamma_A \to Y$, $B \stackrel{\text{def.}}{=} \Pi_X \circ \Pi_Y^{-1}$)) ■

Example 10.9 — running example continues..

A is surjective: for $g \in C$ define $f(t) = \int_0^t g(s)ds$. Then by FTC, $Af = g$.

A is not injective: $Af = A\tilde{f} \implies f = \tilde{f} + c, c \in \mathbb{R}$. Let $D(A) \stackrel{\text{def.}}{=} C_0^1[0,1] = \{f \in C^1[0,1] : f(0) = 0\}$ Then $A : D(A) \to C$ is bijective and has continuous inverse $B = A^{-1}$ by Corollary 10.8. In fact, $Bf(t) = \int_0^t f(s)ds$ with $Bf \in D(A)$.

11 Weak vs. Strong topologies

Let $(X, ||\cdot||_X)$ be a normed linear space with dual space X^* (over $\mathbb{R}$).

Definition 11.1 — Weak Convergence.

A sequence $(x_n) \subset X$ **converges weakly** to $x \in X$, written as $x_n \xrightarrow{\text{w}} x (n \to \infty)$, if $\forall l \in X^*$,

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

(x_n) **converges (strongly/in norm)** to x if $\lim_{n\to\infty} ||x_n - x|| = 0$, write as: $x_n \to x (n \to \infty)$.

Remark 11.2 We have the following remarks about weak convergence.

1) $x_n \to x$ implies $x_n \xrightarrow{w} x$: $|\ell(x_n) - \ell(x)| \leq ||\ell||_* ||x_n - x||$

2) The converse of 1) is false.

   Let $x_n (= e_n) = (0, ..., 0, 1, 0, ...) \in \ell^2$. $||x_n - x_m|| = \sqrt{2}, n \neq m$, so x_n doesn't converge (strongly). But $x_n \xrightarrow{w} 0$: by Riesz Representation, $\ell(\cdot) = \langle y, \cdot \rangle_{\ell^2}$ for some $y \in \ell^2$, Hence with $y = (y^{(n)})_n$(This means $y = (y^{(1)}, y^{(2)}, y^{(3)}, \ldots)$),

$$\ell(x_n) = \langle y, x_n \rangle = y^{(n)} \leq \sqrt{\Sigma_{k\geq n} |y^k|^2} \xrightarrow{n\to\infty} 0$$

   , since $||y||_2 < \infty$.

3) If $x_n \xrightarrow{w} x, x_n \xrightarrow{w} y$, then $x = y$.

   Assume not, by Proposition 7.14, $\exists \ell \in X^* : \ell(x) \neq \ell(y)$ if $x \neq y$. With this ℓ:

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

   ,a contradiction.

4) $x_n \xrightarrow{w} x \Rightarrow \sup ||x_n|| < \infty$.

   Consider $A_n \in \mathcal{L}(X^*, \mathbb{R})$ $(= X^{**}$, the bidual) with $A_n(\ell) := \ell(x_n), \ell \in X^*$. Now $x_n \xrightarrow{w} x$ implies $\sup_n |A_n(\ell)| < \infty$, $\forall \ell \in X^*$, and X^* is complete so by Banach-Steinhaus:

$$\sup_n ||A_n||_{\mathcal{L}(X^*,\mathbb{R})} < \infty$$

But by Corollary 7.19, $||A_n||_{\mathcal{L}(X^*,\mathbb{R})} = \sup_{l\in X^*,||\ell||\le 1} |\ell(x_n)| = ||x_n||_X$.

This naturally leads to:

Definition 11.3 — Bidual. $X^{**} := (X^*)^* \quad (= \mathcal{L}(X^*, \mathbb{R}))$ is called **bidual** of X. X embeds canonically into X^{**} via:

$$\iota : X \to X^{**} : \iota(x)(\ell) := \ell(x) \quad \forall x \in X, \ell \in X^*$$

Remark 11.4 ι is a linear isometry: similarly as in Remark 11.2 4) above. One has:

$$\forall x \in X : ||x||_X = \sup_{l\in X^*,||\ell||\le 1} |\ell(x)| = ||\iota(x)||_{**}$$

Definition 11.5 — Reflexive.

The space X is reflexive if ι in Definition 11.3 is surjective.

Example 11.6 Some examples of reflexive spaces.

1. if $\dim X < \infty$, X is reflexive;
2. H a Hilbert space is reflexive;
3. $L^p, 1 < p < \infty$ is reflexive;
4. L^1, L^∞ are in general not reflexive.

Proposition 11.7 $L^1[-1,1]$ and $L^\infty[-1,1]$ are not reflexive.

Proof. Consider $L^\infty[-1,1]$.

Define the Dirac function

$$\delta_{x_0} : C^0[-1,1] \to \mathbb{R} : f \mapsto \delta_{x_0}(f) = f(x_0)$$

δ_{x_0} is linear and continuous on $(C^0[-1,1], ||\cdot||_\infty)$. By Corollary 7.11, there exists extension

$$\ell \in (L^\infty[-1,1])^* \qquad \text{and} \qquad \ell\mid_{C^0[0,1]} = \delta_{x_0}.$$

with the same norm (Omit "$[-1,1]$" henceforth)

For $g \in L^1$, we define $\forall f \in L^\infty$

$$\ell_g(f) = \int_{-1}^{1} gf\,dx$$

then $\ell_g \in (L^\infty)^*$.

Claim: $\nexists g \in L^1 : \ell = \ell_g$.

The claim implies that $\iota : L^1 \to (L^1)^{**} = (L^\infty)^*, g \mapsto \ell_g$ is not surjective, i.e., L^1 is not reflexive.

For L^∞, use: X reflexive $\Rightarrow$ X^* reflexive (exercise).

Proof of Claim: Suppose $\exists g \in L^1$ s.t. $\ell = \ell_g$. For simplicity let $x_0 = 0$.

Pick a bump function $\phi \in C^\infty[-1,1] : 0 \leq \phi \leq 1$

$$\phi(x) = 1,\ x \in [-\frac{1}{2}, \frac{1}{2}] \qquad \text{and} \qquad \phi(x) = 0,\ x = \pm 1$$

For $n \leq 1 : \phi_n(x) := \phi(nx)$. Then $0 \leq \phi_n \leq 1$, $\phi_n \overset{n\to\infty}{\longrightarrow} 0$ *a.e.*.

This yields:

$$1 = \phi_n(0) = \delta_0(\phi_n) = \ell(\phi_n) = \ell_g(\phi_n) = \int_{-1}^{1} g\phi_n dx \overset{n\to\infty}{\longrightarrow} 0$$

this is a contradiction. ■

An example of a bump function is given below:

$$f_n(x) = \begin{cases} 0 & |x| < \frac{1}{2n} \\ \exp(1 - \frac{1}{1-(2n|x|-1)^2}) & \frac{1}{2n} < |x| < \frac{1}{n} \\ 1 & \frac{1}{n} < |x| \leq 1 \end{cases} \tag{11.1}$$

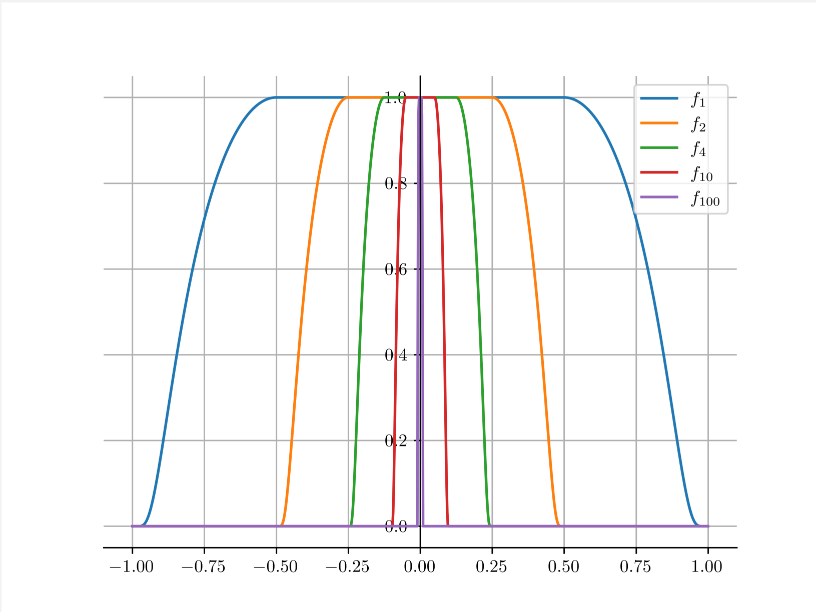

Figure 1: Example of bump functions

Recall that we showed unit balls in ∞-dimensions are never (sequentially) compact (cf. Theorem 4.15). Weak convergence allows us to restore a weak version of this .

For reflexive spaces that is the whole story. Since X may not be reflexive, one must consider an even weaker topology.

In the following, we let $(X, \|\cdot\|)$ be a normed linear space, X^* its dual space, X^{**} its bidual, and the isometry $\iota : X \to X^{**}$.

Definition 11.8 A sequence of linear functionals $(\ell_n) \subset X^*$ is **weak*-convergent** to $\ell \in X^*$ if

$$\lim_{n\to\infty} \ell_n(x) = \ell(x) \qquad \forall x \in X$$

(i.e. pointwise convergence in X) Notation: $\ell_n \xrightarrow{w^*} \ell$

Remark 11.9 1) We now have 3 notions of convergence on X^*:

i) norm/strong convergence: $\|\ell_n - \ell\|_* \xrightarrow{n\to\infty} 0$ (i.e. $\ell_n \to \ell$)

ii) weak convergence: $\ell_n \xrightarrow{w} \ell$, i.e.

$$\forall \xi \in X^{**} : \qquad \lim_{n\to\infty} \xi(\ell_n) = \xi(\ell) \tag{$**$}$$

iii) weak*-convergence: $\ell_n \xrightarrow{w^*} \ell$: equivalent to asking $(**)$ for $\xi \in \iota(X)$ only.

2) If X is reflexive $ii) \iff iii)$ [e.g. Hilbert space]

3) In general, $i) \implies ii) \implies iii)$

Theorem 11.10 — Banach-Alaoglu.

Let X be separable. If $(\ell_n) \subset X^*$ is bounded (in X^*) there exists $\ell \in X^*$ and a subsequence $\Lambda \subset \mathbb{N}$ s.t.

$$\ell_n \xrightarrow{w^*} \ell \qquad n \to \infty, n \in \Lambda$$

Proof. Let $(x_j) \subset X$ be a countable, dense subset. 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 $j \in \mathbb{N}$:

$$\ell_n(x_j) \to a_j \in \mathbb{R} \qquad (n \to \infty, n \in \Lambda_j)$$

$\Lambda \overset{\text{def.}}{=}$ diagonal sequence of $(\Lambda_j)_j$, so $\forall j, \ell_n(x_j) \to a_j, (n \to \infty, n \in \Lambda)$.

Define $\ell(x_j) \stackrel{\text{def.}}{=} a_j$, extend it linearly on $M = span\{x_j : j \in \mathbb{N}\}$ and for all $x \in M$:

$$|\ell(x)| = \lim_{k\to\infty, k\in\Lambda} |\ell_k(x)| \leq \sup_k \|\ell_k\|_* \|x\|_X$$

so $\ell \in M^*$, hence it can be extended to $\ell \in X^*$ by Corollary 7.11.

We now show: $\ell_n \stackrel{w^*}{\to} \ell$ $(n \to \infty, n \in \Lambda)$.

Let $x \in X$, pick $J \subset \mathbb{N}$ s.t. $x_j \to x$ $(j \to \infty, j \in J)$. For such j and $n \geq 1$:

$$\begin{aligned}|\ell_n(x) - \ell(x)| &\leq |\ell_n(x - x_j)| + |\ell(x - x_j)| + |\ell_n(x_j) - \ell(x_j)| \\ &\leq (\sup_n \|\ell_n\|_* + \|\ell\|_*) \|x - x_j\|_X + |\ell_n(x_j) - \ell(x_j)|\end{aligned}$$

Letting first $n \to \infty$ yields

$$\lim_{n\to\infty, n\in\Lambda} |\ell_n(x) - \ell(x)| \leq C \|x - x_j\|_X, \qquad j \in J$$

Now letting $j \to \infty, j \in J$ yields the desired result. ■

Remark 11.11 If X is reflexive, separability can be removed.

Together with Theorem 11.10 and Remark 11.11, this gives immediately:

Corollary 11.12 For H Hilbert. If $(x_n) \subset H$ is bounded ($\sup_n \|x_n\|_H < \infty$), then (x_n) has a weakly convergent subsequence.

Unless $\dim H < \infty$, one **cannot** replace weak by strong in Corollary 11.12.

Example 11.13 i) $X = L^1[0,1]$ is separable, $X^* \cong L^\infty$. If $(f_n) \subset L^\infty$ is bounded, i.e. $\sup \|f_n\|_\infty < \infty$, Theorem 11.10 yields a subsequence $(n_k)_k \subset \mathbb{N}$ and $f \in L^\infty$ s.t.

$$\lim_{k\to\infty} \int f_{n_k} g dx = \int f g dx, \qquad \forall g \in L^1$$

ii) $X = L^\infty (= L^\infty[0,1])$ is not separable (and also not reflexive). The following example shows that the conclusions of Theorem 11.10 fail in this case.

For $0 < \varepsilon \leq 1$ consider,

$$T_\varepsilon : L^\infty \to \mathbb{R} \qquad T_\varepsilon f = \frac{1}{\varepsilon} \int_0^\varepsilon f dx, \qquad f \in L^\infty$$

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

Claim: $\{T_\varepsilon : 0 < \varepsilon \leq 1\}$ is not weak*-sequentially compact.

Proof. Suppose it is, i.e. $\varepsilon \overset{k\to\infty}{\longrightarrow} 0$ and $T \in (L^\infty)^*$ s.t. $T_{\varepsilon_k} \overset{w^*}{\longrightarrow} T$ as $k \to \infty$. By passing to a subsequence, one can assume

$$1 > \frac{\varepsilon_{k+1}}{\varepsilon_k} \to 0 \qquad \text{as } k \to \infty$$

Pick $f \overset{\text{def.}}{=} \sum_{k=1}^\infty (-1)^k \mathbf{1}_{(\varepsilon_{k+1},\varepsilon_k]} \in L^\infty$ with $\|f_n\|_\infty = 1$.

For $k \geq 1$, we have:

$$T_{\varepsilon_k} f = \frac{1}{\varepsilon_k} \sum_{j=k}^\infty (-1)^j (\varepsilon_j - \varepsilon_{j+1}) = (-1)^k \frac{\varepsilon_k - \varepsilon_{k+1}}{\varepsilon_{k+1}} + \frac{1}{\varepsilon_k} \int_0^{\varepsilon_{k+1}} f dx$$

Hence

$$|T_{\varepsilon_k} f - (-1)^k| \leq \frac{1}{\varepsilon_k} \left(\varepsilon_{k+1} + \int_0^{\varepsilon_{k+1}} |f| dx \right) \leq \frac{2\varepsilon_{k+1}}{\varepsilon_k} \overset{k\to\infty}{\longrightarrow} 0$$

so $(T_{\varepsilon_k} f)_k$ accumulates at ± 1 and is thus divergent. ∎

iii) If instead consider $X = C^0[0,1] \subset L^\infty$, a separable closed subspace, then Theorem 11.10 applies to $T_\varepsilon|_X$. Indeed, one immediately sees that

$$T_\varepsilon f \overset{\varepsilon\downarrow 0}{\longrightarrow} f(0), \qquad \textit{i.e. } T_\varepsilon \overset{w^*}{\to} \delta_0 \ (\varepsilon \downarrow 0)$$

where δ_0 is the Dirac delta functional at 0 defined in Proposition 11.7.

Proposition 11.14 X reflexive $\implies$ X^* reflexive

Proof. (Sketch) If X is reflexive, then the isometric embedding $\iota : X \to X^{**}$ is bijective and has a bijectve inverse ι^{-1}.

Consider the dual operator of ι given by $\iota^* : (X^{**})^* \to X^*$. It is also bijective (which follows from the properties of dual operators) and has an inverse $(\iota^*)^{-1}$ which gives the isometric bijection between X^* and its bidual $(X^*)^{**}$. ∎

Definition 11.15 — Weak topology. Let X be a Banach space. We define the weak topology on X as the topology with the following base of neighborhoods: for $x \in X$, $\varepsilon_1, \ldots, \varepsilon_n > 0$, and $f_1, \ldots, f_n \in X^*$:

$$U^{f_1,\ldots,f_n}_{\varepsilon_1,\ldots,\varepsilon_n} := \{y \in X : \forall 1 \leq i \leq n \qquad |f_i(y) - f_i(x)| < \varepsilon_i\}$$

12 Compact operators

Compact operators form a very important class of bounded operators. Roughly: they are the closest thing to a matrix in infinite dimension spaces. (cf. Section 13).

Definition 12.1 — Compact operator.

Let X, Y be normed spaces. $T : X \to Y$ linear. T is compact if for all $B \subset X$ bounded (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 12.2 Let X, Y be Banach spaces. The following are equivalent:

i) T is compact

ii) $\overline{T(B_X(0,1))} \subset Y$ is compact

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

Remark 12.3

The above are true if X, Y are normed and one replaces "Cauchy" by "convergent".

Proof. iii) $\implies$ i). Let $B \subset X$ be bounded. Consider $(y_n) \subset T(B)$ and by iii), (y_n) has a Cauchy subsequence. Hence $\overline{T(B)}$ is compact. Rest is exercise. ■

Example 12.4 1) $T = id : X \to X$ is compact $\iff$ $\dim X < \infty$.

For $\dim X = \infty$, recall the closed unit ball $B = B_X(0,1)$ is not compact.

2) T has **finite rank** if $\dim(im(T)) < \infty$. If $T \in \mathcal{L}(X,Y)$ has finite rank, then T is compact:

Using Lemma 12.2 iii): let $(x_n) \subset X$ be bounded. Then $\|Tx_n\| \le \|T\| \, \|x_n\| \le C$ so $(Tx_n) \subset im(T)$ is bounded. Since $im(T)$ is finite dimensional, one can choose a convergent subsequence.

3) If $\dim X < \infty$, T is compact. (apply 2)

4) (Diagonal Operator) $1 \le p \le \infty$, $\lambda = (\lambda_n)_{n\in\mathbb{N}}, \lambda_n \in \mathbb{R}$ and $\sup_n |\lambda_n| < \infty$. Then

$$T_\lambda : \ell^p \to \ell^p \qquad T_\lambda x \stackrel{\text{def.}}{=} (\lambda_n x_n)_{n\in\mathbb{N}} \qquad \text{for } x = (x_n)_{n\in\mathbb{N}}$$

is well-defined. If T_λ is compact then $\lim_{n\to\infty} \lambda_n = 0$.

For, if $\Lambda \subset \mathbb{N}$ is such that $|\lambda_n| \ge \delta$, $n \in \Lambda$, 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:

$$\forall n \neq m, n, m \in \Lambda : \|T_\lambda e_n - T_\lambda e_m\|_p \geq \delta 2^{1/p}$$

We return to this example after the following.

Theorem 12.5 — Limit of compact operators.

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

$$\|T_n - T\|_{\mathcal{L}(X,Y)} \to 0 \qquad n \to \infty \tag{12.1}$$

Then T is compact.

Remark 12.6 This means the space of compact operators $(\{T \in \mathcal{L}(X,Y) : T \text{ compact}\}, \|\cdot\|_{\mathcal{L}(X,Y)}) \subset (\mathcal{L}(X,Y), \|\cdot\|_{\mathcal{L}(X,Y)})$ is closed hence a Banach space.

Proof. We will use Lemma 12.2 iii) and the diagonal argument.

Let $(x_n) \subset X, \sup_n \|x_n\|_X \leq C$ be bounded.

Goal: Show (Tx_n) has a Cauchy subsequence.

Since T_1 is compact, there exists subsequence $\Lambda_1 \subset \mathbb{N}$ s.t.

$$(T_1 x_n) \subset Y \text{ converges w.r.t. } \|\cdot\|_Y \text{ as } n \to \infty, n \in \Lambda_1$$

By induction, one can find subsequence $\Lambda_1 \supset \Lambda_2 \supset \cdots$ s.t.

$$\forall k \in \mathbb{N} : \qquad (T_k x_n) \subset Y \text{ converges as } n \to \infty, n \in \Lambda_k$$

Let Λ be the diagonal subsequence of $\Lambda_1, \Lambda_2, \ldots$, then $\Lambda \subset \Lambda_k, \forall k$ so

$$\forall k \in \mathbb{N} : \qquad (T_k x_n) \subset Y \text{ converges as } n \to \infty, n \in \Lambda \tag{12.2}$$

Claim: $(Tx_n)_{n \in \Lambda}$ is Cauchy (in fact converges)

For $n, m \in \Lambda$ and $k \in \mathbb{N}$ write

$$\begin{aligned}\|Tx_n - Tx_m\|_Y &\leq \|(T - T_k)x_n\|_Y + \|T_k(x_n - x_m)\|_Y + \|(T - T_k)x_m\| \\ &\leq \|T - T_k\|_{\mathcal{L}(X,Y)} \cdot 2C + \|(T - T_k)x_n\|_Y\end{aligned}$$

Let $\varepsilon > 0$. First pick k s.t. $\|T - T_k\| < \frac{\varepsilon}{4C}$ (use Equation (12.1)).

Then use Equation (12.2) to obtain $\forall n, m \in \Lambda$, with $\min(m, n) \geq N_0(\varepsilon)$:

$$\|T_k x_n - T_k x_m\|_Y < \frac{\varepsilon}{2}$$

■

Example 12.7 Back to Example 12.4 4):

$$T_\lambda \text{ compact} \iff \lim_{n\to\infty} \lambda_n = 0$$

Proof. " $\implies$ ": see Example 12.4.

" $\impliedby$ ": use Theorem 12.5. Define

$$T_n : \ell^p \to \ell^p \qquad x \mapsto T_n x = (\lambda_0 x_0, \ldots, \lambda_n x_n, 0, 0, \ldots)$$

Then $\dim(im(T)) \leq n$, since $im(T) \subset \{x \in \ell^p : x_i = 0, \forall i > n\}$, so T_n has finite rank, hence T_n is compact (cf. Lemma 12.2). Moreover, for $x \in \ell^p$,

$$\|(T - T_n)x\|_p = (\sum_{m>n} |\lambda_m x_m|^p)^{\frac{1}{p}} \leq \sup_{m\geq n} |\lambda_m| \, \|x\|_p$$

so $\|T - T_n\| \to 0$. By Theorem 12.5 T is compact. ■

Example 12.8 — Hilbert-Schmidt integral operator.

Let $X = L^2[0, 1]$ and $a \in C^0[0, 1]^2$. Define operator $A : X \to X$ by,

$$Af(x) = \int_0^1 a(x, y)f(y)dy, \qquad f \in L^2[0, 1]$$

1) A is well-defined and bounded:

$$Af(x) = \int_0^1 |Af(x)|^2 dx \overset{\text{Cauchy Schwartz}}{\leq} \underbrace{\int_0^1 dx \int_0^1 dy |a(x, y)|^2}_{\leq C} \|f\|_2^2$$

2) A is compact:

Let $(f_n) \subset X, \|f_n\|_2 \leq M$. Check (Af_n) is continuous, $\sup_n \|Af_n\|_\infty < \infty$. Moreover, let $\varepsilon > 0$, we can pick $\delta > 0$ s.t. $|a(x, y) - a(x', y')| < \varepsilon$ if $|x - x'| + |y - y'| < \varepsilon$ (uniform-continuity), we have

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

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

13 Spectral Theory

In this section, we consider Banach space over $\mathbb{C}$.

Definition 13.1 — Resolvent & Spectral.

Let X be Banach. $A : D_A \subset X \to X$ linear operator.

The **resolvent set** of A is

$$\varrho(A) = \{\lambda \in \mathbb{C} : (\lambda I - A) \text{ is bijective with } \exists(\lambda I - A)^{-1} \in \mathcal{L}(X)\}$$

The **spectrum** is defined as the complement:

$$\sigma(A) = \mathbb{C} \backslash \varrho(A)$$

The **resolvent** of A is the map $R : \varrho(A) \to \mathcal{L}(X)$, $\varrho(A) \ni \lambda \mapsto R_\lambda = (\lambda - A)^{-1} \in \mathcal{L}(X)$

Remark 13.2

In this section, we define $\lambda - A = \lambda Id - A = \lambda I - A$, where id is the identity map and the third expression uses first-year linear algebra notation.

Example 13.3

Consider $X = \mathbb{C}$, $A \in \mathcal{L}(X)$, i.e. $D_A = X$, $\lambda \in \mathbb{C}$.

$$(\lambda - A) \text{ invertible} \iff p(\lambda) \overset{\text{def.}}{=} \det(\lambda - A) \neq 0.$$

Since $p(\cdot)$ has at least 1 and at most n (distinct) solutions, one get $\sigma(A) \neq \emptyset$, $\sigma(A)$ contains at most n points. Hence $\varrho(A) \neq \emptyset$ and $\varrho(A) \subset \mathbb{C}$ is dense.

Lemma 13.4

If $z_0 \in \varrho(A)$, then

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

Hence $\varrho(A)$ is open, and $\sigma(A)$ is closed.

Proof. Write

$$z - A = (z - z_0) + (z_0 - A) = (1 + (z - z_0)R_{z_0})(z_0 - A) \tag{$*$}$$

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

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

hence also

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

For (1) use: if $A \in \mathcal{L}(X)$, $\|A\| < 1$ then with $A^0 = Id = 1$,

$$\sum_{n=0}^{\infty} A^n \in \mathcal{L}(X)$$

i.e. the sequence

$$\left(S_n = \sum_{k=0}^{n} A^k \right) \subset \mathcal{L}(X)$$

converges, and

$$\sum_{n=0}^{\infty} A^n = (1 - A^{-1})^{-1}$$

■

Example 13.5

cf. Example 12.4 Diagonal operator $T = T_\lambda$ continued. Claim: $\sigma(T) = \overline{\{\lambda_k : k \in \mathbb{N}\}}$.

Proof.

- $\sigma(T) \supset \overline{\{\lambda_k : k \in \mathbb{N}\}}$

  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 13.4 $\overline{\{\lambda_k : k \in \mathbb{N}\}} \subset \sigma(T)$.

- $\sigma(T) \subset \overline{\{\lambda_k : k \in \mathbb{N}\}}$

  If $\widetilde{\lambda} \notin \overline{\{\lambda_k : k \in \mathbb{N}\}}$, then $\exists \delta > 0$ s.t. $|\widetilde{\lambda} - \lambda_k| > \delta$, $\forall k \in \mathbb{N}$.

  Let $x \in \ell^2$,

  $$y \overset{\text{def.}}{=} (\widetilde{\lambda} - T)x = ((\widetilde{\lambda} - \lambda_k)x_k)_k$$

  So $x_k = (\widetilde{\lambda} - \lambda_k)^{-1} y_k$ and $\|x\|_{\ell^2} \leq \delta^{-1} \|y\|_{\ell^2}$ which implies $(\widetilde{\lambda} - T)^{-1} \in \mathcal{L}(H)$ and $\widetilde{\lambda} \in \varrho(T)$

■

Remark 13.6

In finite dimension, $(\lambda - A)$ not invertible $\iff$ $(\lambda - A)$ not injective by rank formula. One may wonder if "lack of injectivity"is the only reason for $\lambda \notin \sigma(A)$.

Definition 13.7

Consider linear operator A with closed graph and spectrum $\sigma(A)$.

- Point spectrum: $\sigma_p(A) \overset{\text{def.}}{=} \{\lambda \in \mathbb{C} : \lambda - A \text{ not injective}\}$
- Continuous spectrum: $\sigma_c(A) \overset{\text{def.}}{=} \{\lambda \in \mathbb{C} \setminus \varrho(A) : (\lambda - A) \text{ injective }, im(\lambda - A) \text{ not dense }\}$
- Residual spectrum: $\sigma_r(A) \overset{\text{def.}}{=} \sigma(A) \setminus (\sigma_p(A) \cup \sigma_c(A))$

Elements of point specturm are called **eigenvalues** of A with **eigenspaces**(null spaces)

$$ker(\lambda - A) = \{x \in D_A : Ax = \lambda x\} \neq \{0\}$$

Example 13.8 — Shift operator. 1) $S : \ell^2 \to \ell^2$, $S(x_1, x_2...) = (0, x_1, x_2...)$. Then $0 \in \sigma(S)$:

Indeed S is not invertible since it's not surjective: $\forall y \in \ell^2$ with $y_1 \neq 0$, $y \notin im(S)$, but $0 \notin \sigma_p(S)$: S is injective.

In fact $Sx = \lambda x \implies 0 = \lambda x_1, x_n = \lambda x_{(n+1)}, \forall n \in \mathbb{N} \implies x_k = 0$ for all $k \in \mathbb{N}$. So $\sigma_p(S) = \emptyset$. (In fact, $\sigma(S) = \overline{D} = \{\xi \in \mathbb{C} : |\xi| \leq 1\}$ the closed unit disk, $\sigma_r(S) = D$, $\sigma_c(S) = \partial D = S^1$, the unit circle.)

2) $X = \mathbb{C}^n$, $\sigma(\cdot) = \sigma_p(\cdot)$

3) T_λ is indicative of a certain class: X Hilbert, $T \in \mathcal{L}(X)$ compact, self-adjoint.

Then by Riesz-Schauder, $\sigma(T) \setminus \{0\} = \sigma_p(T) \setminus \{0\}$

13.1 Spectral Theory in Hilbert space

Consider $(H, \langle \cdot, \cdot \rangle$, Hilbert space over $\mathbb{C}$, $A : D_A \subset H \to H$ linear, with adjoint $A^* : D_{A^*} \subset H \to H$. Recall A^* characterised by

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

and

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

In sequel write $A \subset B$, reads B is extension of A if $D_A \subset D_B$ and $B|_{D_A} = A$.

Definition 13.9

i) A is **symmetric** if $A \subset A^*$, i.e. $D_A \subset D_{A^*}$ and $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in D_A$.

ii) A is **self-adjoint** if $A = A^*$, i.e. A symmetric with $D_{A^*} = D_A$.

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

Lemma 13.10 If A is symmetric, $\sigma_p(A) \subset \mathbb{R}$.

Proof. Let $\lambda \in \sigma_p(A)$ with non-zero eigenvector $x \in ker(\lambda - A)$.

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

So $\lambda = \overline{\lambda} \implies \lambda \in \mathbb{R}$. ■

Is this true for all the spectrum of H (cf. $\mathbb{C}^n$)?

Example 13.11 — Weak derivative.

H=$L^2(0,1)$, $\langle f, g\rangle = \int_0^1 f\bar{g}\mathrm{d}t$. $A \in \frac{\mathrm{d}}{\mathrm{d}t}$.

More precisely, $f \in H$ is said to have a **weak derivative** f' if $f' \overset{\text{def.}}{=} v$ for some $v \in H$,

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

Consider

$$A_\infty = i\frac{\mathrm{d}}{\mathrm{d}t} : C_c^\infty(0,1) \subset H \to H$$

and extensions A_1, A_2, A_3 with

$$\begin{aligned} &D_{A_1} = H^1 \overset{\text{def.}}{=} \{f \in H : f \text{ has a weak derivative f'}\} \\ &\bigcup \\ &D_{A_2} = \{f \in H^1 : f(0) = f(1)\} \text{ (periodic boundary condition)} \\ &\bigcup \\ &D_{A_3} = \{f \in H^1 : f(0) = 0 = f(1)\} \text{ (Dirichlet boundary condition)} \end{aligned} \tag{13.1}$$

Set $A_k(f) = if'$, $\forall f \in D_{A_k}$. Evidently $A_\infty \subsetneq A_3 \subsetneq A_2 \subsetneq A_1$.

One can show that

$$A_3 \subset {A_1}^* \subset {A_2}^* = A_2 \subset {A_3}^*$$

So: A_3 is symmetric but because $A_3 \subsetneq A_2 \subset {A_3}^*$ not self-adjoint, A_2 is self-adjoint.

Claim:

i) $\sigma(A_1) = \sigma_p(A_1) = \mathbb{C}$, $\varrho(A_1) = \emptyset$.

ii) $\sigma(A_2) - \sigma_p(A_2) = a\pi\mathbb{Z}$, $\varrho(A_2) = \mathbb{C}$.

iii) $\sigma(A_3) = \mathbb{C}$, $\sigma_p(A_3) = \emptyset$, $\varrho(A_3) = \emptyset$

So symmetric operators can have imaginary spectrum!

Proof. i) For $\lambda \in \mathbb{C}$, pick $f(t) = e^{-i\lambda t} \in ker(\lambda - A_1)$.

ii) For $k \in \mathbb{Z}$, $f(t) = e^{-2\pi ikt} \in D_{A_2} \cap ker(a\pi k - A_2)$, so $2\pi\mathbb{Z} \subset \sigma_p(A_2) \subset \sigma(A_2)$. Let $\lambda \in \mathbb{C} \setminus a\pi\mathbb{Z}$. Need to show: $\lambda \in \varrho(A_2)$ i.e. $\lambda - A_2 : D_{A_2} \to H$ is invertible and $(\lambda - A_2)^{-1} \in H$. For $g \in H$, the general sd. of $\lambda f - if' = g$ can be obtained via variation of constant formula:

$$(*) \rightsquigarrow f(t) = ae^{-i\lambda t} + i\int_0^t e^{i\lambda} g(s)\mathrm{d}s$$

for some $a \in \mathbb{C}$. But if $\lambda \notin 2\pi\mathbb{Z}$, the b.b. determines a uniquely:

$$a = f(0) = f(1) = ae^{-i\lambda} + i\int_0^1 e^{i\lambda(s-1)} g(s)\mathrm{d}s$$

so

$$a = (1 - e^{-i\lambda})^{-1} i\int_0^1 e^{i\lambda(s-1)} g(s)\mathrm{d}s$$

so $\lambda - A_2$ is invertible and

$$\|f\|_{L^2} \leq |A| + \|g\|_{L^2} \leq (|1 - e^{-i\lambda}|^{-1})\,\|g\|_{L^2}$$

which shows $(\lambda - A_2)^{-1} \in \mathcal{L}(H)$

iii) If $A_3 f = if' = \lambda f$ for some $\lambda \in \mathbb{C}$, then by (*) $f(t) = ae^{-i\lambda t}$ and $a = 0$ since $f(0) = 0$, so $\sigma_p(A_3) = \emptyset$.

On the other hand $(\lambda - A_3)$ for $\lambda \in \mathbb{C}$ is never surjective ($\implies \lambda \in \varrho(A_3)$). Indeed, consider $g(s) = e^{i\lambda s}$, then using $(*)$ and b.c. get $a = 0$ and

$$f(t) = ie^{-i\lambda t}\int_0^1 e^{i\lambda s}\mathrm{d}s = ite^{-i\lambda t}$$

but $f(1) \neq 0$ so $f \notin D_{A_3}$. ∎

We have seen: symmetric operators can have imaginary spectrum, but:

Lemma 13.12 Let $A \subset A^*$ (*i.e.* A is symmetric). Then

$$\forall \xi \in \mathbb{C}\, \forall u \in D_A : \qquad \|(\xi - A)u\|_H \geq |Im(\xi)|\, \|u\|_H$$

(So for $\xi \notin \mathbb{R} \implies (\xi - A)$ injective, i.e. $\xi \notin \sigma_p(A)$)

Proof. For $u \in D_A : \langle u, Au \rangle \overset{A \subseteq A^*}{=} \langle Au, u \rangle = \overline{\langle u, Au \rangle} \in \mathbb{R}$. Hence,

$$|Im(\xi)| \, \|u\|_H^2 = |Im(\langle u, (\xi - A)u \rangle) \, | \leq | \, \langle u, (\xi - A)u \rangle \, | \leq \|u\|_H \, \|(\xi - A)u\|_H$$

■

The example also nicely illustrates:

Proposition 13.13 If $A = A^*$, then $\sigma(A) \subset \mathbb{R}$

Proof. Let $\xi \in \mathbb{C} \backslash \mathbb{R}$. Want to show $\xi \in \varrho(A)$, i.e. $\xi - A : D_A \to H$ is bijective with $(\xi - A)^{-1} \in \mathcal{L}(H)$. We will show:

$$(*) \ \xi - A \text{ is surjective}$$

Once $(*)$ holds, we are done: by previous lemma, $\xi - A$ is injective hence bijective, and surjectivity + same lemma also yields

$$\left\| (\xi - A)^{-1} \right\|_{\mathcal{L}(H)} \leq \frac{1}{|Im(\xi)|}$$

proof of (*): we first show

$$(**) \qquad im(\xi - A)(\subset H) \text{ is closed}$$

Assume $v_k = (\xi - A)u_k \xrightarrow{k \to \infty} v$. By Lemma 13.12,

$$\|u_k - u_l\|_H \leq \frac{1}{|Im(\xi)|} \, \|v_k - v_l\|_H \xrightarrow{k,l \to \infty} 0$$

Hence (u_k) is Cauchy and $u_k \to u$ for some $u \in H$. But $A = A^*$ has a closed graph so $v = (\xi - A)u$, i.e. (**) holds.

every self-adjoint operator is bounded link here.

Back to (*): Due to (**), $M \overset{\text{def.}}{=} Im(\xi - A)$ is closed. Assume $M \neq H$. Pick $v \in M^\perp \setminus \{0\}$. Then

$$\forall u \in D_A : \langle v, (\xi - A)u \rangle = 0 \text{ or } \langle v, Au \rangle = \overline{\xi} \, \langle v, u \rangle$$

Hence, $D_A \ni u \mapsto \langle v, Au \rangle$ is continuous, $v \in D_{A^*} = D_A$ and $Av = A^* v = \overline{\xi} v$ but by Lemma 13.12,

$$|Im(\xi)| \, \|v\|_H \leq \left\| (\overline{\xi} - A)v \right\|_H = 0$$

which yields $v = 0$. Contradiction. ■

13.2 Spectral theorem for compact self-adjoint operators

H: Hilbert space over $\mathbb{C}$, inner product $\langle \cdot, \cdot \rangle$, with $\|x\|_H^2 = \langle x, x \rangle$.

Following is an extension (!) of the familiar result from linear algebra concerning diagonalization of symmetric matrices.

Theorem 13.14 — Riesz-Schauder.

Let $T : H \to H$ be compact and self-adjoint, then:

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

ii) $\sigma_p(T)$ contains at most countably many eigenvalues $\lambda_k \in \mathbb{R} \setminus \{0\}$, which accumulate at most at $\lambda = 0$

iii) One can choose e_k corresponding to λ_k such that $e_k \perp e_l \ \forall k \neq l$ and one has $\forall x \in H$: $Tx = \sum_k \lambda_k e_k \langle x, e_k \rangle$

Example 13.15

Diagonal operator (Example 12.4) $T_\lambda : \ell^2 \to \ell^2$ continued.

- T_λ is compact $\iff \lim_{k\to\infty} \lambda_k = 0$
- T_λ is self-adjoint $\iff \lambda_k \in \mathbb{R}\, , \forall k$

and we know $\sigma_p(T_\lambda) = \overline{\{\lambda_k : k \in \mathbb{N}\}}$.

We start with the following lemma:

Lemma 13.16 — Lemma 1. $T \in \mathcal{L}(H)$, self-adjoint. If $\lambda_1 \neq \lambda_2$, $\lambda_1, \lambda_2 \in \sigma_p(T)$ with eigenvectors e_1, e_2, i.e. $\lambda_1 e_1 = Te_1$ and $\lambda_2 e_2 = Te_2$, then $\langle e_1, e_2 \rangle = 0$

Proof.

$$\begin{aligned} \lambda_1 \langle e_1, e_2 \rangle &= \langle \lambda_1 e_1, e_2 \rangle = \langle Te_1, e_2 \rangle \overset{S.A.}{=} \langle e_1, Te_2 \rangle \\ &= \langle e_1, \lambda_2 e_2 \rangle \overset{\lambda_2 = \overline{\lambda_2}}{=} \lambda_2 \langle e_1, e_2 \rangle \end{aligned} \tag{13.2}$$

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

Henceforth, we always assume $T : H \to H$ is compact and self-adjoint. In particular, Lemma 1(Lemma 13.16) is in force.

Define, for $\lambda \in \sigma_p(T) \setminus \{0\}$,

$$X_\lambda = ker(\lambda - T) \neq \{0\}$$

By Lemma 1(Lemma 13.16):

$$X_\lambda \perp X_{\lambda'} \qquad \forall \lambda \neq \lambda' \text{ and } \lambda, \lambda' \in \sigma_p(T) \setminus \{0\}$$

Lemma 13.17 — Lemma 2.

Let $\lambda \in \sigma_p(T) \setminus \{0\}$.

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

ii) $\forall r > 0$: $\sigma_p(T) \setminus B_r(0)$ is finite.

Proof. i) Let $B_r^{X_\lambda}(0) = \{x \in X_\lambda : \|x\|_H < r\}$, which is a bounded set. By compactness of T, $\overline{T(B_r^{X_\lambda}(0))}$ is compact. But since $Tx = \lambda x \ \forall x \in X_\lambda$, taking $r = 1$

$$T(B_1^{X_\lambda}(0)) = \lambda B_1^{X_\lambda}(0)$$

So

$$\lambda\overline{(B_1^{X_\lambda}(0))} \text{ is compact } \implies \overline{B_1^{X_\lambda}}(0) \text{ is compact } \implies \dim(X_\lambda) < \infty$$

By Theorem 4.15, the unit ball is compact if and only if the space is finite dimensional

ii) Suppose not, then $\exists r > 0$: $\sigma_p(T) \setminus B_r(0)$ is infinite (one can show $\sup_{\lambda \in \sigma_p(T)} |\lambda| < \infty$, no proof provided here).

A proof of $\sup \sigma(A) = \sup_{\|x\|=1} \langle x, Ax \rangle$ can be found on Pg.231(243 in pdf) in THEOREM 5.3.16 of the notes provided on Blackboard.

Then one can pick sequence $(\lambda_k) \subset \sigma_p(T)$ with

$$\lambda_k \neq \lambda_l, \ \forall k \neq l \text{ and } |\lambda_k| > r, \ \forall k \in \mathbb{N}$$

Let $e_k \neq 0$ be eigenvector for λ_k:

$$Te_k = \lambda_k e_k \qquad \forall k \in \mathbb{N}$$

By compactness of T, $\exists \Lambda \subset \mathbb{N}$ such that for some $y \in H$,

$$Te_k \xrightarrow[k \to \infty, k \in \Lambda]{\|\cdot\|_H} y \in H$$

Hence,

$$(\lambda_k e_k) \xrightarrow[k \to \infty, k \in \Lambda]{\|\cdot\|_H} y \in H$$

In particular, $(\lambda_k e_k)_{k\in\Lambda}$ is Cauchy. But for $k \neq l$,

$$\|\lambda_k e_k - \lambda_l e_l\|_H^2 = \langle \lambda_k e_k, \lambda_k e_k\rangle + \langle \lambda_l e_l, \lambda_l e_l\rangle + \underbrace{\langle \lambda_k e_k, \lambda_l e_l\rangle}_{=\lambda_k \overline{\lambda_l}\langle e_k, e_l\rangle \overset{Lemma 1}{=} 0} + \langle \lambda_l e_l, \lambda_k e_k\rangle$$

Assume, $\|e_k\|_H = \|e_l\|_H = 1$, otherwise replace e_k by $e_k/\|e_k\|_H$,

$$\|\lambda_k e_k - \lambda_l e_l\|_H^2 = |\lambda_k|^2 + |\lambda_l|^2 > 2r$$

■

Now return to the proof of Theorem 13.14.

Proof. i) $\sigma(T) \subset \mathbb{R}$ by Proposition 13.13.

ii) By Lemma 2 ii)(Lemma 13.17),

$$A_n = \sigma_p(T) \cap \{z : \frac{1}{n+1} \leq |z| \leq \frac{1}{n}\}(\subset \sigma_p(T) \setminus B_{1/(n+1)}(0))$$

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

iii) By applying Gram-Schmidt to the (finite-dimensional) space X_λ, $\lambda \in \sigma_p(T)\setminus\{0\}$, can ensure that eigenvectors e_k, e_l to eigenvalues $\lambda_k = \lambda_l$ are orthonormal (that is $e_k \perp e_l$ and $\|e_k\| = \|e_l\| = 1$). If they belong to distinct eigenvalues, this is automatic after normalizing them to have norm 1 by Lemma 1(Lemma 13.16).

Remark 13.18 Lemma 2 gives important structural information on the spectrum. In particular, ii) implies that $\sigma_p(T)\setminus\{0\}$ is countable with no accumulation point and by i) each eigenvalue has "finite multiplicity".

So let (λ_k) be the elements of $\sigma_p(T) \setminus \{0\}$, counted with multiplicities(i.e. $\dim(X_{\lambda_k})$ copies of λ_k.

Let

$$X \overset{\text{def.}}{=} \overline{span\{e_k\}} = \overline{\bigoplus_{\lambda\in\sigma_p(T)\setminus\{0\}} X_\lambda}$$

Claim 1: $\forall x \in X$: $x = \sum_k \langle x, e_k\rangle e_k$

Proof of Claim1:

Let $x_n = \sum_{k\leq n} \langle x, e_k\rangle e_k$. Then $\forall n \geq 0$,

$$\|x_n\|_H^2 = \sum_{k\leq n} |\langle x, e_k\rangle|^2 = \langle x_n, x\rangle \leq \|x_n\|_H \|x\|_H$$

so $\|x_n\|_H \leq \|x\|_H$ unless $x_n = 0$. Hence

$$\sum_k |\langle x, e_k \rangle|^2 = \lim_{n\to\infty} \|x_n\|_H^2 \leq \|x\|_H^2 < \infty$$

and $\forall n \geq m \geq 0$:

$$\|x_n - x_m\|_H^2 = \sum_{m\leq k\leq n} |\langle x, e_k \rangle|^2 \xrightarrow{n,m\to\infty} 0$$

Thus $x_n \xrightarrow[\|\cdot\|_H]{n\to\infty} y \in X$. Moreover, $\forall k \geq 0$,

$$\langle x - y, e_k \rangle = \lim_{n\to\infty} \langle x - x_n, e_k \rangle = \langle x, e_k \rangle - \lim_{n\to\infty} \langle x_n, e_k \rangle = 0$$

so $x = y$. With Claim 1 and continuity of T we have

$$\forall x \in X: \qquad Tx = \sum_k \langle x, e_k \rangle Te_k = \sum_k \langle x, e_k \rangle \lambda_k e_k$$

It remains to argue:

Claim 2: $Y \overset{\text{def.}}{=} X^\perp = ker(T)$, which concludes the proof.

Proof of Claim2:

If $Y = \{0\}$, it's trivial. So one can assume $Y \neq \{0\}$. First, note that

$$T(Y) \subset Y \tag{$*$}$$

For, if $y \in Y$ then $\forall k$:

$$\langle e_k, Ty \rangle = \langle Te_k, y \rangle = \lambda \langle e_k, y \rangle \overset{Y\perp\{e_k\}}{=} 0$$

By $(*)$,

$$T_Y = T|_Y : Y \to Y$$

is well defined. T_Y inherits compactness and self-adjointness from T (exercise). We want to show that $T_Y : Y \to Y$: $y \mapsto 0$ is the "0-map" on Y, or equivalently, $\|T_Y\|_{\mathcal{L}(X)} = 0$. If not, one can show (no proof) that

$$\sigma_p(T_Y) \setminus \{0\} \neq \emptyset$$

(in fact $+\lambda$ or $-\lambda$ is an eigenvalue where $\lambda = \|T_Y\|_{\mathcal{L}(X)}$)

But this can't be because if $e \in Y$ is an eigenvector for $\lambda \in \sigma_p(T_Y)$, $\lambda \neq 0$, then

$$Te \overset{e\in Y}{=} T_Y e = \lambda e$$

So $e \in X_\lambda \subset X = Y^\perp$. But $Y^\perp \cap Y = \{0\}$. ■

Remark 13.19 We have actually shown that H admits the orthogonal decomposition:

$$H = ker(T) \oplus \overline{\bigoplus_{\lambda \in \sigma_p(T) \setminus \{0\}} X_\lambda}$$

where there are countably many X_λ and each is of finite dimension.