Functional Analysis

RIF Galaxy

August 2022

On this note: This note is primarily made for personal study purpose. The note is made prior to the course taking place, and will be updated according to the lectures for Imperial College London 3rd year FA course(2022).

Contents

1 List of content by week

1.1 Week 1

- Structure and arrangement of the course
- Linear space, definition and examples
- L^p spaces, proof that it's a normed linear space
- Minkowski's inequality
- Hölder's inequality
- Young's Inequality
- Banach space, definition and method of checking its properties

1.2 Week2

- Metric linear space
- Jensen's inequality and convex function
- Topology: dense, separable space
- Schauder basis, existence of it implies separability
- Inner product, definition
- Hilbert space, definition
- Convex set
- Nearest point property
- Parallelogram law
- Orthogonal complement

1.3 Week3

- Finite dimensional Banach Space
- Compactness, closedness and boundedness
- Compactness and closed unit ball
- Boundedness and continuity of Linear operator
- Operator Norm
- Content On linear functionals, helpful to understanding operator

1.4 Week4

- Riesz representation theory
- Dual space of Hilbert space
- Dual space of Banach space
- Dual space of ℓ^p space
- Dual Operator(Not finished!)

1.5 Week5

- Hahn-Banach Theorem
- Zorn's Lemma
- Sublinear Map
- Separation Theorem
- Dual functional

1.6 Week6

- Uniform boundedness principle
- Baire Category Theorem
- Banach Steinhaus Theorem

1.7 Week7

- Open Map
- Open Mapping Theorem
- Equivalence of Norm for complete spaces
- Closed Operator
- Closed Graph Theorem

2 Preliminaries

This section aims to provide preliminary knowledge to functional analysis. This field of maths is decorated by ideas of both algebra and analysis, particularly linear algebra and real analysis. It is thus important to get familiar with the relevant ideas, as lack of either viewpoint stops you from getting the whole story. At some point, since we are talking about different spaces, topological concepts also comes in. Luckily, they're generally not complicated and presented here as preliminaries.

2.1 Linear space

Mathematicians usually talks about spaces. However they are simply sets with additional structure. Linear spaces, also called vector spaces, are those with linear structure. This means you have vector addition, scalar multiplication, commutativity and distributivity.

Definition 2.1.1 (Linear space)
A linear space $(V, \oplus, (\mathbb{F}, +, \cdot), \odot)$ over a field $\mathbb{F}$, where

- $(V, \oplus)$ is an abelian group
- $(\mathbb{F}, +, \cdot)$ is a field

and multiplication by a scalar $\odot : \mathbb{F} \times V \to V$ satisfies for every $\alpha, \beta \in \mathbb{F}$ with $v, \omega \in V$

- $\alpha \odot (v \oplus \omega) = \alpha \odot v + \alpha \odot \omega$
- $(\alpha + \beta) \odot v = \alpha \odot v + \beta \odot v$
- $\alpha \odot (\beta \odot v) = (\alpha \cdot \beta) \odot v$
- $\mathbf{1} \cdot v = v$, where $\mathbf{1}$ is unit element in $\mathbb{F}$

2.1.1 Examples of linear spaces

In this section we have both example and counter example.

Example 2.1.2 (Vector space over field)
$\mathbb{F}^n$ where $\mathbb{F}$ is a field, $n \in$
natu is a linear space. This includes 3 or 2-dimensional vector space over $\mathbb{R}$, or 3-dimensional vector field over $\mathbb{F}_p$. For example, let's check

$$V = \mathbb{F}_2^2 = \{(v_1, v_2) | v_1, v_2 \in \mathbb{F}_2\}$$

with the natural definition of scalar multiplication and term-wise addition over $\mathbb{F}_2$. Note that this is indeed a space of four elements:

$$V = \{(0,0), (0,1), (1,0), (1,1)\}$$

With scalars only 1 or 0. Thus it's easy to closure under scalar multiplication. Now consider vector addition, one just have to check every pair of addition and see if still falls into V, like

$$(1,0) + (1,1) = (0,1) \in V$$

Example 2.1.3
Consider set of convergent sequence:

$$V = \left\{ \{x_n\}_1^\infty : x_i \in \mathbb{F} \, \forall i \in \mathbb{N}, \, and \lim_{n\to\infty} x_n \to C \right\}$$

First we consider the addition to be term-wise and multiplication applied to whole sequence:

$$x + y = \{x_i + y_i\}_1^\infty,\ \alpha \odot x = \{\alpha \cdot x_i\},\ \forall x, y \in V,\ \alpha \in \mathbb{F}$$

When C is fixed, this is a vector space if and only if $C = 0$ When C is not fixed, this becomes a vector space.

Example 2.1.4 (Polynomials)
V is set of all polynomials:

$$f(z) = \sum_{j=0}^{n} a_j z^j,\ n \in \mathbb{N}$$

with $z \in \mathbb{C}$ and $a_j \in \mathbb{Q}$ with addition and multiplication of polynomials. Unfortunately this is not a vector space, since the field is set to be $\mathbb{C}$. When we multiply an element in V by a complex number, say $1 + 2i$, we could end up in some polynomials with complex coefficient. But this would be a vector space if we change to $a_j \in \mathbb{C}$ or $z \in \mathbb{Q}$

Example 2.1.5 (Analytic functions)
Consider set of all analytic functions $f : \mathbb{C} \to \mathbb{C}$ satisfying:

$$\frac{d^2}{dz^2} f - \frac{d}{dz} f - 2z = 0$$

This is not a vector space. Take a non-trivial f, consider $g = 2f$:

$$\begin{aligned}
&\frac{d^2}{dz^2} g - \frac{d}{dz} g - 2z \\
=&\frac{d^2}{dz^2}(2f) - \frac{d}{dz}(2f) - 2z \\
=&2\frac{d^2}{dz^2} f - 2\frac{d}{dz} f - 2z \\
=&4z - 2z = 2z \neq 0
\end{aligned}$$

However, removing $-2z$ will make this a vector space.

Example 2.1.6 (Weak L^p space)
Define $D_f(t) = \{\lambda x \in \mathbb{R} : |f(x)| > t\}$, where λ is Lebesgue measure on $\mathbb{R}$. Consider:

$$L^{p,w} = \left\{ f : \mathbb{R} \to \mathbb{R} : f,\ measurable,\ and\ \exists C > 0,\ s.t.\ D_f(t) < \frac{C^p}{t_p}\ \forall t > 0 \right\}$$

This is a vector space for $p > 0$ and $L^p(\mathbb{R}) \subset L^{p,w}(\mathbb{R})$. Distributivity and commutativity of scalar operation follows immediately from their definition. The point is to check closure under scalar multiplication and addition.
$\mathbf{f + g \in L^{p,w}}$:
Let $f, g \in L^{p,w}$, then we can find $C_1 > 0$ and $C_2 > 0$ with

$$\begin{aligned}
C_1^p > t^p \cdot \lambda\{x : |f(x)| > t\},\ \forall t > 0 \\
C_2^p > t^p \cdot \lambda\{x : |g(x)| > t\},\ \forall t > 0
\end{aligned}$$

Now by triangular inequality we have $|f(x)| + |g(x)| \geq |f(x) + g(x))|$.
Thus $|(f + g)(x)| > t \implies |f(x)| + |g(x)| > t \implies 2 \cdot max(|f(x)|, |g(x)|) > t$. So consider set A, B, C:

$$\begin{aligned}
X &:= \{x \in \mathbb{R} : |f(x) + g(x)| > t\} \\
Y &:= \{x \in \mathbb{R} : |f(x)| + |g(x)| > t\} \\
Z &:= \{x \in \mathbb{R} : 2max(|f(x)|, |g(x)|) > t\} = \{x \in \mathbb{R} : max(|f(x)|, |g(x)|) > t/2\}
\end{aligned}$$

We have $X \subseteq Y \subseteq Z$, thus $\lambda(X) \leq \lambda(Y) \leq \lambda(Z)$.
Thus $t^p \cdot \lambda(A) \leq 2^p(t/2)^p\lambda\{x : max(|f(x)|, |g(x)|) > t/2\} \leq 2^p \cdot max(C_1^p, C_2^p) < \infty$ for all $t > 0$.
Showing closure under scalar multiplication is a bit easier. Take any $t > 0$, we have

$$\begin{aligned}
&t^p \cdot \lambda\{x \in \mathbb{R} : |2f(x)| > t\} \\
=&2^p \left(\frac{t}{2}\right)^p \cdot \lambda\{x \in \mathbb{R} : |f(x)| > t/2\} \\
<&2^p C_1^p
\end{aligned}$$

Which completes the proof.

2.2 Metric Linear space

Definition 2.2.1 (Metric linear space)
A metric space (V, d) is called metric linear space if its vector addition and scalar multiplication $\oplus$, $\odot$ are continuous

Remark 2.2.2 (Equivalence of metrics)
$\oplus$ can be considered as a function: $\oplus : V \times V \to V$, endowed with metric $\rho_1 : V \times V \to V$ defined as $\rho_1(x_1+y_1, x_2+y_2) = max(d(x_1, x_2), d(y_1, y_2))$ or sum $\rho_2 : V \times V \to V$ defined as $\rho_2(x_1+y_1, x_2+y_2) = d(x_1, x_2) + d(y_1, y_2)$. The two metrics are topologically equivalent, i.e. they induces same topology.
Similarly, $\odot$ can be considered as a function: $\odot : \mathbb{F} \times V \to V$, endowed with metric $\rho_1 : \mathbb{F} \times V \to V$ defined as $\rho_1(k_1x_1, k_2x_2) = max(|k_1-k_2|, d(x_1, x_2))$ or sum $\rho_2 : V \times V \to V$ defined as $\rho_2(x_1+y_1, x_2+y_2) = |k_1 - k_2| + d(x_1, x_2)$. Again, the two metrics are topologically equivalent.

Definition 2.2.3 (Translation invariant)
A metric ρ is translation invariant if for all $x, y, z \in V$, we have $\rho(x, y) = \rho(x - z, y - z)$

Proposition 2.2.4
Addition is continuous with respect to translation invariant metic.

Proposition 2.2.5 (Metric induced by norm)
Let $\|\cdot\|$ be a norm on V, then its induced metric $\rho(x, y) = \|x - y\|$ is translation invariant. See definition of norm here.

2.3 Topology

2.3.1 Separability

Definition 2.3.1 (dense)
A set in S metric space (V, d) is dense if there it intersects with any open subset of V. Equivalently this is $\forall x \in V, \forall \varepsilon > 0$, $D \cap B_{x,\varepsilon} \neq \emptyset$.

Definition 2.3.2 (separable)
A metric space (V, d) is separable if it contains a countable dense subset.

2.3.2 Schauder Basis and Hamel basis

Definition 2.3.3 (Schauder basis)
A Schauder basis of Normed vector space $(v, \|\cdot\|)$ is a set $B \subset V$ that is linearly independent, and that $\forall x \in V$, we have a sequence $\{a_n\}$ with $\lim_{n\to\infty} \sum_{k=1}^{n} a_k b_k \to x$. In plain language, this means that every element of the set can be expressed as a infinite linear combination of the basis.

Proposition 2.3.4 (Schauder implies separability)
If a normed vector space has a Schauder basis, then it's separable.

Definition 2.3.5 (Hamel basis)
A set $H \subset X$ is a Hamel basis for linear space X, if and only if H is linearly independent and $\forall x \in X$, $\exists! a_i \in \mathbb{R}$ for all $1 \leq i \leq n$ for some n and

$$x = \sum_{i=1}^{n} a_i h_i, \;\; h_i \in H$$

Remark 2.3.6 (Hamel basis and Schauder basis)
Having Hamel basis means that one can express any element by a necessarily finite linear combination of the basis. For Schauder basis, it could be infinite linear combination.

2.3.3 Compactness

Definition 2.3.7 (Compactness)
Let (X, d) be a metric space. A subset $S \subset X$ is **compact** if any sequence (x_n) in S has a convergent subsequence converges in to some $x \in S$

Remark 2.3.8
The definition is "sequential compactness". There're many versions of definition of compactness. One shall really pay attention to the definition in linear algebra or real analysis that compactness is equivalence to closedness and boundedness. We'll see later that this is guaranteed to true only for finite-dimensional spaces. However, it is true that compactness always implies closedness and boundedness. Following examples we prove the implications in detail and present the fact that compactness of space can studied by looking at closed unit ball.

Proposition 2.3.9 (Compact $\implies$ closed and bounded)
If a set K is compact, then it's closed and bounded. Here, bounded means $\exists L > 0$ such that for all $x, y \in K$, we have $d(x, y) \leq L$.

Proof

First we show that compact implies closed. Let (x_n) be a convergent sequence in K, then by compactness, there is a subsequence of (x_n), name it (y_n) which converges to $y \in K$, but by uniqueness of limit (x_n) also converges to $y \in K$.
Now we show boundedness by contradiction. Assume K is not bounded. Then we fix $a \in K$, and by unboundedness we have that $\forall n \in \mathbb{N}$, $\exists x \in K$ with $d(x, a) > n$. Now consider a sequence (x_n) where $d(x_n, a) > n \; \forall n > 0$. This sequence has no convergent subsequence.

Theorem 2.3.10 (F.Riesz)
Let $(X, \|\cdot\|)$ be a normed vector space. Following are equivalent:

- $dim(X) \leq \infty$
- Closed unit ball $B_U = \{x \in X : \|x\| \leq 1\}$ is compact.

proof: Proof uses a lemma which is shown laterNOOOOOOT COMPLEEEEEEET! COOOOOOOM BAAAAAACK!!!

Lemma 2.3.11 (F.Riesz)
Let $(X, \|\cdot\|$ be a normed linear space with a subspace $Y \subset X, Y \neq X$. Then for all $\varepsilon \in (0, 1)$, there exists $x \in X$ with $\|x\| = 1$ and $d(x, Y) \equiv \inf_{y \in Y} \|x - y\| > 1 - \varepsilon$ *proof*: NOOOOOOT COMPLEEEEEEET! COOOOOOOM BAAAAAACK!!!

3 Banach Space

Banach space is the one of the core concepts of functional analysis. It is a special type of **vector space**, with a **norm** working on the space as well as the property that every **Cauchy sequence converges** in the space.
One should pay attention that a considerable portion of content in this chapter is not based on completeness but only requires a norm on the space. However, it is obvious that they can be applied to Banach spaces and most importantly, these results do relate themselves to Banach spaces.

3.1 Definitions

Definition 3.1.1 (Normed vector space)
A normed vector space $\mathbf{X}$ is a vector space (over $\mathbb{F}$, usually $\mathbb{C}$ or $\mathbb{R}$), equipped with a norm function $||\cdot|| : \mathbf{X} \times \mathbf{X} \to \mathbb{R}$ satisfying following:

- $||x|| \geq 0$, $\forall x \in \mathbf{X}$
- $||x|| = 0$ if and only if $x = 0$
- $||ax|| = a||x||$, $\forall x \in \mathbf{X}$, $\forall a \in \mathbb{F}$
- $||x + y|| \leq ||x|| + ||y||, \forall x, y \in \mathbf{X}$

A normed vector space is also called a normed linear space.

Definition 3.1.2 (Cauchy Sequence)
A Cauchy sequence in a normed vector space V is a sequence $\{a_n\}_1^\infty$, where each $a_n \in V$, satisfying the following: $\epsilon > 0, \exists\, N \in \mathbb{N}\ s.t.\ \forall\, m, n > N, ||a_m - a_n|| < \epsilon$

Definition 3.1.3 (Convergence)
A sequence $\{a_n\}_1^\infty$in a normed vector space V is convergent if $\exists a \in V$ s.t. $\forall \epsilon > 0, \exists N > 0 s.t. \forall n > N, ||a_n - a|| < \epsilon$.

Definition 3.1.4 (Completeness)
A normed vector space is complete if every Cauchy Sequence converges to a point in the space.

Theorem 3.1.5 (Every metric space can be completed)
Result is put in completion of metric space in appendix.

Definition 3.1.6 (Banach space)
A Banach space is a complete normed vector space.

3.2 Examples

Example 3.2.1 ($\mathbb{R}^n$)
Our acquainted three-dimensional vector space over $\mathbb{R}$ is a Banach space under the standard vector norm, this is the case when $n = 3$. This is a trivial result, since such norm gives the "length" of a vector, and a Cauchy sequence of vectors indicates that 'endpoints' of vectors come arbitrarily close. It's easy to prove that such a sequence converges to a three-dimensional vector. In fact, any n-dimensional vector space over $\mathbb{R}$ is a Banach space.

Example 3.2.2 (Real valued functions)
The set of all real-valued function on [0,1] with norm $\|f\| = \max_{t\in[0,1]} |f(t)|$ is a Banach space. To verify this (thoroughly) we shall first show that this is a vector space. This is trivial since sum and and scalar multiplication of a real-valued function is also real-valued. Then we should show that this

function is indeed a norm. Finally, we should check that every Cauchy sequence under this norm is convergent. Intuitively, this norm gives the 'maximal pointwise difference' between two functions, hence if a sequence is Cauchy, the 'maximal pointwise difference' converges to zero. Rigorous proof is left to readers.

Example 3.2.3 (Continuous function under supremum norm)
Consider $C[0,1]$, set of continuous real-valued defined on $[0,1]$ with supremum norm:

$$\|f\|_\infty = \sup_{t\in[0,1]} |f(t)|$$

This is a Banach space (why?). Now a relevant example is $C^1[0,1]$,set of complex-valued function defined on $[0,1]$ with continuous first derivative. Unfortunately this is not a Banach space under supremum norm, however, if we equip the space with a new norm: $\|f\| = \|f\|_\infty + \|f'\|_\infty$, we end up having a Banach space.

Example 3.2.4 (L-p spaces)
L-p spaces are function spaces with finite $p-norm$. Let (S,Σ,μ) be measure space, $p\in[1,+\infty]$. L-p space consists of functions $S\to\mathbb{C}$ with $p-norm$:

$$\|f\|_p \equiv \left(\int_S |f|^p \, d\mu\right)^{\frac{1}{p}} < \infty$$

It is not obvious how $p-norm$ really gives a norm, the difficulties lie in the part of proving the triangular inequality. In the specific background of L-p spaces, the inequality is precisely *Minkowski's inequality.* Detailed material on Hölder's inequality and Minkowski's inequality are put in appendix.

Proposition 3.2.5 (Completeness of product of Banach spaces)
Let X, Y be Banach spaces, then $X\times Y = \{(x,y) : x\in X,\ y\in Y\}$ endowed with norm $\|(x,y)\|_{X\times Y} = \|x\|_X + \|y\|_Y$ is complete. Proof of this thoerem is left as an exercise.

3.3 Finite dimensional Normed space

Bare in mind that in this section we do not assume that spaces are complete, so please pay attention which results are based on completeness. However, we shall see that all finite dimensional vector space over complete fields are complete.
Some main theorems to focus on in this section, presented in plain language:

- All finite dimensional vector space over a complete field are complete
- All norms on finite dimensional vector spaces are equivalent
- $Bounded + Closed = Compact$ if and only if dimension is finite.

Reference: norm-equivalence note

3.3.1 Equivalence of Norms and topology

There are many norms. Some acts similarly, some acts differently. One may have seen different types of matrix norm, for example, Frobinius norm and operator norm. Computational mathematicians don't seem to care about this, why is it? Maybe they don't make much difference!

Definition 3.3.1 (Equivalence of Norm)
Let X be a vector space. Two norms $\|\cdot\|_a$, $\|\cdot\|_b$ on X are **equivalent** if their exists $m,n>0$ satisfying following equation:

$$m\|x\|_a \le \|x\|_b \le M\|x\|_a,\ \forall x\in X$$

Theorem 3.3.2 (Equivalence of finite dimensional norms)
Let X be a finite dimensional vector space, then any two norms $\|\cdot\|_a$ and $\|\cdot\|_b$ are equivalent.
proof: The proof are divided into four steps:

- Showing that equivalence of norm is transitive
- Showing that equivalence of equivalence on unit sphere implies equivalence on X
- Showing that any norm is continuous with respect to $\|\cdot\|_1$
- Showing that any norm is equivalent to $\|x\|_1 \equiv \sum_{s=1}^{n} |x_s|$

Proof : STEP I

Let $\|\cdot\|_a$ be equivalent to $\|\cdot\|_b$ and $\|\cdot\|_b$ equivalent to $\|\cdot\|_c$. Then $\exists m_1, m_2, M_1, M_2 > 0$ with

$$m_1 \|x\|_a \le \|x\|_b \le M_1 \|x\|_a \,, \ \forall x \in X$$
$$m_2 \|x\|_b \le \|x\|_c \le M_2 \|x\|_b \,, \ \forall x \in X$$

Then

$$m_1 m_2 \|x\|_a \le m_2 \|x\|_b \le \|x\|_c \,, \ \forall x \in X$$
$$\|x\|_c \le M_2 \|x\|_b \le M_1 M_2 \|x\|_c \ \ \forall x \in X$$

Which gives $m_1 m_2 \|x\|_a \le \|c\| \le M_1 M_2 \|x\|_a$ for arbitrary x. Hence $\|\cdot\|_a$ and$\|\cdot\|_c$ are equivalent.

Proof : STEP II

Now let us assume that $\|\cdot\|_a$ is equivalent to $\|\cdot\|_b$ on $U_a = \{s \in X : \|s\|_a = 1\}$. Then let $x \in X$ be non-zero. Then we have

$$m \left\| \frac{x}{\|x\|_a} \right\|_a \le \left\| \frac{x}{\|x\|_a} \right\|_b \le M \left\| \frac{x}{\|x\|_a} \right\|_a$$

So

$$m \|x\|_a \frac{1}{\|x\|_a} \le \|x\|_b \frac{1}{\|x\|_a} \le M \|x\|_a \frac{1}{\|x\|_a}$$

And since $\|x\|_a$ is non-zero,we have that

$$m\|x\|_a \le \|x\|_b \le M \|x\|_a$$

Now since x is arbitrary, the proof is completed.

Proof : STEP III

Now we shall proof continuity of any norm under $\|\cdot\|_1$. This can be done by showing for a sequence (x_n) converging to x under the metric induced by $\|\cdot\|_1$, the norm if its terms under $\|\cdot\|_a$ converges to $\|x\|_a$. So let (x_n) be a sequence in X with $x_n \xrightarrow{n\to\infty} x$. We have

$$|\, \|x_n\|_a - \|x\|_a \,| \le \|x_n - x\|_a \le M \|x_n - x\|_1 \to 0 \quad \text{when } n \to \infty$$

Thus

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

Proof : STEP IV

Now we shall apply extreme value theorem to obtain our final result here. Using the theorem requires unit sphere to be a compact set. Proof of this fact is given later, one should realise that the proof does not depend on equivalence of norms, as we only require compactness in $X, \|\cdot\|_1$. However, it is true that unit sphere is compact under any norm in finite dimensional cases. So we have that $U_1 = \{s \in X : \|s\|_1 = 1\}$ is compact with a function $\|\cdot\|_a$ continuous on it, so by extreme value theorem it attains maximum $M_U = max\{\|x\|_a : x \in U_1\}$ and minimum $m_u = min\{\|x\|_a : x \in U_1\}$ on U_1, thus for any $x \in U_1$ we have

$$m_u \|x\|_1 = m_u \leq \|x\|_a \leq M_u = M_u \|x\|_1$$

Hence we show that any norm is equivalent to $\|\cdot\|_1$ on unit sphere.

By combining the results of the four steps, we finish the proof of the theorem.

Remark 3.3.3

The proof is not unique. We can also choose other norms to be the "bridging" norm, say supremum norm which in finite dimensional case becomes the max norm: $\|x\| \equiv max\{|x_i|\}$. About the meaning of equivalence here, in fact, equivalent norms are equivalent in the sense that they induces same topology, so it is also called "topologically"equivalent. Generally speaking, this means that topological properties such as open, close, compact, convergence, continuity which holds for one norm will hold in its equivalent norms.

Following results are simple exercises to check statements above.

Proposition 3.3.4 (Equivalence of openness)

Open sets in $(X, \|\cdot\|_a)$ are open in $(X, \|\cdot\|_b)$ (Following notation in definition of equivalent norms).

Proof

It suffices to check open balls. Let $B^a_x(r) \equiv \{s \in X : \|x - s\|_a < r\}$ be open balls with radius $r > 0$ centered at x, which is an open ball in $(X, \|\cdot\|_a)$. Choose $p \in B^a_x(r)$, we should show that $\exists \varepsilon > 0$ with $B^b_p(\varepsilon) \equiv \{s \in X : \|p - s\|_b < \varepsilon\} \subset B^a_x(r)$.
By openness of $B^a_x(r)$, we have that $\exists \varepsilon_a > 0$ with

$$\begin{aligned} B^a_p(\varepsilon_a) &\equiv \{s \in X : \|p - s\|_a < \varepsilon_a\} \\ &= \{s \in X : m\|p - s\|_a < m\varepsilon_a\} \\ &\supseteq \{s \in X : \|p - s\|_b < m\varepsilon_a\} \\ &= B^b_p(m\varepsilon_a) \end{aligned}$$

Note that $B^b_p(m\varepsilon_a) \subseteq B^a_p(\varepsilon_a) \subset B^a_x(r)$, hence $\varepsilon = m\varepsilon_a$.

3.4 Subspace of Normed linear space

As shown previously, finite dimensional vector spaces are all complete (we will assume that the fields of vector spaces are complete, $\mathbb{R}$ or $\mathbb{C}$). This is also true for finite dimensional subspace of normed linear spaces. Note that we don't need the space to be complete when asserting completeness of their finite dimensional subspace.

Proposition 3.4.1 (Completeness of finite-dimensional subspace)

Let $(x, \|\cdot\|)$ be a normed linear space. Then a linear subspace $Y \subset X$ with $dim(Y) < \infty$ endowed with $\|\cdot\|$ induced in Y is a Banach space.

Also, we have closedness of finite-dimensional subspace.

Proposition 3.4.2 (Closedness of finite-dimensional subspace)
Let $(x, \|\cdot\|)$ be a normed linear space. Then a linear subspace $Y \subset X$ with $dim(Y) < \infty$ endowed with $\|\cdot\|$ induced in Y is closed.

Proof of this two proposition is simple use of the results of finite dimensional normed linear spaces begin Banach, which is left to readers. We shall now take a look at few examples showing different subspaces of normed vector spaces.

Example 3.4.3 (Finite-dim subspace being complete)
NOOOOOOT COMPLEEEEEEET! COOOOOOOM BAAAAAACK!!!

Example 3.4.4 (Infinite-dim subspace not complete)
Consider $X = C[0,2] \subset L^1[0,2]$, set of all real-valued continuous function on $[0,1$, endowed with 1-norm: $\|f\|_1 = \int_0^1 |f(t)|dt$. Consider sequence of function (f_n):

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

Now f_n is Cauchy in 1-norm, but its limit is not in $C[0,1]$:

$$f(t) = \begin{cases} 0 & 0 \leq t < 1 \\ 1 & 1 \leq t \leq 2 \end{cases}$$

Thus X is not a complete subspace of $L^1[0,2$, as we have a Cauchy sequence that does not converge to a point in X.

Example 3.4.5 (Completeness depends on choice of norm)
We continue considering the settings in our last example, but endow X with supremum norm: $\|\cdot\|_\infty$. Now our (f_n) is no longer Cauchy. One can prove that $(X, \|\cdot\|_\infty$ is actually complete. Intuitively, supremum norm is sensitive to "discontinuity at points", so convergence in supremum norm ensures no "jump" of value at any point.

3.5 Fixed point

Contraction mapping theorem, sometimes called Banach fixed point theorem, guarantees the existence and uniqueness of fixed points of certain self-maps of metric spaces, and provides a constructive method to find those fixed points.

Theorem 3.5.1 (Fixed point)
Let X be Banach space, $f : X \to X$ is a contraction mapping, i.e. a mapping satisfying $d(f(x), f(y)) \leq qd(x, y)$, $\forall x, y \in X$, where $q < 1$ is a fixed constant. Then f admits a unique fixed point $x^* \in X$, which can be found by defining sequence x_n in X, with $x_{n+1} = f(x_n)$, then $\lim_{n\to\infty} x_n = x^*$
proof:
By assumption, we have $\|x_n + 1 - x_n\| = \|f(x_n) - f(x_{n+1})\| \leq q\|x_n - x_{n+1}\|$.
So if we let $k = \|x_2 - x_1\|/q$, we have that $\|x_{n+1} - x_n\| \leq kq^n$ with $q < 1$. This sequence is clearly Cauchy. To see this, choose $m, n \in \mathbb{N}$ with $m < n$

$$\begin{aligned} \|x_m - x_n\| &\leq \sum_{i=m}^{n-1} \|x_{i+1} - x_i\| \\ &\leq \sum_{i=m}^{n-1} kq^i \\ &= \frac{kq^m(1 - q^{(n-m)})}{1 - q} \\ &< \frac{kq^m}{1 - q} \end{aligned} \tag{1}$$

Fix $\varepsilon > 0$, we can find large N so that

$$q^N \leq \frac{\varepsilon(1 - q)}{k}$$

Then for all $m, n > N$ we have

$$\|x_m - x_n\| < \frac{kq^m}{1 - q} < \frac{\varepsilon(1 - q)}{k}\frac{k}{1 - q} = \epsilon$$

Hence x_n is Cauchy, thus convergent to a unique point $x^* \in X$

Remark 3.5.2
This theorem is easy to prove, however it's very useful. We may see examples of fixed point by playing with calculator: choose any real number and calculate its cos value, and continue input the result to *cos*, we may find that the result become stable around 0.7390....... This is a fixed point of *cos*. *sin* also has fixed point, which is zero. It can be used to give sufficient condition where Newton method for finding root converges. In study of ODE contraction mapping can be used to guarantee that Picard iteration converges to a certain function. See Picard–Lindelöf theorem on Wikipedia.

3.6 Exam

3.6.1 Proof of completeness

Generally there are three steps in proving completeness.

- Find a candidate for the "limit" of a Cauchy sequence.
- Show that it's indeed the "limit'.
- Show that it's still in the space.

4 Hilbert Space

Hilbert space is a special class of Banach space. Apart from completeness and norm, it is also equipped with an additional structure, **inner product**. This allow us to explore nice geometric properties of the space,like orthogonality and angle. We'll see later that this structure resemble Euclidean space in many ways. A Hilbert space is naturally Banach, while the reverse may not be true.

4.1 Definitions and notations

Definition 4.1.1 (inner product)
Let X be a vector space over $\mathbb{C}$. An **inner product** is a function $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{C}$ satisfying following: $\forall x, y, z \in X, \alpha$ a scalar,

1 $\langle x, y \rangle = \overline{\langle y, x \rangle}, \forall x, y \in X$

2 $\langle x, x \rangle \geq 0$

3 $\langle x, x \rangle = 0$ iff $x = 0$

4 $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$

5 $\langle ax, z \rangle = a \langle x, z \rangle$

1 is complex conjugation. 2 and 3 is positive-definiteness. 4 and 5 is left-linearity.

Proposition 4.1.2
Let X be a vector space, $\langle \cdot, \cdot \rangle$ an inner product on the space, $x, y, z, w \in X$ and α a scalar. Then:

- $\langle x, ay \rangle = \overline{a} \langle x, y \rangle$
- $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$
- $\langle x, ay + z \rangle = \overline{a} \langle x, y \rangle + \langle x, z \rangle$
- $\langle x + y, z + w \rangle = \langle x, z \rangle + \langle y, z \rangle + \langle y, w \rangle + \langle x, w \rangle$

Proof is trivial. Left to readers.

Proposition 4.1.3 (Cauchy-Schwartz)
If $\langle \cdot, \cdot \rangle$ is an inner product on X, then $\forall x, y \in X$,

$$| \langle x, y \rangle |^2 \leq \langle x, y \rangle \cdot \langle y, y \rangle$$

Proposition 4.1.4 (induced norm)
If $\langle \cdot, \cdot \rangle$ is an inner product on X, then

$$\|x\| \equiv \langle x, x \rangle^{\frac{1}{2}}$$

defines a norm on X.

Definition 4.1.5 (Hilbert space)
A (complex) Hilbert space $\mathscr{H}$ is a vector space over $\mathbb{C}$ that is complete in the metric

$$d(x, y) = \|x - y\| = \langle x - y, x - y \rangle^{\frac{1}{2}}$$

One may also say, Hilbert space is a complete inner product space. The followings are a few examples of Hilbert space.

Example 4.1.6 (Euclidean Space over $\mathbb{R}$)
It happens that 2-D or 3-D vector space over $\mathbb{R}$ is an example of Hilbert space, under the standard definition of vector inner product.

Example 4.1.7 (Euclidean Space over $\mathbb{C}$)
$\mathbb{C}^n$ with inner product

$$\langle x, y\rangle = \sum_{i=1}^{n} \overline{x_i}\, y_i$$

is a Hilbert space. This is a generalization of last example. Still, this is a finite-dimensional Hilbert space.

Example 4.1.8 (Sequence space)
Complex sequence space:

$$\ell^2 = \left\{ \{x_n\}_1^\infty : \sum_{k=1}^{\infty} |x_k|^2 < \infty \right\}$$

with inner product,denoting $x = \{x_n\}_1^\infty$ and $y = \{y_n\}_1^\infty$

$$\langle x, y\rangle = \sum_{k=1}^{\infty} \overline{x_k}\, y_k$$

4.2 Hilbert space Geometry

Structure of inner product allows discussion for nice geometric property of Hilbert spaces. This includes orthogonality, angles and nearest distance etc.

4.2.1 Orthogonality

Orthogonality is the generalization of two lines being perpendicular. In euclidean geometry, we have Pythagorean theorem closely related to such property. Results on orthogonality in Hilbert spaces in many ways resemble their Euclidean version.

Definition 4.2.1 (Orthogonality)
Let $\mathscr{H}$ be a Hilbert space and $f \in \mathscr{H}$. Say g is orthogonal to f if $\langle f, g\rangle = 0$, writes $f \perp g$. For two sets $A, B \in \mathscr{H}$, write $A \perp B$ if $a \perp b$ for all $a \in A$ and $b \in B$

Definition 4.2.2 (Orthogonal Complement)
Let A be a set in Hilbert space $\mathscr{H}$. Its orthogonal complement, $A^\perp$ is the set of all vectors $f \in \mathscr{H}$ such that $f \perp g\, , \forall g \in A$. $A^\perp$ is always a subset of $\mathscr{H}$. Moreover, $A^\perp = \cup_{a\in A}(a)^\perp$, and we can prove that $A^\perp$ is always a closed subset of $\mathscr{H}$.

Proposition 4.2.3 (Pythagorean Theorem)
Let f_1, f_2, ..., f_n be pairwise orthogonal vectors in Hilbert space $\mathscr{H}$. Then

$$\left\| \sum_{k=1}^{n} f_k \right\|^2 = \sum_{k=1}^{n} \|f_k\|^2$$

proof:
It suffices to show for n=2, and proceed with induction. Consider $a, b \in \mathscr{H}$ with $a \perp b$. Then

$$\begin{aligned} \|a+b\|^2 &= \langle a+b, a+b\rangle \\ &= \langle a, a\rangle + \langle a, b\rangle + \langle b, a\rangle + \langle b, b\rangle \\ &= \langle a, a\rangle + 0 + 0 + \langle b, b\rangle \\ &= \|a\|^2 + \|b\|^2 \end{aligned} \tag{2}$$

Remark 4.2.4 (Parallelogram equality)
Let $a, b \in \mathscr{H}$ be arbitrary vectors. Then

$$\begin{aligned} \|a+b\|^2 &= \langle a+b, a+b\rangle \\ &= \langle a,a\rangle + \langle a,b\rangle + \langle b,a\rangle + \langle b,b\rangle \\ &= \langle a,a\rangle + \langle a,b\rangle + \overline{\langle a,b\rangle} + \langle b,b\rangle \\ &= \|a\|^2 + 2Re(\langle a,b\rangle) + \|b\|^2 \end{aligned} \tag{3}$$

Similarly,

$$\begin{aligned} \|a-b\|^2 &= \langle a-b, a-b\rangle \\ &= \langle a,a\rangle - \langle a,b\rangle - \langle b,a\rangle + \langle b,b\rangle \\ &= \langle a,a\rangle - \langle a,b\rangle - \overline{\langle a,b\rangle} + \langle b,b\rangle \\ &= \|a\|^2 - 2Re(\langle a,b\rangle) + \|b\|^2 \end{aligned} \tag{4}$$

Adding up, we obtain ***parallelogram equality***:

$$\|a+b\|^2 + \|a-b\|^2 = 2\|a\|^2 + 2\|b\|^2$$

This equation holds for all Hilbert spaces including 2-d vector space over $\mathbb{R}$. Readers may find this form identical to the parallelogram equality in that vector space. It is also where the name comes from.

4.2.2 Nearest Point Property

In Euclidean geometry, choosing a point and a line we can find the minimal distance form the point to any point on the line, defined as the distance from the point to the line. This can be generalized, with line extending to a ***closed convex set*** and space becoming a Hilbert space.

Definition 4.2.5 (Convexity)
A set $S \subset \mathscr{H}$ is convex if $\forall f, g \in S, \forall t \in [0,1]$, we have $(tf + (1-t)g) \in S$.

Intuitively, this means that given any two points in the a convex set, the "segment" connecting the points stays inside the set. Hence a closed convex set is such a set with the property that every convergent sequence in the set converges to a point in the set.

Definition 4.2.6 (Set-point distance)
Let $\mathscr{H}$ be a Hilbert space, $S \in \mathscr{H}$ be a closed subset, $x \in \mathscr{H}$ be arbitrary vector. We define the **distance** from x to S to be the infimum of distance between x and element of S:

$$\text{dist}(S,x) \equiv \inf_{s\in S}(\|s-x\|)$$

Proposition 4.2.7 (Nearest Point Property)
Let $\mathscr{H}$ be a Hilbert space, $S \in \mathscr{H}$ be a closed subset, $x \in \mathscr{H}$ be arbitrary vector. Then there exists a **unique** $s_0 \in S$ such that

$$\|s_0 - x\| = \text{dist}(S,x)$$

Proof

The result is trivial when $x \in S$, we can choose $s_0 = x$ and $\text{dist}(S, x) = 0$.
When $x \notin S$, the idea of the proof is, by definition of distance especially the "infimum" we have a sequence of points in S whose distance with x converges to $\text{dist}(S, x) = 0$. Then from the convergence of distance we can show that this sequence is Cauchy, hence by completeness, convergent. Finally by cussedness we conclude that it converges to a point in S and convexity can be used to show uniqueness.
So we choose $x \notin S$ and a sequence $\{y_n\}_1^\infty$ in S such that $\|y_n - x\| \to 0$. And let $d = \text{dist}(S, x)$. Fix $\varepsilon_0 > 0$, we can find $N_0 \in \mathbb{N}$ such that for all $m, n > N_0$, we have

$$\|x - y_n\| < \frac{\varepsilon_0}{2} \quad \textit{and} \quad \|x - y_m\| < \frac{\varepsilon_0}{2}$$

Then for all $m, n > N_0$,

$$\begin{aligned} \|y_m - y_n\| &= \|(y_m - x) + (x - y_n)\| \\ &\leq \|y_m - x\| + \|y_n - x\| \\ &< \frac{\varepsilon_0}{2} + \frac{\varepsilon_0}{2} \\ &= \varepsilon_0 \end{aligned} \tag{5}$$

Thus we conclude that $\{y_n\}_1^\infty$ is Cauchy. By completeness of $\mathscr{H}$ and closedness of S we conclude that it converges to a point y in S.
It still remains to check uniqueness. We assume that exists y_1 and y_2 that satisfies the definition. Then by parallelogram equality:

$$\begin{aligned} \|(y_1 - x) - (y_2 - x)\|^2 &= \|y_1 - x\|^2 + \|y_2 - x\|^2 - \|y_1 + y_2 - 2x\|^2 \\ &= 2d^2 - \|y_1 + y_2 - 2x\|^2 \\ &= 2d^2 - 2\left\|\frac{y_1 + y_2}{2} - x\right\|^2 \end{aligned} \tag{6}$$

Now by convexity, $\frac{1}{2}(y_1 + y_2) \in S$. Thus

$$\left\|\frac{y_1 + y_2}{2} - x\right\| \geq d$$

and so

$$\|y_1 - y_2\|^2 = \|(y_1 - x) - (y_2 - x)\|^2 \leq 0$$

Hence $y_1 = y_2$, giving the uniqueness.

4.2.3 Projection Theorem

Definition 4.2.8 (Projection mapping)
Let $\mathscr{H}$ be a Hilbert space. A mapping $P : \mathscr{H} \to \mathscr{H}$ is a projection mapping if

$$P(Px) = Px, \forall x \in \mathscr{H}$$

Proposition 4.2.9 (Projection Theorem)
Let $\mathscr{H}$ be a Hilbert space and M a closed subspace. Then there exists unique pair of projection mapping $P : \mathscr{H} \to M$ and $Q : \mathscr{H} \to M^\perp$ satisfying $x = Px + Qx$ for any $x \in \mathscr{H}$, with the following property:

(1) $x \in M$ if and only if $Px = x$, $Qx = 0$

(2) $x \in M^{\perp}$ if and only if $Px = 0$, $Qx = x$

(3) $\|Px\|^2 + \|Qx\|^2 = \|x\|^2$

(4) P and Q are linear maps

(5) Px is the closest vector in M to x.

(6) Qx is the closest vector in $M^{\perp}$ to x.

proof:
We shall use nearest point property. Since M is a closed subspace, it is clearly a convex subset of $\mathscr{H}$. Thus for each $x \in \mathscr{H}$ there exists a unique point in M that is closest to x. So we define Px to be the unique nearest point in M for each x. Uniqueness of nearest point gives the uniqueness of mapping P. Q is then defined as $x - Px$. We should show that this definition of Q indeed gives a mapping from $\mathscr{H}$ to $M^{\perp}$.
Fix x, let $m \in M$ be such that $\|m\| = 1$. We must have, for all $a \in \mathbb{C}$:

$$\begin{aligned} \|Qx\|^2 &\leq \|Qx + am\|^2 \\ &= \|Qx\|^2 + |a|^2 \|m\|^2 + \langle Qx, m \rangle + \langle m, Qx \rangle \end{aligned} \tag{7}$$

Then we choose $a = -\langle Qx, m \rangle$ and simplify the equation, we have

$$\begin{aligned} 0 &\leq |a|^2 \|m\|^2 + \langle Qx, am \rangle + \langle am, Qx \rangle \\ &= |\langle Qx, m \rangle|^2 \|m\|^2 + \langle Qx, -\langle Qx, m \rangle m \rangle + \langle -\langle Qx, m \rangle m, Qx \rangle \\ &= |\langle Qx, m \rangle|^2 - \overline{\langle Qx, m \rangle} \langle Qx, m \rangle - \overline{\langle m, Qx \rangle} \langle m, Qx \rangle \\ &= -|\langle Qx, m \rangle|^2 \end{aligned} \tag{8}$$

Thus we must have $\langle Qx, m \rangle = 0$, so $Qx \in M^{\perp}$.
We proceed to prove (1) and (2).
First $x \in M$. We have $Qx \in M^{\perp}$. Note that $x \in M$ and $Px \in M$ we have $Qx = x - Px \in M$. Hence $Qx \in (M \cap M^{\perp}) = 0$. Thus $Qx = 0$ and $Px = x - 0 = x$.
To see other direction, we simply notice that $x = Px \in M$ by definition.
Proof of (2) is similar.
Now let's prove (3). We shall use the fact that $Px \perp Qx$, giving $\langle Px, Qx \rangle = 0$

$$\begin{aligned} \|x\|^2 &= \langle x, x \rangle \\ &= \langle Px + Qx, Px + Qx \rangle \\ &= \langle Px, Px \rangle + \langle Qx, Px \rangle + \langle Px, Qx \rangle + \langle Qx, Qx \rangle \\ &= \langle Px, Px \rangle + \langle Qx, Qx \rangle \\ &= \|Px\|^2 + \|Qx\|^2 \end{aligned} \tag{9}$$

Proof of (4) is a routine work verifying linearity. Left as an exercise.
Now let's prove (5) and (6).
(3) is given by the construction of P. To see (4), consider $y \in M^{\perp}$, we have:

$$\|x - y\| = \|Px + Qx - y\| = \|Px\| + \|Qx - y\| \geq \|Px\|$$

So minimal distance is $\|Px\|$, Obtained at $y = Qx$

5 Operator and Functional

5.1 Linear functionals

Linear functionals are a special class of linear mapping from a vector space to its scalar field.

Definition 5.1.1 (Linear mapping)
Let X and Y be vector spaces over the same scalar field. A mapping $\Lambda : X \to Y$ is called a **linear mapping** if

$$\Lambda(\alpha x + \beta y) = \alpha \Lambda x + \beta \Lambda y$$

for all x, y vectors and α, β scalars.
When Y is the scalar field, Λ is called a **linear functional**. In the following text, we assume $Y = \mathbb{C}$ unless mentioned otherwise.

Definition 5.1.2 (Image, preimage, kernel)
Let T be linear mapping :$T : X \to Y$, let $M \subset X$, $N \subset Y$.
Then the **Image** of X is

$$T(M) \equiv \{y \in Y : y = T(x), x \in M\} \subset Y$$

The **Preimage** of Y is

$$T^{-1}(Y) \equiv \{x \in X : y = T(x), y \in N\} \subset X$$

Specially, **Kernel** of T is defined to be the preimage of $\{0_Y\}$,a set containing only the null-element of Y (0_Y is a scalar 0 when Y is the scalar field:

$$\text{Ker}(Y) \equiv T^{-1}(\{0_Y\}) = \{x \in X : T(x) = 0_Y, y \in N\} \subset X$$

5.1.1 Boundedness and continuity

In this section we shall see a very nice result about linear functional, which links their boundedness and continuity

Definition 5.1.3 (Boundedness)
Let X be a nomred vector space. $\Lambda : X \to \mathbb{C}$ a linear functional. Then Λ is **bounded** if there exists a constant M, such that for all $x \in X$, we have

$$\|\Lambda(x)\| \leq M \|x\|$$

Definition 5.1.4 (Functional norm)
For a bounded linear functional $\Lambda : X \to \mathbb{C}$, we define its norm as follows:

$$\|\Lambda\| = \sup_{x \in X} \{\|\Lambda(x)\|, \|x\| \leq 1\}$$

Proposition 5.1.5 (An inequality of no name)
Let X be a nomred vector space. $\Lambda : X \to \mathbb{C}$ a bounded linear functional. Then

$$\|\Lambda(x)\| \leq \|\Lambda\| \|x\|$$

Proof

Let $e \in X$ be such that $\|e\| = 1$. Then:

$$\begin{aligned}\|\Lambda\| \|e\| &= \sup_{x\in X} \{\|\Lambda(x)\|, \|x\| \leq 1\} \|e\| \\ &\leq \|\Lambda(e)\| \|e\| \\ &= \|\Lambda(e)\|\end{aligned} \tag{10}$$

Then for arbitrary non-zero vector $x \in X^*$ we write $x = \|x\| \frac{x}{\|x\|}$ and exploit linearity to finish the proof.

Definition for continuous linear functional is similar to the that of continuous function in real-analysis, changing absolute value to corresponding norms. It is a routine exercise to check that this indeed defines a norm.

Definition 5.1.6 (Continuous linear functional)
Let $\Lambda : X \to \mathbb{C}$ be a linear functional. It is continuous at $y \in X$ if for all $\varepsilon > 0$, there exists $\delta > 0$ such that whenever $x \in \{x \in X, \|x - y\| < \delta\}$, we have $\|\Lambda(x) - \Lambda(y)\| < \varepsilon$. If Λ is continuous at all $y \in X$, then it is called a **continuous linear map**.

Proposition 5.1.7 (Continuity $\iff$ Continuity at 0)
Let $\Lambda : X \to \mathbb{C}$ be a linear functional. Then λ is continuous if and only if it is continuous at 0. Proof of this proposition is trivial, left as an exercise.

Proposition 5.1.8 (Continuity $\iff$ Boundedness)
Let $\Lambda : X \to \mathbb{C}$ be a linear functional. Then λ is continuous if and only if it is bounded.

Proof : Part I: Continuity implies Boundedness

($\Longrightarrow$)We first show that continuity implies boundedness. We know Λ is continuous at 0, so fix $\varepsilon > 0$, we may find $\delta_0 > 0$ such that $\|x\| < \delta_0 \Longrightarrow \|\Lambda(x)\| < \varepsilon$. So for any $y \in X$ we have that:

$$\begin{aligned}\Lambda(y) &= \Lambda\left(\frac{2\|y\|}{\delta_0}\left(\frac{\delta_0}{2}\frac{y}{\|y\|}\right)\right) \\ &= \frac{2\|y\|}{\delta_0}\Lambda\left(\frac{\delta_0}{2}\frac{y}{\|y\|}\right) \\ &\leq \frac{2\|y\|}{\delta_0}\varepsilon = \frac{2\varepsilon}{\delta_0}\|y\|\end{aligned} \tag{11}$$

Hence $M = 2\varepsilon/\delta_0$ gives a bound of the functional.

Proof : Part II:Boundedness implies Continuity

($\Longleftarrow$)Now we show that boundedness implies continuity. By last proposition we may reduce this to showing that boundedness implies continuity at 0. Let K be such that

$$\Lambda(x) \leq K \|x\|, \quad \forall x \in X$$

Now fix $\varepsilon > 0$, let $\delta = \frac{\varepsilon}{2K}$. Then when $x < \delta$, we have

$$\begin{aligned} \Lambda(x) &\leq K \|x\| \\ &< K\delta \\ &= K\frac{\varepsilon}{2K} \\ &= \frac{\varepsilon}{2} < \varepsilon \end{aligned} \tag{12}$$

Hence Λ is continuous at 0, which finishes the proof.

Proposition 5.1.9 (Sequential continuity) Let $\Lambda : f \to Y$ be a mapping, X is a metric space and Y is topological space. Then it is continuous if and only if, for any sequence $\{x_n\}$ in X such that $x_n \longrightarrow x_0 \in X$, we have $f(x) \longrightarrow f(x_0) \in Y$

5.2 Basic Operator Theory

5.2.1 Bounded linear operator

Definition 5.2.1 (Linear operator)
Let X, Y be normed vector spaces. A mapping $\Gamma : X \to Y$ is a linear operator if

$$\Gamma(k_1 x_1 + k_2 x_2) = k_1 \Gamma(x_1) + k_2 \Gamma(x_2)$$

for all $x_1, x_2 \in X$ and k_1, k_2 scalars.

Definition 5.2.2 (Bounded linear operator)
Let X, Y be normed vector spaces. Linear operator $\Gamma : X \to Y$ is bounded if there is a finite constant $C > 0$ such that

$$\|\Gamma(x)\|_Y \leq \|x\|_X$$

holds for all $x \in X$.

Definition 5.2.3 (Operator norm)
Let X, Y be normed vector spaces, $\Gamma : X \to Y$ a bounded linear operator, then

$$\|\Gamma\| \equiv \sup_{x \in X} \{\|\Gamma(x)\|_Y, \|x\|_X \leq 1\}$$

The definition is very similar to that for functionals.

Proposition 5.2.4 (Boundedness and continuity)
Let X, Y be normed vector spaces, $\Gamma : X \to Y$ a linear operator. Then the following three are equivalent:

- Γ is bounded
- Γ is continuous
- Γ is continuous at 0
- **Also:** Γ is Lipschitz, i.e.$\exists C > 0$ with $\|Aa - Ab\|_Y \leq C \|a - b\|_X$, $\forall a, b \in X$
- **Also:** Γ is continuous at any $x \in X$

Definition 5.2.5 ($\mathscr{B}$)
We define $\mathscr{B}(X, Y)$ to be the collection of all bounded linear operators from X to Y. We also write $\mathscr{B}(X, X)$ as $\mathscr{B}(X)$

Proposition 5.2.6
$\mathscr{B}(X, Y)$ is normed linear space in operator norm. Specially if Y is Banach then $\mathscr{B}(X, Y)$ is also Banach.

Theorem 5.2.7
Every finite dimensional linear operator is bounded.

Theorem 5.2.8
A bounded linear operator attains its inf and sup on a compact set.

5.3 Duality

5.3.1 Dual space: A nice self-symmetry

One may notice, that linear maps also form a vector space: multiple of a linear map are also linear, sum of linear maps are also linear. We'll formalize this idea to the concept of dual space.

Definition 5.3.1 (Dual Space)
Let X be a Banach space. Define its dual space X^* as follows:

$$X^* = \{T : T \text{ is a bounded linear functional on } X\}$$

Proposition 5.3.2 (Dual space of Banach Space)
Dual space of a Banach space is also Banach, under functional norm.

Proof

Here we assume scalar field to be $\mathbb{C}$.
First,to check that X^* is a vector space, it suffices to show that multiple and sum of bounded linear functional remains to be linear and bounded. This part of proof is trivial.
Second, we shall check that X^* is complete under functional norm. To show this, we let $\{T_n\}$ to be a Cauchy sequence in X^*, which means given $\varepsilon > 0$, there exists a positive constant N such that for all $a, b > N$, we have

$$\|T_a - T_b\| < \varepsilon$$

So given any point $x \in X$, we have that

$$\begin{aligned} \|T_a(x) - T_b(x)\| &= \|(T_a - T_b)x\| \\ &\leq \|T_a - T_b\| \, \|x\| \\ &< \varepsilon \, \|x\| \end{aligned} \tag{13}$$

This implies that $\{T_n(x)\}$ is a Cauchy sequence on $\mathbb{C}$. By completeness of $\mathbb{C}$, it converges to a point on $\mathbb{C}$. Note that this works for arbitrary $x \in X$, we can define T to be the pointwise limit in following form: $T(x) \equiv \lim_{n\to\infty} T_n(x), \quad \forall x \in X$ We can verify that T is linear:

$$T(ax) = \lim_{n\to\infty} T_n(ax) = \lim_{n\to\infty} aT_n(x) = a \lim_{n\to\infty} T_n(x) = aT(x) \tag{14}$$

and

$$\begin{aligned} T(x) + T(y) &= \lim_{n\to\infty} T_n(x) + \lim_{n\to\infty} T_n(y) \\ &= \lim_{n\to\infty} T_n(x) + T_n(y) \\ &= \lim_{n\to\infty} T_n(x + y) \\ &= T(x + y) \end{aligned} \tag{15}$$

It remains to show that T is bounded. Fix $r > 0$, we can find $a \in \mathbb{N}$ such that $\|T - T_a\| < r$,

$$\|T\| = \|T - T_a + T_a\| \leq \|T - T_a\| + \|T_a\| = r + \|T_a\| \tag{16}$$

Hence T is also bounded, so $T \in X^*$, which finishes the proof for completeness.

We have shown that dual space of a Banach space is Banach, what about dual space of a Hilbert space? In the following result, we'll see that the inner product of a Hilbert space allows great elegance in structure of its dual space, which is identity in the sense of isomorphism.

5.3.2 Riesz representation Theorem

Theorem 5.3.3 (Riesz representation Theorem) A bounded linear functional T on Hilbert space $\mathscr{H}$ is uniquely associated with a vector $h_0 \in \mathscr{H}$ in the sense that

$$T(h) = \langle h, h_0 \rangle, \quad \forall h \in \mathscr{H} \quad \text{and } \|T\| = \|h\|$$

Proof

If $T = 0$ we simply choose $h_0 = 0$.
When $T \neq 0$,let $S = Ker(T)$, pick a non-zero vector $w \in S^{\perp}$, without loss of generality, we may assume $T(w) = 1$. Then, for any $x \in \mathscr{H}$, we observe for vector $(T(x)w - x)$ that

$$T\left(T(x)w - x\right) = T(w)T(x) - T(x) = T(x) - T(x) = 0$$

Therefore $(T(x)w - x) \in S$, which means $(T(x)w - x) \perp w$. Hence

$$\langle (T(x)w - x), w \rangle = 0, \quad \forall w \in \mathscr{H}$$

By (left) linearity of inner product,

$$\langle x, w \rangle = T(x) \langle w, w \rangle = T(x) \|w\|^2$$

Thus

$$T(x) = \frac{\langle x, w \rangle}{\|w\|} = \left\langle x, \frac{w}{\|w\|^2} \right\rangle$$

So $h = w / \|w\|^2$ is the desired vector. Also,

$$\begin{aligned} \|T\| &= \sup_{x \in \mathscr{H}} \{\|T(x)\|, \|x\| \leq 1\} \\ &= \sup_{x \in \mathscr{H}} \{\|\langle x, h \rangle\|, \|x\| \leq 1\} \\ &= \left\| \left\langle \frac{h}{\|h\|}, h \right\rangle \right\| \\ &= \|h\| \end{aligned} \tag{17}$$

The result shows that $\mathscr{H} = \mathscr{H}^*$ in the sense that the ma from h to corresponding linear functional is an isometry: $\|h\| = \|T\|$

Remark 5.3.4 (Alternative proof of Riesz Representation theorem)
The proof above is purely algebraic, in the sense that no analysis is used. However, one may realise that the proof relies heavily on the construction of $T(x)w - x$, which may not be that easy to thought of. Hence another completely different proof is presented here.

Proof : Riesz Representation theorem

The idea of the proof is to construct a sequence of points that finally converges to the desired h_0.
If $\|T\| = 0$, we simply choose $h_0 = 0$.
If not, without loss of generality, let's consider operators with norm equal to 1.We first show existence of such h_0. Let $T \in \mathscr{H}^*$ be such that $\|T\| = 1$.
Now by definition of operator norm, specializes to 1-Euclidean norm on $\mathbb{R}$:

$$\|T\| = \sup_{x \in \mathscr{H}} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\|$$

The existence of **supremum** in the definition guarantees that we can find a sequence of points $(x_n)_0^\infty$ in $\{x \in \mathscr{H} : \|x\| = 1\}$ that gives

$$\lim_{n\to\infty} T(x_n) = 1$$

The norm can be removed here as we're discussing functionals $T : \mathscr{H} \to \mathbb{R}$, and for each $y \in \mathscr{H}$ if $T(y) < 0$ there is $(-y) \in \mathscr{H}$ and $T(-y) = -T(y) > 0$ by linearity.
We claim that $(x_n)_1^\infty$ is Cauchy. To prove this, we first notice parallelogram equality in Hilbert space:

$$2\|a\|^2 + 2\|b\|^2 = \|a+b\|^2 + \|a-b\|^2$$

we have that

$$\begin{aligned}
\|x_m - x_n\|^2 &= 2\|x_m\|^2 + 2\|x_n\|^2 - \|x_m + x_n\|^2 \\
&= 2 \times 1 + 2 \times 1 - \|x_m + x_n\|^2 \\
&= 4 - \|x_m + x_n\|^2 \\
&= 4 - \|T\|^2 \|x_m + x_n\|^2 \\
&\leq 4 - \|Tx_m + Tx_n\|^2 \\
&=\leq 4 - (Tx_m + Tx_n)^2
\end{aligned} \tag{18}$$

Taking n and m to infinity we have that $\|x_m - x_n\| \to 0$, which means that $(x_n)_1^\infty$ is Cauchy.

Theorem 5.3.5 (Dual space of Hilbert space)
Let $\mathscr{H}$ be a Hilbert space, then its dual space $\mathscr{H}^*$ is isomorphic to itself: $\mathscr{H} \cong \mathscr{H}^*$, moreover, the natural norm on them is an isometry, which means that the operator norm of a linear functional equals the element that generates this functional by inner product. This is an immediate corollary from results of Riesz representation theory.

5.3.3 Dual space of lp space

In this section, we consider sequence spaces, l_p space. Specially, when $p = 2$, the space is Hilbert. In other cases, when $p \neq 2$, what happens? One may realise that $q = p = 2$ is precisely a solution to $1/p + 1/q = 1$, and $p = q$ gives the self-duality. We will investigate this intuition here and prove that the dual space of l_p is precisely l_q, where $1/p + 1/q = 1$,$p, q \in \mathbb{R}$. One should pay attention here, that $p = \infty$ is slightly different, the result may not hold in given $p = \infty$.

Theorem 5.3.6 (Dual of ℓ^p)
Let $p \in (1, \infty)$, then $(\ell^p)^* \cong \ell^q$, where $1/p + 1/q = 1$.
proof:
The proof has two parts. First we show that every element $y \in \ell^q$ uniquely defines a linear linear functional $\Lambda_y : X \to \mathbb{R}$, then we prove surjectivity, which means that every functional $l \in (\ell^p)^*$ can

be uniquely represented by an element in ℓ^q. Equation $1/p + 1/q = 1$ and its variation like $p + q = pq$ and $1 + q/p = q$ will be useful at some steps.

Proof : STEP I

Let $y = (y_n)_1^\infty \in \ell^q$,,define $\Lambda_y : X \to \mathbb{R}$, by

$$\Lambda_y(x) = \sum_{k=1}^{\infty} x_n y_n, \ \forall x = (x_n)_1^\infty \in \ell^p$$

We claim that:

- $\Lambda_y \in (\ell^p)^*$, which means that Λ_y is a bounded linear functional.
- $\|\Lambda_y\| \equiv \sup\{|\Lambda_y(x)| : \|x\|_p = 1\} = \|y\|_q$, i.e. the map $y \to \Lambda_y$ is an isometry.

Linearity immediately follows from definition. To check boundedness, we notice that

$$|\Lambda_y(x)| = \left|\sum_{k=1}^{\infty} x_n y_n\right| \leq \sum_{k=1}^{\infty} |x_n y_n| \leq \|xy\|_1 \quad \leq \|x\|_p \|y\|_q \tag{19}$$

Now, plugging in any $x \in \ell^p$ with norm 1 we get boundedness, and to see isometry we need to find a proper x in ℓ^p that saturates the supremum.
The candidate here is $x = (x_n)_1^\infty$, with $x_n = sign(y_n) \cdot (y_n)^{q-1}$. We need to check followings:

- $x \in \ell^p$
- $|\Lambda_y(x)| = \|y\|_q \|x\|_p$

First, we compute the p-norm of x:

$$\begin{aligned} \|x\|_p &= \left(\sum_{k=1}^{\infty} |x_n|^p\right)^{1/p} = \left(\sum_{k=1}^{\infty} |sign(y_n) \cdot (y_n)^{q-1}|^p\right)^{1/p} \\ &= \left(\sum_{k=1}^{\infty} |(y_n)^{q-1}|^p\right)^{1/p} = \left(\sum_{k=1}^{\infty} |(y_n)|^q\right)^{1/p} \\ &= \left(\sum_{k=1}^{\infty} |(y_n)|^q\right)^{(1/q)\cdot(q/p)} = \|y\|_q^{q/p} < \infty \end{aligned} \tag{20}$$

This gives that $x \in \ell^p$. Then

$$\begin{aligned} \Lambda_y(x) &= \sum_{k=1}^{\infty} x_n y_n \\ &= \sum_{k=1}^{\infty} (sign(y_n) \cdot (y_n)^{q-1} \cdot y_n) \\ &= \sum_{k=1}^{\infty} |y_n|^q \\ &= \|y\|_q^q \end{aligned} \tag{21}$$

Thus we have that

$$\begin{aligned} \frac{|\Lambda_y(x)|}{\|x\|} &= \frac{|\,\|y\|_q^q\,|}{\|y\|_q^{q/p}} \\ &= \|y\|_q^{q-q/p} \\ &= \|y\|_q \end{aligned} \tag{22}$$

Which shows isometry.

Proof : STEP II

Now let us assume $l \in (\ell^p)^*$. We define natural basis of an ℓ^q space: $e_n = ((e_n)_k)_1^\infty$, where $(e_n)_k = 1$ if $k = n$, otherwise $(e_n)_k = 0$. Each e_n is a sequence with all entries zero, except at $k - th$ entry the value is 1. Now consider a sequence of $y_n = \sum_{k=1}^n l(e_k) \cdot e_k$,i.e. y_n is a sequence where $(y_n)_k = l(e_k)$ for $k \le n$. By denoting x^n to be the sequence with first n-terms equals to corresponding term of x and any term afterwards set to zero, we claim the following three things:

- $y_n \in \ell^q$ for all $n \in \mathbb{N}$.
- $l(x^{(n)}) = \sum_{k=1}^n x_k (y_n)_k$
- $\lim_{n\to\infty} \|l(x) - xy_n\| = 0$

First, note that each y_n is a finite sum, its q-norm is finite, hence $y_n \in \ell^q$ for all $n \in \mathbb{N}$. Then we notice that we can express $x^(n)$ as a finite summation:

$$x^{(n)} = \sum_{k=1}^n x_k e_k \implies l(x^{(n)}) = \sum_{k=1}^n l(x_k e_k) = \sum_{k=1}^n x_k \cdot l(e_k)$$

Then the second statement is clear.

$$\sum_{k=1}^n x_k (y_n)_k = \sum_{k=1}^n x_k \cdot l(e_k) = l(x^{(n)})$$

For the last statement, we use separability of any ℓ^p space and express x uniquely by canonical Schauder basis e_k: $x = \sum_{k=1}^\infty x_k e_k$. Thus for any $\varepsilon > 0$, an index $s \in \mathbb{N}$ such that

$$\left\| y_n x - y_n x^{(n)} \right\| \le \left\| x - x^{(n)} \right\| \cdot \|y_n\| < \varepsilon/2$$

Then, by continuity (boundedness) of l, we can choose another index $t \in \mathbb{N}$ such that

$$\left\| l(x) - l(x^{(n)}) \right\| < \varepsilon/2$$

Then

$$\begin{aligned} \|l(x) - y_n x\| &= \left\| l(x) - l(x^{(n)}) + y_n x^{(n)} - y_n x \right\| \\ &\le \left\| l(x) - l(x^{(n)}) \right\| + \left\| y_n x^{(n)} - y_n x \right\| \\ &< 2\varepsilon/2 = \varepsilon \end{aligned} \tag{23}$$

By letting $n \to \infty$ we have that that $x \lim_{n\to\infty} y_n = l(x)$ for all $x \in \ell^p$. Thus we can define $y = \lim_{n\to\infty} y_n$. y is an element of ℓ^q since $(y_n)_1^\infty$ is Cauchy, since the different of any y_m and y_n is the partial sum from m-th term to n-th term of a sequence that converges in q-norm, and thus can be arbitrarily small when n and m goes to infinity. Existence and uniqueness of y is then follows from completeness of ℓ^q. This shows that every linear functional in $(\ell^p)^*$ uniquely gives an element in ℓ^q, which proves that Λ_y is a surjection, hence we have an isomorphism as stated: $(\ell^p)^* \cong \ell^q$.

Theorem 5.3.7 (results on l0 and l inf)

NOOOOOOT COMPLEEEEEEET! COOOOOOOM BAAAAAACK!!!

5.3.4 Dual operator

Definition 5.3.8 (dual operator)
Let X, Y be Banach spaces, with norm $\|\cdot\|_X$ and $\|\cdot\|_Y$. Let A be an operator from X to Y : $A : X \to Y$. Then dual operator of A, $A^* : Y^* \to X^*$ is defined as follows:

$$A^* y^* = y^* A : x \to \mathbb{R}, \quad \forall y^* \in Y^*$$

Adopting the notation of bracket, which means to write a linear functional l acting on an element x: $l(x)$ to be $\langle l, x\rangle$, dual operator can be rewritten in this way:

$$\langle A^* y^*, x\rangle = \langle y^*, Ax\rangle$$

Remark 5.3.9
This bracket here does not stand for inner product in Banach space. However it's used in a way such that this notation coincides with the adjoint under Hilbert space. This means that under Hilbert space context, one can think of this as an inner product, as every linear functional can be uniquely represented by an element in Hilbert space by Riesz representation theorem.

6 Hahn-Banach Theorem

One of the most import results in functional analysis, **Hahn-Banach theorem** is a theorem dealing with extending linear maps from a subspace to the whole space. The theorem says that any bounded linear functional defined on a subspace can be extended to the whole space, while preserving the norm. The result does not rely on completeness of the space, so it's a result for all normed linear spaces. The proof of this theorem involves using a version of **axiom of choice (AC)** , Zorn's lemma. We shall review this lemma first.

Definition 6.0.1 (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 6.0.2 (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 $y \leq x \ \forall y \in Y$.

Definition 6.0.3 (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 $m = x$.

Lemma 6.0.4 (Zorn's lemma)
If X is a nonempty partially ordered set with the property that every totally ordered subset of X has an upper bound in X, then X has a maximal element.

Theorem 6.0.5 (Hahn-Banach Theorem)
Let X be a normed vector space over $\mathbb{F}$ ($\mathbb{C}$ or $\mathbb{R}$), Y is a proper subspace of X. If $T_0 : Y \to \mathbb{F}$ is a bounded linear functional, then there exists a bounded linear functional $T : X \to \mathbb{F}$ satisfying:

- $T(y) = T_0(y)$ for all $y \in Y$
- $\|T\| = \|T_0\|$

To prove the theorem, the idea is first to show that we can extend linear functional by one dimension, with induction to show that extension can be done to "arbitrarily high dimension". Then by using Zorn's lemma we show that such extension "reaches" every dimension of the space. We first provide real version of the theorem.

Lemma 6.0.6 (one-dimensional extension)
Let X be a normed vector space over $\mathbb{F}$ ($\mathbb{C}$ or $\mathbb{R}$), Y_n is a proper subspace of X. Let $v \in X \backslash Y$, $X_{n+1} = \{x + hv : x \in X_n,\ h \in \mathbb{C}\}$. If $T_n : Y \to \mathbb{F}$ is a bounded linear functional, then there exists a bounded linear functional $T_{n+1} : X_{n+1} \to \mathbb{F}$ satisfying:

- $T_{n+1}(x) = T_n(x)$ for all $x \in X_n$
- $\|T_{n+1}\| = \|T_n\|$

Proof : One-dimensional extension

Define linear functional $P : X_{n+1} \to \mathbb{R}$ by

$$P(x + kv) = T_n(x) - Ck, \, \forall\, x \in X_n, \, k \in \mathbb{R}$$

where C is a constant to be determined. First we shall check linearity, which is left as an exercise. Then we shall show that we can find a proper constant C so that $\|P\| = \|T_n\|$. Note that $X_n \subset X_{n+1}$, so we have

$$\begin{aligned} \|P\| &= \sup_{x \in X_{n+1}} (\{|Px| : \|x\| = 1\}) \\ &\geq \sup_{x \in X_n} (\{|Px| : \|x\| = 1\}) \\ &= \sup_{x \in X_n} (\{|T_n x| : \|x\| = 1\}) \\ &= 1 \end{aligned} \tag{24}$$

So by choosing C such that $P(x + kv) \leq \|x + kv\|$ for any $x \in X_n$ and $k \in \mathbb{R}$, we will have that $\|P\| \leq 1$, giving $\|P\| = 1$. Thus it remains to show that we can find such a constant C.
We aim to find C such that

$$|P(x + kv)| = |T_n(x) - Ck| \leq \|x + kv\| \, , \, \forall x \in X_n, \, \forall k \in \mathbb{R}$$

Hence,

$$T_n(x) - \|x + kv\| \leq Ck \leq T_n(x) + \|x + kv\| \, , \, \forall x \in X_n, \, \forall k \in \mathbb{R}$$

Note that for all $x, y \in X_n$ we have:

$$\begin{aligned} T_n x - T_n y &= T_n(x - y) \\ &\leq \|x - y\| \\ &= \|(x + kv) - (kv + y)\| \\ &\leq \|x + kv\| + \|y + kv\| \end{aligned} \tag{25}$$

Thus

$$l^- = \sup_{x \in X_n, k \in \mathbb{R}} (T_n(x) - \|x + kv\|) \leq \inf_{x \in X_n, k \in \mathbb{R}} (T_n(x) + \|x + kv\|) = l^+$$

Hence we can always find a C such that

$$T_n(x) - \|x + kv\| \leq l^- \leq Ck \leq l^+ \leq T_n(x) + \|x + kv\| \, , \, \forall x \in X_n, \, \forall k \in \mathbb{R}$$

Which finishes the proof.

6.1 Proof of Hahn-Banach theorem,real case

Starting from $T_0 : Y \to \mathbb{R}$, by 6.0.6 we can define T_{n+1} to be the one-dimensional extension of T_n for any $n \in \mathbb{N}$, with domain Y_{n+1} extended from Y_n, for convenience we let$Y_0 = Y$. Then consider the set

$$M = \{(T_n, Y_n), n \in \mathbb{N}\}$$

which can be partially ordered by $\leq$ defined as

$$(T_a, Y_a) \leq (T_b, Y_b) \text{ if } Y_a \subset Y_b, \text{ and } T_b = T_a \text{ on } Y_a$$

Now let $S = \{(T_i, Y_i), i \in I\}$ (where I is the index set) be a totally ordered subset of M. Consider $Y' = \cup_{i \in I} Y_i$ with $T'(x) = T_i(x)$ if $x \in Y_i$, we have that $(T', Y') \in M$ is an upper bound of S. By Zorn's lemma, we know that M has a maximal element, denoted as (T_∞, Y_∞). We claim that $Y_\infty = X$, because if not, we can do one-dimensional extension to Y_∞, resulting in $X \subset Y_\infty + 1$, contradicting with maximality. Thus we have $Y_\infty = X$, and T_∞ is the desired extension to X.

6.2 Proof of Hahn-Banach theorem,complex case

To prove the statement for complex case, we shall exploit a connection between real valued functional and complex one.

Proposition 6.2.1
Let $T : X \to \mathbb{C}$ be a complex linear functional. Define $u(x) = Re(T(x))$ for all $x \in X$. Then

- $u(x)$ is a real-valued linear functional
- $T(x) = u(x) - iu(ix)$
- $\|u\| = \|T\|$

Moreover, given any linear functional $u(x)$, $T(x) = u(x) - iu(ix)$ defines a complex linear functional

Proof

The first two are very easy to show. The hardest part is on the third statement. We first show that $\|T\| \geq \|u\|$:

$$\begin{aligned} \|T\|^2 &= \sup\{|Tx|^2, \|x\| = 1\} \\ &= \sup\{|u(x) - iu(ix)|^2, \|x\| = 1\} \\ &= \sup\{[u(x)]^2 + [u(ix)]^2, \|x\| = 1\} \\ &\geq \sup\{[u(x)]^2, \|x\| = 1\} \\ &= \|u\|^2 \end{aligned} \tag{26}$$

On the other hand, pick any $x \in X$ with $\|x\| = 1$, denote $T(x) = re^{i\theta}$, we have

$$|T(x)| = |e^{-i\theta}||T(x)| = |T(e^{-i\theta}x)| = |u(e^{-i\theta}x) - iu(e^{-i\theta}x)|$$

$$\begin{aligned} |T(x)| &= |e^{-i\theta}||T(x)| \\ &= |T(e^{-i\theta}x)| \\ &= |u(e^{-i\theta}x) - iu(e^{-i\theta}x)| \end{aligned} \tag{27}$$

But we have that $T(e^{-i\theta}x) = r \in \mathbb{R}$, thus

$$|T(e^{-i\theta}x)| = |Re(T(e^{-i\theta}x))| = |u(e^{-i\theta}x)| \leq \|u\|$$

Hence $\|T\| = \|u\|$.

The remaining part of the proof is simply combining last result and the proof of real case. However, last result also gives a insight on complex bounded linear functional

6.3 Hahn-Banach with sublinear function

Previous discussion of Hahn Banach theorem considers extension preserving norm. However, we can consider extension in a more general setting, namely sublinear.

Definition 6.3.1 (sublinear function)
Let X be a vector space. $p : X \to \mathbb{R}$ is a sublinear function if the following holdsL

- $p(kx) = kp(x)$, for all $x \in X$ and $k \geq 0$
- $p(x+y) = p(x) + p(y), \ \forall x, y \in X$

Definition 6.3.2 (Restriction)
Let $f : X \to Y$. Let $S \subset X$, if $g : M \to Y$ is such that $g(x) = f(x), \ \forall x \in M$ then we say f restricted to M is g, denoted by $g = f|_M$

Theorem 6.3.3 (Hahn-Banach Theorem, sublinear)
Let $M \in X$ be a linear subspace of X, where X is a linear space. $p : X \to \mathbb{R}$ is sublinear, and $f : M \to \mathbb{R}$ is a linear map such that $f(x) \leq p(x), \ \forall x \in M$.
Then, there exists a linear map $F : X \to \mathbb{R}$ with $F|_M = f$ and $F(x) \leq p(x), \ \forall x \in M$.

Proof : Hahn-Banach Theorem, sublinear

NOOOOOOT COMPLEEEEEEET! COOOOOOOM BAAAAAACK!!!

6.4 Results of Hahn-Banach

Theorem 6.4.1 (Dual functional)
Let X be a normed linear space. $\forall x \in X, \exists x^* \in X^*$ s.t. $\langle x^*, x\rangle \equiv x^*(x) = \|x\|_X^2 = \|x^*\|_{X^*}^2$
Note that here $\langle \cdot, \cdot \rangle$ does not necessarily stand for inner product.

Proof : Dual functional

Let $M = span(x)$. Define $f(tx) = t\,\|x\|_X^2\,, \ \forall t \in \mathbb{R}$. Clearly f is linear, and that $\|f\|_{M^*} = \|x\|_X$. Then we apply Hahn-Banach theorem to extend f to $x^* = F \in X^*$, with $\|x^*\|_{X^*} = \|f\|_M = \|x\|_X$

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.

Proposition 6.4.2 (Point-point separation)
Let x and y be points in X NOOOOOOT COMPLEEEEEEET! COOOOOOOM BAAAAAACK!!! (Question here is do we need X to be normed? or even Banach?) There exists $l \in X^*$ such that $l(x) \neq l(y)$.

Proof

Choose $l \in X^*$ to be the dual functional of $(y - x) \in X$.
Then $l(y - x) = \|y - x\|_X^2 > 0 \implies l(y) \neq l(x)$.

Proposition 6.4.3 (separation from proper closed subspaces)
$M \subset X$ is a linear space, closed and (naturally) convex, Assume $x_0 \notin M$ such that $d \equiv dist(x_0, M)$, then $\exists \ell \in X^*$ with $\ell|_M = 0$ and $\|\ell\|_X^* = 1$ and $\ell(x_0) = d$.

Proof

Let $M_0 = \{xt + x_0 : x \in M\}$. Define $f : M_0 \to \mathbb{R}$ which is linear, $f(x + tx_0) = td$. Then NOOOOOOT COMPLEEEEEEET! COOOOOOOM BAAAAAACK!!!

7 Uniform Boundedness principle

Uniform boundedness principle is sometimes called Banach–Steinhaus theorem. In its basic form, it asserts that for a family of bounded linear operators whose domain is a Banach space, pointwise boundedness is equivalent to uniform boundedness in operator norm.

Theorem 7.0.1 (Uniform Boundedness principle)
Let X be a Banach space, Y a normed vector space. Let F be a collection of bounded linear operators from X to Y. Then if

$$\sup_{f\in F}\|fx\| < \infty, , \forall x \in X$$

Then

$$\sup_{f\in F}\|f\| < \infty$$

To prove this theorem we shall use **Baire category theorem**.

Definition 7.0.2 (Nowhere dense)
A set S in metric space X is **nowhere dense** if its closure has empty interior. i.e. $\overline{S}^{\circ} = \emptyset$.

Theorem 7.0.3 (Baire Category Theorem)
A complete metric space is not countable union of nowhere dense sets.

Proof : Baire Category Theorem

The idea of the proof is to construct a Cauchy sequence in the space with no limit point, giving contradiction. First, let M be a complete metric space. Assume

$$M = \bigcup_{n=1}^{\infty} A_n$$

We know that A_1 is nowhere dense, which means $\overline{A_1}^{\circ} = \emptyset$. Since $\overline{A_1}$ is closed, $M\backslash\overline{A_1}$ is open, we may find open ball B_1 with radius $r_1 < 1$ such that $B_1 \cap \overline{A_1} = \emptyset$. Clearly we have that $B_1 \not\subset \overline{A_2}$ otherwise $\overline{A_2}$ has non-empty interior. So we have that $(M\backslash\overline{A_2}) \cap B_1$ is open and non-empty. Now we may choose open ball $B_2 \subset ((M\backslash\overline{A_2}) \cap B_1) \subset B_1$ with radius $r_2 < 1/2$. We repeat this process, so that $B_n \subset ((M\backslash\overline{A_n}) \cap B_{n-1}) \subset B_{n-1}$ is an open ball with radius $r_n < 2^{1-n}$. and name the center of the open ball B_i to be x_i. Clearly, $\{x_i\}$ gives a Cauchy sequence (why?). Thus it converges to a point x. Since x is a limit point in open ball B_j, it has the property that

$$x \in \overline{B_{j+1}} \subset B_j \subset (M\backslash\overline{A_j}) \subset (M\backslash A_j),\ \forall j \in \mathbb{N}$$

So we have that $x \notin A_j$, $\forall j \in \mathbb{N}$. So

$$x \notin \bigcup_{j=1}^{\infty} A_j = M$$

Which leads to contradiction.

Proof : Uniform Boundedness principle

Let $A_n = \{x \in X, \|fx\| \leq n, \forall f \in F\}$, $n \in \mathbb{N}$. By assumption we have $\bigcap_{n=1}^{\infty} A_n = X$.
We claim that there exists some $j \in \mathbb{N}$ such that A_j is non-empty and closed. To see this, first by by Baire category theorem, there is some A_j such that $\overline{A_j}^{\circ} \neq \emptyset$. Then let $\{x_m\}$ be a Cauchy sequence in A_j with $x_n \to x$, then by continuity of f, $\|fx\| = \lim_{n\to\infty} \|fx_m\| \leq n$, $\forall f \in F$. So $x \in A_j$, hence A_j is closed, thus $\overline{A_j} = A_j$, $A_j^{\circ} = \overline{A_j}^{\circ} \neq \emptyset$. So we can choose a point p from interior of A_j, and $\varepsilon > 0$ such that open ball $B_\varepsilon(p) \subseteq A_j$.
Now for any $x < \|\varepsilon\|$ with any $T \in F$ we have

$$\|T(x)\| = \|T(x + p - p)\| = \|T(x + p) - T(p)\| \leq \|T(x + p)\| + \|T(p)\| \leq n + n = 2n$$

So for any non-zero vector $x \in X$, we have

$$\|T(x)\| = \frac{\|x\|}{\varepsilon} \left\| T(\varepsilon \frac{x}{\|x\|}) \right\| \leq \frac{2n}{\varepsilon} \|x\|$$

This holds for any $T \in F$, thus

$$\sup_{f \in F} \|f\| \leq \frac{2n}{\varepsilon} < \infty$$

A simple corollary of the theorem is Banach limit.

Corollary 7.0.4 (Banach Limit)
Let $T_n : X \to Y$ be a sequence of operators, where X and Y are Banach spaces. Suppose $\{T_n\}$ converges pointwise, then these pointwise limits define a bounded linear operator T.

8 Open Mappiing and Closed Graph

8.1 Open mapping theorem

Definition 8.1.1 (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 8.1.2 (Open map)
Let X, Y be linear spaces. $A : X \to Y$ is open if $A(U) \subset Y$ is open.

Remark 8.1.3

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

Theorem 8.1.4 (Open Mapping Theorem)
Let X, Y be Banach, $A \subset \mathcal{L}(X, Y)$. 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 8.1.5
ii) important in application. If $A \in \mathcal{L}(X, Y)$ is bijective then $A^{-1} : X \to Y$ liner is easy (why?). The point is A^{-1} is also bounded, or equivalently continuous.

The main step of the proof is the following:

Lemma 8.1.6 (A as in i))
$\exists r > 0$ s.t. $B_Y(r) \subset \overline{A(B_x(1))}$

Proof

Since A is suerjective

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

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

$$int(\overline{A(B_X(1))}) \neq \emptyset$$

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

Proof of theorem:

Proof

i) Pick r as in Lemma.
Claim: $B_Y(r/2) \subset A(B_X(1)))$.
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 $y_0 \overset{def}{=} Ax_0$, get

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

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). $\square$

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

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

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

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

By lemma above,

$$\forall s > 0, B_Y(sr) \subset \overline{A(B_X(s))} \; (*)$$

$s = 1/2$. Pick $x_1 \in B_X(1/2)$ s.t. $\|Ax_1 - y\| < r/2$. Now set $y_1 = y - Ax(\in B_X(r/2)$. Iterate. Assume that for some ≥ 1 have $x_1, ...x_k, y_1, ...y_k$ s.t.

$$\forall 1 \leq \tilde{k} \leq k: \; \|\tilde{x_k}\|_X < 2^{-k},\, y_{\tilde{k}} = y_{\tilde{k}-1} - Ax_{\tilde{k}} \in B_Y(2^{-\tilde{k}}r$$

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

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

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

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

Example 8.1.7 (Equivalence of Norm)
Let $X = Y$, with norms $\|\cdot\|_1$ and $\|\cdot\|_2$ and assume $\exists C > 0$ s.t.

$$\|x\|_2 \leq C\,\|x\|_1\,, \forall x \in X \quad (1)$$

If X is complete, with respect to both $\|\cdot\|_1$ and $\|\cdot\|_2$ then consider $A = id : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ is open by Theorem (indeed thm applies b/c A is bounded by (1). Since A is bijective, ii) gives that $A^{-1} = id : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$ is bounded, i.e.

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

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

Example 8.1.8 (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:

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

This shows Y needs to be complete in theorem.

Example 8.1.9 (Completeness of X)
This example shows completeness of X is also required. Take

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

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 $A : X \to X$,

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

Then A is linear, bijective with

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

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 $x^{(n)} = (\overbrace{1, 1, 1, 1}^{n}, 0, ...)$ 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 i), A^{-1} would be bounded.

8.2 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 8.2.1 (Running Example)
$Y = X = C = C^0[0,1]$ with $\|\cdot\|_X = \|\cdot\|_\infty$ and $A = \frac{d}{dt}$, with $D(A) \stackrel{eg}{=} C^1[0,1] \subset X$ or subspaces thereof. Prime example of (<u>unbounded</u>) operator with dense domain $D(A)$: indeed $C^1[0,1] = C$ using e.g. Weierstrass Approximation Theorem (Polynomials are already $\|\cdot\|_\infty$-dense in C).

Definition 8.2.2 (Graph)
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 8.2.3 (Closed Operator)
A is called <u>closed</u> if Γ_A is closed in $(X \times Y, \|\cdot\|_{X \times Y})$

Example 8.2.4 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 $(x,y) \in X \times Y$
NTS: $(x,y) \in \Gamma_A$ i.e. $y = Ax$. Know $y_k = Ax_k$ and $\|x_k - x\|_X \xrightarrow{k\to\infty} 0$, $\|Ax - y\|_Y \xrightarrow{k\to\infty} 0$
But $\forall k \geq 1$

$$\|y - Ax\|_Y \leq \|y - Ax\|_Y + \|Ax_k - ax\|_Y \leq \|y - Ax\|_Y \, \|A\| \, \|x_k - x\|_X$$

Thus

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

Theorem 8.2.5 (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:

$$\begin{aligned} \Pi_X :\ \Gamma_A &\to X \\ (x, Ax) &\mapsto x \end{aligned} \qquad\qquad \begin{aligned} \Pi_Y : \Gamma_A &\to Y \\ (x, Ax) &\mapsto Ax \end{aligned} \tag{28}$$

Π_X, Π_Y are continuous with $\|\Pi_X\|, \|\Pi_Y\| \le 1$, Π_X is injective, and surjective. By OMT, 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 8.2.6
ii) is simpler than i), but equivalent.
i) says A is continuous, i.e. if $(x_n) \subset X$, $x \in X$

$$\|x_n - x\|_X \to n \to \infty 0 \implies \|Ax_n - Ax\|_Y \to n \to \infty 0$$

This contains two things to check: (Ax_n) converges and limit is Ax.
ii) says A is closed, i.e.

$$\begin{cases} \|x_n - x\|_X \to n \to \infty 0 \\ \|Ax_n - y\|_Y \to n \to \infty 0 \end{cases} \implies Ax = y \tag{29}$$

Which is only one condition to check.

Example 8.2.7 (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

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

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

Corollary 8.2.8 (Continuous Inverse)
X, Y Banach, $A : (DA) \subset X \to Y$ linear, closed and bijective. Then $\exists B = A^{-1} \in \mathcal{L}(Y, X)$ with $AB = id_Y$ and $BA = id_{D(A)}$. Proof is left as an exercise. Hint: similar to CGT, consider $\Pi_Y : \Gamma_A \to Y$, $B \overset{def.}{=} \Pi_X \circ \Pi_Y^{-1}$

Example 8.2.9 (Inverse exists on a subset)
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 $D(A) \overset{def.}{=} C_0^1[0,1] = \{f \in C^1[0,1] : f(0) = 0\}$
Then $A : D(A) \to C$ is bijective and has continuous inverse $B = A^{-1}$ by corollary. In fact, $Bf(t) = \int_0^t f(s)ds$ with $Bf \in D(A)$.

9 Weak Topology

9.1 Weak Convergence and weak* convergence

One annoying thing about our definition of convergence in norms has a problem, which is that it sometimes give different results in infinite dimensional spaces, for example, unit ball in infinite dimensional spaces are not compact. One way to "restore" the good properties is weak topology.

Definition 9.1.1 (Weak convergence)
Let X be normed. A sequence $(x_n) \subset X$ converges weakly to $x \in X$ if $l(x_n) \to l(x)$ $\forall l \in X^*$ writes $x_n \to wx$.

Proposition 9.1.2 (Convergence implies weak convergence)
If $x_n \to x$ then $x_n \to wx$.

Proof

Let $(x_n) \subset X$ be such that $x_n \to x \in X$. Then for any $l \in X^*$, we have

$$|l(x_n) - l(x)| = |l(x_n - x)| \leq \|l\|_* \|x - x_n\|$$

where $\|l\| < \infty$ by boundedness of l and $\|x - x_n\| \to 0$ by assumption. Thus we conclude $|l(x_n) - l(x)| \to 0$, and since this holds for arbitrary bounded linear functional, we have $x_n \to wx$.

Definition 9.1.3 (Weak* convergence)
Let X be normed space X^* be its dual space. A sequence $(l_n) \subset X^*$ is weak* convergent to $l \in X^*$ if $l_n(x)] \to l(x)$ $\forall x \in X$. (Point-wise convergence of linear functional)

Definition 9.1.4 (Bidual)
Bidual of X is $X^{**} = (X^*)^*$.

Proposition 9.1.5 (Embedding X with canonical map)
X always embeds into X^{**}, in a sense that there exists a canonical map $\iota : X \to X^{**}$, $\iota(x)(l) \equiv l(x)$ for any $x \in X$, $l \in X^*$, which is linear and is an isometry.

Definition 9.1.6 (Reflexive)
A normed space X is reflexive if the canonical map ι given above is surjective, equivalently this means X is reflexive if $X \cong X^{**}$.

Remark 9.1.7 (On X^*)

- Strong Convergence: $\|l_n - l\|_* \to n \to \infty 0$
- Weak convergence: $|T(l_n) - T(l)| \to n \to \infty 0$, $\forall T \in X]^{**}$
- Weak * convergence: $\|l_n(x) - l(x)\|_X \to n \to \infty 0$, $\forall x \in X$
- Strong $\implies$ Weak $\implies$ Weak *.
- When $X^{**} \cong X$, weak* and weak convergence are equivalent.

Example 9.1.8 (Reflexive space)
Typical examples include all Hilbert spaces, since $X \cong X^*$ by corollary of Riesz Rep. Theorem. Also, for any $p \in (1, \infty)$, L^p space is reflexive since, with its dual being L^q satisfying $1/p + 1/q = 1$, and bidual congruent to L^p itself. Unfortunately when $l = 1$ or $l = \infty$ the space is not reflexive.

9.2 Consequences of weak convergence

Proposition 9.2.1 (Weak convergence $\implies$ bounded)
Let $(x_n) \subset X$ be such that $x_n \xrightarrow{w} x$. Then (x_n) is bounded, i.e.

$$\sup_{n \in \mathbb{N}} \|x_n\| < \infty$$

Theorem 9.2.2 (Banach-Alaoglu Theorem)
Let X be separable. If $(l_n) \subset X^*$ is bounded in X^*, then $\exists l \in X^*$ and a subsequence $\Lambda \subset \mathbb{N}$ such that $l_n \xrightarrow{w^*} l$. In plain language, this means that any bounded sequence in dual of separable space has a weakly convergent subsequence.

Remark 9.2.3
If X is reflexive, then we don't need separability. A special case of this is when X is Hilbert.

Corollary 9.2.4 (Banach-Alaoglu in Hilbert space)
If $(x_n) \subset \mathscr{H}$ is bounded, there are two consequences, equivalent:

- (x_n) has a weakly convergent subsequence.
- Unit ball in $\mathscr{H}$, $B(0, 1)$ is weakly compact.

If $(x_n) \subset \mathscr{H}$ is bounded, then it has a convergent subsequence.

10 Compact Operators

In this section, we'll define and investigate compact operators, the class of operators closest to the concept of finite rank operators. The main theorem of this chapter is that compact operators form a Banach space.

Definition 10.0.1 (Compact Operators)
X, Y normed, $T : X \to Y$ linear. Then T is a compact operator if $\forall B \subset X$ bounded, $T(B)$ is sequentially Compact.

Remark 10.0.2
Compact Operators maps bounded set to (sequentially) compact set.

Lemma 10.0.3 (Equivalent definitions)
For X, Y Banach, the following are equivalent:

- T is compact operator.
- $\forall B$ bounded, $\overline{T(B)}$ is compact.
- $\overline{T(B_1)}$, image of unit ball under T is compact
- $\forall\,(x_n)_1^\infty \subset X$ bounded, $T(x_n)$ has a Cauchy subsequence.

Proposition 10.0.4 (Compact $\implies$ bounded)
If T is a compact linear operator, it's bounded.

Example 10.0.5 (Examples of compact operators)
NOOOOOOT COMPLEEEEEEET! COOOOOOOM BAAAAAACK!!!

Theorem 10.0.6 (Limit of Compact Operators)
Let X , Y be Banach. If $T_n : X \to Y$ is a sequence of compact operators and for some $T \in \mathcal{L}(X, Y)$,

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

Then T is compact.

Remark 10.0.7
The theorem above means that the normed space formed by compact linear operators endowed with operator norm

$$\Big(\{T \in \mathcal{L}(X, Y) : T \textit{ compact}\}, \|\cdot\|_{\mathcal{L}(X,Y)}\Big)$$

is closed, i.e. Banach space.

11 Spectrum

In this section, we consider normed linear spaces over $\mathbb{C}$, for the reason that it provides a simpler picture of the spectrum problem. In linear algebra, we know that Hermitian matrices over $\mathbb{C}$ has real eigenvalues. This is the spectral theory of finite dimensional space. In infinite dimensional spaces, the settings are different, however, are extensions of finite dimensional cases.

11.1 Spectrum and Resolvent

Definition 11.1.1 (Resolvent & Spectral)
Let X be Banach. $T : D_T \subset X \to X$ be a linear operator, its resolvent is

$$\varrho(T)\{\lambda \in \mathbb{C} : (\lambda I - T) \text{ is bijective and } \exists (xI - T)^{-1} \in X^*\}$$

The spectrum is defined as the complement:

$$\sigma(T) = \mathbb{C} \backslash \varrho(T)$$

The <u>resolvent</u> of A is the map $R : \varrho(A) \to \mathcal{L}(X)$,$\varrho(A) \ni \lambda \mapsto \in \mathcal{L}(X)$

Remark 11.1.2
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 11.1.3 *Consider $X = \mathbb{C}$, $A \in \mathcal{L}(X)$, i.e. $D_A = X$, $\lambda \in \mathbb{C}$, $(\lambda - A)$ invertible $\iff p(\lambda) \overset{def.}{=} det(\lambda - A) \neq 0$. 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 11.1.4
If $z_0 \in \varrho(A)$, then

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

Hence $\varrho(A)$ is open, and $\sigma(A)$ is closed.

Proof

Write

$$z - A = (z - z_0) + (z_0 - A) = (1 + (z - z_0)R_{z_0})(z_0 - A)$$

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

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

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 $A^0 = Id = 1$,

$$(1 - A^{-1})^{-1} = \sum_{n=0}^{\infty} A^n \in \mathcal{L}(X)$$

proof left as exercise.

Example 11.1.5
Diagonal operator $T = T_\lambda$ continued. Claim: $\sigma(T) = \overline{\{\lambda_k : k \in \mathbb{N}\}}$.

Proof : $\sigma(T) \supset \overline{\{\lambda_k : k \in \mathbb{N}\}}$

$Tx = \lambda_k x$ so $\lambda_k - T$ is not injective, so $\{\lambda_k : k \in \mathbb{N}\} \subset \sigma(T)$, hence by lemma $\sigma(T) = \overline{\{\lambda_k : k \in \mathbb{N}\}}$.

Proof : $\sigma(T) \subset \overline{\{\lambda_k : k \in \mathbb{N}\}}$

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 $x \in \ell^2$, $y \overset{def.}{=} (\widetilde{\lambda} - T)x = ((\widetilde{\lambda} - \lambda_k)x_k)_k$. 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 11.1.6
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 11.1.7
Consider linear operator A with closed graph and spectrum $\sigma(A)$.

- Point spectrum: $\sigma_p(A) \overset{def.}{=} \{\lambda \in \mathbb{C} : \lambda - A \text{ not injective}\}$
- Continuous spectrum: $\sigma_c(A) \overset{def.}{=} \{\lambda \in \mathbb{C} \setminus \varrho(A) : (\lambda - A) \text{ injective }, Im(\lambda - A) \text{ dense }\}$
- Residual spectrum: $\sigma_r(A) \overset{def.}{=} \sigma(A) \setminus (\sigma_p(A) \cup \sigma_c(A))$

Elements of point spectrum are called Eigenvalues of A with eigenspace $ker(\lambda - A) = \{x \in D_A : Ax = \lambda x\} \neq \{0\}$

Example 11.1.8 (Shift operator)
$S : \ell^2 \to \ell^2$, $S(x_1, x_2 ...) = (0, x_1, x_2 ...)$. 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. So $\sigma_p(S) = \emptyset$. In fact, $\sigma(S) = \overline{D} = \{\zeta \in \mathbb{C} : |\zeta| \leq 1\}$, $\sigma_p(S) = D$, $\sigma_c(S) = \partial D = S^1$

Example 11.1.9
$X = \mathbb{C}^n$, $\sigma(\cdot) = \sigma_p(\cdot)$

Example 11.1.10
T_λ is indicative of a certain class : X Hilbert, $T \in \mathcal{L}(X)$ compact, self-adjoint, then by Riesz-Schauder, $\sigma(T) \setminus \{0\} = \sigma_p(T) \setminus \{0\}$

11.2 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^* characterized by $\forall x \in D_A$, $y \in D_A^*$: $\langle A^* y, x \rangle = \langle y, Ax \rangle$ and $D_A^* = \{y \in H : l_y : D_A \to \mathbb{C}, x \mapsto \langle y, Ax \rangle$ 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 11.2.1
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 $D_A^* = D_A$.

Remark 11.2.2 (Subtleness of "symmetric")
In linear algebra, we defined symmetric matrices: $A^T = A$. However we should notice that "symmetric operator" is the conceptual extension of "Hermitian matrix". To see this, consider $H \in \mathbb{C}^{n\times n}$, Hermitian, i.e. $A \in \mathbb{C}^{n\times n}$. Then $\forall x, y \in \mathbb{C}^n$:

$$\langle Hx, y\rangle = \overline{x^T H^T} y = (\overline{x^T})(Hy) = \langle x, Hy\rangle$$

However, an arbitrary symmetric matrix with complex entries: $S \in \mathbb{C}^{n\times n}$ and $S^T = S$, is not necessarily a symmetric operator:

$$\langle Sx, y\rangle = \overline{x^T S^T} y = (\overline{x^T})(\overline{S}y) = \langle x, \overline{S}y\rangle$$

In fact, in Quantum Mechanics, the class of self-adjoint operators are called Hermitian operator. The idea of this definition is being "self-adjoint", at least to an extend. One may show that every "symmetric operator" can be extended to a self-adjoint operator.

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

Lemma 11.2.3 (Symmetric operator has real point spectrum)
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 $\lambda \|X\|_H^2 = \langle Ax, x\rangle \overset{symm.}{=} \langle x, Ax\rangle = \overline{\langle Ax, x\rangle} = \overline{\lambda}\|x\|_H^1$. So $\lambda = \overline{\lambda} \implies \lambda \in \mathbb{R}$.

Is this true for all the spectrum of $\mathcal{H}$ (cf. $\mathbb{C}^n$)?

Example 11.2.4
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, $f \in \mathcal{H}$ is said to have a <u>weak derivative</u> f' if $f' \overset{def.}{=} v$ for some $v \in \mathcal{H}$, $\int_0^1 fg'\mathrm{d}t = -\int_0^1 vg\mathrm{d}t$, $\forall g \in C_c^\infty(0,1)$.
Consider

$$A_\infty = i\frac{\mathrm{d}}{\mathrm{d}t} : C_c^\infty(0,1) \subset H \to H$$

and extensions A_1.....A_3 with

$$\begin{aligned} &D_{A_1} = H^1 \overset{def.}{=} \{f \in H : f \text{ has a weak derivative f'}\} \\ &\cup \\ &D_{A_2} = \{f \in H^1 : f(0) = f(1)\} \text{ periodic b.c.} \\ &\cup \\ &D_{A_3} = \{f \in H^1 : f(0) = 0 = f(1)\} \text{ Dirichlet b.c.} \end{aligned} \tag{30}$$

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 $k \in \mathbb{Z}$, $f(t) = e^{-2\pi i k t} \in D_{A_2} \cap ker(a\pi k - A_2$, 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 $(\lambda - A_2)^{-1} \in \mathcal{H}$.
For $g \in H$, the general sd. of $\lambda f - i f' = g$ 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 $A_3 f = i f' = \lambda f$ 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 11.2.5
Let $A \subset A^*$. Then $\forall \zeta \in \mathbb{C} \, \forall u \in D_A$: $\|(\zeta - A)u\|_H \geq |Im(\zeta)| \, \|u\|_H$.
(So for $\zeta \notin \mathbb{R} \implies (\zeta - A)$ injective, i.e. $\zeta \notin \sigma_p(A)$)

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,

$$|Im(\zeta)| \, \|u\|_H^2 = |Im(\langle u, (\rangle \, \zeta - A)u)| \leq | \, \langle u, (\zeta - A)u \rangle \leq \|u\|_H \, \|(\zeta - A)u\|_H$$

The example also nicely illustrates:

Proposition 11.2.6 (Self-adjoint operator has real specturm) *If $A = A^*$, then $\sigma(A) \subset \mathbb{R}$*

Proof

Let $\zeta \in \mathbb{C} \setminus \mathbb{R}$. *Want to show* $\zeta \in \varrho(A)$, *i.e.* $zeta - A : D_A \to H$ *is bijective with* $(\zeta - A)^{-1} \in \mathcal{L}(H)$.
We will show:
(*) $\zeta - A$ *is surjective.*
Once (*) holds, we are done: by previous lemma, $\zeta - A$ *is injective hence bijective, and surjectivity + same lemma also yields*

$$\left\|(\zeta - A)^{-1}\right\|_{\mathcal{L}(H)} \leq \frac{1}{|Im(\zeta)}$$

proof of (*): we first show

$$(**) \qquad im(\zeta - A)(\subset H) \text{ is closed}$$

Assume $v_k = (\zeta - A)u_k \xrightarrow{k\to\infty}$. *BY Lemma p.65,*

$$\|u_k - v_k\|_H \leq \frac{1}{|Im(\zeta)|} \|v_k - v_l\|_H \xrightarrow{k,l\to\infty}$$

Hence (u_k) *is Cauchy and* $u_k \to u$ *for some* $u \in H$. *But* $A = A^*$ *has closed graph so* $v = (\zeta - A)u$, *i.e. (**) holds.*
Back to (*):
Due to (**), $M \overset{def.}{=} Im(\zeta - A)$ *is closed. Assume* $M \neq H$. *Pick* $v \in M^\perp \setminus \{0\}$. *Then*

$$\forall u \in D_A : \langle v, (\zeta - A)u\rangle = 0 \text{ or } \langle v, Au\rangle = \overline{\zeta}\,\langle v, u\rangle$$

Hence, $D_A \ni u \mapsto \langle v, Au\rangle$ *is continuous,* $v \in D_A^* = D_A$ *and* $Av = A^*v = \overline{\zeta}v$ *but by lemma p.65,*

$$|Im(\zeta)|\,\|v\|_H \leq \left\|(\overline{\zeta} - A)v\right\|_H = 0$$

which yields $v = 0$

11.3 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 11.3.1 (Reisz-Schauder)
Let $T : H \to H$ be compact and self-adjoint, then:

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

- $\sigma_p(T)$ 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 11.3.2
Shift operator $T_\lambda : \ell^2 \to \ell^2$ continued. is compact $\iff \lim_{k\to\infty} \lambda_k = 0$, 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 11.3.3 (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

$$\begin{aligned} \lambda_1 \langle e_1, e_2 \rangle &= \langle \lambda_1 e_1, e_2 \rangle = \langle Te_1, e_2 \rangle \overset{S.A.}{=} \langle e_1, Te_2 \rangle \\ &= \langle e_1, \lambda_2 e_2 \rangle \overset{\lambda_2 = \overline{\lambda_2}}{=} \lambda_2 \langle e_1, e_2 \rangle \end{aligned} \tag{31}$$

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 is in force.
Define, for $\lambda \in \sigma_p(T) \setminus \{0\}$, $X_\lambda = ker(\lambda - T) \neq \{0\}$. By lemma 1:

$$X_\lambda \perp X_\lambda \; \forall \lambda \neq \lambda', \lambda\lambda' \in \sigma_p(T) \setminus \{0\}$$

Lemma 11.3.4 (Lemma 2)
$\lambda \in \sigma_p(T) \setminus \{0\}$.

- $dim(X_\lambda) < \infty$
- $\forall r > 0$: $\sigma_p(T) \setminus B_r(0)$ 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 $Tx = \lambda x \, \forall x \in X_\lambda$, $T(B_r^{X_\lambda}(0)) = \lambda B_1^{X_\lambda}(0)$. So $\lambda \overline{(B_1^{X_\lambda}(0))}$ is compact $\implies \overline{B_1^{X_\lambda}(0)}$ is compact $\implies dim(X_\lambda) < \infty$.
ii) Suppose not, then $\exists r > 0$: $\sigma_p(T) \setminus B_r(0)$ is infinite (one can show $\sup_{\lambda \in \sigma_p(T)} |\lambda| < \infty$). Then one can pick sequence $(\lambda_k) \subset \sigma_p(T)$ with $\lambda_k \neq \lambda_l \, \forall k \neq l$ and $\lambda_k| > r \, \forall k$.
Let $e_k \neq 0$ be eigenvector for λ_k: $Te_k = \lambda_k e_k \, \forall k$. By compactness of T, $\exists \Lambda \subset \mathbb{N}$ such that for some $y \in H$,

$$Te_k \xrightarrow[k \to \infty, k \in \Lambda]{wrt. \|\cdot\|_H} y \in H$$

Hence,

$$(\lambda_k e_k) \xrightarrow[k \to \infty, k \in \Lambda]{wrt. \|\cdot\|_H} y \in H$$

In particular, $(\lambda_k e_k))k \in \Lambda$ is Cauchy. But for $k \neq l$,

$$\|\lambda_k e_k - \lambda_l e_l\|_H^2 = + \langle \lambda_l e_l, \lambda_l e_l \rangle + \underbrace{\langle \lambda_k e_k, \lambda_l e_l \rangle}_{= \lambda_k \overline{\lambda_l} \langle e_k, e_l \rangle \overset{Lemma 1}{=} 0} + \langle \lambda_l e_l, \lambda_k e_k \rangle$$

Assume, $\|e_k\|_H = \|e_l\|_H = 1$, otherwise replace e_k by $e_k / \|e_k\|_H$,

$$\|\lambda_k e_k - \lambda_l e_l\|_H^2 = |\lambda_k|^2 + |\lambda_l|^2 > 2r$$

Proof : Riesz-Schauder

Proof of i) see last week.
ii) By Lemma 2 ii),

$$A_n = \sigma_p(T) \cap \{z : \frac{1}{n+1} \leq |z| \leq \frac{1}{n}\}(\subset \sigma_p(T) \setminus B_{1/(n+1)}(0))$$

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. If they belong to distinct eigenvalues $\lambda + k \neq \lambda_l$ this is automatic after normalizing them to have norm 1 by lemma 1.
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

$$X \overset{def.}{=} \overline{span\{e_k\}} = \overline{\bigoplus_{\lambda \in \sigma_p(T)\setminus\{0\}} X_\lambda}$$

Claim 1: $\forall x \in X$: $x = \sum_k \langle x, e_k \rangle e_k$
Claim 2: $Y \overset{def.}{=} X^\perp = ker(T)$
If the two claim holds,
NOOOOOOT COMPLEEEEEEET! COOOOOOOM BAAAAAACK!!!
which concludes the proof. Proof of the two claims follows.

Proof : Claim 1

: 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, $\forall k \geq 0$,

$$\langle x - y, e_k \rangle = \lim_{n\to\infty} \langle x - x_k, e_k \rangle = \langle x, e_k \rangle - \lim_{n\to\infty} \langle x_n, e_k \rangle = 0$$

so $x = y$. With claim 1 and continuity of T we have

$$\forall x \in X \ : Tx = \sum_k \langle x, e_k \rangle Te_k = \sum_k \langle x, e_k \rangle \lambda_k e_k$$

Proof : Claim 2

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

$$\sigma_p(T) \setminus \{0\} \neq \emptyset$$

(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 $\lambda \in \sigma_p(T)$, $\lambda \neq 0$, then

$$Te \overset{e \in Y}{=} T_Y e = \lambda e$$

So $e \in X_\lambda \subset X = Y^\perp$. But $Y^\perp \cap Y = \{0\}$. Thus $Y \overset{def.}{=} X^\perp = ker(T)$.

Remark 11.3.5 *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 $\sigma_p(T) \setminus \{0\} X_\lambda$ is countable.

12 Appendix

12.1 Young's Inequality

For non-negative real numbers $a \geq 0$ and $b \geq 0$ with positive real numbers $p > 1$ and $q > 1$ satisfying

$$\frac{1}{p} + \frac{1}{q} = 1$$

Then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}, \quad \text{equality holds when } a^p = b^q$$

More details at Wikipedia.

12.2 Minkowski's inequality

Minkowski's inequality gives the triangular inequality in ℓ^p and L^p spaces:

$$\|x\|_p + \|y\|_p \geq \|x + y\|_p$$

which can be written equivalently, for ℓ^p with $p \in (1, \infty)$:

$$\left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{1}{p}} + \left(\sum_{n=1}^{\infty} |y_n|^p\right)^{\frac{1}{p}} \geq \left(\sum_{n=1}^{\infty} |x_n + y_n|^p\right)^{\frac{1}{p}}$$

More details at Wikipedia.

12.3 Hölder's Inequality

Theorem 12.3.1 (Hölder's inequality,)
Let $p, q \in [1, \infty]$, satisfying:

$$\frac{1}{p} + \frac{1}{q} = 1$$

Then if $x \in \ell^p$ and $y \in \ell^q$ then

$$\|x\|_p \|y\|_q \geq \|xy\|_1 \iff \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p} \left(\sum_{n=1}^{\infty} |y_n|^q\right)^{1/q} \geq \sum_{n=1}^{\infty} |x_n y_n|$$

Then if $g \in L^p(S)$ and $g \in L^q(S)$ then

$$\|f\|_p \|g\|_q \geq \|fg\|_1 \iff \left(\int_S |f|^p dx\right)^{1/p} \left(\int_S |g|^q dx\right)^{1/q} \geq \int_S |fg| dx$$

Remark 12.3.2 (Short version)
Here's the key information of the above theorem:

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

Remark 12.3.3
When supremum occurs, which is one of p and q becomes infinity and another becomes 1, the inequality still holds. One should pay attention that for L^p spaces the supremum norm is actually essential supremum.

Remark 12.3.4 (Proof of Hölder's inequality)
The proof uses young's inequality. See this link for detailed proof.

12.4 Jensen's Inequality (convex function)

Let $\mathbb{I} \subset \mathbb{R}$. A function $f : \mathrm{I} \to \mathbb{R}$ is **convex** if the following holds for all $t \in [0, 1]$:

$$t \cdot f(x) + (1-t) \cdot f(y) \geq f(tx + (1-t)y),\ \forall x,\, y \in \mathbb{R} \qquad \text{(Jensen)}$$

Similarly, a function $f : \mathrm{I} \to \mathbb{R}$ is **concave** if the following holds for all $t \in [0, 1]$:

$$t \cdot f(x) + (1-t) \cdot f(y) \leq f(tx + (1-t)y),\ \forall x,\, y \in \mathbb{R}$$

For convex function, the inequality may be rewritten as

$$\frac{1}{t}(f(y + t = (x-y)) - f(y)) \leq f(x) - f(y)$$

For concave function with $f(0) \geq 0$ we have **sub-additivity**,which is for $t \in [0, 1]$,

$$f(tx) = f(tx + (1-t)0) \geq tf(x) + (1-t)f(0) \geq tf(x)$$

and thus,for $a, b \in \mathbb{R}^+$

$$f(a) + f(b) = f(\frac{a}{a+b}(a+b)) + f(\frac{b}{a+b}(a+b)) \geq \frac{a}{a+b}f(a+b) + \frac{b}{a+b}f(a+b) = f(a+b)$$

Remark 12.4.1 (Generalised Jensen's Inequality)
Jensen's inequality can be generalized to a sequence variables with weight. Consider a convex function evaluated at $x_1, x_2 ...,\ f(x_1), f(x_2)...f(x_n)$, with weight $w_1, w_2...$ with $\sum_{j \in \mathbb{N}} w_j = 1$, where

$$\sum_{j \in \mathbb{N}} w_j f(x_j) < \infty \quad \text{and} \quad \sum_{j \in \mathbb{N}} f(w_j x_j) < \infty$$

Then,

$$\sum_{j \in \mathbb{N}} f(w_j x_j) \leq \sum_{j \in \mathbb{N}} w_j f(x_j)$$

Remark 12.4.2 ($f'' > 0$)
For a single variable twice differentiable function, the second derivative being non-negative implies convexity.

12.5 Completion of metric space

Theorem 12.5.1 (Completion of metric space)
Every metric space has a completion.

Idea of proof: First construct a space of Cauchy sequence and define a metric on this space, then show that the original space can be embedded to the space of Cauchy sequence as a dense subset by an isometry. Proof given step by step.

Lemma 12.5.2 (Step I:Space of Cauchy sequence)
Let (X, d) be a metric space. Let $C[X]$ denote set of all Cauchy sequence in X. Define equivalence relation $\sim$ on $C[X]$: $x \sim y \iff \lim_{n\to\infty} d(x_n, y_n) = 0$. Define set $X^* = \{[(x_n)], (x_n) \in C[X]\}$ and metric on this set $d^*((x_n), (y_n)) = \lim_{n\to\infty} d(x_n, y_n)$. One can check that d^* is indeed a well-defined metric on X^*

12.6 L^p space and l^p space

L^p space and ℓ^p space are very important for both Lebesgue measure and also functional analysis. We shall fist introduce ℓ^p space by introducing its $p-norm$. Here the definition is valid on both $\mathbb{R}$ and $\mathbb{C}$.

Definition 12.6.1 (p-norm,discrete)
Let $p \in \mathbb{R}$ with $1 \leq p \leq \infty$.
For $p < \infty$,we define p-norm of sequence $(x_n)_1^\infty$ to be:

$$\|(x_n)_1^\infty\|_\infty \equiv \left(\sum_{s=1}^{\infty} |x_n|^p\right)^{1/p}$$

And for when "$p = \infty$" the p-norm becomes ∞-norm, defined as

$$\|(x_n)_1^\infty\|_p \equiv \sup_{n\in\mathbb{N}} |x_n|$$

And ℓ^p space is defined to be the set of all sequences with finite p-norm:

$$\ell^p \equiv \left\{(x_n)_1^\infty : \|(x_n)_1^\infty\|_p < \infty\right\}$$

Definition 12.6.2 (p-norm,continuous)
Let $p \in \mathbb{R}$ with $1 \leq p \leq \infty$. Again,∞ gives supremum.(why?)
Consider measurable functions $f : [a, b] \to \mathbb{R}$.
For $p < \infty$,we define p-norm of function f to be:

$$\|f\|_p \equiv \left(\int_a^b |f(x)|^p dx\right)^{1/p}$$

And for when "$p = \infty$" the p-norm again becomes ∞-norm, defined as

$$\|f\|_\infty \equiv \sup_{n\in[a,b} |f(x)|$$

And L^p space is defined to be the set of measurable all functions with finite p-norm from $[a, b]$ to $\mathbb{R}$ or $\mathbb{C}$:

$$L^p \equiv \left\{f : \|f\|_p < \infty\right\}$$

Remark 12.6.3
There are very typical ℓ^p spaces, like

- ℓ^1, space of absolutely convergent sequences
- ℓ^2, gives a set of square-summable sequences, forms a Hilbert space
- ℓ^∞, set of all bounded sequences.