MA3G8 Functionalanalysis II

① Metric spaces as topological spaces

1.1 Metric spaces
1.2 Topological spaces. Basic facts
1.3 Topological spaces. Convergence and Continuity

② Normed spaces

2.1 Definitions (revision)
2.2 Completeness
2.3 Separability and bases
2.4 Linear operators

③ Duality and main theorems

3.1 The Hahn-Banach theorem
3.2 The Baire Category theorem and its consequences
3.3 The Uniform Boundedness Principle

④ Hilbert spaces

4.1 Basic definitions

⑤ Bounded operators in Hilbert spaces

⑥ Unbounded operators

1 Metric spaces as topological spaces

The concepts of limit, convergence, and continuity are central to all of analysis, and it is useful to have a general framework for studying them which includes their classical manifestations as special cases.
One such framework, which is closely related to analysis on Euclidean space, is that of ~~a~~ metric spaces. However, metric spaces are not sufficiently general to describe even some very classical modes of convergence, for example, pointwise convergence of functions on $\mathbb{R}$.
A more flexible theory can be built by taking the open sets, rather than a metric, as the primitive data. This will lead to the notion of topological spaces.

1.1 Metric Spaces (revision)

Metric spaces are an abstract setting in which to discuss basic analytical concepts such as convergence of sequences and continuity of functions.
The fundamental tool required for this is a distance function or "metric".

Definition 1.1: A metric d on a set X is a function $d: X \times X \to \mathbb{R}$ with the following properties.
For all $x, y, z \in X$

(a) $d(x,y) \geq 0$, and $d(x,y) = 0 \iff x = y$ (positivity)

(b) $d(x,y) = d(y,x)$ (symmetry)

(c) $d(x,z) \leq d(x,y) + d(y,z)$ (triangle inequality)

If d is a metric on X, then the pair (X,d) is called a metric space.

Example 1.2:
Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. For any integer $n \in \mathbb{N}$, the function $d: \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{R}$ defined by

$$d_p(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p} \quad \text{for } p \in [0, \infty)$$

respectively

$$d_\infty(x,y) = \max \{ |x_i - y_i| : 1 \leq i \leq n \}$$

for $p = \infty$,

is a metric on the set $\mathbb{K}^n$.

$\left.\begin{matrix} p=2 \\ \mathbb{K}=\mathbb{R} \end{matrix}\right\}$ Euclidean metric on $\mathbb{R}^n$

Example 1.3:

(a) $X \neq \emptyset$, $x, y \in X$

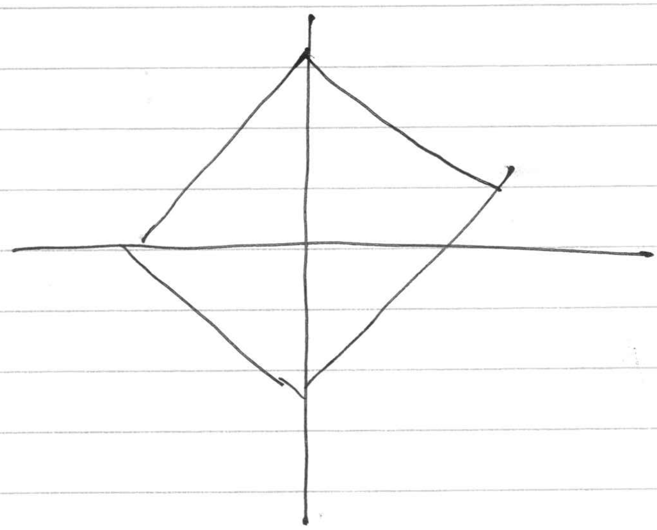
$$d^*(x,y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases},$$

d^* is a metric on X, and it is called the discrete metric

(b) The taxicab metric on $\mathbb{R}^2$

$X = \mathbb{R}^2$; and $d(x,y) = |x_1 - y_1| + |x_2 - y_2|$,

$x, y \in \mathbb{R}^2$ $(x = (x_1, x_2),\ y = (y_1, y_2))$



$\{(x,y) : d(x$

$\{x \in \mathbb{R}^2 : d(x,0) \leq 1\}$

(c) (X, d_X), (Y, d_Y) two metric spaces, and let $Z = X \times Y$ be the Cartesian product of X and Y

Then

$$d_Z\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) = d_X(x_1, x_2) + d_Y(y_1, y_2)$$

$(x_1, y_1), (x_2, y_2) \in Z$

defines a metric on Z.

$d_Z((\cdot),(\cdot)) \geq 0$ is clear

$$d_Z((\cdot),(\cdot)) = 0 \iff d_X(x_1, x_2) = 0 \wedge d_Y(y_1, y_2) = 0$$

$$\iff x_1 = x_2 \wedge y_1 = y_2$$

$$\iff \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

triangle ineq. clear from d_X and d_Y

Definition 1.4: Sub metric spaces

(X,d) metric space, $Y \subset X$

Define $d_Y : Y \times Y \to \mathbb{R}$ by $d_Y(x,y) = d(x,y)$ for all $x,y \in Y$, that is, d_Y is the restriction of d to the subset Y.

Corollary 1.5: d_Y is a metric on Y, and is called the metric induced on Y by d.

Definition 1.6:

(a) A sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space (X,d) converges to $x \in X$ (or sequence $(x_n)_{n\in\mathbb{N}}$ is convergent) if, for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ s.t.

$$d(x, x_n) < \varepsilon \quad \forall\ n \geq N.$$

As usual, we write $\lim_{n\to\infty} x_n = x$ or $x_n \to x$

(b) A sequence $(x_n)_{n\in\mathbb{N}}$ in (X,d) is a Cauchy sequence if, for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$d(x_m, x_n) < \varepsilon, \text{ for all } m,n \geq N$$

Note that, the above definitions are equivalent to

$$d(x, x_n) \to 0, \text{ as } n \to \infty; \quad d(x_m, x_n) \to 0, \text{ as } m,n \to \infty.$$

Notation 1.7: Let (X, d) be a metric space.

For any $x \in X$ and any number $\varepsilon > 0$, the set

$$B_\varepsilon(x) = \{y \in X : d(x,y) < \varepsilon\}$$

will be called the open ball with centre x and radius $\varepsilon > 0$, or ε-neighbourhood of x. The set

$$\{y \in X : d(x,y) \leq \varepsilon\}$$ will be

called the closed ball with centre x and radius ε,
~~or the close~~

Definition 1.8: Let (X, d) be a metric space and $Y \subset X$

(a) Y is bounded if there is a number $b > 0$ s.t. $d(x,y) < b$ for all $x, y \in Y$

(b) Y is open if, for each point $x \in Y$, there is an $\varepsilon > 0$ s.t. $B_\varepsilon(x) \subset Y$

(c) Y is closed if the set $X \setminus Y$ is open

(d) A point $x \in Y$ is a closure point of Y if, for every $\varepsilon > 0$, there is a point $y \in Y$ with $d(x, y) < \varepsilon$
(equivalently, if there exists a sequence $(y_n)_{n \in \mathbb{N}} \subset Y$ s.t. $y_n \to x$.

(e) The closure of ~~A~~ Y, denoted by $\overline{Y}$, is the set of all closure points of Y.

<u>**Remark 1.9:**</u> The "open" and "closed" balls in notation 2.7 are open and closed. Furthermore,

$$\overline{B_\varepsilon(x)} \subset \{y \in X : d(x, y) \leq \varepsilon\},$$

but these sets need not be equal in general (but for most spaces ~~th~~ considered in the course these sets are equal)

__Proposition 1.10:__ Let (X, d) be a metric space. Then the following statements are equivalent

(1) $\emptyset$ and X are open

(2) an arbitrary union of open sets is open

(3) any finite union of open sets is open

__Proof:__ (1) clear

(2) $M = \bigcup_{i \in I} M_i$ and $x \in M$. $\exists\ j \in I$ with $x \in M_j$, and as M_j is open, $\exists\ \delta > 0$ with $B_\delta(x) \subset M_j \subset M$

$\Rightarrow M$ open

(3) $M = \bigcap_{i=1}^{k} M_i$, M_i open, $x \in M$

$\Rightarrow x \in M_i$ and $\exists\ \delta_i > 0$ with $B_{\delta_i}(x) \subset M_i$

for all $i \in \{1, \dots, k\}$. $\delta := \min\{\delta_i : i = 1, \dots, k\}$

$\Rightarrow B_\delta(x) \subset B_{\delta_i}(x) \subset M_i \Rightarrow B_{\delta_i}(x) \subset M.$

Definition 1.11: Let (X, d) be a metric space:

A set $Y \subset X$ is compact if every sequence $(x_n)_{n \in \mathbb{N}}$ in Y contains a subsequence that converges to an element of Y.

relatively compact if $\overline{Y}$ is compact.

If X itself is compact we say (X, d) is a compact metric space.

1.2 Topological spaces. Basic facts

~~We call~~
Recall : the power set of a set X is the collection of all subsets of X, and is denoted by 2^X or $\mathcal{P}(X)$.

Definition 1.12: Let X be a nonempty set.

A topology τ on a set X is a collection of subsets of X (i.e., $\tau \subset 2^X$, $\tau \subset \mathcal{P}(X)$) satisfying

(1) $\phi, X \in \tau$

(2) τ is closed under finite intersections
(3) τ is closed under arbitrary unions

A nonempty set X equipped with a topology τ is called a topological space, and is denoted by (X, τ).
We call a member of τ an open set in X.
The complement of an open set is a closed set.

Examples 1.13:

(a) $X \neq \phi$, $\tau = \mathcal{P}(X)$, then τ is called the discrete topology

(b) $X = \mathbb{R}^n$, d_2 Euclidean metric,

$\tau = \{Y \subset \mathbb{R}^n : Y \text{ open}\}$

τ is called the Euclidean topology (it is induced by the Euclidean metric)

(c) $X = \{x_1, x_2, x_3\}$ x_1, x_2, x_3 pairwise disjoint

$\tau = \{\emptyset, X, \{x_1, x_2\}, \{x_2, x_3\}, \{x_2\}\}$

is a topology on X.

Further examples are

$\tau_1 = \{\emptyset, X\}$

$\tau_2 = \{\emptyset, X, \{x_1\}\}$

$\tau_3 = \{\emptyset, X, \{x_1\}, \{x_2\}, \{x_1, x_2\}\}$

$\tau_4 = \{\emptyset, X, \{x_1, x_2\}\}$, $\tau_5 = \{\emptyset, X, \{x_1\}, \{x_1, x_2\}\}$

$\tau_6 = \{\emptyset, X, \{x_3\}, \{x_1, x_2\}\}$

$\tau_7 = \{\emptyset, X, \{x_1\}, \{x_2\}, \{x_1, x_2\}, \{x_2, x_3\}\}$

$\tau_8 = \{Q \mid Q \subset X\}$

Definition + Notation 1.14: $(X, \mathcal{T})$ top. space

(a) $Y \subset X$, the union of all open sets contained in Y is called the interior of Y, and the intersection of all closed sets containing Y is called the closure of A

interior of $Y = Y^\circ$
closure of $Y = \overline{Y}$

(b) The difference $\overline{Y} \setminus Y^\circ = \overline{Y} \cap \overline{Y^c}$ is called the boundary of Y and is denoted by ∂Y.

(c) If $\overline{Y} = X$, Y is called dense in X

If $(\overline{Y})^\circ = \emptyset$, Y is called nowhere dense.

(d) If $x \in X$, a neighborhood of x is a set $Y \subset X$ such that $x \in Y^\circ$.

Thus, a set Y is open $\Leftrightarrow$ it is a neighbourhood of itself.

(e) If $x \in X$ and $Y \subset X$, x is called an accumulation point of Y if $Y \cap (\mathcal{N} \setminus \{x\}) \neq \emptyset$ for every neighborhood $\mathcal{N}$ of x.
(we write $\mathcal{N}_x$ for neighbourhood of x)

(X, τ) top. space, $Y \subset X$

$x \in X$ is a boundary point of Y (i.e, $x \in \partial Y$)

if all neighbourhoods U of x contain a point from Y and a point from $X \setminus Y$.

Definition 1.15: A topology on X is called Hausdorff (or separated) if any two distinct points can be separated by disjoint neighbourhoods of the points.

1.3 Topological spaces. Convergence and continuity

Definition 1.16: (X, τ) top. space

A sequence $(x_n)_{n \in \mathbb{N}}$ in a top. space (X, τ) is said to converge to $x \in X$ if for every neighbourhood $\mathcal{N}$ of x there exists $N \in \mathbb{N}$ such that

$$x_n \in \mathcal{N} \quad \text{for all } n \geq N.$$

Topological space are the natural setting for the concept of continuity, which can be described either global or local terms, as follows.

Definition 1.17: (X, τ_X), (Y, τ_Y) top. spaces

(a) A map $f : X \to Y$ is called continuous if $f^{-1}(V)$ is open in X for every open $V \subset Y$, or equivalently, if $f^{-1}(K)$ is closed in X for every closed $K \subset Y$

that is for all $Q \in \tau_Y : f^{-1}(Q) \in \tau_X$

(b) If $x \in X$, f is called continuous at x if for every neighbourhood $\mathcal{N}$ of $f(x)$ there is a neighbourhood $\mathcal{U}$ of x such that $f(\mathcal{U}) \subset \mathcal{N}$.

Remark 1.18: $(X,\tau_X), (Y,\tau_Y), (Z,\tau_Z)$ top. spaces

Clearly, if $f: X \to Y$ and $g: Y \to Z$ are continuous, then
$g \circ f$ is continuous.

The set of continuous maps from (X,τ_X) to (Y,τ_Y) shall be denoted by $\mathcal{C}(X,Y)$

Question: (X,τ_1), (X,τ_2)

When is the identity continuous

$id: X \longrightarrow \quad , id(x) = x$

$\Longleftrightarrow \tau_2 \subset \tau_1$.

Definition 1.19: (X,τ) top. space
is a compact (top) space, if it is seperated (Hausdorff)

and for any open covering of X $\exists$ a finite

subcovering, that is.

$\forall \tau^* \subset \tau$ with $X \subset \bigcup_{Q \in \tau^*} Q$

$\exists Q_1, \dots, Q_m \in \tau^*$

with $X \subset \bigcup_{k=1}^{m} Q_k$

② Normed Spaces

$\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ with absolute value $|\cdot|$

2.1 Definitions (revision)

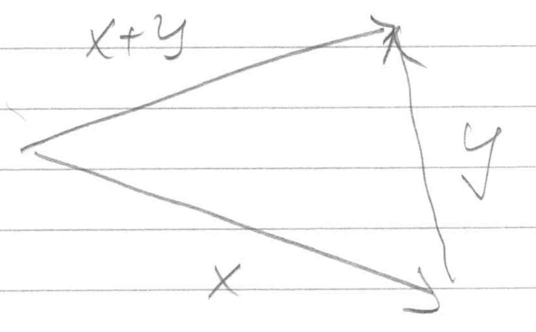
Definition 2.1: Let X be a vector space over $\mathbb{K}$.
A norm on X is a function $\|\cdot\| : X \to \mathbb{R}$ such that for all $x, y \in X$ and $\alpha \in \mathbb{K}$

(i) $\|x\| \geq 0 \ \forall x \in X$, and $\|x\| = 0$ iff $x = 0$ (positivity)

(ii) $\|\alpha x\| = |\alpha| \, \|x\| \ \forall \alpha \in \mathbb{K}$ (homogeneity)

(iii) $\|x + y\| \leq \|x\| + \|y\|$ (triangle ineq.)

If $\|\cdot\|$ is a norm on X, then the pair (X, d) is called a normed vector space, or normed linear space, or normed space.



Example 2.2: $X = \mathbb{K}^n$

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

$p = \infty$ $\quad \|x\|_\infty = \max_{1 \le i \le n} |x_i|$

$\mathbb{R}^n$, $\|x\|_2$ Euclidean norm of $x \in \mathbb{R}^n$.

Exercise: check that $\|\cdot\|_p$ is a norm on $\mathbb{K}^n$, and that $\|x\|_\infty = \lim_{p \to \infty} \|x\|_p \quad \forall x \in \mathbb{K}^n$

Definition 2.3: $(X, \|\cdot\|)$ a normed space.
The induced metric / distance is

$d : X \times X \to \mathbb{R}$, $(x, y) \longmapsto d(x,y) = \|x - y\|$

Hence, any normed space is a metric space.

Example 2.4: Let (X, d) be a compact metric space and let $\mathcal{C}(X; \mathbb{K})$ be the vector space of continuous $\mathbb{K}$-valued functions defined on X.

Then $\|\cdot\|_\infty = \sup\{|f(x)| : x \in X\}$

is a norm on $\mathcal{C}(X; \mathbb{K})$

Solution:

i) $\|f\|_\infty \geq 0$ ✓

$f(x) = 0 \quad \forall x \in X \Rightarrow \|f\|_\infty = 0$

$\|f\|_\infty = 0 \Rightarrow \sup\{|f(x)| : x \in X\} = 0 \Rightarrow f(x) = 0$

~~$f(x) \leq |f(x)| \leq \sup$~~

(ii) ✓

(iii) $|(f+g)(y)| \leq |f(y)| + |g(y)| \leq$

$\|f\|_\infty + \|g\|_\infty$

$\Rightarrow \|f+g\|_\infty = \sup\{|(f+g)(x)| : x \in X\}$

$\leq \|f\|_\infty + \|g\|_\infty$

2.2 Completeness

Definition 2.5: (a) (X, d) metric space

A metric space (X, d) is complete if every Cauchy sequence in (X, d) is convergent.

(b) $(X, \|\cdot\|)$ is complete if

(X, d) is a complete metric space where

d is the induced metric.

(c) A Banach space is a normed space that is complete w.r.t. to the induced metric

Examples 2.6. (a) $(\mathbb{K}^n, \|\cdot\|_2)$ Banach space

(b) $\ell^p = \ell^p(\mathbb{K}) = \ell^p(\mathbb{N}; \mathbb{K})$

$x = (x_1, \dots,)$ $\qquad \sum |x_i|^p < \infty$

$1 \leq p < \infty$. ℓ^∞ space of bounded sequences,

i.e. $\sup |x_i| < \infty$

We now introduce L^p-spaces without using much of measure theory.
Recall that function spaces have a natural linear structure. If f, g are functions $[a,b] \to \mathbb{R}$, we define addition and multiplication by numbers as follows:

$$(f+g)(x) = f(x) + g(x) \quad , x \in [a,b]$$
$$(\alpha f)(x) = \alpha f(x) \quad , \alpha \in \mathbb{K}.$$

<u>Notation</u> $a \leq b$

The functions in $\mathcal{C}([a,b];\mathbb{K})$ are Riemann integrable (in case $\mathbb{K} = \mathbb{C}$ consider real and imaginary part).
The L^p-norm is defined as

$$\|f\|_p := \left(\int_a^b |f(x)|^p \, dx \right)^{1/p} \quad , p \in [1, \infty)$$

$$\|f\|_\infty := \max_{a \leq x \leq b} |f(x)|$$

<u>Proposition 2.7</u>: (a) If $f, g \in \mathcal{C}([a,b];\mathbb{K})$ and $p \in [1,\infty)$, then $\|f+g\|_p \leq \|f\|_p + \|g\|_p$

(b) The normed space $(\mathcal{C}([a,b]); \|\cdot\|_p)$ is not complete for $p \in [1, \infty)$.

Proof: Functionalanalysis I

Definition 2.7: $p \in [1, \infty)$

$L^p([a,b])$ is the completion of $\mathcal{C}([a,b])$ with respect to the metric d_p induced by the L^p-norm $\|\cdot\|_p$.

Proposition 2.8: The normed space $(\mathcal{C}([a,b]), \|\cdot\|_\infty)$ is complete

Proof: See functionalanalysis I or supervision class. □

Remark: The Lebesgue spaces L^p are normally constructed in measure theory / analysis. One introduces the Lebesgue measure on the Borel-σ-algebra, and defines the Lebesgue integral (via simple step functions and a convergence concept). $L^p(X;\mathbb{K})$ is then the space of all functions such that

$$\int_X |f(x)|^p < \infty .$$

However, this definition brings a difficulty, namely the first norm property is violated:
$\|f\|_p = 0 \not\Rightarrow f \equiv 0$ for some $f \in L^p(X;\mathbb{K})$.
This can easily be resolved by considering the quotient space, i.e, by considering equivalence classes of functions. The resulting space can be identified with the completion of $\mathcal{C}(X;\mathbb{K})$ with respect to the metric d_p

The advantage of the approach via Lebesgue measure is that we have the quite useful convergence theorems (monotone convergence, dominated convergence) at hand that allow interchange of limit and integral, and also Fubini's theorem that allows to interchange order of integration.

2.3 Separability and bases

Definition 2.9: Two norms, $\|\cdot\|_1$ and $\|\cdot\|_2$, on a vector space X are equivalent if there exist constants $c_1, c_2 > 0$ s.t.

$$c_1 \|x\|_1 \leq \|x\|_2 \leq c_2 \|x\|_1 \quad \text{for all } x \in X.$$

We defined (Def 1.14) dense subsets of topological spaces. If a topological space has a dense subset which is countable it will be called as follows

Definition 2.10: $(X, \mathcal{T})$ topological space. X is called separable if there ~~does~~ exists a dense countable subset, i.e. $\exists\ M = \{x_n \in X : n \in \mathbb{N}\}$ with $\overline{M} = X$.

Equivalently, a normed space is separable if there exists a dense countable subset.
Separability is a very important property.
It holds for: $\mathbb{R}^n$, ℓ^p and L^p $(p \in [1, \infty)$

But ℓ^∞ and L^∞ are not separable

Why is this property so important?
Recall the situation for finite dimensional vector spaces and the concept ~~and~~ of bases.
We are going to generalise this concept to arbitrary ("infinite dimensional") normed spaces.

Definition 2.11: Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a ~~the~~ normed space $(X, \|\cdot\|)$, i.e. $x_n \in X \ \forall n \in \mathbb{N}$.

The series $\sum_{n=1}^{\infty} x_n$ is said to converge to $x \in X$ if

$$\lim_{N\to\infty} \left\| \sum_{n=1}^{N} x_n - x \right\| = 0$$

The series is absolutely convergent if $\sum_{n=1}^{\infty} \|x_n\| < \infty$

Proposition 2.12:

A normed space $(X, \|\cdot\|)$ is complete $\Longleftrightarrow$ every absolutely convergent series converges

Proof: „$\Rightarrow$" any ~~all~~ Cauchy series converges. Suppose $(x_n)_{n\in\mathbb{N}}$ satisfies $\sum_{n=1}^{\infty} \|x_n\| < \infty$. Let $Y_N = \sum_{n=1}^{N} x_n$

Then, if $N \le M$,

$$\|Y_M - Y_N\| = \left\| \sum_{n=N+1}^{M} x_n \right\| \le \sum_{n=N+1}^{M} \|x_n\| \le \sum_{n=N+1}^{\infty} \|x_n\|$$

$$\sum_{n=N+1}^{\infty} \|x_n\| \leq \sum_{n=1}^{\infty} \|x_n\| < \infty$$
(assumption)

$\Rightarrow \forall \varepsilon > 0 \ \exists N_0$ s.t. $\displaystyle\sum_{n=N+1}^{\infty} \|x_n\| < \varepsilon \ \forall N \geq N_0$

$\Rightarrow (Y_N)_{N\in\mathbb{N}}$ Cauchy sequence, and, as $(X, \|\cdot\|)$ is complete, it converges to some $x \in X$, and the series converges to x.

"$\Leftarrow$" Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence.

$\forall n \in \mathbb{N} \ \exists N_n \in \mathbb{N}$ s.t. $\|x_k - x_\ell\| < 2^{-n}$

$\forall k, \ell > N_n$

W.l.o.g assume $N_{n+1} > N_n$.
Define the subsequence $(y_n)_{n\in\mathbb{N}} = (x_{N_n})_{n\in\mathbb{N}}$.

$(y_0 = 0) \qquad y_n = \displaystyle\sum_{k=1}^{n} (y_k - y_{k-1})$

then $\displaystyle\sum_{k=1}^{\infty} \|y_k - y_{k-1}\| \leq \sum_{k=1}^{\infty} 2^{-(k-2)} = 2 \sum_{k=0}^{\infty} 2^{-k}$

$= 4 < \infty$

then the series converges, hence $(y_n)_{n\in\mathbb{N}}$ convergence which implies that $(x_n)_{n\in\mathbb{N}}$ converges □

Definition 2.13: Let $(X, \|\cdot\|)$ be a Banach space

A sequence $(x_n)_{n\in\mathbb{N}}$ is a Schauder basis of X if for every $x \in X$ there is a unique sequence $(\alpha_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ ($\alpha_n \in \mathbb{K}\ \forall n \in \mathbb{N}$) s.t.

$$x = \sum_{n=1}^{\infty} \alpha_n x_n$$

Remarks:

(a) For the spaces $\ell^p, p \in [1,\infty)$, the sequences

$$x_n = (0, \dots, \underset{n\text{-th place}}{1}, 0 \dots) \quad \text{form a Schauder basis.}$$

(b) $(C([0,1]), \|\cdot\|_\infty)$ has Schauder bases.

(c) There exist separable Banach spaces that possess no Schauder bases! (Enflo 1973)

(d) Any separable Hilbert space has Schauder bases (ONB - orthonormal base).

(e) Any vector space (norm not required) has a Hamel basis (or algebraic basis) s.t. each vector is uniquely expressed a a finite linear combination. In ∞-dim. vector spaces, the construction requires the axiom of choice.

2.4 Linear operators (revision)

We now look at functions which map one normed space into another. The simplest maps between two vector spaces are the ones which respect the linear structure, that is, the linear transformations.

Definition 2.14 $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ normed spaces.

(a) A linear map, or linear operator, is a map $T: X \to Y$ s.t.

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y) \quad \forall \alpha, \beta \in \mathbb{K}, \forall x, y \in X$$

(b) The operator norm of the operator $T: X \to Y$ is defined by

$$\|T\| = \sup_{x \neq X} \frac{\|T(x)\|_Y}{\|x\|_X} = \sup_{\|x\|=1} \|Tx\|_Y$$

T is bounded if $\|T\| < \infty$, it is unbounded if $\|T\| = \infty$.

(c) Notation: $\mathcal{L}(X, Y)$ = the set of linear operators $X \to Y$
$\mathcal{B}(X, Y)$ = " bounded "

(Both are vector spaces.)

Examples 2.15:

(a) The space $\mathcal{C}^\infty([0,1];\mathbb{R})$ with the $\|\cdot\|_\infty$-norm

is uncomplete. $\qquad M: \mathcal{C}^\infty([0,1];\mathbb{R}) \to Y$

$$(Mu)(x) = x\,u(x)$$
$$(Du)(x) = u'(x)$$

M is bounded and D is unbounded.

(b) On the space $\ell^\infty(\mathbb{R})$ we define the following operator A given a matrix (infinite matrix)

$(a_{ij})_{i,j\geq 1}$

$$(Ax)_i = \sum_{j=1}^{\infty} a_{ij} x_j \qquad (*)$$

It is necessary that $\displaystyle\sum_{j=1}^{\infty} |a_{ij}| < \infty \;\; \forall i \in \mathbb{N}$

in order that the series on the right hand side in $(*)$ converges for all $x \in \ell^\infty(\mathbb{R})$. Furthermore,

$$\|A\| = \sup_{\|x\|_\infty = 1} \|Ax\|_\infty = \sup_{\|x\|_\infty = 1} \sup_{i \geq 1} |(Ax)_i|$$

$$= \sup_{i\geq 1} \sup_{\|x\|_\infty = 1} \Big|\sum_{j\geq 1} a_{ij} x_j\Big| = \sup_{i \geq 1} \sum_{j\geq 1} |a_{ij}|$$

Hence, if $\|A\| < \infty$, $Ax \in \ell^\infty(\mathbb{R})$ $\forall x \in \ell^\infty(\mathbb{R})$

and the operator is well-defined. But if $\|A\| = \infty$, the action of A is defined only for a subset of $\ell^\infty(\mathbb{R})$. This subset will be called domain in the theory of unbounded operators.

(c) $a, b \in \mathbb{R}$, $k : [a,b] \times [a,b] \to \underset{(\mathbb{K})}{\mathbb{C}}$ be continuous

(i) If $g \in (\mathcal{C}([a,b]), \|\cdot\|_\infty)$, then $f : [a,b] \longrightarrow \mathbb{K}$ defined by

$$f(s) = \int_a^b k(s,t)\, g(t)\, dt \quad \text{is in } \mathcal{C}([a,b], \|\cdot\|_\infty)$$

(ii) Let $K : \mathcal{C}([a,b]) \longrightarrow \mathcal{C}([a,b])$ be defined by

$$(K(g))(s) = \int_a^b k(s,t)\, g(t)\, dt.$$

Then $K \in \mathcal{B}\big((\mathcal{C}([a,b]), \|\cdot\|_\infty), \mathcal{C}([a,b]), \|\cdot\|_\infty)\big)$

and

$$\|K\| = \max_{a \le t \le b} \int_a^b |k(t,s)|\, ds.$$

k is called the kernel of K.

<u>**Lemma 2.16:**</u> Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces and $T: X \to Y$ linear transformation / linear operator. The following statements are equivalent

(a) T is uniformly continuous

(b) T is continuous

(c) T is continuous at 0

(d) $\exists\, C > 0$ such that $\|T(x)\|_Y \leq C$ whenever $x \in X$ and $\|x\|_X \leq 1$ (i.e. $\forall x \in B_1(0)$)

(e) $\exists\, C > 0$ s.t. $\|T(x)\|_Y \leq C\|x\|_X \quad \forall x \in X$.

<u>Proof:</u> exercise (sheet 2) :
$(a) \to (b) \to (c)$ is clear.

It is enough to define a continuous / bounded operator on a dense subspace.
There exists a unique extension to the whole space.

Proposition 2.17: Let $(X, \|\cdot\|_X)$ be a normed space and $(Y, \|\cdot\|_Y)$ be a Banach space. Further, let M be a dense subset of X (i.e. $\overline{M} = X$), and $T: M \to Y$ a bounded operator.
Then there exists a unique operator $\overline{T}: X \to Y$ s.t.

$$\overline{T}x = Tx \qquad \text{for all } x \in M.$$

In addition, $\|\overline{T}\| = \|T\|$.

Proof: M dense subset of $X \Rightarrow \forall x \in X \ \exists$ sequence $(x_n)_{n\in\mathbb{N}}$ in M that converges to x.
Then $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, and so is $(Tx_n)_{n\in\mathbb{N}}$ because T is bounded ($\|Tx_m - Tx_k\|_Y \leq C\|x_m - x_k\|_X$). Since $(Y, \|\cdot\|_Y)$ is complete, the sequence $(Tx_n)_{n\in\mathbb{N}}$ converges to some $y \in Y$.
Define $\overline{T}x = y$. Does this depends on the choice of the sequence $(x_n)_{n\in\mathbb{N}}$? No, because if $(\tilde{x}_n)_{n\in\mathbb{N}}$ is another converging sequence with limit x, we have

$$\|T\tilde{x}_n - y\|_Y \leq \|T\tilde{x}_n - Tx_n\|_Y + \|Tx_n - y\|_Y$$

$$\leq \|T\| \underbrace{\|\tilde{x}_n - x_n\|_X}_{\leq \|\tilde{x}_n - x\|_X + \|x - x_n\|_X} + \|Tx_n - y\|_Y \to 0 \text{ as } n \to \infty$$

Then $T\tilde{x}_n = y$.

$$\|\bar{T}\| \geq \sup_{\substack{x \in \Pi \\ x \neq 0}} \frac{\|\bar{T}x\|_Y}{\|x\|_Y} = \|T\|.$$

If $(x_n)_{n \in \mathbb{N}}$ with $x_n \to x$ as $n \to \infty$

$$\|\bar{T}x\|_Y = \lim_{n \to \infty} \|Tx_n\|_Y \leq \|T\| \lim_{n \to \infty} \|x_n\|_X = \|T\| \, \|x\|_X$$

$\Rightarrow \|\bar{T}\| \leq \|T\|$. Hence, $\|\bar{T}\| = \|T\|$.

Finally, suppose that $\tilde{T}$ is another extension of T, then

$$\tilde{T}x = \lim_{n \to \infty} \tilde{T}x_n = \lim_{n \to \infty} \bar{T}x_n = \bar{T}x. \quad \blacksquare$$

Definition 2.18: X, Y be vector space and $T : X \to Y$ a linear map

(a) The kernel of T, or null space, is

$$\ker T = \{ x \in X : Tx = 0 \}$$

The dimension of $\ker T$ is the nullity of T.

(b) The range of T, or image, is

$$\operatorname{ran} T = \{ \cancel{y=} Tx : x \in X \}$$

The dimension of $\operatorname{ran} T$ is the rank of T

Lemma 2.19: $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ normed spaces. $T : X \to Y$ linear operator. If $\|T\| < \infty$, then

$\ker T$ is closed.

Proof: Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in $\ker T$ with $\lim\limits_{n \to \infty} x_n = x \in X$. Then $Tx = \lim\limits_{n \to \infty} T x_n = 0$ (continuity), hence $x \in \ker T$. $\blacksquare$

Let us look in more detail at the normed space $\mathcal{B}(X,Y)$, where $(X,\|\cdot\|_X)$ and $(Y,\|\cdot\|_Y)$ are normed spaces.

Is $(\mathcal{B}(X,Y),\|\cdot\|)$ a Banach space?

A natural guess might suggest that it would be related to completeness of $(X,\|\cdot\|_X)$ and $(Y,\|\cdot\|_Y)$.

This is only half correct.

<u>**Theorem 2.20:**</u> $(X,\|\cdot\|_X)$ a normed space and $(Y,\|\cdot\|_Y)$ a Banach space.

Then $(\mathcal{B}(X,Y),\|\cdot\|)$ is a Banach space (here $\|\cdot\|$ is the above defined operator norm)

<u>Proof:</u> Let $(T_n)_{n\in\mathbb{N}}$ be a Cauchy sequence. $(\mathcal{B}(X,Y), d_{\|\cdot\|})$ is also a metric space ($d_{\|\cdot\|}$ = induced metric), and in a metric space any Cauchy sequence is bounded, therefore $\exists$ $M>0$ s.t.

$$\|T_n\| \leq M \quad \forall\, n\in\mathbb{N} \quad . \text{ Pick } x\in X.$$

$$\|T_n x - T_m x\|_Y \leq \|T_n - T_m\|\, \|x\|_X$$

$(T_n)_{n\in\mathbb{N}}$ Cauchy sequence implies that also

$(T_n x)_{n\in\mathbb{N}}$ is a Cauchy sequence in Y. As Y is complete,

$$T(x) = \lim_{n\to\infty} T_n x \quad \text{exists}.$$

We have thus defined a map $T: X\to Y$.

We will show that $T \in \mathcal{B}(X,Y)$ and that $\lim\limits_{n\to\infty} T_n = T$

$$T(x+y) = \lim_{n\to\infty} T_n(x+y) = \lim_{n\to\infty} T_n(x) + \lim_{n\to\infty} T_n(y) = Tx + Ty$$

$$T(\alpha x) = \lim_{n\to\infty} T_n(\alpha x) = \alpha \lim_{n\to\infty} T_n x = \alpha T(x).$$

$$\|Tx\|_Y = \lim_{n\to\infty} \|T_n x\|_Y \le M \|x\|_X$$

implies that $T \in \mathcal{B}(X,Y)$.

Let $\varepsilon > 0$. $\exists\ N \in \mathbb{N}$ s.t. $\|T_n - T_m\| < \frac{\varepsilon}{2}\ \forall m,n \geq N$

$\forall x \in X$ with $\|x\|_X \le 1\ \forall m,n \geq N$

$$\|T_n x - T_m x\|_Y \le \|T_n - T_m\| \|x\| < \frac{\varepsilon}{2}$$

$\exists\ N_1 \in \mathbb{N}$ s.t. $\|Tx - T_m x\| < \varepsilon/2 \quad \forall m \geq N_1$

Then ($n \geq N$ and $m \geq N_1$)

$$\|Tx - T_n x\|_Y \le \|Tx - T_m x\|_Y + \|T_n x - T_m x\|$$
$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2}\|x\| \le \varepsilon$$

$\Longrightarrow \lim\limits_{n\to\infty} T_n = T$ $\blacksquare$

We conclude this chapter with a few facts about finite-dimensional Banach spaces. (revision)

Proposition 2.20:

(a) Any finite-dimensional normed space is complete.

(b) Any finite-dimensional subspace of a normed space is closed.

(c) Any operator on a finite-dimensional normed space is bounded

(d) Any two norms on a finite-dimensional linear space are equivalent.

(3) Duality and main theorems

3.1 The Hahn-Banach theorem

Definition 3.1: Let $(X, \|\cdot\|)$ be a normed space

Linear transformations $\lambda : X \to \mathbb{K}$ are called linear functionals.
The space $\mathcal{B}(X, \mathbb{K})$ is called the dual space of X and denoted by X^* or X'

Corollary 3.2 (recall Theorem 2.20)

If $(X, \|\cdot\|_X)$ is a normed space then $(X', \|\cdot\|)$ is a Banach space. This follows from Theorem 2.20 as $\mathbb{K}$ is complete. (here $\|\cdot\|$ is again the operator norm).

In this chapter we give various formulations of the so called Hahn-Banach theorem.
This theorem is one of the most important theorems in analysis. It will give us many conclusions on dual spaces and some applications.

We start with the first version. Suppose X is a real or complex vector space. It often happens that we have a linear functional

$$f_W : W \to \mathbb{K}$$

defined in a natural way on a subspace $W \subset X$,

(W is often finite dimensional)

but to make use of this functional we require it to be defined on the whole of X.
Can we extend the domain of the functional to the whole of X?

To describe the 'size' of a functional we introduce the following concepts.

Definition 3.3:

(a) Let X be a real vector space. A sublinear functional on X is a function $p: X \to \mathbb{R}$ such that:

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

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

(b) Let X be a real or complex vector space. A seminorm on X is a real-valued function $p: X \to \mathbb{R}$ such that:

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

(ii) $p(\alpha x) = |\alpha| p(x), \quad x \in X, \alpha \in K.$

Examples 3.4:

(a) $f \in B(X, \mathbb{R}) \Rightarrow p(x) = |f(x)|$ is not linear but sublinear (when $f \neq 0$)

(b) $(X, \|\cdot\|)$ normed space, then $p(x) = \|x\|$ is sublinear

(c) $X = \mathbb{R}^2$; $p(x_1, x_2) = |x_1| + x_2$

p is sublinear

Note that (a) and (b) are seminorms, even on complex spaces.

We will prove the following versions of the Hahn-Banach theorem

Theorem 3.5	real vectorspaces	
Theorem 3.9	complex vector spaces	
Theorem 3.10	for normed space	— main result

<u>Theorem 3.5</u> The Hahn-Banach Theorem (version: real vector spaces)

Let X be a real vector space, with a sublinear functional p defined on X.
Suppose that W is a vector subspace of X and

$f_W : W \to \mathbb{R}$ linear functional satisfying

$$f_W(w) \leq p(w) \quad \forall w \in W.$$

Then f_W has an extension $f_X : X \to \mathbb{R}$ s.t.

$$f_X(x) \leq p(x), \quad x \in X.$$

The proof is difficult. It is based on Zorn's lemma. We need to define the following concepts.

__Definition 3.6:__ Suppose $M \neq \emptyset$ is a non-empty set and $\prec$ is an ordering on M. Then $\prec$ is a partial order on M if

(a) $x \prec x$ for all $x \in M$;

(b) if $x \prec y$ and $y \prec x$ then $x = y$;

(c) if $x \prec y$ and $y \prec z$ then $x \prec z$.

M is the a partially ordered set.
If, in addition, $\prec$ is defined for all pairs of elements (that is, for any $x, y \in M$, either $x \prec y$ or $y \prec x$ holds), then $\prec$ is a total order on M.

If M is a partially ordered set, then

$y \in M$ is a maximal element of M if $y \prec x \Rightarrow y = x$.

An element $x \in X$ is an upper bound of a subset $N \subset M$ if $z \prec x \ \forall z \in N$.

__Remarks:__ (i) The usual order $\leq$ on $\mathbb{R}$ is a total order

(ii) A partial order on $\mathbb{R}^2$ is given by $(x_1, x_2) \prec (y_1, y_2)$

$\iff \quad x_1 \leq y_1$ and $x_2 \leq y_2$.

Lemma 3.7 Zorn's Lemma

Let M be a non-empty, partially ordered set, such that every totally ordered subset of M has an upper bound. Then there exists a maximal element in M.

Axiom 3.8 Axiom of choice

For any non-empty set M, there exists a map $c: \mathcal{P}(X) \longrightarrow X$ ("choice function) such that $c(Y) \in Y$ for any non-empty subset $Y \in \mathcal{P}(X)$.

Proof of Theorem 3.5:

Let $\mathcal{E} \subset \mathcal{L}(X, \mathbb{R})$ the set of linear functionals $f: X \to \mathbb{R}$ satisfying

(a) f is defined on a subspace D_f, such that $W \subset D_f \subset X$;

(b) $f(w) = f_W(w)$, $w \in W$;

(c) $f(x) \leq p(x)$, $x \in D_f$

$\mathcal{E}$ is the set of all extensions f of f_W, to general subspaces $D_f \subset X$.

<u>Idea</u>: Apply Zorn's lemma to the set $\mathcal{E}$, and show that the resulting maximal element of $\mathcal{E}$ is the desired functional

<u>Verification of the hypotheses of Zorn's lemma</u>:

- $f_W \in \mathcal{E} \Rightarrow \mathcal{E} \neq \emptyset$

- define for any $f, g \in \mathcal{E}$

$$f \prec g \iff D_f \subset D_g \text{ and } f(x) = g(x) \ \forall x \in D_f$$

($f \prec g \iff g$ is an extension of f); $\prec$ is a partial order.

- $\mathcal{G} \subset \mathcal{E}$ totally ordered:
we construct an upper bound for $\mathcal{G}$ in $\mathcal{E}$.

Define $\displaystyle Z_{\mathcal{G}} = \bigcup_{f \in \mathcal{G}} D_f \quad \subset X.$

$Z_{\mathcal{G}}$ is a subspace (using the total ordering of $\mathcal{G}$), and we now define a linear functional $f_{\mathcal{G}}$ on $Z_{\mathcal{G}}$ as follows:
Pick $z \in Z_{\mathcal{G}}$, then $\exists\ \xi \in \mathcal{E}$ s.t. $z \in D_\xi$;
then define

$$f_{\mathcal{G}}(z) = \xi(z)$$

(independent on ξ. Why?) $\rightarrow$ because $\mathcal{G}$ is totally ordered that is, $\exists\ \xi'$ with either $\xi' \prec \xi$ or $\xi \prec \xi'$, i.e. $\xi'(z) = \xi(z)$)

$f_{\mathcal{G}}$ is linear, and

$f_{\mathcal{G}}(z) = \xi(z) \leq p(z)$, and if $z \in W$ then $f_{\mathcal{G}}(z) = f_W(z)$

$\Rightarrow f_{\mathcal{G}} \in \mathcal{E}$ and $f \prec f_{\mathcal{G}} \quad \forall f \in \mathcal{G}$.

Thus $f_{\mathcal{G}}$ is an upper bound for $\mathcal{G}$.

Now Zorn's lemma gives the existence of a maximal element f_{max} of $\mathcal{E}$.
Suppose $D_{f_{max}} \neq X$. We can then again apply our arguments to get an extension of f_{max}, which also lies in $\mathcal{E}$.
However, this contradicts the maximality of f_{max} in $\mathcal{E}$, so $D_{f_{max}} = X$, and hence $f_X = f_{max}$. ∎

"one dimensional extension" process
X real vector space, $W \subset X$ proper subspace of X, p sublinear functional on X, f_W linear functional on W with $f_W(w) \leq p(w) \ \forall w \in W$
$z_1 \notin W$

$W_1 = \{\alpha z_1 + w : \alpha \in \mathbb{R}, w \in W\}$

$\Rightarrow \exists \ \xi_1 \in \mathbb{R}$ and $f_{W_1} : W_1 \to \mathbb{R}$

$$f_{W_1}(\alpha z_1 + w) = \alpha \xi_1 + f_W(w) \leq p(\alpha z_1 + w)$$

Hint: $\left(\xi_1 = \inf_{v \in W} \{-f_W(v) + p(v + z_1)\} > -\infty \right)$

Theorem 3.9 The Hahn-Banach Theorem (version: complex vector spaces)

Let X be a complex vector space and p a seminorm on X. Suppose that W is a vector subspace of X and $f_W : W \to \mathbb{C}$ a complex linear functional satisfying

$$|f_W(w)| \leq p(w) \qquad \forall w \in W.$$

Then f_W has an extension $f_X : X \to \mathbb{C}$ such that

$$|f_X(x)| \leq p(x) \quad \forall x \in X.$$

<u>Proof.:</u> We write $f_W(x) = f_{W,1}(x) + i f_{W,2}(x)$,

where $f_{W,1}$ and $f_{W,2}$ are real valued linear functionals. Check that

$$i\,(f_{W,1}(x) + i f_{W,2}(x)) = i f_W(x) = f_W(ix) = f_{W,1}(ix)$$

$+ i f_{W,2}(ix)$ implies $f_{W,2}(x) = -f_{W,1}(ix)$, so that

$$f_W(x) = f_{W,1}(x) - i f_{W,1}(ix).$$

We have $f_{W,1}(x) \leq |f_W(x)| \leq p(x)$, and a seminorm is a sublinear functional. Hence, using Theorem 3.5 we get an extension $f_{X,1}$ of $f_{W,1}$ on X.
Then define

$$f_X(x) := f_{X,1}(x) - i f_{X,1}(ix) \quad \forall x \in X,$$

and check that f_X is a complex extension of f_W on X, Additional comments to the proof of Theorem 3.9:

Recall $f_{W,1}(x) \leq |f_W(x)| \leq p(x)$

p seminorm ; and a seminorm is a sublinear functional

$\Rightarrow f_{W,1}(x)$ has an extension on X_R, called $f_{X,1}^{(R)}$.

Here X_R is the real vectorspace obtained from X by simply restricting the scalar multiplication to real numbers.

We have $|f_{X,1}^{(R)}(x)| \leq p(x) \quad x \in X_R$.

Then f_W has an extension f_X on X such that

$$|f_X(x)| \leq p(x) \quad \forall x \in X$$

To see this, define

$$f_X(x) = f_{X,1}^{(R)}(x) - i f_{X,1}^{(R)}(ix),$$

then clearly f_X is an extension. Moreover, with $\alpha = \arg(f_X(x))$ we get (see page 46):

and with $\alpha = \arg(f_X(x))$

$$|f_X(x)| = (\cos\alpha - i\sin\alpha)\,|f_X(x)|\,(\cos\alpha + i\sin\alpha)|$$

$$= e^{-i\alpha} f_X(x) = \underbrace{f_X(e^{-i\alpha}x)}_{\in \mathbb{R} \ (\Leftrightarrow f_{X,1}(ie^{-i\alpha}x)=0)} = f_{X,1}(e^{-i\alpha}x) \le p(x)$$

for any $x \in X$. (The last inequality uses the fact that p is a seminorm, a sublinear functional is not enough) ∎

Note further that complex linear functionals $f: X \to \mathbb{C}$ satisfy $\|f\| = \|f_1\|$ where

$f(x) = f_1(x) - i f_1(ix)$. Indeed, the

inequality $\|f_1\| \le \|f\|$ is obvious. And since

$|f(x)| = f_1(e^{-i\alpha}x)$, where $\alpha = \arg(f(x))$,
we have $|f(x)| \le \|f_1\|\,\|x\| \quad \forall x \in X$, hence $\|f\| \le \|f_1\|$ ∎

The next result is the main result of this chapter. In principle, it follows immediately from Theorem 3.9. However, in all application we mostly use the following general version. We will only sketch the proof for this general statement. Note that in general we have not specified if the normed space has to be separable or not. The Hahn-Banach Theorem holds for both, only the techniques of the proofs are different.

Theorem 3.10: The Hahn-Banach Theorem for Normed spaces

Let $(X, \|\cdot\|)$ be a real or complex normed space and W a linear subspace of X.

For any $f_W \in W'$ there exists an extension $f_X \in X'$ of f_W such that $\|f_X\| = \|f_W\|$

Proof.: (Sketch)

$p(x) = \|f_W\| \|x\|$ for $x \in X$. Then p is a seminorm on X.

W.l.o.g. $W \neq X$ and W is closed (Why?)

We get an extension f_X with $|f_X(x)| \leq \|f_W\| \|x\| \; \forall x \in X$

This shows that $\|f_X\| \leq \|f_W\|$, but we also have $\|f_X\| \geq \|f_W\|$ because f_X is an extension of f_W. Hence $\|f_X\| = \|f_W\|$.

(alternative method: the "one-dimensional extension" process mentioned earlier). ∎

Remarks: The Hahn-Banach theorem was proved independently by Hahn (1926) and by Banach (1929).

For main applications of this theorem see E. Zeidler: Applied Functional Analysis

There are many applications of the Hahn-Banach Theorem. See the following figure

An important consequence of the Hahn-Banach theorem is that the ~~dual~~ dual space of a normed space is big enough to separate (distinguish) its elements.

<u>**Corollary 3.11:**</u> (Separability of a normed space)

Let $(X, \|\cdot\|_X)$ be a normed space, and $(X^*, \|\cdot\|)$ its dual space.
If $x \neq y$, $x, y \in X$, then there exists $f \in X^*$ such that

$$f(x) \neq f(y)$$

<u>Proof:</u> If $x \neq y$, we can consider the linear space W spanned by $x-y$, and the linear functional $f_W : W \to \mathbb{K}$ such that $f_W(x-y) = 1$.
By Theorem 3.10, there exists an extension $f_X \in X^*$, and

$$f_X(x) - f_X(y) = f_X(x-y) = f_W(x-y) = 1 \neq 0$$

■

There exist a general separation Theorem which is based on geometrical ideas. We will only outline this direction briefly.

<u>**Definition 3.12:**</u> Let X be a vector space. A hyperplane in X (through $x_0 \in X$) is a set of the form

$H = x_0 + \ker h \subset X$, where $h \neq 0$ is a linear functional on X.
Equivalently, $H = h^{-1}(\gamma)$, where $\gamma = h(x_0)$.

A hyperplane in $\mathbb{R}^n$ is an $(n-1)$-dimensional plane.
The general "Separation Theorem" shows that convex subsets can be separated by hyperplanes.

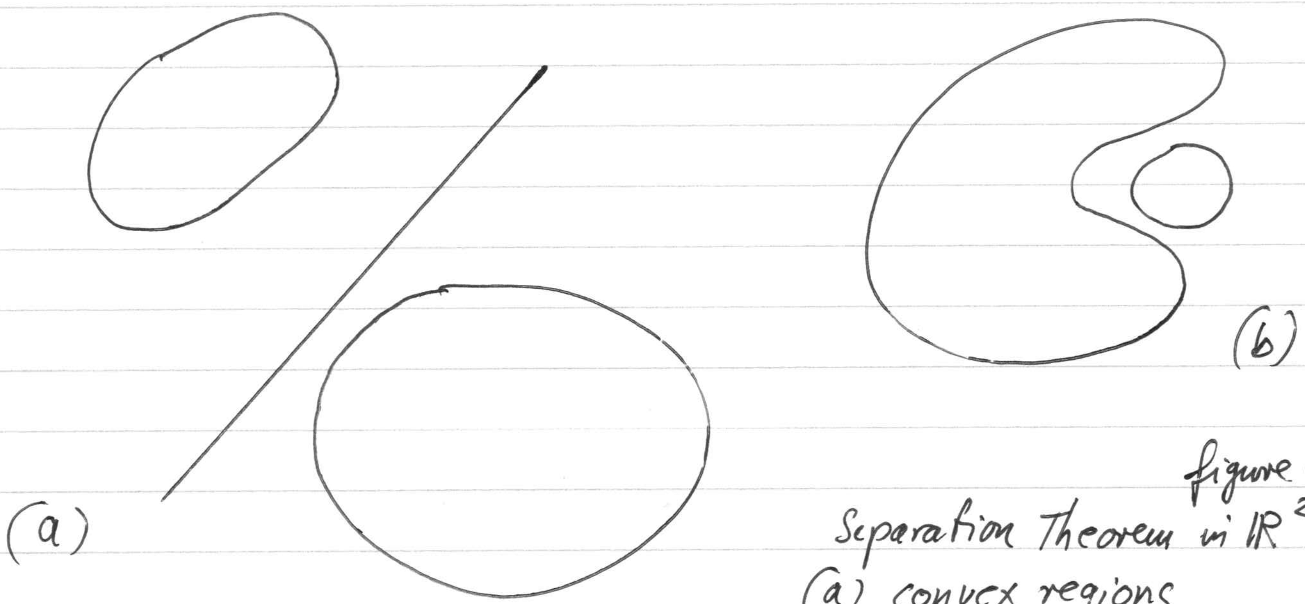


figure ②
Separation Theorem in $\mathbb{R}^2$:
(a) convex regions
(b) non-convex regions

We already know that (see Corollary 3.2) the dual $(X', \|\cdot\|)$ of a normed space $(X, \|\cdot\|_X)$ is a Banach space.

Definition 3.12: For any normed space $(X, \|\cdot\|_X)$, the space X'' (or X^{**}) is called the second dual of X, or bidual. X'' is the dual of the dual space X'

Lemma 3.13: $(X, \|\cdot\|_X)$ normed space

For any $x \in X$, define $F_x : X' \to \mathbb{R}$

$$F_x(f) = f(x), \quad f \in X'$$

Then $F_x \in X''$ and $\|F_x\| = \|x\|_X$

Proof: $\alpha, \beta \in \mathbb{K}$ and $f, g \in X'$

- $F_x(\alpha f + \beta g) = \alpha f(x) + \beta g(x) = \alpha F_x(f) + \beta F_x(g)$
- $\|F_x(f)\| = |f(x)| \leq \|f\| \|x\|_X$, and so $F_x \in X''$, with $\|F_x\| \leq \|x\|_X$.
- $\|x\|_X = \sup\limits_{\|f\|=1} |f(x)| = \sup\limits_{\|f\|=1} |F_x(f)| = \|F_x\|$

$\blacksquare$

Remark: formally we have $X \subset X''$.

Definition 3.14

A normed space is called ~~refl~~ reflexive if all elements of X'' $(=X^{**})$ can be isometrically identified with a unique element of X.

Lemma 3.15: $(X, \|\cdot\|_X)$ normed space, $(X'', \|\cdot\|)$ its bidual.

(a) The mapping $J_X : X \to X''$, $x \mapsto J_X x$ with

$$(J_X x)(f) = f(x) \quad , f \in X'$$

is a linear isometry, that is X is isometrically isomorphic to a subset of X''.

(b) X is isometrically isomorphic to a dense subset of a Banach space

Proof: (a) is clear from Lemma 3.13.
(b) $(X'', \|\cdot\|)$ is a Banach space, and hence the closure of $J_X(X)$ is a Banach space. ∎

Examples 3.16:

(a) The dual of $(\ell^p(\mathbb{K}), \|\cdot\|_p)$, $p \in [1, \infty)$, is $(\ell^q(\mathbb{K}), \|\cdot\|_q)$ with $\frac{1}{p} + \frac{1}{q} = 1$:

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

$$f_y(x) = \sum_{n=1}^{\infty} x_n y_n \qquad \text{for } x = (x_n)_{n\in\mathbb{N}} \in \ell^p(\mathbb{K})$$

$f_y \in (\ell^p(\mathbb{K}))'$ with $\|f_y\| = \|y\|_q$

(note that $q = \infty$ respectively $q = 1$ if $p = 1$ resp. $p = \infty$)

$1 \leq p < \infty$: $(\ell^q(\mathbb{K}), \|\cdot\|_q) \to ((\ell^p(\mathbb{K}))', \|\cdot\|)$

is isometrically isomorphic

$((\ell^\infty(\mathbb{K}))', \|\cdot\|)$ and $(\ell^1(\mathbb{K}), \|\cdot\|_1)$ are <u>not</u>

isometrically isomorphic, (see ~~exer~~ exercise).
that is the dual of $\ell^\infty(\mathbb{K})$ is much larger than $\ell^1(\mathbb{K})$.

(6) Let Ω be any measure space, and consider the Lebesgue spaces $L^p(\Omega)$.
The dual of $L^p(\Omega)$ is $L^q(\Omega)$ with $1/p + 1/q = 1$, $p \in [1, \infty)$.
The dual of $L^\infty(\Omega)$ is much larger than $L^1(\Omega)$, unless Ω is a finite set.

<u>Proposition 3.17:</u> If $1 < p < \infty$, then $\ell^p(\mathbb{K})$ is reflexive.

<u>Proof:</u> exercise (example sheet) ■

We conclude this section with a discussion of weak and weak-$*$ convergence:
The importance of compactness in analysis is well-known

(closed bounded sets in finite dimensional spaces are compact).
This does not hold in general in infinite dimensional spaces. The idea is to adopt a weaker definition of convergence of a sequence, so far we considered $x_n \to x$ as $n \to \infty$ if $\|x_n - x\| \to 0$ as $n \to \infty$.

<u>Definition 3.18:</u> Let $(X, \|\cdot\|_X)$ be a Banach space and $(x_n)_{n\in\mathbb{N}}$ a sequence ~~and~~ in X, and $(f_n)_{n\in\mathbb{N}}$ a sequence in X'.

(a) $(x_n)_{n\in\mathbb{N}}$ is weakly convergent to $x \in X$ if

$$\lim_{n\to\infty} f(x_n) = f(x) \quad \text{for all } f \in X'$$

(b) $(f_n)_{n\in\mathbb{N}}$ is weak-$*$ convergent to $f \in X'$

if $\lim\limits_{n\to\infty} f_n(x) = f(x)$ for all $x \in X$.

<u>**Notation:**</u> $x_n \rightharpoonup x$, $f_n \overset{*}{\rightharpoonup} f$

<u>Remark:</u> If $\|x_n - x\| \to 0$ as $n \to \infty$, then $x_n \rightharpoonup x$;

both notions of convergence are equivalent in finite-dimensional normed spaces. If $x_n \rightharpoonup x$ and $x_n \rightharpoonup y \Rightarrow x = y$ (exercise).

The notion of weak-* convergence is important because of the following compactness result.

Theorem 3.19 (Banach–Alaoglu)

Let $(X, \|\cdot\|_X)$ be a normed space. Then the closed unit ball

$$B_1(0) = \{ f \in X' : \|f\| \leq 1 \} \text{ in } X'$$

is weak-* compact (i.e. compact with respect to the weak-* topology)

Remarks: A slightly weaker but more transparent statement is that,
if $(f_n)_{n\in\mathbb{N}}$ is a bounded sequence in X', i.e,

$\exists\, c>0$ such that $\|f_n\| \leq c \;\; \forall n \in \mathbb{N}$,
then there exists a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ of $(f_n)_{n\in\mathbb{N}}$ that converges pointwise
(i.e. $f_{n_k}(x) \to f(x)$ as $k\to\infty$ for all $x\in X$).

This theorem is very useful in minimisation problems.

<u>Proof:</u> It uses Tychonoff's theorem (if $(X_\alpha)_{\alpha\in A}$ is a family of compact topological space, then $\prod_{\alpha\in A} X_\alpha$ is compact in the product topology), whose proof requires the axiom of choice.

$\forall x \in X$ define the closed ball

$$B_x = \{ y \in \mathbb{K} : |y| \leq \|x\|_X \}$$

B_x is compact in $(\mathbb{K}, |\cdot|)$, then $\mathcal{B} = \prod_{x\in X} B_x$ is compact.

$B_1(0) \subset \mathcal{B}$ because $\circledast$ $f \in \mathbb{K}^X$

$$|f(x)| \leq \|f\| \|x\|_X \leq \|x\|_X$$

and hence $f : X \to \mathbb{K}^{\circledast}$ with $f(x) \in B_x \ \forall x \in X$

The topology in $B_1(0)$ inherited from $\mathcal{B}$ is equivalent to the topology of pointwise convergence, i.e. to the weak-$*$ topology.

It is easy to see that $B_1(0)$ is closed:
$(f_n)_{n\in\mathbb{N}}$ with $f_n \to f$, then

$$f(\alpha x + \beta y) = \lim_{n\to\infty} f_n(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

$\Rightarrow f \in B_1(0)$

because an element $f \in \mathcal{B} \subset \mathbb{K}^X$ is in $B_1(0)$

$\iff$ f is linear. $\square$

3.2 The Baire category Theorem and its consequences

The Baire category theorem is an abstract and technical theorem that applies to complete metric spaces. It has surprising, asthonishing applications to Banach spaces:
the open and inverse mapping theorem, the closed graph theorem, and the uniform boundedness theorem.

As a preparation we discuss the following geometric property first.

<u>Proposition 3.20:</u> Let (X,d) be a complete metric space. For any $n \in \mathbb{N}$ let K_n be a closed ball with radius $r_n \geq 0$ and center $x_n \in X$. If $K_{n+1} \subset K_n$, $n \in \mathbb{N}$, and $\lim_{n\to\infty} r_n = 0$, then $K := \bigcap_{n=1}^{\infty} K_n \neq \emptyset$ contains exactly one element.

<u>Proof:</u>
Step 1: $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence

$\forall \varepsilon > 0 \ \exists n_0 \in \mathbb{N}$ with $r_{n_0} < \frac{\varepsilon}{2}$

$\forall n \geq n_0$ we have $x_n \in K_n \subset K_{n_0}$ and thus

$\forall n, m \geq n_0$: $d(x_n, x_m) \leq d(x_n, x_{n_0}) + d(x_{n_0}, x_m)$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

$\Rightarrow (x_n)_{n \in \mathbb{N}}$ Cauchy sequence with limit $\hat{x} \in X$

Pick $k_0 \in \mathbb{N}$.

Step 2: ✓ Subsequence $(x_{n+k_0})_{n \in \mathbb{N}}$ also converges to $\hat{x}$ and it is element of K_{k_0}, hence $\hat{x} \in K_{k_0}$ as K_{k_0} is closed.

As $k_0 \in \mathbb{N}$ was arbitrary ~~that~~ we get $\hat{x} \in K_n \ \forall n \in \mathbb{N}$ and hence

$$\hat{x} \in \bigcap_{n=1}^{\infty} K_n = K \neq \emptyset$$

Assume $x, y \in K \Rightarrow x, y \in K_n \ \forall n \in \mathbb{N}$

$d(x,y) \leq d(x, x_n) + d(x_n, y) \leq r_n + r_n = 2r_n$

$\Rightarrow d(x,y) = 0 \quad \forall x, y \in K.$

Definition 3.21 : Let (X,d) be a metric space.

A subset $M \subset X$ is called

- nowhere dense (or rare) in X if its closure $\overline{M}$ has no interior points;
- meager (or of the first category) in X if M is given by a countable union of nowhere dense sets;
- nonmeager (or of the second category) in X if M is not meager in X.

Examples 3.22

(a) Any finite subset of $\mathbb{R}$ is nowhere dense:

$M \subset \mathbb{R}$ finite. You know from analysis that M is closed, so we are left to show that $\mathring{M} = \emptyset$.
Assume $x \in \mathring{M}$
$\Rightarrow \exists\, r > 0 : B_r(x) = (x-r, x+r) \subset M$. But the open interval $B_r(x)$ contains ∞ many points.
Remember: $\mathring{\overline{M}}$ is the "biggest" open subset contained in M.

(b) $\mathbb{Q}$ is meager in $\mathbb{R}$, because $\mathbb{Q}$ is equal to the countable union of its elements; $\mathbb{Q} = \bigcup_{q \in \mathbb{Q}} \{q\}$.

More generally (compare (a)), any countable set in $\mathbb{R}$ is meager in $\mathbb{R}$.

(c) Let (X, τ) be a topology space. Then the following statements are equivalent

(i) X is of the second category

(ii) For any sequence $(Y_n)_{n\in\mathbb{N}}$ of closed nowhere dense subsets of X:
$$\bigcup_{n\in\mathbb{N}} Y_n \neq X$$

Proof: (i) $\Rightarrow$ (ii) clear

(ii) $\Rightarrow$ (i): Let $(M_n)_{n\in\mathbb{N}}$ a sequence of nowhere dense subsets of X.
Define $\quad Y_n = \overline{M_n} \quad \forall n \in \mathbb{N}$.

$\bigcup_{n\in\mathbb{N}} M_n \subset \bigcup_{n\in\mathbb{N}} Y_n \neq X$. Hence, we conclude if

Y_n is nowhere dense $\forall n \in \mathbb{N}$.

$$\emptyset = \overset{\circ}{\overline{M_n}} = \overset{\circ}{Y_n} = \overset{\circ}{\overline{Y_n}} \qquad \text{as } M_n \text{ is nowhere dense.}$$

The following theorem is a deep result in metric space theory which is frequently used in functional analysis. It is one of the reasons why complete metric spaces ~~are~~ are so important in functional analysis.

Theorem 3.23 (Baire category)
Let (X,d) be a metric space which is complete.

(a) If $M_1, M_2, \dots$ are open dense subsets of X, then

$$\bigcap_{n=1}^{\infty} M_n \text{ is dense in } X.$$

(b) X is not a countable union of nowhere dense sets; in other words, X is nonmeager in itself.

As a consequence, the complement of a meager set is nonmeager. Indeed, if $X = A \cup B$ with A meager, then B cannot be meager, since this would imply that X is meager.

Proof: (a) We show that if $A \subset X$ is any nonempty open set, then A intersects $\bigcap_{n\in\mathbb{N}} M_n$.

M_1 is dense, hence $A \cap M_1 \neq \emptyset$ and open, $\exists\ x_1 \in X$ and $\varepsilon_1 > 0$ such that

$$\overline{B_{\varepsilon_1}(x_1)} \subset M_1 \cap A.$$

Next $B_{\varepsilon_1}(x_1) \cap M_2 \neq \emptyset$ and open $\exists\, x_2 \in X$ and $\varepsilon_2 > 0$ such that $\overline{B_{\varepsilon_2}(x_2)} \subset M_2 \cap B_{\varepsilon_1}(x_1)$,

and so ... W.l.o.g $\varepsilon_n \to 0$ as $n \to \infty$

$N \in \mathbb{N}$, $(x_n)_{n\in\mathbb{N}}$ sequence in X with $x_n \in B_{\varepsilon_N}(x_N)$ $\forall n > N$;

$\Rightarrow$ $(x_n)_{n\in\mathbb{N}}$ Cauchy sequence, and it converges to some

$x \in X$ because X is complete. $x \in \overline{B_{\varepsilon_n}(x_n)}$ $\forall n \in \mathbb{N}$

$\Rightarrow$ $x \in M_n \cap A$ $\forall n \in \mathbb{N}$, and so A intersects

$\bigcap_{n=1}^{\infty} M_n$.

(b) If $A_1, A_2, \ldots$ are nowhere dense sets, then

$(\overline{A_1})^c, (\overline{A_2})^c, \ldots$ are dense open sets. Then,

by (a), $\bigcap_{n\in\mathbb{N}} (\overline{A_n})^c$ is dense in X and nonempty,

so that its complement $\bigcup_n \overline{A_n}$ cannot be equal to X,

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

<u>Remark:</u> A nowhere dense $\Leftrightarrow$ $\overline{A}$ has no interior points

$\Leftrightarrow$ $\forall x \in X, \forall \varepsilon > 0, \exists y \in B_\varepsilon(x)$ s.th. $y \notin \overline{A}$

$(\overline{A}^c)$ dense $\Leftrightarrow$ $\forall x \in X \, \forall \varepsilon > 0$, $\exists y \in B_\varepsilon(x)$ s.t. $y \in (\overline{A})^c$

A nowhere dense $\Leftrightarrow$ $(\overline{A})^c$ dense

The Baire category theorem is often used to prove existence results: one shows that objects having a certain property exist by showing that the set of objects not having the property (within a suitable complete metric space) is meager.

We turn to the applications of the Baire category theorem. We first need some terminology.

<u>Definition 3.24:</u> (a) Let (X, τ_X) and (Y, τ_Y) be topology space and $f: X \to Y$.

- f is called open if $f(Q)$ is open for any open set $Q \subset X$, i.e.

$$\forall\, Q \in \tau_X \implies f(Q) \in \tau_Y$$

(b) $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ normed spaces and $f: X \to Y$ linear.

f is open $\iff$ $f(B_1(0))$ contains a ball centered at 0 in Y.

<u>Remark</u>: The function $f: \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} 1-|x| & \text{if } |x| \leq 1 \\ |x|-1 & |x| > 1 \end{cases}$$

$f(\mathbb{R}) = \mathbb{R}_0^+$. Pick $Q =]-3/2, 3/2[$

is obviously continuous. Then $f(Q) = [0,1]$

Under special assumption continuity implies that a mapping is open.

<u>Theorem 3.25:</u> (Open Mapping)
Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces. If $T \in \mathcal{B}(X,Y)$ is surjective, i.e. $T(X) = Y$, then T is open

<u>Proof:</u>

Let $B_n = B_n(0)$ ball of radius n in X.
We show that $T(B_1)$ contains a ball around 0 in Y:
Since $X = \bigcup_{n=1}^{\infty} B_n$, and $T(X) = Y$, we have $Y = \bigcup_{n=1}^{\infty} T(B_n)$. Since $(Y, \|\cdot\|_Y)$ is complete, it is nonmeager by Theorem 3.23 (Baire category). Then there exists $n \in \mathbb{N}$ such that $T(B_n)$ is not nowhere dense.
So there exists $y \in T(B_n)$ and $\varepsilon > 0$ such that the ball $B_\varepsilon(y) \subset \overline{T(B_n)}$.

Let us see that $\overline{T(B_n)}$ contains a ball ~~centred~~ centered at 0.

Any $z \in Y$ with $\|z\|_Y < \varepsilon$

$$z = -Tx + \underbrace{(z + y)}_{\in B_\varepsilon(y) \subset \overline{T(B_n)}} \in \overline{T(-x + B_n)} \subset \overline{T(B_{2n})}$$

Here, $x \in B_n$ such that $Tx = y$.

This shows that $\overline{T(B_{2n})}$ contains a ball of radius ε centered at 0. By rescaling, we obtain that $\overline{T(B_1)}$ contains a ball of radius $\eta = \varepsilon/2n$ (centered at 0).
We are left to replace $\overline{T(B_1)}$ by $T(B_1)$.

Let $y \in B_{\eta/2}(0)$. There exists $y_1 \in T(B_{1/2})$ s.t

$\| y - Tx_1 \|_Y < \frac{1}{4}\eta$ with $x_1 \in B_{1/2}(0)$ and $Tx_1 = y_1$.

Pick $\tilde{y} \in Y$ with $\|\tilde{y}\|_Y < \eta/2$, i.e. $\tilde{y} \in B_{\eta/2}(0)$;

we want to show that $\tilde{y}$ is the image of some $x \in B_1(0)$.
If we do so we clearly have shown that $B_{\eta/2}(0) \subset T(B_1)$.
How to get this x?
We will construct it. T continuous implies $\exists\, y_1 \in T(B_{1/2}(0))$

s.t $\| \tilde{y} - Tx_1 \|_Y < \frac{1}{4}\eta$ with $x_1 \in B_{1/2}(0)$ and $Tx_1 = y_1$

Then there exists $x_2 \in B_{1/4}(0)$ s.t.

$\| \tilde{y} - Tx_1 - Tx_2 \|_Y < \frac{1}{8}\eta$. By induction,

$\exists$ sequence $(x_n)_{n\in\mathbb{N}}$ with $x_n \in B_{2^{-n}}(0)$, s.t.

$$\| \tilde{y} - \sum_{i=1}^{n} Tx_i \|_Y < 2^{-n-1}\eta$$

Since $\sum_{n=1}^{\infty} \|x_n\|_X < \sum_{n=0}^{\infty} 2^{-n} - 1 = 1$,

the series $\sum_{n\in\mathbb{N}} x_n$ converges to some element $x \in X$

with $\|x\|_X < 1$ and $Tx = y$. ∎

<u>Corollary 3.26</u> (Inverse Mapping)
Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces. Let $T: X \to Y$ be a bijective (one-to-one and onto) linear mapping. If T is bounded, then the inverse operator

T^{-1} is also bounded.

<u>Proof:</u> T^{-1} exists since T is bijective. Continuity of T^{-1} is equivalent to T being open, which is true by Theorem 3.25. ∎

Recall the notion of product norm: If $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are normed spaces, we define the product norm on $X \times Y$ by

$$\|(x,y)\|_{X,Y} = \max\{\|x\|_X, \|y\|_Y\}, \quad x \in X, y \in Y.$$

<u>Definition 3.27</u>: Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces.

(a) The graph of an operator $T: X \to Y$, is the subset

$$G(T) = \{(x,y) \in X \times Y : y = Tx\} \text{ of } X \times Y$$

(b) A linear operator $T \in \mathcal{L}(X,Y)$ is called closed if its graph $G(T)$ is a closed subspace of $X \times Y$.

Remarks: $G(T)$ is a vector subspace of $X \times Y$.

Let $\pi_1 : X \times Y \to X$ be the projection onto the first element. The restriction $\pi_1 : G(T) \to X$ is bijective. Conversely, if $G \subset X \times Y$ such that $\pi_1 : G \to X$ is bijective, then there exists a unique linear ~~optor~~ operator $T : X \to Y$ s.t $G(T) = G$.

Closedness of the graph $G(T)$ means that if $((x_n, Tx_n))_{n \in \mathbb{N}}$ converges in $X \times Y$, then the limit lies in the graph $G(T)$.
$x_n \to x$ and $Tx_n \to y$ as $n \to \infty \Rightarrow Tx = y$.
It is clear that continuous operators are necessarily closed.
Now, if X and Y are complete, the converse holds true.

Theorem 3.28 (Closed Graph Theorem)
If $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are Banach spaces and $T \in \mathcal{L}(X,Y)$ is closed, then T is bounded.

Proof: $X \times Y$ is Banach space (with the product norm $\|\cdot\|_{X \times Y}$). Since $G(T)$ is a closed subspace of a Banach space it is complete.

$\pi_1 : G(T) \to X$, $(x, Tx) \mapsto \pi_1(x, Tx) = x$

$\pi_2 : G(T) \to Y$, $(x, Tx) \mapsto \pi_2(x, Tx) = Tx$

~~Both projections are onto~~

Π_1 is a bounded bijection, and Π_2 is a bounded operator.

$T = \Pi_2 \circ \Pi_1^{-1} : X \to Y$

Corollary 3.26 implies that Π_1^{-1} is bounded.

Then $\|T\| = \|\Pi_2 \circ \Pi_1^{-1}\| \leq \|\Pi_2\| \|\Pi_1^{-1}\| < \infty$ □

Revision $T = G \circ F$, then $\|T\| \leq \|G\| \|F\|$

$$\|T\| = \sup_{\|x\|_X \leq 1} \underbrace{\|Tx\|_X}_{\|G(Fx)\|_X \leq \|G\| \|Fx\|_X \leq \|G\| \|F\|}.$$

Corollary 3.29 $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ normed spaces and $T \in \mathcal{B}(X,Y)$ invertible then, for all $x \in X$

$$\|Tx\|_Y \geq \|T^{-1}\|^{-1} \|x\|_X$$

Proof: $\|x\|_X = \|T^{-1}(Tx)\|_X \leq \|T^{-1}\| \|Tx\|_Y$ ■

~~2.4~~ 3.3 The Uniform Boundedness Principle

The last cornerstone result about Banach spaces is the uniform boundedness theorem. It states that if $(T_n)_{n \in \mathbb{N}}$ is a sequence of operators such that $\|T_n x\|$ is bounded uniformly in n for each fixed x, then $\|T_n x\| / \|x\|$ is actually bounded uniformly in n and x.

Let us examine an example first.

<u>Example 3.30:</u> Let $\varphi = \{x=(x_n)_{n\in\mathbb{N}} : \exists k\in\mathbb{N} \text{ s.t. } x_n = 0 \ \forall n\geq k\}$

$(\varphi, \|\cdot\|_\infty)$ normed space. $\forall n\in\mathbb{N}$ define

$$T_n : \varphi \longrightarrow \mathbb{K} \quad , \quad x=(x_k)_{k\in\mathbb{N}} \longmapsto T_n x = \sum_{k=1}^{n} x_k$$

and $\Phi = \{T_n : n\in\mathbb{N}\}$ (family of linear functionals)

- pointwise boundedness:

  Φ pointwise bounded because $\forall x\in\varphi \ \exists m_x\in\mathbb{N}$ with $x_k=0$ for all $n \geq m_x$; and

$$|T_n x| \leq \sum_{k=1}^{m_x} |x_k| \leq m_x \|x\|_\infty \quad \forall n\in\mathbb{N}$$

- uniform boundedness: Φ is uniform bounded if $\exists C>0$ $\|T_n\| < C \ \forall n\in\mathbb{N}$

We show that Φ is not uniformly bounded.

$$x^{(n)} = (x_k^{(n)})_{k\in\mathbb{N}} \quad \text{with} \quad x_k^{(n)} := \begin{cases} 1 & k\leq n \\ 0 & k>n \end{cases}$$

$\|x^{(n)}\|_\infty = 1$ and $T_n x^{(n)} = n$, and hence

$\|T_n\| \geq n \ \forall n\in\mathbb{N}$. Moreover

$$|T_n x| \leq \sum_{k=1}^{n} |x_k| \leq n\|x\|_\infty \quad \text{implies} \quad \|T_n\|\leq n,$$

and hence $\|T_n\| = n$.

Theorem 3.31 (Uniform Boundedness, Banach-Steinhaus)
Let $(X, \|\cdot\|_X)$ be a Banach space, and $(Y, \|\cdot\|_Y)$ be a normed space. Let $(T_n)_{n\in\mathbb{N}}$ be a sequence in $B(X,Y)$ such that $\forall x \in X \ \exists\ c_x > 0$ such that

$$\|T_n x\|_Y \leq c_x \quad \forall n \in \mathbb{N} \text{ (pointwise bounded)}$$

Then $\exists\ C > 0$ such that

$$\|T_n\| \leq C \quad \forall n \in \mathbb{N}.$$

Proof. Again an application of Baire category theorem!

Pick $k \in \mathbb{N}$, define

$$E_k = \{x \in X : \sup_{n\in\mathbb{N}} \|T_n x\|_Y \leq k\} = \bigcap_{n\in\mathbb{N}} \{x \in X : \|T_n x\| \leq k\}$$

Then E_k is closed, and $X = \bigcup_{k\in\mathbb{N}} E_k$. By Theorem 3.23 (Baire category)
there exists $k \in \mathbb{N}$ such that E_k has an interior point.
Since E_k is closed, E_k contains a closed ball $\overline{B_r(x_0)}$
W.l.o.g. $r \leq k$. Then $\overline{B_r(0)} \subset E_{2k}$

Check: $\underset{\overline{B_r(0)}}{\underset{\cap}{x}} = \underset{E_k}{\underset{\cap}{x_0}} + \underset{\overline{B_r(x_0)} \subset E_k}{\underset{\cap}{(x - x_0)}}$, so that

$$\|T_n x\|_Y \leq \|T_n x_0\|_Y + \|T_n(x - x_0)\|_Y \leq 2k$$

Then, for any $n \in \mathbb{N}$,

$$\sup_{\|x\|_X = 1}\{\|T_n x\|_Y\} = \frac{1}{r} \underbrace{\sup_{\|x\|_X=1}\{\|T_n(\underbrace{rx}_{\in E_{2k}})\|_Y\}}_{\leq 2k} \leq \frac{2k}{r}$$

Thus $\|T_n\| \leq \frac{2k}{r}$ $\forall n \in \mathbb{N}$. $\square$

<u>Corollary 3.32</u>

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ normed space be Banach spaces, and $(T_n)_{n\in\mathbb{N}}$ be a sequence in $B(X,Y)$ such that $\lim_{n\to\infty} T_n x$ exists for all $x \in X$. Let $Tx = \lim_{n\to\infty} T_n x$

Then T is a bounded oprator.

<u>Proof:</u> T is clearly linear. We need to show that T is bounded. $\forall x \in X$, $\forall \varepsilon > 0$ $\exists$ $n_0 \in \mathbb{N}$ s.t.

$$\|T_n x - Tx\|_Y < \varepsilon \quad \forall n > n_0.$$

$$c_x := \max\left\{ \max_{1\leq n\leq n_0}\{\|T_n x\|_Y\}, \|Tx\|_Y + \varepsilon \right\}.$$ Then $c_x < \infty$

and $\|T_n x\|_Y \leq C_x$ $\forall n \in \mathbb{N}$. Now apply Theorem 3.31. $\blacksquare$

<u>Proposition 3.33</u> Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces and $T \in \mathcal{B}(X,Y)$.
There exists a unique operator $T' \in \mathcal{B}(Y', X')$ such that

$$T'(f)(x) = f(Tx) \quad , x \in X, f \in Y'.$$

<u>Proof</u>. Pick $f \in Y'$ and define $T'(f) = f \circ T$.
Then $T'(f) \in X'$ because $f \in \mathcal{B}(Y, \mathbb{K})$ and $T \in \mathcal{B}(X,Y)$, and composition of bounded (continuous) operators is bounded (continuous).
Hence, T' is a mapping $Y' \to X'$ such that $T'(f)(x) = f(Tx)$.
Uniqueness: Suppose $\exists\ S \in \mathcal{B}(Y', X')$ with $S(f)(x) = f(Tx)$ $\forall x \in X, f \in Y'$.
Then $S(f) = T'(f)$ $\forall f \in Y'$, so $S = T'$.
Linearity of T':
$f, g \in Y'$, $\lambda, \mu \in \mathbb{K}$

$$(\lambda f + \mu g) \circ T = \lambda (f \circ T) + \mu (g \circ T),$$

so $\quad T'(\lambda f + \mu g) = \lambda T'(f) + \mu T'(g)$.
Boundedness
$\|T'(f)\| = \|f \circ T\| \leq \|f\| \|T\|$, hence $\|T'\| \leq \|T\|$. $\square$

<u>Definition 3.34</u>: Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces and $T \in \mathcal{B}(X,Y)$. The operator $T' \in \mathcal{B}(Y', X')$ is called the dual of T.

Summary : main theorems of functional analysis

- linear forms / functionals $(X', \|\cdot\|)$

Hahn-Banach: $(X, \|\cdot\|_X)$ normed space, $W \subset X$ subspace
$\forall f_W \in W'$ $\exists$ extension
$f_X \in X'$ s.t. $\|f_X\| = \|f_W\|$

enables us to separate points in $(X, \|\cdot\|_X)$ with linear functionals
second dual, dual spaces,
weak convergence, weak-$*$ convergence

Banach-Alaoglu: $(X, \|\cdot\|_X)$ normed space

$$\Rightarrow \{ f \in X' : \|f\| \leq 1 \} \subset X'$$

is weak-$*$ compact.

nowhere dense, meager, non-meager

Baire category: (X, d) complete metric space
M_n open dense
$\Rightarrow \bigcap_{n=1} M_n$ is dense

Open Mapping Theorem: $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ Banach spaces. $T \in B(X, Y)$
surjective $\Rightarrow$ T is open

Inverse Mapping Theorem: $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ Banach spaces
$T \in \mathcal{B}(X,Y)$ bijective $\Rightarrow T^{-1}$ is bounded

Closed Graph Theorem: $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ Banach spaces, $T \in \mathcal{L}(X,Y)$ closed $\Rightarrow T$ is bounded

Banach-Steinhaus Theorem: (Uniform Boundedness Principle)
$(X, \|\cdot\|_X)$ Banach space, $(Y, \|\cdot\|_Y)$ normed space
$(T_n)_{n\in\mathbb{N}}$ seq. in $\mathcal{B}(X,Y)$ pointwise bounded, i.e. $\forall x \in X \; \exists c_x > 0$ such that
$\|T_n x\|_Y \leq c_x \quad \forall n \in \mathbb{N}$
$\Rightarrow \exists C > 0$ such that
$\|T_n\| \leq C \quad \forall n \in \mathbb{N}.$

4 Hilbert spaces

4.1 Basic definitions

The most important Banach spaces, and the ones on which the most refined analysis can be done, are the Hilbert spaces, which are a direct generalisation of finite-dimensional Euclidean spaces.
We review basic definitions which have been discussed in Functionalanalysis I.

Inverse Mapping Theorem: $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ Banach spaces
$T \in \mathcal{B}(X,Y)$ bijective $\Rightarrow$ T^{-1} is bounded

Closed Graph Theorem: $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ Banach spaces, $T \in \mathcal{L}(X,Y)$ closed $\Rightarrow$ T is bounded

Banach-Steinhaus Theorem: (Uniform Boundedness Principle)
$(X, \|\cdot\|_X)$ Banach space, $(Y, \|\cdot\|_Y)$ normed space
$(T_n)_{n\in\mathbb{N}}$ seq. in $\mathcal{B}(X,Y)$ pointwise bounded, i.e. $\forall x \in X$ $\exists$ $c_x > 0$ such that
$\|T_n x\|_Y \leq c_x \quad \forall n \in \mathbb{N}$
$\Rightarrow$ $\exists$ $C > 0$ such that
$\|T_n\| \leq C \quad \forall n \in \mathbb{N}.$

4 Hilbert spaces

4.1 Basic definitions

The most important Banach spaces, and the ones on which the most refined analysis can be done, are the Hilbert spaces, which are a direct generalisation of finite-dimensional Euclidean spaces.
We review basic definitions which have been discussed in Functionalanalysis I.

Definition 4.1: Let X be a real or complex vector space.

A scalar product (or inner product) on X is a map

$(\cdot,\cdot) : X \times X \longrightarrow \mathbb{K} \qquad (\mathbb{K} = \mathbb{C} \text{ or } \mathbb{K} = \mathbb{R})$

such that

(i) $\quad (y,x) = \overline{(x,y)} \quad , x,y \in X$ (here $\overline{}$ is the complex conjugation)

(ii) $\quad (x, \alpha y + \beta z) = \alpha(x,y) + \beta(x,z) \quad , x,y,z \in X \quad , \alpha,\beta \in \mathbb{K}$

(iii) $\quad (x,x) \geq 0$, and $(x,x) = 0$ iff $x = 0$.

Remarks: It follows from (i) and (ii) that

$$(\alpha x + \beta y, z) = \bar{\alpha}(x,z) + \bar{\beta}(y,z).$$

Hence, the scalar product (for X complex vector space) is antilinear in its first argument and linear in its second argument.

Definition 4.2 (a) A vector space (over $\mathbb{K}$) $(X, (\cdot,\cdot))$ equipped with a scalar product (inner product) is called a pre-Hilbert space. Notation: We will write $(H, (\cdot,\cdot))$ instead of $(X, (\cdot,\cdot))$.

(b) $\forall x \in X: \quad \|x\| := \sqrt{(x,x)}$

Proposition 4.3 Let $(H, (\cdot,\cdot))$ be a pre-Hilbert space

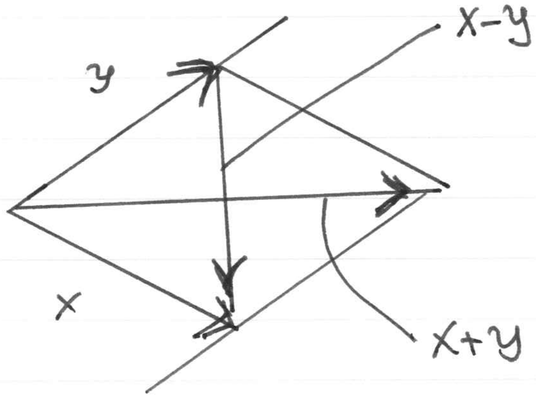
(a) The map $\|\cdot\|: H \to \mathbb{K}$, $x \mapsto \|x\| = (x,x)$ is a norm on H, and it is called the induced norm. We write $\|\cdot\|$ or $\|\cdot\|_H$ as before.

(b) $|(x,y)| \leq \|x\| \, \|y\| \qquad x, y \in H$

(Cauchy-Schwarz inequality)

(c) The induced norm satisfies the parallelogram identity:

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



Proof: See Functional Analysis I $\square$

A scalar product (inner product) always induces a norm. The converse is true iff the norm satisfies the parallelogram identity. The scalar product (inner product) is given by the polarization identity:

$$(x,y) = \tfrac{1}{4}\|x+y\|^2 - \tfrac{1}{4}\|x-y\|^2 - \tfrac{i}{4}\|x+iy\|^2 + \tfrac{i}{4}\|x-iy\|^2 \quad \forall x, y$$

(See example sheet 5)

Proposition 4.4: Let $(H,(\cdot,\cdot))$ be a pre-Hilbert space with induced norm $\|\cdot\|$.
Then for all $x, y \in H$:

(a) if H is real then $4(x,y) = \|x+y\|^2 - \|x-y\|^2$

(b) if H is complex then

$$4(x,y) = \|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2$$

Proof: See Functional ~~Anal~~ Analysis I or example sheet 5. Hint: Assume a given norm is not induced by an inner product. Then show that it does not satisfy the parallelogram rule. $\square$

Definition 4.5: A Hilbert space $(H,(\cdot,\cdot))$ is a pre-Hilbert space who is complete with respect to the induced norm.

Remark: Hilbert spaces are Banach spaces.

Examples 4.6:

(a) $\mathbb{R}^n$, $\mathbb{C}^n$ with scalar product $(x,y) = \sum_{i=1}^{n} \overline{x_i} y_i$

(b) $(\ell^2(\mathbb{K}), (\cdot,\cdot)_2)$ with $(x,y)_2 = \sum_{i=1}^{\infty} \overline{x_i} y_i$

Notice that $(\ell^p(\mathbb{K}), \|\cdot\|_p)$ with $p \neq 2$

cannot be turned into a Hilbert space, i.e. the $\|\cdot\|_p$-norm is not induced by an inner product.

(c) The Lebesgue space $(L^2(\mathbb{R}^n), (\cdot,\cdot)_2)$ with

$$(f,g)_2 = \int_{\mathbb{R}^n} \overline{f(x)}\, g(x)\, dx$$ is a Hilbert space.

$L^p(\mathbb{R}^n)$, $p \neq 2$, are not Hilbert spaces.

Definition 4.7: Let $(X, (\cdot,\cdot))$ be a (pre-) Hilbert space

(a) Two elements $x, y \in X$ are orthogonal if $(x,y) = 0$, written as $x \perp y$.

(b) Let $M \subset X$ subset. The orthogonal complement $M^\perp$ of M is the set

$$M^\perp = \{x \in X : (x,y) = 0 \quad \forall y \in M\}.$$

Proposition 4.8: Let $(X, (\cdot,\cdot))$ be a pre-Hilbert space.

(a) The orthogonal complement $M^\perp$ of a subset $M \subset X$ is a closed subspace.

(b) $M \subset (M^\perp)^\perp$. If M is a closed subspace, then $M^{\perp\perp} = M$.

Proof. See Functional Analysis I. $\square$

The Pythagorean Theorem

Theorem 4.9: Let $(X,(\cdot,\cdot))$ be a Hilbert space

If $x_1,\dots,x_n \in X$ and $x_i \perp x_k$ for $i \neq k$, $i,k=1,\dots,n$, then
$$\left\| \sum_{k=1}^{n} x_k \right\|^2 = \sum_{k=1}^{n} \|x_k\|^2$$

Proof.: $$\left\| \sum_{k=1}^{n} x_k \right\|^2 = \left(\sum_{k=1}^{n} x_k, \sum_{k=1}^{n} x_k \right)$$
$$= \sum_{k,i=1}^{n} (x_k, x_i) = \sum_{k=1}^{n} \|x_k\|^2 ,$$

where $\|\cdot\|$ is the induced norm. ☐

Theorem 4.10: $(X,(\cdot,\cdot))$ be a Hilbertspace, and $M \subset X$ a closed subspace. Then $X = M \oplus M^\perp$; that is,

each $x \in X$ can be expressed as $x = y + z$ where $y \in M$ and $z \in M^\perp$.
Moreover, y and z are the unique elements of M and $M^\perp$ whose distance to x is minimal

Proof. Pick $x \in X$, and let $\delta := \inf\{\|x-y\| : y \in M\}$

and $(y_n)_{n\in\mathbb{N}}$ sequence in M such that $\|x-y_n\| \to \delta$.

By the parallelogram law/identity,

$$2(\|y_n-x\|^2+\|y_m-x\|^2)=\|y_n-y_m\|^2+\|y_n+y_m-2x\|^2,$$

or, since $1/2\,(y_n+y_m)\in M$,

$$\|y_n-y_m\|^2=2\|y_n-x\|^2+2\|y_m-x\|^2-4\|1/2(y_n+y_m)-x\|^2$$

$$\leq \underbrace{2(\|y_n-x\|^2+\|y_m-x\|^2)-4\delta^2}_{\to 0 \text{ as } n,m\to\infty}$$

$\Rightarrow$ $(y_n)_{n\in\mathbb{N}}$ Cauchy sequence with $y=\lim y_n$ (Why?).

Let $z=x-y$. Then $y\in M$ because M is closed, and $\|x-y\|=\delta$.
We are left to show that $z\in M^\perp$
Pick $u\in M$ (and eventually multiply it by a non-zero scalar) with $(z,u)\in\mathbb{R}$

$$\Rightarrow f(t)=\|z+tu\|^2=\|z\|^2+2t(z,u)+t^2\|u\|^2\in\mathbb{R}$$

$\forall t\in\mathbb{R}$, and it has a minimum (namely δ^2) at $t=0$, because $z+tu=x-(y-tu)$ and $(y-tu)\in M$

Thus $2(z,u)=f'(0)=0$, so $z\in M^\perp$. Moreover, if $z'\in M^\perp$, $z'\neq z$, by the Pythagorean theorem

$$\|x-z'\|^2=\|x-z\|^2+\|z-z'\|^2\geq\|x-z\|^2$$

($x-z'=(x-z)+(z-z')$ and $(x-z,z-z')=0$ because $(x-z)\in M$ as $y\in M$ and $z=x-y$), hence "="

iff $z = z'$. Similar one shows that y is the unique element of M closest to x.
Finally, if $x = y' + z'$ with $y' \in M$, $z' \in M^\perp$, then $y - y' = z' - z \in M \cap M^\perp$,

so $y - y'$ and $z' - z$ are orthogonal to themselves and hence are zero. ■

<u>**Definition 4.11.**</u> $M \subset X$ subspace ; $(X, (\cdot,\cdot))$ Hilbert space.

Then $P: X \to X$, $x \mapsto Px \in M$,
and $Q: X \to X$, $x \mapsto Qx \in M^\perp$

such that $x = Px + Qx$,

i.e. $Px = y$ (Theorem 4.10) and $Qx = z$ (Theorem 4.10),

are linear mappings called orthogonal projections on M respectively on $M^\perp$.

<u>**Corollary 4.12**</u> $M \subset X$, $(X, (\cdot,\cdot))$

$P: X \to M$, $Q: X \to M^\perp$ orthogonal projections

$\|x\|^2 = \|Px\|^2 + \|Qx\|^2$

$\|P\| = 1$, $\|Q\| = 1$

Proof: $x = y + z$ with $y = Px$ and $z = Qx$

$$\|x\|^2 = (x,x) = (y+z, y+z) = (y,z) + (z,y) + (y,y) + (z,z)$$
$$= \|y\|^2 + \|z\|^2$$

$$\|P\| = \sup_{\|x\|_X \le 1} \|Px\|_X \le \|x\|_X \quad \Rightarrow \|P\| \le 1$$

$$x \in M \Rightarrow \|Px\|_X = \|x\|_X \Rightarrow \|P\| \ge 1. \qquad \square$$

We already observed that, for every $y \in X$, $(X, (\cdot,\cdot))$ a Hilbert space, the map
$x \to (y,x)$ is linear. Is it also continuous? Are all continuous linear functionals on X of this type? We answer both questions in the following theorem.

Theorem 4.13: Let $(X, (\cdot,\cdot))$ be a Hilbert space

(a) For any fixed $y \in X$, the mappings

$$x \to (y,x) \quad \text{and} \quad x \to (x,y), \quad x \to \|x\|$$

are continuous functions on X.

(b) If $f \in X^*$ (continuous linear functional), then there is a unique $y \in X$ such that

$$f(x) = (y,x), \quad x \in X.$$

Proof: (a)

$$|(y,x_1)-(y,x_2)| = |(y,x_1-x_2)| \leq \|x_1-x_2\| \, \|y\|$$

$\Rightarrow$ $x \to (y,x)$ and $x \to (x,y)$ are continuous (infact uniformly continuous)

$\|x_1\| \leq \|x_1-x_2\| + \|x_2\|$ yields

$\|x_1\| - \|x_2\| \leq \|x_1-x_2\|$; interchanging x_1 and x_2 we get

$|\|x_1\| - \|x_2\|| \leq \|x_1-x_2\|$, hence $x \to \|x\|$ is uniformly continuous.

(b) If $f(x)=0$ $\forall x \in X$, take $y=0$.

Otherwise, define $M=\{x \in X : f(x)=0\} \subset X$ subspace which is closed (continuity). Since $f(x) \neq 0$ for some $x \in X$ we have that $M^\perp$ does not consists of 0 alone.

$\exists z \in M^\perp$ with $\|z\|=1$

$u = f(x)z - f(z)x$

$f(u)=0 \Rightarrow u \in M$. Thus $(u,z)=0$

$f(x) = f(x)(z,z) = f(z)(x,z)$

$\Rightarrow f(x) = (\overline{f(z)}z, x)$ $\square$

Remark: The last theorem is also called Riesz representation theorem for Hilbert spaces. It implies that $(\ell^2)' = \ell^2$ and $(L^2)' = L^2$.
There are several Riesz representation theorems, that always relate the dual of a space with some space defined differently.
For example, the dual of the space of continuous functions with compact support is represented by the space of regular complex Borel measures (Riesz–Markov theorem).

4.2 Orthonormal bases

Definition 4.14: Let $(X, (\cdot,\cdot))$ be a pre-Hilbert space, and $\emptyset \neq S \subset X$

Then the set S is called

(a) orthogonal if $\forall x, y \in S, x \neq y: \quad x \perp y$

(b) orthonormal if S is orthogonal and $\|x\| = 1 \; \forall x \in S$, that is
$$\forall x, y \in S: \quad (x,y) = \begin{cases} 0 & \text{if } x \neq y \\ 1 & \text{if } x = y \end{cases}$$

(c) a complete orthonormal system if S is orthonormal and maximal, that is, $\tilde{S}$ orthonormal set of X with $S \subset \tilde{S}$, then $\tilde{S} = S$.

Then, formally one can consider

$$x = \sum_{n \in I} (x, e_n)\, e_n$$

Does this series converge? Does it converge to x?

<u>Proposition 4.17:</u> Let $(X, (\cdot\,;\cdot))$ be a pre-Hilbert space, I an index set and $S = \{x_n : n \in I\}$ an orthonormal system in X.
Then the following statements hold for any $x \in X$:

(a) $(x, x_n) \neq 0$ for at most countable many $n \in I$

(b) the series $\sum_{n \in I} |(x, x_n)|^2$ converges and

$$\sum_{n \in I} |(x, x_n)|^2 \leq \|x\|^2 \quad \text{(Bessel's Inequality)}$$

(c) $$x = \sum_{n \in I} (x, x_n)\, x_n$$

$$\Longleftrightarrow \quad \sum_{n \in I} |(x, x_n)|^2 = \|x\|^2 \quad \text{(Parseval's identity)}$$

<u>Proof:</u> Functionalanalysis I. $\square$

Corollary 4.15: $(X,(\cdot,\cdot))$ pre-Hilbert space, $\emptyset \neq S \subset X$.

S complete orthonormal system $\Leftrightarrow$ $S^{\perp} = \{0\}$

Proof.: exercise, functional analysis I. $\square$

Example 4.16: (a) The set of functions

$E = \{e_n : n \in \mathbb{Z}\}$ with $e_n : [-\pi, \pi] \to \mathbb{C}$

$x \mapsto (2\pi)^{-1/2} e^{inx}$, $n \in \mathbb{Z}$, is an

orthonormal set (sequence) / family in the space $L^2([-\pi,\pi]; \mathbb{C})$.

This follows from $(e_m, e_n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(m-x)x} dx = \begin{cases} 1, & \text{if } m = n \\ 0, & \text{if } m \neq n \end{cases}$.

(b) Recall $(\cdot,\cdot) : \ell^2(\mathbb{K}) \times \ell^2(\mathbb{K}) \to \mathbb{K}$, $(x,y) = \sum_{k=1}^{\infty} \overline{x_k} y_k$

$S = \{e^k : k \in \mathbb{N}\}$ where

$e^k = (0, \dots, \underset{i\text{-th entry/index}}{1}, 0, \dots 0 \dots)$.

Moreover, one can show that the orthonormal set S is complete, i.e. $S^{\perp} = \{0\}$. (exercise).

Let $\{e_n : n \in I\}$ be an orthonormal system, set, family in a Hilbert space. We assume here that the index set I is countable, and that $n, m \in I$ with $n \neq m$ implies $e_n \neq e_m$.

<u>Proposition 4.18</u> : Let $(X, (\cdot,\cdot))$ be a Hilbert space and $S = \{x_n : n \in I\}$ an orthonormal system (family) in X with index set $I \neq \emptyset$.
Then the following statements are equivalent.

(a) S is a complete orthonormal system

(b) $S^{\perp} = \{0\}$

(c) $x = \sum_{n\in I} (x, x_n) x_n \quad \forall x \in X$

(d) $\|x\|^2 = \sum_{n \in I} |(x, x_n)|^2 \quad \forall x \in X$

<u>Proof.</u> Functionalanalysis I or a combination of Proposition 4.17, Corollary 4.15, Theorem 4.9 (Pythagorean), and completeness of X (Cauchy sequences do converge) $\square$

Property (c) of Proposition 4.18 is the desired infinite-dimensional version of finite dimensional spaces.

<u>Definition 4.19</u> Let $(X, (\cdot,\cdot))$ be a Hilbert space and let $\{e_n : n \in I\}$ be an orthonormal system / family in X.
Then $\{e_n : n \in I\}$ is called an orthonormal basis for X if any of the conditions (a)−(d) in Proposition 4.18 hold.

Proposition 4.20: Every Hilbert space has an orthonormal basis.

Proof.: Zorn's Lemma again. Orthonormal sets can be ordered by inclusion. Every countable set of orthonormal sets has an upper bound (take the union of all elements of that set). Then there exists a maximal element, which is an orthonormal set S with $S^\perp = \{0\}$, because if $S^\perp \neq \{0\}$, this would contradict the maximality of S. $\square$

Theorem 4.21: An infinite-dimensional Hilbert space $(X, (\cdot,\cdot))$ is separable if and only if it has an orthonormal basis

Proof: Suppose $(X, (\cdot,\cdot))$ is ∞-dimensional and separable, and let $\{x_n\}_{n\in\mathbb{N}}$ be a countable, dense sequence in X.

$(y_n)_{n\in\mathbb{N}}$ new sequence obtained by omitting every member of the sequence (x_n) which is a linear combination of the preceding members of the sequence.

Gram-Schmidt algorithm $\longrightarrow$ an orthonormal system $(e_n)_{n\in\mathbb{N}}$ in X with the property that for each $k \geq 1$

$$\text{Linear span}\{e_1, \dots, e_k\} = \text{Linear span}\{y_1, \dots, y_k\}$$

$$\Rightarrow \text{Linear span}\{e_n : n\in\mathbb{N}\} = \dots \{y_n : n\in\mathbb{N}\}$$
$$= \dots \{x_n : n\in\mathbb{N}\}$$

$$\Rightarrow \overline{LS\{e_n : n\in\mathbb{N}\}} = X \Rightarrow \{e_n\}_{n\in\mathbb{N}}$$

orthonormal basis

Example 4:22 The Hilbert space $(\ell^2(\mathbb{K}), (\cdot,\cdot)_2)$ is separable

An important property of every separable Hilbert space is that is is isomorphic to ℓ^2.

<u>Definition 4.23</u>: Let $(X_1, (\cdot,\cdot)_1)$, $(X_2, (\cdot,\cdot)_2)$ be Hilbert spaces with inner products $(\cdot,\cdot)_1$ and $(\cdot,\cdot)_2$ respectively.
A unitary map from X_1 to X_2 is an invertible linear map
$U: X_1 \longrightarrow X_2$ such that

$$(Ux, Uy)_2 = (x,y)_1 \quad \forall x,y \in X_1.$$

Notice that $\|Ux\|_2 = \|x\|_1 \quad \forall x \in X_1.$

<u>Proposition 4.24</u>: Let $(X,(\cdot,\cdot))$ be a Hilbert space (over $\mathbb{K}$)
Let $\{e_n : n \in \mathbb{N}\}$ be a orthonormal basis for $(X,(\cdot,\cdot))$.
Then the map

$$U: X \longrightarrow \ell^2(\mathbb{K})$$
$$x \longmapsto x_n = (e_n, x)$$

is a unitary map.

<u>Proof</u>: See FA I $\square$

4.3 Fourier Series

In this subsection we will study briefly the most important example for basis in certain L^2-spaces.

Theorem 4.25: The set of functions

$$C = \{ c_0, c_n ; n \in \mathbb{N} \} \text{ with}$$

$$c_0(x) = (1/\pi)^{1/2} \quad ; \quad c_n(x) = \left(\frac{2}{\pi}\right)^{1/2} \cos nx, \; n \in \mathbb{N},$$

is an orthonormal basis in $L^2([0,\pi];\mathbb{R})$.

Proof: exercise or FA I $\square$

What about $L^2([0,\pi];\mathbb{C})$?

$$f \in L^2([0,\pi];\mathbb{C}) \Rightarrow \exists f_R, f_I \in L^2([0,\pi];\mathbb{R})$$

with $f = f_R + i f_I$ $\square$

Corollary 4.26: $(L^2([0,\pi];\mathbb{K}), (\cdot,\cdot)_2)$ is separable.

Theorem 4.27:

(a) The set of functions

$$S = \{ s_n : n \in \mathbb{N} \} \text{ with } s_n(x) = \left(\frac{2}{\pi}\right)^{1/2} \sin nx, \quad x \in [0,\pi],$$

is an orthonormal basis in $L^2([0,\pi];\mathbb{R})$.

(b) The set of functions

$$E = \{ e_n : n \in \mathbb{Z} \} \text{ with } e_n(x) = (2\pi)^{-1/2} e^{inx}, \quad n \in \mathbb{Z},$$

$x \in [-\pi,\pi]$, is an orthonormal basis in $L^2([-\pi,\pi];\mathbb{C})$

Remark: In (b) also the set of functions

$$\mathcal{F} = \{ 2^{-1/2} c_0, 2^{-1/2} c_n, 2^{-1/2} s_n : n \in \mathbb{N} \}$$

is a orthonormal basis in $L^2([-\pi,\pi];\mathbb{C})$ and $L^2([-\pi,\pi];\mathbb{R})$.

5 Bounded operators in Hilbert spaces

5.1 Adjoint of an operator

Definition 5.1: Let $T \in \mathcal{B}(X,Y)$ be a bounded linear operator $T: X \to Y$, where $(X, (\cdot,\cdot)_X)$ and $(Y, (\cdot,\cdot)_Y)$ are Hilbert spaces. The adjoint of T is the operator $T^*: Y \to X$ such that

$$(y, Tx)_Y = (T^*y, x)_X \quad \forall y \in Y, x \in X$$

In general, it is not clear of the adjoint exists or is unique. But the Riesz representation theorem gives the answer.

Theorem 5.2: Let $(X, (\cdot,\cdot)_X)$ and $(Y, (\cdot,\cdot)_Y)$ be Hilbert spaces. Any bounded operator $T: X \to Y$ has a unique adjoint T^*. In addition, $\|T^*\| = \|T\|$.

Proof: FA I ∎

Remark: From the definition it follows that $T^{**} = T$ (note: only for bounded operators)

Example 5.3: Let $K: L^2([0,1]) \to L^2([0,1])$ be the integral operator with kernel $k \in \mathcal{C}([0,1]\times[0,1]; \mathbb{C})$

$$(Kf)(x) = \int_0^1 k(x,y) f(y)\, dy$$

Then
$$(K^* f)(x) = \int_0^1 \overline{k(x,y)} f(y) dy$$

Why?

$$(f, Kg) = \int_0^1 \overline{f(x)}\, (Kg)(x)\, dx = \int_0^1 dx \int_0^1 dy\, \overline{f(x)}\, k(x,y)\, g(y)$$
$$(K^* f, g) = \int_0^1 \overline{(K^* f)(x)}\, g(x)\, dx = \int_0^1 dx \int_0^1 dy\, k(y,x) \overline{f(y)}\, g(x).$$

Proposition 5.4: Let $(X, (\cdot,\cdot)_X)$ and $(Y, (\cdot,\cdot)_Y)$ be Hilbert spaces and $T, S \in \mathcal{B}(X,Y)$. Then the following properties hold

(a) $(S+T)^* = S^* + T^*$

(b) $(\alpha T)^* = \overline{\alpha}\, T^*$

(c) $T^* T = 0 \iff T = 0$

(d) $(ST)^* = T^* S^*$

(e) $\| T^* T \| = \| T T^* \| = \| T \|^2$

Proof: exercise. (e) below ∎

Theorem 5.5: Let $(X, (\cdot,\cdot)_X)$ and $(Y, (\cdot,\cdot)_Y)$ be Hilbert spaces and $T: X \longrightarrow Y$ be a bounded operator. Then

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

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

Remarks: An alternate statement is

$$X = \ker T \oplus \overline{\operatorname{ran} T^*}$$

$$Y = \ker T^* \oplus \overline{\operatorname{ran} T}$$

(T, T^* can be interchanged, hence statement of the theorem is symmetric)

Proof: Pick $x \in \operatorname{ran} T$. $\exists y \in X$ s.t. $x = Ty$

$\Rightarrow \forall z \in \ker T^*$

$$(x, z) = (Ty, z) = (y, T^* z) = 0$$

$\Rightarrow \operatorname{ran} T \subset (\ker T^*)^{\perp}$. Since $(\ker T^*)^{\perp}$ is closed, we actually have $\overline{\operatorname{ran} T} \subset (\ker T^*)^{\perp}$.

Conversely, if $x \in (\operatorname{ran} T)^{\perp}$, then $\forall y \in X$, we have

$$0 = (x, Ty) = (T^* x, y)$$

$\Rightarrow T^* = 0$, and $(\operatorname{ran} T)^{\perp} \subset \ker T^*$

$\Rightarrow (\ker T^*)^{\perp} \subset \overline{\operatorname{ran} T}$.

The second claim follows by interchanging the role of T and T^*. ∎

<u>Definition 5.6a:</u> $(X, (\cdot,\cdot))$ be a Hilbert space. A bounded linear operator $T \in B(X, X)$, $T: X \to X$, is <u>self-adjoint</u> if $T^* = T$.

<u>Remark:</u> Symmetric real matrices, hermitian complex matrices are s.a. operators.

K from example is self-adjoint $\Leftrightarrow k(t,s) = k(s,t)$ $\forall s,t$

<u>Definition 5.6b</u> A self-adjoint operator T s.t. $(x, Tx) \geq 0$ $\forall x \in X$ is called <u>non negative</u>.

If $(x, Tx) > 0$, $\forall x \neq 0$, it is positive, or positive-definite.

Proposition 5.7: If $T \in \mathcal{B}(X, X)$ and self-adjoint, then $\|T\| = \sup_{\|x\|=1} |(x, Tx)|$

Proof. FAI ■

Corollary 5.8: Let $(X, (\cdot,\cdot))$ be a Hilbert space and $T \in \mathcal{L}(X)$.

If T is bounded, then $\|T^* T\| = \|T\|^2$.

If T is bounded and self-adjoint, then $\|T^2\| = \|T\|^2$.

Proof: By Proposition 5.7, $\|T^*T\| = \sup_{\|x\| = 1} |(x, T^*Tx)|$

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

5.2 Spectrum of bounded operators

If T is a linear transformation on $\mathbb{K}^n$, then the eigenvalues of T are the numbers λ such that the determinant of $\lambda 1\!\!1 - T$ is equal to zero.
The set of such λ is called the spectrum of T. It can consist of at most n points since $\det(\lambda 1\!\!1 - T)$ is a polynomial of degree n. If λ is not an eigenvalue, then $\lambda 1\!\!1 - T$ has an inverse since $\det(\lambda 1\!\!1 - T) \neq 0$.

The spectral theory of operators on infinite-dimensional spaces is more complicated, more interesting, and very important for an understanding of the operators themselves.

Recall that $T \in B(X, Y)$, X and Y two normed spaces, means that the domain of definition $D(T) = X$.

Definition 5.9: Let $(X, \|\cdot\|)$ be a normed space. $T: X \to X$ linear

Then the set of numbers $\varrho(T) = \{\lambda \in \mathbb{K} : \lambda 1\!\!1 - T$ is injective, $\overline{\operatorname{ran}(\lambda 1\!\!1 - T)} = X$, $(\lambda 1\!\!1 - T)^{-1}$ bounded$\}$

is called the resolvent set of T. The operator

$R_\lambda(T) = (\lambda 1\!\!1 - T)^{-1}$ is called the resolvent of T at $\lambda \in \varrho(T)$.

The set $\sigma(T) = \mathbb{K} \setminus \varrho(T)$ is called the spectrum of T.

Remark: (1) If $(X, \|\cdot\|)$ is a Banach space, $T \in \mathcal{B}(X)$ and T bijective, then the Inverse Mapping
(Corollary 3.26) Theorem automatically implies that $R_\lambda(T) = (\lambda \mathbb{1} - T)^{-1}$, $\lambda \in \varrho(T)$, is bounded.

(2) $X = \{0\} \Rightarrow \varrho(T) = \mathbb{K}$ and $\sigma(T) = \emptyset$, hence we shall assume $X \neq \{0\}$ in the following.

In infinitesimal dimensional spaces, the spectrum of an operator is usually more complicated than a set of eigenvalues.
We distinguish ~~two subsets~~ the following subsets of the spectrum.

Definition 5.10: Let $(X, \|\cdot\|)$ be a normed space, and $T \in \mathcal{L}(X)$ a linear operator in X.

(a) The subset of the spectrum

$$\sigma_p(T) = \{\lambda \in \mathbb{K} : \lambda \mathbb{1} - T \text{ is not injective}\}$$

is called the point spectrum of T.

(b) The subset of the spectrum

$$\sigma_c(T) = \{\lambda \in \mathbb{K} : \lambda \mathbb{1} - T \text{ is injective}, \overline{\operatorname{ran}(\lambda \mathbb{1} - T)} = X, \text{ and } (\lambda \mathbb{1} - T)^{-1} \text{ is not bounded}\}$$

is called the continuous spectrum of T.

(c) The subset of the spectrum

$$\sigma_r(T) = \{\lambda \in \mathbb{K} : \lambda\mathbb{1} - T \text{ injective but } \overline{\operatorname{ran}(\lambda\mathbb{1}-T)} \neq X\}$$

is called the residual spectrum of T.

(d) An $x \neq 0$ which satisfies $Tx = \lambda x$ for some $\lambda \in \mathbb{K}$ is called an eigenvector of T; λ is called the corresponding eigenvalue.

Remark: (1) λ eigenvalue of $T \Rightarrow \ker(\lambda\mathbb{1} - T) \neq \{0\}$, so $\lambda\mathbb{1}-T$ is not injective, and hence

eigenvalues $\subset \sigma_p(T)$. In general eigenvalues $= \sigma_p(T)$ but for unbounded operator we have a further distinction.

(2) $\mathbb{K} = \rho(T) \cup \sigma_c(T) \cup \sigma_r(T) \cup \sigma_p(T)$

and $\sigma(T) = \sigma_c(T) \cup \sigma_r(T) \cup \sigma_p(T)$, and these set partitions are disjoint

In most settings we are given a closed operator T. Is λ no eigenvalue then with T also the operators $\lambda\mathbb{1} - T$ and $(\lambda\mathbb{1}-T)^{-1}$ are closed, and the closed graph theorem implies that the boundedness of $(\lambda\mathbb{1}-T)^{-1}$ is equivalent to the fact that $\operatorname{ran}(\lambda\mathbb{1}-T)$ is closed.

Hence, we summarize the new definitions for these important operators

Definition 5.11: Let $(X, \|\cdot\|)$ be a Banach space and $T \in \mathcal{L}(X)$ be a closed operator. Then

$$\rho(T) = \{\lambda \in \mathbb{K} : \lambda \mathbb{1} - T \text{ is bijective (from } D(T) \to X)\}$$

$$= \{\lambda \in \mathbb{K} : (\lambda \mathbb{1} - T)^{-1} \in \mathcal{B}(X)\}$$

$$\sigma(T) = \{\lambda \in \mathbb{K} : \lambda \mathbb{1} - T \text{ is not bijective}\}$$

$$\sigma_c(T) = \{\lambda \in \mathbb{K} : \lambda \mathbb{1} - T \text{ is injective}, \overline{\operatorname{ran}(\lambda \mathbb{1} - T)} = X, \operatorname{ran}(\lambda \mathbb{1} - T) \neq X\}$$

$$\sigma_r(T) = \{\lambda \in \mathbb{K} : \lambda \mathbb{1} - T \text{ is injective}, \overline{\operatorname{ran}(\lambda \mathbb{1} - T)} \neq X\}$$

$$\sigma_p(T) = \{\lambda \in \mathbb{K} : \lambda \mathbb{1} - T \text{ is not injective}\}$$

Remark: $T \in \mathcal{B}(X)$, then $\rho(T) = \{\lambda \in \mathbb{K} : (\lambda \mathbb{1} - T)^{-1} \in \mathcal{B}(X)\}$

Example 5.12: Multiplication operator

$(L^2([0,1]), \|\cdot\|_2)$, $M \in \mathcal{B}(L^2([0,1]))$, $(Mf)(x) = x f(x)$

check that $\|M\| = 1$, so M is bounded.

$$((\lambda \mathbb{1} - M) f)(x) = (\lambda - x) f(x)$$

If $(\lambda - x) f(x) = 0$ a.e. $\Rightarrow f(x) = 0$ a.e. Hence, $\ker M = \{0\}$; i.e.

$(\lambda \mathbb{1} - M)$ is one-to-one, i.e. $\sigma_p(M) = \emptyset$.

<u>Claim:</u> $\left(\left(\lambda \mathbb{1} - T\right)^{-1} f\right)(x) = \dfrac{1}{\lambda - x} f(x)$

provided the right hand side belongs to $L^2([0,1])$
If $\lambda \notin [0,1]$, then $\left|\dfrac{1}{\lambda - x}\right| < C \;\; \forall\, x \in [0,1]$

Then $(\lambda \mathbb{1} - M)^{-1}$ is a well-defined bounded operator, and

$\lambda \in \rho(M)$. Thus $\rho(M) \supset \mathbb{C} \setminus [0,1]$

If $\lambda \in [0,1]$, then $(\lambda \mathbb{1} - M)^{-1}$ is unbounded since
$\|(\lambda \mathbb{1} - M)^{-1} f\|_2 = \infty$ when f is a constant function.
Indeed, there are no functions $f \in L^2([0,1])$ s.t.

$$(\lambda - x) f(x) = c \quad \text{a.e,}$$

hence $(\lambda \mathbb{1} - M)$ is injective but not onto. Pick $f \in L^2([0,1])$

$$f_n(x) = \begin{cases} f(x) & \text{if } |\lambda - x| > 1/n \\ 0 & \text{if } |\lambda - x| < 1/n \end{cases}$$

$f_n \longrightarrow f$ as $n \to \infty$ in $L^2([0,1], \|\cdot\|_2)$,

and $f_n \in \operatorname{ran}(\lambda \mathbb{1} - M)$, since the preimage

is $h_n(x) = \dfrac{1}{\lambda - x} f_n(t)$ with $h_n \in L^2([0,1])$.

$\rho(M) = \mathbb{C} \setminus [0,1]$; $\sigma_p(M) = \sigma_r(M) = \emptyset$, and

$\sigma_c(M) = [0,1]$.

Recall the following Theorem from FA I.

<u>Theorem 5.13:</u> Let $(X, \|\cdot\|_X)$ be a Banach space. If $T \in \mathcal{B}(X)$ is an operator with $\|T\| < 1$ then

$(\mathbb{1} - T)$ is invertible and the inverse is given by

$$(\mathbb{1} - T)^{-1} = \sum_{n=0}^{\infty} T^n$$

Proof: FA I $\blacksquare$

<u>Proposition 5.14:</u> Let $(X, \|\cdot\|_X)$ be a ~~normed~~ Banach space and $T \in \mathcal{L}(X)$ be a linear operator.
Then the resolvent set $\rho(T)$ is open and the spectrum $\sigma(T)$ is closed.

<u>Proof:</u> $X = \{0\} \Rightarrow \rho(T) = \mathbb{K}$, and $\rho(T) = \emptyset$ clear.

Hence, $\rho(T) \neq \emptyset$, $X \neq \{0\}$

Pick $\lambda_0 \in \rho(T)$.
If T is closed, then (see def. 5.11 and Inverse Mapping Theorem)

$(\lambda_0 \mathbb{1} - T)^{-1} \in \mathcal{B}(X)$, and

$$(\lambda \mathbb{1} - T)^{-1} = (\lambda \mathbb{1} - \lambda_0 \mathbb{1} + \lambda_0 \mathbb{1} - T)^{-1} \in \mathcal{B}(X)$$

$$= \left(\frac{1}{\lambda_0 - T}\right) \left[1 - \left(\frac{(\lambda_0 - \lambda)\mathbb{1}}{(\lambda_0 - T)}\right)\right]^{-1} \quad \left(\frac{1}{\lambda_0 - T} \hat{=} (\lambda_0 \mathbb{1} - T)^{-1}\right)$$

with Theorem 5.13 we get

$$=(\lambda_0\mathbb{1}-T)^{-1}\left[\mathbb{1}+\sum_{n=1}^{\infty}(\lambda_0\mathbb{1}-\lambda\mathbb{1})^{n}(\lambda_0\mathbb{1}-T)^{-n}\right],$$ hence set

$\tilde{R}_\lambda(T)=R_{\lambda_0}(T)\left[\mathbb{1}+\sum_{n=1}^{\infty}(\lambda_0-\lambda)^n R_{\lambda_0}(T)^n\right]$. We shall show that $\tilde{R}_\lambda(T)=R_\lambda(T)$ for certain values of λ (dependent of λ_0). For this we need to check for which λ the series converges.

Note $\|R_{\lambda_0}(T)^n\| \le \|R_{\lambda_0}(T)\|^n \Rightarrow$ the series converges if

$|\lambda-\lambda_0| < \|R_{\lambda_0}(T)\|^{-1}$ (because $|\lambda_0-\lambda|\,\|R_{\lambda_0}(T)\| < 1$) (Note that $\|(\lambda_0\mathbb{1}-T)^{-1}\|>0$). For such λ the r.h.s for $\tilde{R}_\lambda(T)$ is well defined, and it is easily checked that

$(\lambda\mathbb{1}-T)\tilde{R}_\lambda(T)=\mathbb{1}=\tilde{R}_\lambda(T)(\lambda\mathbb{1}-T)$ (use that $[\dots]=\left(\frac{1}{1-(\lambda_0-\lambda)R_{\lambda_0}(T)}\right)$),

hence $\lambda\in\rho(T)$ if $|\lambda-\lambda_0|<\|R_{\lambda_0}(T)\|^{-1}$, and $\tilde{R}_\lambda(T)=R_\lambda(T)$), and thus we conclude $\rho(T)$ open.

If $T\in\mathcal{L}(X)$ is not closed, the proof goes similar but one has first to check that $(\lambda\mathbb{1}-T)^{-1}$ ~~exists~~ exists and that $\overline{\operatorname{ran}(\lambda\mathbb{1}-T)}=X$.

For the first check, pick $f\in\mathcal{D}(T)$

$(\lambda\mathbb{1}-T)f=(\lambda-\lambda_0)f+(\lambda_0 f-Tf)$

$\|(\lambda\mathbb{1}-T)f\|_X \ge \|(\lambda_0\mathbb{1}-T)f\|_X - |\lambda-\lambda_0|\,\|f\|_X$

$\Rightarrow \|f\|_X=\|(\lambda_0\mathbb{1}-T)^{-1}(\lambda_0\mathbb{1}-T)f\|_X \le \|(\lambda_0\mathbb{1}-T)^{-1}\|\,\|(\lambda_0\mathbb{1}-T)f\|_X$

$$\Rightarrow \|(\lambda\mathbb{1}-T)f\|_X \ge \underbrace{\left(\|(\lambda_0\mathbb{1}-T)^{-1}\|^{-1}-|\lambda-\lambda_0|\right)}_{>0}\|f\|_X$$

because $|\lambda-\lambda_0|<\|(\lambda_0\mathbb{1}-T)^{-1}\|^{-1}$.

We skip the remaining part. If we have take more time we'll finish the last piece for not necessarily closed operators.

The proof of the last proposition easily implies the following representation of the resolvent.

<u>Corollary 5.15</u>: Let $(X, \|\cdot\|_X)$ be a Banach space, and let $T \in \mathcal{L}(X)$ be a closed operator.

For any $\lambda_0 \in \rho(T)$, and any $\lambda \in \mathbb{K}$ with

$|\lambda - \lambda_0| < \|R_{\lambda_0}(T)\|^{-1}$, we have $\lambda \in \rho(T)$

and

$$R_\lambda(T) = R_{\lambda_0}(T) \sum_{n=0}^{\infty} (\lambda - \lambda_0)^n R_{\lambda_0}(T)^n$$

$$= \sum_{n=0}^{\infty} (\lambda - \lambda_0)^n R_{\lambda_0}(T)^{n+1}$$

<u>Proof</u>: Prop. 5.14 and its proof. ∎

<u>Corollary 5.16</u>: Let $(X, \|\cdot\|_X)$ be a Banach space and $T \in \mathcal{B}(X)$. Then

$$\sigma(T) \subset B_{\|T\|}(0).$$

<u>Proof</u>: $\lambda \in \mathbb{K}$ with $|\lambda| > \|T\|$

then $\|\frac{1}{\lambda} T\| < 1$, so $(\mathbb{1} - \frac{1}{\lambda} T)^{-1}$ exists

by Theorem 5.13.
But $(\mathbb{1} - 1/\lambda\, T)^{-1} = +1/\lambda\,(\lambda\mathbb{1} - T)^{-1}$, and thus $\lambda \in \rho(T)$. ■

<u>Proposition 5.17</u>: Let $(X, \|\cdot\|_X)$ be a Banach space and $T \in B(X)$. Then the following holds

(a) $\{\lambda \in \mathbb{K} : |\lambda| > \|T\|\} \subset \rho(T)$

and for $|\lambda| > \|T\|$

$$R_\lambda(T) = \sum_{k=0}^{\infty} \lambda^{-k-1} T^k$$

(b) $\sigma(T)$ is compact

(c) X complex and $X \neq \{0\}$, then $\sigma(T) \neq \emptyset$

<u>Proof</u>: (a) – (b) follow from the previous proposition.

ad(c): Suppose $\sigma(T) = \emptyset$, that is, $\rho(T) = \mathbb{C}$.

Then $R_{\cdot}(T) : \mathbb{C} \to B(X)$ has to be bounded if $|\lambda| > \|T\|$, then we estimate

$$\|(\lambda 1\!\!1 - T)f\|_X \geq (|\lambda| - \|T\|)\|f\|_X \quad ,$$

and

$$\|R_\lambda(T)\| \leq (|\lambda| - \|T\|)^{-1} \to 0 \text{ as } |\lambda| \to \infty .$$

But $R_\cdot(T)$ and $\|R_\cdot(T)\|$ are continuous on $\mathbb{C}$, hence $R_\cdot(T)$ is bounded.
But the Theorem of Liouville says, that any analytic function on $\mathbb{C}$ is bounded iff it is constant.
Why analytic? See Prop. 5.17 (a).

$\Rightarrow$ this constant has to be zero as $\|R_\lambda(T)\| \to 0$ as $|\lambda| \to \infty$

$\Rightarrow R_\lambda(T) = 0 \ \forall \lambda \in \mathbb{C}$

$\Rightarrow \operatorname{ran}((\lambda 1\!\!1 - T)^{-1}) = D(\lambda 1\!\!1 - T) = X \ \forall \lambda \in \mathbb{C}$

$\Rightarrow X = \{0\} \longrightarrow$ contradiction. ■

<u>Proposition 5.18</u>: Let $(X, \|\cdot\|_X)$ be a Banach space and $T \in \mathcal{B}(X)$. Then the following holds

(a) The sequence $(\|T^n\|^{\frac{1}{n}})_{n\in\mathbb{N}}$ is convergent

(b) If $|\lambda| > \lim_{n\to\infty} \|T^n\|^{1/n}$, then $\lambda \in \rho(T)$, and

$$R(\lambda,T) = \sum_{n=0}^{\infty} \lambda^{-n-1} T^n .$$

(c) If X is complex and $X \neq \{0\}$, then

$$r_\sigma(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}$$
$$= \lim_{n\to\infty} \|T^n\|^{1/n}$$

$r_\sigma(T)$ is called the spectral radius of T.

Proof. (a) Pick $m \in \mathbb{N}$, then $\forall\, k \geq m$ we have

$k = m q_k + r_k$, $q_k, r_k \in \mathbb{N}\cup\{0\}$ and $r_k < m$

$$\|T^k\| = \|T^{mq_k + r_k}\| \leq \|T^m\|^{q_k} \|T^{r_k}\|$$
$$\leq \|T^m\|^{q_k} \underbrace{\max\{1, \|T\|, \dots, \|T^{m-1}\|\}}_{=:\alpha}$$

$$\|T^k\|^{1/k} \leq \|T^m\|^{\frac{q_k}{k}} \alpha^{1/k} = \|T^m\|^{\frac{k-r_k}{km}} \alpha^{1/k} = \|T^m\|^{\frac{1}{m} - \frac{r_k}{km}} \alpha^{1/k}$$

$$\Rightarrow \limsup_{k\to\infty} \|T^k\|^{1/k} \leq \|T^m\|^{\frac{1}{m}} \limsup_{k\to\infty} \left(\|T^m\|^{-\frac{r_k}{m}} \alpha\right)^{1/k} = \|T^m\|^{1/m}$$

because of $-1 < -\frac{r_k}{m} \leq 0$. Hence,

$$\limsup_{k\to\infty} \|T^k\|^{1/k} \leq \liminf_{m\to\infty} \|T^m\|^{\frac{1}{m}}$$

$\Rightarrow$ (a).

(b) $\sum_{k=0}^{\infty} \lambda^{-k-1} T^k$ Power series in λ^{-1} with values in $B(X)$; convergence radius is $\left(\limsup_{k\to\infty} \|T^k\|^{1/k}\right)^{-1} = \left(\lim_{k\to\infty} \|T^k\|^{1/k}\right)^{-1}$

(rest see analysis).

(c) $\sup\{|\lambda| : \lambda \in \sigma(T)\} \leq \lim_{k\to\infty} \|T^k\|^{1/k}$,

Laurent-series (see complex analysis)
(We skip this part!)

Note that $R_\lambda(T) = \frac{1}{\lambda}\left(1\!\!1 + \sum_{n=1}^{\infty} \left(\frac{T}{\lambda}\right)^n\right)$ is a Laurent-series about ∞. ■

The resolvent set is connected with the inverse of operators. There are some characterizations for ~~the~~ the existence of the inverse of an operator. Roughly speaking, if one has a lower bound on the image under some circumstances one can conclude that T invertible (see one example on sheet 6).

<u>Proposition 5.19:</u> $(X, \|\cdot\|_X)$ be a Hilbert space and $T \in B(X)$ be self-adjoint.

$\lambda \in \rho(T) \iff \exists c > 0$ s.t. $\|(\lambda 1\!\!1 - T)x\|_X \geq c\|x\|_X \quad \forall x \in X$

„$\Rightarrow$": okay.

Proof. „$\Leftarrow$": It follows that $\ker(\lambda \mathbb{1} - T) = \{0\}$,

so $(\lambda\mathbb{1}-T)$ is injective. Because of $\|(\bar\lambda\mathbb{1}-T)x\|_X = \|(\lambda\mathbb{1}-T)x\|_X$

$\ker(\bar\lambda\mathbb{1}-T) = \{0\}$ as well.

By Theorem 5.5 $(\overline{\mathrm{ran}(T)} = (\ker T^*)^\perp)$ $\overline{\mathrm{ran}(\lambda\mathbb{1}-T)} = \left(\ker(\bar\lambda\mathbb{1}-T)\right)^\perp = X$.

There exists an inverse map $(\lambda\mathbb{1}-T)^{-1}: \mathrm{ran}(\lambda\mathbb{1}-T) \to X$,

and $$\|(\lambda\mathbb{1}-T)^{-1}\| = \sup_{\substack{y\neq 0\\ y\in\mathrm{ran}(\lambda\mathbb{1}-T)}} \frac{\|(\lambda\mathbb{1}-T)^{-1}y\|_X}{\|y\|_X} \leq \frac{1}{c}$$

because $\|(\lambda\mathbb{1}-T)^{-1}y\|_X = \|x\|_X < \frac{1}{c}\|(\lambda\mathbb{1}-T)x\|_X$.

Actually, the range of $(\lambda\mathbb{1}-T)$ is the whole of X:
$x_n \in \mathrm{ran}(\lambda\mathbb{1}-T)$ and $x_n \to x$, then $y_n = (\lambda\mathbb{1}-T)^{-1}x_n$
is a Cauchy sequence with some limit y that satisfies (why?
– continuity) $x = (\lambda\mathbb{1}-T)y$)

$\lambda \in \rho(T)$. ■

<u>Remark:</u> (1) Banach space $(X, \|\cdot\|_X)$ and normed space $(Y, \|\cdot\|_Y)$
(a) $T \in \mathcal{B}(X,Y)$
$\|Tx\|_Y \geq \alpha\|x\|_X \;\forall x\in X \Rightarrow \mathrm{Im}(T)$ closed

(b) $(Y, \|\cdot\|_Y)$ also Banach space. Then:
T invertible $\Leftrightarrow \exists \alpha>0$ s.t. $\|Tx\|_Y \geq \alpha\|x\|_X \;\forall x\in X$
and $\mathrm{Im}(T)$ is dense in Y.

recall the estimation (Corollary 3.29):
$$\|Tx\|_Y \geq \|T^{-1}\|^{-1}\|x\|_X.$$

Equivalent statement is that

$$\lambda \in \sigma(T) \iff \inf_{\|x\|_X = 1} \|(\lambda \mathbb{1} - T)x\|_X = 0.$$

This generalises the notion of eigenvalues, since

$$\lambda \in \sigma_p(T) \iff \ker(\lambda \mathbb{1} - T) \neq \{0\}.$$

Hence, we introduce a new notion of eigenvalues in the following definition.

<u>Definition 5.20</u>: Let $(X, \|\cdot\|_X)$ be a normed space and $A \in \mathcal{L}(X)$. A number $\lambda \in \mathbb{K}$ is called generalized eigenvalue of A, if there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathcal{D}(A)$ (domain of definition) with $\|x_n\|_X = 1$ $\forall n \in \mathbb{N}$, such that

$$\lim_{n\to\infty} (\lambda \mathbb{1} - A) x_n = 0$$

$$\sigma_K(A) = \{\lambda \in \mathbb{K} : \lambda \text{ generalized eigenvalue}\},$$

and it is called spectral kernel of A.

Without proof we state the following property:

$$\boxed{\partial \sigma(A) \subset \sigma_K(A).}$$

Recall that for $A \in \mathcal{B}(X)$ self-adjoint we have (see Prop. 5.4(e))

$\|A\|^2 = \|A^2\|$, and with (c) of Proposition 5.18 we get that $r_\sigma(A) = \|A\|$.

Therefore we have a natural bound on the spectrum. Moreover, the spectrum of a self-adjoint operator is ~~...~~ real. We summarize these important properties in the next proposition.

<u>Proposition 5.21:</u> Let $(X, \|\cdot\|_X)$ be a Hilbert space ($\|\cdot\|_X$ induced by $(\cdot,\cdot)$), and $A \in \mathcal{B}(X)$ self-adjoint. Then

$$\sigma(A) \subset [-\|A\|, \|A\|] \quad \text{, respectively}$$

$$\sigma(A) \subset [m, M] \quad \text{where}$$

$$m = \inf_{\|x\|_X = 1} (x, Ax) \quad ; \quad M = \sup_{\|x\|_X = 1} (x, Ax)$$

<u>Proof.</u> First we show $\sigma(A) \subset \mathbb{R}$. Let $\lambda = a + ib$, $a, b \in \mathbb{R}$.

$$\|(\lambda \mathbb{1} - A)x\|_X^2 = ((\lambda \mathbb{1} - A)x, (\lambda \mathbb{1} - A)x)$$

$$= ((a\mathbb{1} - A)x, (a\mathbb{1} - A)x) + ((ib)x, (ib)x) + ((a\mathbb{1} - A)x, (ib)x) + ((ib)x, (a\mathbb{1} - A)x)$$

$$= \|(a\mathbb{1} - A)x\|_X^2 + b^2\|x\|_X^2 \geq b^2 \|x\|_X^2$$

With Prop. 5.12 we conclude that $\lambda \in \rho(A)$ whenever $b \neq 0$.

Suppose $\lambda = M + \varepsilon$, $\varepsilon > 0$. $(x, Ax) \leq M\|x\|_X^2$

$$\varepsilon\|x\|_X^2 = (\lambda - M)\|x\|_X^2 \leq \lambda(x, x) - (x, Ax)$$

$= (x, (\lambda 1\!\!1 - A)x) \underset{\text{Schwarz}}{\leq} \|x\|_X \|(\lambda 1\!\!1 - A)x\|_X$,

then again Prop. 5.12 implies $\lambda \in \rho(A)$. Same with $\lambda = m - \varepsilon$. ■

Proposition 5.22: Let $(X, \|\cdot\|_X)$ be a Hilbert space, and $A \in \mathcal{B}(X)$.

If $\lambda \in \sigma_r(A)$, then $\bar{\lambda} \in \sigma_p(A^*)$.

Proof. If $\lambda \in \sigma_r(A)$, then $\operatorname{ran}(\lambda 1\!\!1 - A)$ is not dense in X.

$X = \overline{\operatorname{ran}(\lambda 1\!\!1 - A)} \oplus M$, and $M \neq \{0\}$.

$X = \overline{\operatorname{ran}(\lambda 1\!\!1 - A)} \oplus \ker(\bar{\lambda} 1\!\!1 - A^*)$ (see Theorem 5.5).

$\Rightarrow \ker(\bar{\lambda} 1\!\!1 - A^*) \neq \{0\}$, so $\bar{\lambda} \in \sigma_p(A^*)$. ■

Proposition 5.23: Let $(X, \|\cdot\|_X)$ be a Banach space, and $A \in \mathcal{L}(X)$ be closed.

Let $\alpha, \beta \in \mathbb{K}$ and $X \neq \{0\}$.

Then:

(i) $\alpha = 0$ and $\overline{\mathcal{D}(A)} \neq X$: $\sigma(\alpha A + \beta 1\!\!1) = \mathbb{K}$

(ii) $\alpha = 0$ and $\overline{\mathcal{D}(A)} = X$: $\sigma(\alpha A + \beta 1\!\!1) = \{\beta\}$

(iii) $\forall \alpha \neq 0$:

$$\sigma(\alpha A + \beta 1\!\!1) = \{\alpha\lambda + \beta : \lambda \in \sigma(A)\}$$

Proof. exercise / sheet ■

5.3 Compact operators

We revise some important properties of a special class of linear operators which are close to matrices.

Definition 5.24: Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. $T \in \mathcal{L}(X, Y)$ is compact if, for any bounded sequence $(x_n)_{n\in\mathbb{N}}$ in X, the sequence $(Tx_n)_{n\in\mathbb{N}}$ in Y contains a convergent subsequence.
The set of all compact operators is denoted by $K(X, Y)$.

Theorem 5.25: Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. $T \in \mathcal{L}(X, Y)$. Then the following holds

(a) T is compact $\iff$ $\forall$ bounded sets $A \subset X$ $T(A)$ is relatively compact in Y

(b) T compact $\Rightarrow$ $\mathrm{Im}(T)$ and $\overline{\mathrm{Im}(T)}$ are separable.

(c) $K(X, Y) \subset \mathcal{B}(X, Y)$

(d) If X is an infinite-dimensional normed space then the identity $\mathbb{1}$ on X is not compact.

Proof. FA I ■

Theorem 5.26: Let $(X,(\cdot,\cdot))$ be an infinite-dimensional Hilbert space. $T \in K(X)$, then $0 \in \sigma(T)$.
If X is separable then either $0 \in \sigma_p(T)$ or $0 \in \sigma(T) \setminus \sigma_p(T)$ may occur. If X is not separable then $0 \in \sigma_p(T)$.

Proof. FA I. ■

Spectral properties of compact self-adjoint operators are particularly simple.

Theorem 5.27: Let $(X,(\cdot,\cdot))$ be a Hilbert space and $T \in K(X)$ self-adjoint.

(a) Either $\sigma(T) = \sigma_p(T)$; or $\sigma(T) = \sigma_p(T) \cup \{0\}$ with $\sigma_c(T) = \{0\}$.

(b) Eigenvalues have finite multiplicity, except possibly 0.

(c) If T has infinite (countable) eigenvalues, then 0 is the only accumulation point.

(d) Spectral theorem
$\lambda_1, \lambda_2 \ldots$ eigenvalues, P_n the corresponding projections onto the eigenspaces
(If $\lambda_n \neq 0$, rank $P_n < \infty$). Then

$$T = \sum_{n \in \mathbb{N}} \lambda_n P_n ,$$

convergence of this series w.r.t. operator norm.

5.4 The Spectral Theorem

We have seen an eigenfunction expansion of any compact operator (Theorem 5.27). In this section we'll derive an analogous expansion for bounded self-adjoint operators.

If A is compact$^{\circledast}$, we get $\circledast$ and self-adjoint

$X = N_0 \oplus N_1 \oplus \cdots$, that is X is the orthogonal sum of subspaces N_k, all of which, except N_0, are finite-dimensional.

$\forall N_k \quad \exists \lambda_k$ with $\lambda_0 = 0$ and $\lambda_k \neq \lambda_j$

$$Af = \lambda_k f \qquad \text{if } f \in N_k$$

P_k orthogonal projection onto N_k, then we can write

$$\left.\begin{aligned} &\mathbb{1} = P_0 + P_1 + P_2 + \cdots \\ &\text{and} \\ &A = \lambda_1 P_1 + \lambda_2 P_2 + \cdots \end{aligned}\right\} \circledast$$

(recall that the only accumulation point of $(\lambda_n)_{n\in\mathbb{N}}$ is 0).

G_t = Linear span $\{ v_k : v_k$ eigenvektor for eigenvalues $\lambda < t \}$ subspace of X; and let E_t denote the corresponding orthogonal projection.

~~Then~~ It is easy to see that ~~that~~ both limits E_{t-0} and E_{t+0} exist, and that E_t is a left-continuous operator-valued function. If λ_k is an eigenvalue, then

$$E_{\lambda_k+0} - E_{\lambda_k} = P_k$$

The idea is to write the sums in (*) as certain operator-valued integrals. In the following we'll introduce an operator-valued measure, called spectral measure.

Recall that for $A, B \in \mathcal{B}(X)$ self-adjoint:
$A \leq B \iff (Ax, x) \leq (Bx, x) \quad \forall x \in X.$

<u>Definition 5.28</u>: Let $(X, \|\cdot\|_X)$ be a Hilbert space. A resolution E of the identity is a 1-parameter family of bounded operators, i.e.

$E: \mathbb{R} \to \mathcal{B}(X)$, with the following properties

(a) $E(\lambda)$ is orthogonal projection for all $\lambda \in \mathbb{R}$

(b) $E(\lambda_1) \leq E(\lambda_2)$ for all $\lambda_1 \leq \lambda_2$

(c) $\lim\limits_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} E(t-\varepsilon) = E(t) \quad \forall t \in \mathbb{R}$ (convergence in the strong sense)

(shorthand $E(t-0) = E(t) \; \forall t \in \mathbb{R}$)

(d) $\lim\limits_{t \to -\infty} E(t) = 0$ and $\lim\limits_{t \to +\infty} E(t) = \mathbb{1}$

<u>Remarks:</u> (1) $A, A_n \in \mathcal{B}(X)$ $A_n \to A$ as $n \to \infty$ in the strong sense if $\forall f \in X$ $\|A_n f - A f\|_X \to 0$ as $n \to \infty$.

(2) Property (b) is equivalent with $E(\lambda_1)E(\lambda_2) = E(s)$ with $s = \min\{\lambda_1, \lambda_2\}$. Let $s = \lambda_1$, then
$(E(\lambda_1)x, x))^2 = (E(\lambda_2)x, E(\lambda_1)x)^2 \leq (E(\lambda_2)x, x)(E(\lambda_1)x, x)$
$\Rightarrow E(\lambda_1) \leq E(\lambda_2)$.

First $(E(\lambda_1)x, x) = (E(\lambda_1)E(\lambda_1)x, E(\lambda_1)x) = (E(\lambda_2)E(\lambda_1)x, E(\lambda_1)x)$

$\Rightarrow (E(\lambda_1)x, x) \le (E(\lambda_2)x, x) \quad \forall x \in \operatorname{ran}(E(\lambda_1)) \cup \operatorname{ran}(E(\lambda_1))^\perp$

$\Rightarrow E(\lambda_1) \le E(\lambda_2)$

(3) $\forall \lambda \in \mathbb{R}$ the limit

$$E(\lambda+) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} E(\lambda + \varepsilon) \text{ is an orthogonal projection.}$$

__Example 5.29:__ $(X, \|\cdot\|_X)$ be a Hilbert space and $P \in \mathcal{B}(X)$ some orthogonal projection, and $\lambda_1 \le \lambda_2$, $\lambda_1, \lambda_2 \in \mathbb{R}$, then $E: \mathbb{R} \to \mathcal{B}(X)$,

$$\lambda \longmapsto E(\lambda) = \begin{cases} 0 & \text{if } \lambda \le \lambda_1 \\ P & \text{if } \lambda_1 < \lambda \le \lambda_2 \\ \mathbb{1} & \text{if } \lambda_2 < \lambda \end{cases}$$

is a resolution of the identity.

__Lemma 5.30:__ $(X, \|\cdot\|_X)$ be a Hilbert space and E a resolution of the identity. It follows that, for any $f \in X$, the function

$$\varphi_f : \mathbb{R} \to \mathbb{R}, \quad \lambda \longmapsto \varphi_f(\lambda) = (E(\lambda)f, f) = \|E(\lambda)f\|_X^2$$

is left-continuous, and ~~is~~ non-decreasing and bounded.

__Proof.__ ~~Straigh~~ Easy exercise from the definitions. ∎

__Remark:__ φ_f defines a measure on $\mathbb{R}$ (similar to distribution functions). We will not explore this further.

We shall now define an operator valued measure on $\mathbb{R}$.

Definition 5.31: Spectral measure

Let $(X, \|\cdot\|_X)$ be a Hilbert space and $E: \mathbb{R} \to \mathcal{B}(X)$ be a resolution of the identity.

(a) For any bounded intervall $J \subset \mathbb{R}$ the E-measure or spectral measure of J is defined by

$$E(J) = \begin{cases} E(\beta) - E(\alpha+) & \text{if } J = (\alpha, \beta) \\ E(\beta) - E(\alpha) & \text{if } J = [\alpha, \beta) \\ E(\beta) - E(\alpha+) & \text{if } J = (\alpha, \beta] \\ E(\beta) - E(\alpha) & \text{if } J = [\alpha, \beta] \end{cases}.$$

(b) If $u: \mathbb{R} \to \mathbb{R}$ is a step function, $u = \sum_{k=1}^{m} \alpha_k \chi_{J_k}$,

$\left(\chi_A(x) = 1 \text{ if } x \in A, \ \chi_A(x) = 0 \text{ if } x \notin A \right)$,

then the spectral integral of u is defined as

$$\int u(\lambda)\, dE(\lambda) := \sum_{k=1}^{m} \alpha_k \, E(J_k).$$

Lemma 5.32: (Properties) $(X, \|\cdot\|_X)$ Hilbert space, E resolution of the identity, then for all $f \in X$, $\lambda \in \mathbb{R}$ and all bounded intervalls $J, J_1, J_2 \subset \mathbb{R}$:

(a) $\varphi_f(J) = \|E(J) f\|^2$

(b) $E(\lambda) E(J) = E(J) E(\lambda)$

(c) $E(J_1)E(J_2) = E(J_2)E(J_1) = E(J_1 \cap J_2)$

and $E(J_1)E(J_2) = 0$ if $J_1 \cap J_2 = \emptyset$

Proof. (a) and (b) are easy.
(c) follows from property (b) of Definition 5.28 and its equivalence with $E(\lambda_1)E(\lambda_2) = E(s)$ with $s = \min\{\lambda_1, \lambda_2\}$.
$J_1 = [\alpha, \beta)$, $J_2 = [\alpha', \beta')$ with $\alpha < \alpha' < \beta$, $\alpha' < \beta' < \beta$.

$$E(J_1)E(J_2) = \big(E(\beta) - E(\alpha)\big)\big(E(\beta') - E(\alpha')\big)$$

$$= E(\beta)E(\beta') - E(\beta)E(\alpha') - E(\alpha)E(\beta') + E(\alpha)E(\alpha')$$
$$= E(\beta') - E(\alpha') - E(\alpha) + E(\alpha) = E(J_1)$$
$$= E(J_1 \cap J_2).$$ ■

We are now in the position to define a spectral integral. For our purpose we only outline the main ideas, proofs and details are omitted.

Recall that φ_f is a measure on $\mathbb{R}$ (and is a positive real number for any bounded interval.
$\forall f \in X$: u be a step function, $u = \sum_{k=1}^{n} \alpha_k \chi_{J_k}$,

$$\left\| \int u(\lambda)\, dE(\lambda) f \right\|^2 = \left\| \sum_{k=1}^{n} \alpha_k E(J_k) f \right\|^2$$

$$= \int |u(\lambda)|^2 \, d\varphi_f(\lambda) = \int |u(\lambda)|^2 \, d(E(\lambda)f, f) = \|u\|_2^2,$$

where $\|\cdot\|_2$ in $L_2(\mathbb{R}, \varphi_f)$. (Note that $u: \mathbb{R} \to \mathbb{R}$).

Note that step functions are dense in $L_2(\mathbb{R}, \varphi_f)$ (for any $f \in X$). This allows to define an integral exactly as you know for the Lebesgue integral.

__Definition 5.33__: Let $(X, \|\cdot\|_X)$ be a Hilbert space and E be a resolution of the identity. Let $u : \mathbb{R} \to \mathbb{K}$ be an E-measurable function and $(u_k)_{k \in \mathbb{N}}$ be a sequence of step functions with $\|u_k - u\|_2 \to 0$ as $k \to \infty$. Then a linear operator $\mathcal{E}(u)$ is defined by

$$\mathcal{D}(\mathcal{E}(u)) = \{ f \in X : u \in L_2(\mathbb{R}, \varphi_f) \} = \{ f \in X : \int |u(\lambda)|^2 \, d(E(\lambda) f, f) < \infty \}$$

and

$$\mathcal{E}(u) f = \int u(\lambda) \, dE(\lambda) f = \lim_{k \to \infty} \int u_k(\lambda) \, dE(\lambda) f .$$

$\mathcal{E}(u)$ is called the E-integral or spectral integral of u.

The E-integral is a linear operator on its domain of definition $\mathcal{D}(\mathcal{E}(u))$. For general <u>measurable</u> functions $u : \mathbb{R} \to \mathbb{K}$ one can show that $\overline{\mathcal{D}(\mathcal{E}(u))} = X$.
But we restrict ourself to bounded operators in this chapter. Hence, we focus on bounded measurable functions $u : \mathbb{R} \to \mathbb{K}$. As one expects, in this case $\mathcal{E}(u)$ is a bounded operator.

__Proposition 5.34:__ Let $(X, \|\cdot\|_X)$ be a Hilbert space and E be a resolution of the identity. Let $u : \mathbb{R} \to \mathbb{K}$ be a bounded E-measurable function. Then the following holds:

(a) $\mathcal{E}(u) \in \mathcal{B}(X)$, i.e. $\mathcal{D}(\mathcal{E}(u)) = X$ and $\mathcal{E}(u)$ bounded.

(b) $\forall f, g : \left(\mathcal{E}(u) f, g\right) = \int u(\lambda)\, d(E(\lambda) f, g)$

(c) if $u : \mathbb{R} \to \mathbb{R}$ is real-valued the operator $\mathcal{E}(u)$ is self-adjoint.

<u>Proof.</u> (a) $\displaystyle\int |u(\lambda)|^2\, d(E(\lambda) f, f) \leq \sup_{\lambda \in \mathbb{R}} |u(\lambda)|^2 \int d(E(\lambda) f, f)$

$$= \sup_{\lambda \in \mathbb{R}} |u(\lambda)|^2\, \|f\|_X^2 \text{ , hence}$$

$$\|\mathcal{E}(u)\| \leq \sup_{\lambda \in \mathbb{R}} |u(\lambda)| \quad \text{and} \quad \mathcal{D}(\mathcal{E}(u)) = X .$$

(b) Clear from the definition

(c) (via the sequence of step functions). First not that $\mathcal{D}(\mathcal{E}(\bar{u})) = \mathcal{D}(\mathcal{E}(u))$.

$$(\mathcal{E}(\bar{u}) f, g) = \int \overline{u(\lambda)}\, d(E(\lambda) f, g) = \int \overline{u(\lambda)}\, d(f, E(\lambda) g)$$

$$= \overline{\int u(\lambda)\, d(E(\lambda) g, f)} = \overline{(\mathcal{E}(u) g, f)} = (f, \mathcal{E}(u) g) \quad \forall f, g \in X \; \blacksquare$$

We have seen that a resolution of the identity defines an operator-valued integral, where the integral is a bounded self-adjoint operator when u is real-valued and bounded. To get an expansion / integral for a given bounded self-adjoint operator we first remark that for $A \in \mathcal{B}(X)$ s.-a. the spectrum $\sigma(A) \subset [-\|A\|, \|A\|]$ is a bounded set in $\mathbb{R}$.

<u>Remark</u>: $E : [\alpha, \beta) \longrightarrow \mathcal{B}(X)$ resolution of the identity if we have $E(\alpha + \varepsilon) \to E(\alpha) = 0$ as $\varepsilon \downarrow 0$ and $E(\beta - \varepsilon) \to \mathbb{1}$ as $\varepsilon \downarrow 0$.

What is the resolution of the identity such that its E-integral

over the identity, i.e. $\int_\alpha^\beta \lambda \, dE(\lambda)$, is exactly a given self-adjoint bounded operator A?
Note that identity on $[\alpha, \beta]$ is a bounded function.

Definition 5.35: Let $(X, \|\cdot\|_X)$ be a Hilbert space, $T \in \mathcal{L}(X)$, and P orthogonal projection on to a subspace $Y \subset X$.

Then Y is said to reduce the operator if $Pf \in \mathcal{D}(T)$ for all $f \in \mathcal{D}(T)$ and $PTf = TPf \quad \forall f \in \mathcal{D}(T)$.

Remark: If Y reduce the operator T we get

$Tf = T_Y f_Y + T_Z f_Z$ where $f_Y = Pf$ and

$f_Z = Qf$ are the projections of f on to Y and Z, where

$Y \oplus Z = X$. T_Y and T_Z are the restrictions of the operator T to Y and Z.

Definition 5.36: Let $(X, \|\cdot\|_X)$ be a Hilbert space, $A \in \mathcal{B}(X)$ be self-adjoint. A resolution of the identity $E: \mathbb{R} \to \mathcal{B}(X)$ is said to belong to the operator A if

$$A = \mathcal{E}(id), \text{ i.e. } Af = \int \lambda \, dE(\lambda) f \quad \forall f \in X.$$

Remark: If E belongs to $A \in \mathcal{B}(X)$ self-adjoint it follows that $\lim_{\varepsilon \to 0} E(-\|A\| - \varepsilon) = 0$ and $\lim_{\varepsilon \to 0} E(\|A\| - \varepsilon) = 1\!\!1$.

Spectral Theorem

<u>Theorem 5.37</u>: Let $(X, \|\cdot\|_X)$ be a Hilbert space and $A \in \mathcal{B}(X)$ be a bounded linear operator.
If a resolution of the identity $E : \mathbb{R} \to \mathcal{B}(X)$ satisfies

(i) $E(\Delta)$ reduces A for any interval $\Delta \subset \mathbb{R}$

(ii) $f \in \big(E(t) - E(s)\big) X \; (-\infty < s < t < \infty)$
always implies
$$s\|f\|^2 \leq (Af, f) \leq t\|f\|^2$$, then it

belongs to A, i.e.

$$Af = \int \lambda \, dE(\lambda) f \quad \text{for all } f \in X.$$

<u>Proof.</u> Let E satisfy condition (i) and (ii). Then for $\alpha < \beta$ we get $\big(E(\beta) - E(\alpha)\big) f \in \mathcal{D}(A) = X$.
Subdivide now the intervall

$$\alpha = \alpha_0 < \alpha_1 < \alpha_2 < \dots < \alpha_{n-1} < \alpha_n = \beta$$, and

express

$$\big(E(\beta) - E(\alpha)\big) f = \sum_{k=0}^{n-1} \big(E(\alpha_{k+1}) - E(\alpha_k)\big) f$$

Now condition (i) allows to commute A with any $E(\Delta)$, thus

$$(E(\beta) - E(\alpha)) A f = \sum_{k=0}^{n-1} A\big(E(\alpha_{k+1}) - E(\alpha_k)\big) f$$

Condition (ii) gives for $f \in \big(E(t) - E(s)\big) X$ the estimation

$$-\tfrac{1}{2}(t-s)(f,f) \leq \left(Af - \tfrac{1}{2}(s+t)f, f\right) \leq \tfrac{1}{2}(t-s)(f,f).$$

In other words, the restriction of the self-adjoint operator

$A - \tfrac{1}{2}(s+t)\mathbb{1}$ to $(E(t)-E(s))X$ has a norm $\leq \tfrac{1}{2}(t-s)$.

Adding + substracting gives

$$\left(E(\beta)-E(\alpha)\right)Af = \sum_{k=0}^{n-1} \tfrac{1}{2}(\alpha_k+\alpha_{k+1})\,(E(\alpha_{k+1})-E(\alpha_k))f$$

$$+\sum_{k=0}^{n-1}\left(A-\tfrac{1}{2}(\alpha_k+\alpha_{k+1})\mathbb{1}\right)(E(\alpha_{k+1})-E(\alpha_k))f. \qquad (*)$$

Now (Δ-inequality and operatornorm $\leq \tfrac{1}{2}(\alpha_{k+1}-\alpha_k)$)

$$\Big\|\sum_{k=0}^{n-1}\left(A-\tfrac{1}{2}(\alpha_k+\alpha_{k+1})\mathbb{1}\right)(E(\alpha_{k+1})-E(\alpha_k))f\Big\|_X^2$$

$$\leq \sum_{k=0}^{n-1}\left(\tfrac{1}{2}(\alpha_{k+1}-\alpha_k)\right)^2\|(E(\alpha_{k+1})-E(\alpha_k))f\|_X^2 \leq \varepsilon\|f\|_X^2$$

with $\varepsilon = \max\limits_{0\leq k\leq n-1}\tfrac{1}{2}(\alpha_{k+1}-\alpha_k)$. If the diameter 2ε of the subdivision shrinks to zero we get for the first term on the right hand side of $(*)$ an integral (note it is a 'Riemann sum'), hence

$$\left(E(\beta)-E(\alpha)\right)Af = \int_\alpha^\beta \lambda\, dE(\lambda)f \qquad (**)$$

If $\beta \to \|A\|$ (or $\beta \to +\infty$) and $\alpha \to -\|A\|$ (or $\alpha \to -\infty$) we finally get

$$Af = \int \lambda \, dE(\lambda) f \qquad \text{for all } f \in X \quad \blacksquare$$

Remark: One can also prove the converse, that is, if E belongs to A it must satisfy conditions (i) and (ii).

Example 5.38: We study again the operator

$A: L_2([0,1]) \longrightarrow L_2([0,1])$, $f \longmapsto Af$
with $(Af)(x) = x\, f(x)$ for $x \in [0,1]$.
Then $\sigma(A) = [0,1]$. This can be seen as follows.

Pick $\lambda \in [0,1]$ and $\varepsilon > 0$, and we consider that interval from $[\lambda, \lambda+\varepsilon]$ and $[\lambda-\varepsilon, \lambda]$ which lies in $[0,1]$.
W.l.o.g we hence consider $[\lambda, \lambda+\varepsilon] \subset$

Define $f_\varepsilon(x) = \begin{cases} \frac{1}{\sqrt{\varepsilon}} & \text{if } x \in [\lambda, \lambda+\varepsilon] \\ 0 & \text{if } x \notin [\lambda, \lambda+\varepsilon] \end{cases}$

$$\int_0^1 f_\varepsilon^2(x)\, dx = \int_\lambda^{\lambda+\varepsilon} \frac{1}{\varepsilon}\, dx = 1 \Rightarrow f_\varepsilon \in L_2([0,1])$$

and $\|f_\varepsilon\|_2 = 1$.

$$\|(\lambda \mathbb{1} - A) f_\varepsilon\|_2^2 = \frac{1}{\varepsilon} \int_\lambda^{\lambda+\varepsilon} (\lambda - x)^2\, dx = \frac{\varepsilon^2}{3}.$$

As $\varepsilon \to 0$ we get $\| (\lambda 1\!\!1 - A) f_\varepsilon \|_2 \to 0$, and hence $\lambda \in [0,1]$ element of $\sigma(A)$
(Compare the remark after Proposition 5.19, page 108/109),
thus $\sigma(A) = [0,1]$.
A has no eigenvalues because $(\lambda - x) f(x) = 0$ a.e. implies $f(x) = 0$ a.e.

<u>Claim:</u> $E : \mathbb{R} \to \mathcal{B}(X),\ \lambda \mapsto E(\lambda)$

with $$(E(\lambda) f)(x) = \begin{cases} 0 & x \geq \lambda \\ f(x) & x < \lambda \end{cases} \quad ;\ x \in [0,1] \text{ and}$$

$f \in L_2([0,1]) =: X$, is a resolution of the identity that belongs to the operator A.

(i) E is a resolution of the identity. This is easily proved and left as an (important!) exercise.

(ii) E belongs to the operator A, that is (ii) of Theorem 5.37 and

$E(\Delta) A = A E(\Delta)$ for any bounded interval.
This follows because

$$(E(\lambda) A f)(x) = \begin{cases} 0 & ; x \geq \lambda \\ Af(x) & ; x < \lambda \end{cases} = \begin{cases} 0 & ; x \geq \lambda \\ x f(x) & ; x < \lambda \end{cases}$$

$$= x \begin{cases} 0 & ; x \geq \lambda \\ f(x) & ; x < \lambda \end{cases} = (A E(\lambda) f)(x) .$$

We are left to show the estimation ((iii) of Theorem 5.37)

$$s \|f\|_2^2 \leq (Af, f) \leq t \|f\|_2^2 \quad \text{for any } s < t$$

and any $f \in (E(t)-E(s))X$.
$\exists\, h \in L_2([0,1])$ such that

$$f(x) = (E(t)-E(s))h(x) = \begin{cases} h(x) & \text{for } x \in [s,t) \\ 0 & \text{for } x \geq t \\ 0 & \text{for } x < s \end{cases}$$

$$s\|f\|_2^2 = s\int_s^t f^2(\tau)\,d\tau \leq \int_s^t \tau f^2(\tau)\,d\tau$$

$$= \int_0^1 \tau f^2(\tau)\,d\tau = (Af,f) \leq t\int_s^t f^2(\tau)\,d\tau,$$

because of the first mean-value theorem for the integral.

(recall: $f,g: [a,b] \to \mathbb{R}$, $m = \inf f(x)$, $M = \sup f(x)$
g non-negative

$$\int_a^b (f\cdot g)(\tau)\,d\tau = \mu \int_a^b g(\tau)\,d\tau \quad \text{with}$$

$\mu \in [m,M]$. Note f,g are assumed to be integrable) ∎

<u>Notation 5.39</u>: $(X, \|\cdot\|_X)$ be a Hilbertspace and $E: \mathbb{R} \to \mathcal{B}(X)$ be a resolution of the identity.

(a) A point $t \in \mathbb{R}$ is called a <u>point of constancy of E</u> if there is an $\varepsilon > 0$ such that

$$E(t+\varepsilon) - E(t-\varepsilon) = 0;$$

$\geq \varepsilon^2(f,f)$ because $(\lambda-t)^2 \geq \varepsilon^2$ on both intervals.
To see this pick $\eta>0$, then

$$(\lambda-(\lambda+\varepsilon+\eta))^2 = (\varepsilon+\eta)^2 \geq \varepsilon^2,$$
$$(\lambda-(\lambda-\varepsilon-\eta))^2 = (-\varepsilon-\eta)^2 = (\varepsilon+\eta)^2 \geq \varepsilon^2.$$

Hence, we get $\|(\lambda 1\!\!1-A)f\|_X^2 \geq \varepsilon^2(f,f)$, i.e.

$\|(\lambda 1\!\!1-A)f\|_X \geq \varepsilon\|f\|_X$ which implies with Proposition 5.19 that $\lambda\in\varrho(A)$.

"$\Rightarrow$": $\lambda\in\varrho(A)$ (again with Proposition 5.19) implies that there is an $\varepsilon>0$ with

$\|(\lambda 1\!\!1-A)f\|_X \geq \varepsilon\|f\|_X$ for all $f\in X$. This implies

(**) $\int(\lambda-t)^2(dE(t)f,f) \geq \varepsilon^2\int(dE(t)f,f)$. We have to show that λ is a point of constancy of E:
Assume λ is not a point of constancy of E. Choose $g\in X$ and $\eta<\varepsilon$ such that

$$\big(E(\lambda+\eta)-E(\lambda-\eta)\big)g \neq 0.$$

Applying inequality (**) to $f=(E(\lambda+\eta)-E(\lambda-\eta))g$ gives

$$\int_{\lambda-\eta}^{\lambda+\eta}(\lambda-t)^2\,(dE(t)(E(\lambda+\eta)-E(\lambda-\eta))g,(E(\lambda+\eta)-E(\lambda-\eta))g) \geq \varepsilon^2\int_{\lambda-\eta}^{\lambda+\eta}(dE(t)g,g),$$

(some tedious computations are involved using $E(s)E(t)=E(s)$ if $s\leq t$)

which is impossible ($\eta<\varepsilon$!). Hence, our assumption was wrong and λ is a point of constancy of E.

otherwise it is a point of growth.

(b) A point $t \in \mathbb{R}$ is called a point of discontinuity (a jump point) if $E(t+0) - E(t) \neq 0$, and a point of continuity if $E(t+0) - E(t) = 0$.

Recall that $E(t-0) - E(t) = 0$ everywhere, by definition.

Our last statement in this chapter is the following characterization of eigenvalues and point in the resolvent set.

Lemma 5.40: Let $(X, \|\cdot\|_X)$ be a Hilbert space and $A \in \mathcal{B}(X)$ be self-adjoint, and $E : \mathbb{R} \to \mathcal{B}(X)$ be a resolution of the identity which belongs to the operator A, then:

(a) $\lambda \in \varrho(A) \cap \mathbb{R} \iff \lambda$ is a point of constancy of E

(b) $\lambda \in \sigma_p(A) \subset \mathbb{R} \iff \lambda$ is a point of discontinuity of E

Proof: We use that $\|(\lambda \mathbb{1} - A) f\|_X^2 = \int (\lambda - t)^2 (dE(t) f, f)$ (*) for all $f \in X$.

(a) "$\Leftarrow$": Let λ be a point of constancy of E, then $(E(t) f, f)$ (as a function $\mathbb{R} \to \mathbb{R}$) is constant in some neighbourhood of λ (with radius $\varepsilon > 0$ w.l.o.g.)
Then (*) implies that

$$\|(\lambda \mathbb{1} - A) f\|_X^2 \geq \int_{-\infty}^{\lambda - \varepsilon} (\lambda - t)^2 (dE(t) f, f) + \int_{\lambda + \varepsilon}^{\infty} (\lambda - t)^2 (dE(t) f, f)$$

(b) „$\Rightarrow$": Let λ be an eigenvalue of A with eigenvector $f \in X$, then

$$0 = \|(\lambda 1\!\!1 - A) f\|_X^2 = \int (\lambda - t)^2 (dE(t) f, f)$$

This shows that only the point $t = \lambda$ can be a point of growth of the function $(E(t) f, f)$; and since $0 \neq (f, f) = \int (dE(t) f, f)$, the point $t = \lambda$ must be a point of growth. But an isolated point of growth is a point of discontinuity, and therefore

$$(E(\lambda + 0) f, f) \neq (E(\lambda) f, f) \quad \text{i.e.} \quad E(\lambda + 0) \neq E(\lambda)$$

„$\Leftarrow$": λ a point such that $E(\lambda + 0) \neq E(\lambda)$,

hence for some vector $(E(\lambda + 0) - E(\lambda)) g = f \neq 0$

$$\|(\lambda 1\!\!1 - A) f\|_X^2 = \int (\lambda - t)^2 (dE(t) f, f).$$

But $(E(t) f, f) = (E(t)(E(\lambda + 0) - E(\lambda)) g, f)$

is equal to 0 for $t < \lambda$, and does not depend on t for $t > \lambda$. Therefore

$\|(\lambda 1\!\!1 - A) f\|_X = 0$, i.e. λ is an eigenvalue of A with eigenvector $f \in X$. ■

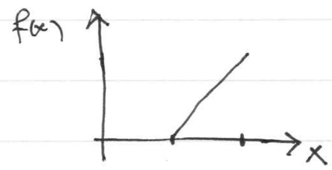
(6) Unbounded Operators

An important class of operators are differential operators. They are unbounded, so that none of the theory so far applies, at least not directly. But many concepts can be extended, and many properties remain true. In this chapter we discuss a few general properties of unbounded operators. We will focus on closed operators (they are "almost continuous"), and discuss as an example the Laplace operator.

6.1 Domain and adjoint

Consider the differential operator $D = \frac{d}{dx}$ in $L^2([0,1])$. If $f \in L_2([0,1])$ is differentiable, then $Df = f'$

Pick $f(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq 1/2 \\ x - 1/2 & 1/2 < x \leq 1 \end{cases}$



It is tempting to define $Df = g$ with $g(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq 1/2 \\ 1 & \text{if } 1/2 < x \leq 1 \end{cases}$ which is a nice $L^2([0,1])$ function.
Thus the operator D can apply to more general functions than differentiable functions. Can we extend D so that Df exists for all $f \in L^2([0,1])$? The answer is no. This is due to the Hellinger-Toeplitz theorem below, an easy conclusion from our closed graph Theorem.
Note that iD is symmetric, in the sense that

$$(f, iDg) = (iDf, g) \qquad \forall f,g \text{ s.t. } Df, Dg \text{ make sense.}$$

This can be seen by integration by parts

Also, D is unbounded.

(Hellinger–Toeplitz theorem)

Theorem 6.1: Let $(X, \|\cdot\|_X)$ be a Hilbert space, and $T \in \mathcal{L}(X)$ (everywhere defined) such that

$$(x, Ty) = (Tx, y) \qquad \text{for all } x, y \in X.$$

Then T is bounded.

Proof.: We will prove that the graph $G(T)$ of the operator is closed.
Suppose $x_n \to x$ and $Tx_n \to y$ as $n \to \infty$.
We need to show that $(x, y) \in G(T)$, that is, that $y = Tx$.
But, for any $z \in X$

$$(z, y) = \lim_{n\to\infty} (z, Tx_n) = \lim_{n\to\infty} (Tz, x_n)$$
$$= (Tz, \overset{x}{\cancel{y}}) = (z, Tx).$$

Thus $y = Tx$ and $G(T)$ is closed ■

This theorem is the cause of much technical pain because in quantum mechanics there are operators which are unbounded but which we want to obey $(x, Ty) = (Tx, y)$ in some sense.
But the Hellinger–Toeplitz theorem tells us that such operators cannot be everywhere defined.
Thus such operators are defined on subspaces $D(A) \subset X$ and defining what one means by $A+B$ or AB may be difficult.
For example, $A+B$ is a priori only defined on $D(A) \cap D(B)$

which may equal $\{0\}$ even in the case where both $D(A)$ and $D(B)$ are dense.

Definition 6.2: Let $(X, \|\cdot\|_X)$ be a Hilbert space

(a) A pair $(T, D(T))$ where $D(T) \subset X$ is a subspace, called domain of T, and where $T: D(T) \to X$ is a linear map, is called an operator in X with domain $D(T)$.

(b) The operator $\tilde{T}$ $((\tilde{T}, D(\tilde{T})))$ is an extension of the operator T $((T, D(T)))$, and T is a restriction of $\tilde{T}$, if
$D(\tilde{T}) \supset D(T)$, and $\tilde{T}x = Tx$ for all $x \in D(T)$.
(in symbols : $T \subset \tilde{T}$)

Remark: ① We have seen that a bounded operator defined on a dense subspace has a unique extension as an operator on the whole space
But if an operator is unbounded, the extension may not be unique. This will be illustrated with the Laplace operator (see example 6.3).

② We always consider densely defined operators T, i.e. operators $(T, D(T))$ with $\overline{D(T)} = X$.

Example 6.3: $(L^2([0,1]), \|\cdot\|_2)$, and let $T_k = \frac{d^2}{dx^2}$, $k = 1,2,3,4$

$D(T_1) = \{u \in \mathcal{C}^2([0,1]) : u(0) = u(1) = 0\}$

② If T is bounded and $D(T)$ is dense in X, then the adjoint T^* is equal to the adjoint of the extension of T on X.

<u>Definition 6.5</u>: Let $(X, \|\cdot\|_X)$ be a Hilbert space and T a densely defined operator in X

(a) T is called <u>symmetric</u> if T^* is an extension of T, i.e. if $D(T) \subset D(T^*)$ and $Tx = T^*x$ for all $x \in D(T)$.

(b) T is <u>self-adjoint</u> if $T = T^*$, that is, $D(T) = D(T^*)$ and T is symmetric.

<u>Remarks</u> ① It is clear from the definition that T self-adjoint $\Rightarrow$ T symmetric

② An alternate definition is that T is symmetric iff

$$(*) \quad \boxed{(Tx, y) = (x, Ty) \quad \text{for all } x, y \in D(T)}$$

(If T satisfies $(*)$ it follows immediately that $D(T) \subset D(T^*)$ and that $T^* = T$ on $D(T)$).

It is usually easy to check if an operator is symmetric, but it is harder to check that the domain of the adjoint is not bigger.

③ If T is bounded and $D(T) = X$, then self-adjoint $\Leftrightarrow$ symmetric

$$\mathcal{D}(T_2) = \mathcal{C}^2([0,1])$$

$$\mathcal{D}(T_3) = \{u \in H^2([0,1]) : u(0) = u(1) = 0\}$$

$$\mathcal{D}(T_4) = H^2([0,1])$$

(H^2 Sobolev space of L^2 functions with first and second (weak) derivatives that are L^2 functions. Sobolev embedding theorem states that $H^2([0,1]) \subset \mathcal{C}^1([0,1])$. More later.)

Notice that $\mathcal{D}(T_1) \subset \mathcal{D}(T_2) \subset \mathcal{D}(T_4)$
and $\mathcal{D}(T_1) \subset \mathcal{D}(T_3) \subset \mathcal{D}(T_4)$.

__Definition 6.4:__ Let $(X, \|\cdot\|_X)$ be a Hilbert space, and T be a densely defined operator in X (i.e. $T : \mathcal{D}(T) \to X$).
Let $\mathcal{D}(T^*)$ be the set of $\varphi \in X$ for which there is an $\eta \in X$ such that

$$(T\psi, \varphi) = (\psi, \eta) \quad \text{for all } \psi \in \mathcal{D}(T).$$

For such $\varphi \in \mathcal{D}(T^*)$, we define $T^*\varphi = \eta$.
T^* is called the adjoint of T.

In other words, T^* is an operator such that $(Tx, y) = (x, T^*y)$ for all $x \in \mathcal{D}(T)$ and for all $y \in \mathcal{D}(T^*)$.

__Remarks:__ ① If $\mathcal{D}(T) = X$ and T bounded, then $\mathcal{D}(T^*) = X$ (Riesz representation theorem), and T^* is defined as in Definition 5.1 on page 92.

Example 6.6: Laplace operator (in one dimension) of example 6.3 above. Integration by parts gives

$$\left(f, \frac{d^2}{dx^2} g\right) = \int_0^1 \overline{f(x)} \left(\frac{d^2}{dx^2} g\right)(x)\, dx = \overline{f(x)} \frac{d}{dx} g(x) \Big|_0^1 - \int_0^1 \overline{\frac{df}{dx}(x)} \frac{dg}{dx}(x)\, dx$$

$$= \overline{f(x)} \frac{d}{dx} g(x) \Big|_0^1 - \overline{\frac{df}{dx}(x)} g(x) \Big|_0^1 + \left(\frac{d^2}{dx^2} f, g\right)$$

(valid for all $f, g \in H^2([0,1])$).

T_1 is symmetric ($D(T_1) = \{u \in C^2([0,1]) : u(0) = u(1) = 0\}$)

T_2 is not symmetric ($D(T_2) = C^2([0,1])$)

T_3 is symmetric

T_4 is not symmetric

We know that $D(T_1) \subset D(T_1^*)$. Are the domains identical?

No, because the relation $(T_1 f, g) = (f, T_1^* g)$ holds for all $g \in H^2([0,1])$ with $g(0) = g(1) = 0$. Then $D(T_1^*)$ is bigger than $D(T_1)$, and T_1 is not self-adjoint.

We shall see that $T_1^* = T_3$, and that $T_3^* = T_3$.

T_2 and T_4 are not self-adjoint, since they are not symmetric.

Theorem 6.7: Let $(X, \|\cdot\|_X)$ be a Hilbert space and S, T be densely-defined operators in X. Then the following holds

(a) $S \subset T \Rightarrow T^* \subset S^*$

(b) If $D(T^*)$ is dense, then T^{**} is an extension of T ($T \subset T^{**}$).

Proof: (a) $(Tx, y) = (x, T^*y)$ for all $x \in D(T)$, $y \in D(T^*)$; and $S \subset T$ implies that $(Sx, y) = (x, T^*y)$ for all $x \in D(S)$, $y \in D(T^*)$. Hence, it follows that $D(T^*) \subset D(S^*)$, and also that $S^*y = T^*y$ for $y \in D(T^*)$. Thus $T^* \subset S^*$

(b) T^{**} exists since $D(T^*)$ is dense. T^* satisfies

$$(+) \qquad (T^*y, x) = (y, T^{**}x) \text{ for all } y \in D(T^*), x \in D(T^{**})$$

The relation (+) holds also for all $x \in D(T)$ ($(T^*y, x) = (y, Tx)\ \forall x \in D(T)$). Since $D(T^{**})$ contains all x for which the relation (+) is satisfied, we have $D(T^{**}) \supset D(T)$. We conclude $T^{**}x = Tx$ for all $x \in D(T)$.

In Theorem 5.5 we show that for a bounded operator $T \in \mathcal{B}(X)$ we have $X = \ker T^* \oplus \overline{\operatorname{ran} T}$. For this important property exists also a version for densely defined operators in X.

Proposition 6.8: Let $(X, \|\cdot\|_X)$ be a Hilbert space and $T: D(T) \to X$ an operator in X with $\overline{D(T)} = X$. Then

$$X = \ker T^* \oplus \overline{\operatorname{ran} T}$$

Proof.: Adaption of the proof of Theorem 5.5.

If $x \in \operatorname{ran} T$, there exists $y \in D(T)$ such that $x = Ty$.

Proof of Remark ①: If T is closable, then $G(\overline{T}) = \overline{G(T)}$

1. Suppose S is closed extension of T. Then

$$\overline{G(T)} \subset G(S)$$

$$D(R) = \{x : \underset{\text{for some } y}{(x, y)} \in \overline{G(T)}\}$$

$Rx = y$, where $y \in X$ unique vector so that $(x, y) \in \overline{G(T)}$

$G(R) = \overline{G(T)}$ so R is a closed extension of T.
But $R \subset S$ which is an arbitrary closed extension, so $R = \overline{T}$ ∎

Then for any $z \in \ker T^*$, we have

$$(x, z) = (Ty, z) = (y, T^* z) = 0 \quad \text{, that is}$$

$\operatorname{ran} T \subset (\ker T^*)^\perp$. Since the orthogonal complement of a subspace is closed we have

$$\overline{\operatorname{ran} T} \subset (\ker T^*)^\perp$$

Conversely, let $x \in \operatorname{ran}(T)^\perp$. Then for all $y \in \mathcal{D}(T)$

$0 = (x, Ty) = (T^* x, y)$ for $y \in \mathcal{D}(T)$ and
as $\mathcal{D}(T)$ is dense we get $T^* x = 0$ (and $x \in \mathcal{D}(T^*)$)
$\Rightarrow x \in \ker(T^*)$, so that $(\operatorname{ran} T)^\perp \subset \ker T^*$, and
finally $(\ker T^*)^\perp \subset \overline{\operatorname{ran} T}$. ∎

6.2 Closed operators

We have seen that, for operators, continuity is equivalent to boundedness. Thus unbounded operators are not continuous.
A closed operator is an operator that is "almost continuous".

Recall the definition of closed operators (see chapter 3): $(X, \|\cdot\|_X)$ Hilbert space.
A densely defined operator T in X is closed
if for any sequence $(x_n)_{n \in \mathbb{N}}$ in $\mathcal{D}(T)$ such that
$x_n \to x$ and $Tx_n \to y$ as $n \to \infty$,
we have $x \in \mathcal{D}(T)$ and $Tx = y$.
Alternate definition is that the graph $G(T) = \{(x, Tx) : x \in \mathcal{D}(T)\}$
is a closed subspace in $X \times X$.

The Closed Graph Theorem (Theorem 3.28) states that closed operators between Banach spaces are bounded. It does not apply here, because $T: D(T) \to X$ and not $T: X \to X$.

Definition 6.9: Let $(X, \|\cdot\|_X)$ be a Hilbert space and $T: D(T) \to X$ be an operator in X (densely defined).

(a) T is closable if there exists a closed operator that is an extension of T. The minimal closed extension of a closable operator T is its closure $\overline{T}$

(b) A symmetric operator is called essentially self-adjoint if its closure is self-adjoint.

Remarks: ① If T is closable, then $G(\overline{T}) = \overline{G(T)}$. If $\overline{G(T)}$ is not the graph of an operator, the operator T is not closable.

② We will show that symmetric operators are closable (recall that T symmetric if $T \subset T^*$).

We summarize important properties of closed operators in the following theorem.

Theorem 6.10: Let $(X, \|\cdot\|_X)$ be a Hilbertspace and $T: D(T) \to X$ be a densely defined operator in X. Then:

(a) T^* is closed

(b) T is closable $\iff$ $D(T^*)$ is dense in which case $\overline{T} = T^{**}$

(c) If T is closable, then $(\overline{T})^* = T^*$

(d) If T is symmetric, then T is closable, and $(\bar{T})^* = T^*$

<u>Proof.</u> (a) We define a unitary operator $V: X\times X \to X\times X$ by
$V(x,y) = (-y, x)$ (exercise) ($V^* = V^{-1}$, $V^{-1}(x,y) = (y,-x)$)

$$(x,y) \in V(G(T))^{\perp} \iff ((x,y), (-Tz, z)) = 0 \ \forall z \in D(T)$$

$\iff (x, Tz) = (y, z) \ \forall z \in D(T)$
$\iff (x,y) \in G(T^*)$
Thus $G(T^*) = V(G(T))^{\perp}$. Since the orthogonal complement of any subspace is closed (here closed subspace in $X\times X$), this proves (a).

(b) Note that $G(T) \subset X\times X$ is a (linear) subspace

$$\overline{G(T)} = \left(G(T)^{\perp}\right)^{\perp} \underset{V^2=-1\!\!1}{=} \left(V^2 G(T)^{\perp}\right)^{\perp}$$

$$= (V(VG(T))^{\perp})^{\perp} = (VG(T^*))^{\perp}.$$ Hence, if

T^* is <u>densely defined</u> we can consider T^{**} and we get that $\overline{G(T)} = (VG(T^*))^{\perp} = G(T^{**})$. $T^{**} \supset T$ is an extension with a closed graph, hence it is a closed extension.

Conversely, suppose that $D(T^*)$ is not dense and that $z \in D(T^*)^{\perp}$, $z \neq 0$

$(z, 0) \in G(T^*)^{\perp}$ so $V(G(T^*))^{\perp}$ is not the graph of an operator.

$\overline{G(T)} = \left(VG(T^*)\right)^{\perp}$, we see that T is not closable.

(c) If T is closable, $T^* = \overline{(T^*)} = T^{***} = (\overline{T})^*$

T closable $\Rightarrow$ T^* closed $\Rightarrow$ $T^* = \overline{T^*} \underset{(b)}{=} T^{***} = (\overline{T})^*$.

(d) T symmetric $\Rightarrow$ $T \subset T^*$ and with (a) it follows that T^* is closed, hence T is closable and part (c) concludes the proof.

$\blacksquare$

Proposition 6.11: Let $(X, \|\cdot\|_X)$ be a Hilbert space and $T: D(T) \to X$ be a densely defined operator.
If T has a bounded inverse, then $(T^*)^{-1} = (T^{-1})^*$

Proof:
T^{-1} bounded implies that $(T^{-1})^*$ is bounded.
$x \in D(T^*)$ and $y \in X$

$$\left((T^{-1})^* T^* x, y\right) = (T^* x, T^{-1} y) = (x, T T^{-1} y) = (x, y)$$

Then $(T^{-1})^* T^* x = x$ for all $x \in D(T^*)$. Conversely, pick $x \in X$ and $y \in D(T)$

$$((T^{-1})^* x, Ty) = (x, T^{-1}Ty) = (x,y)$$

Then $(T^{-1})^* x \in D(T^*)$ $\forall x \in X$, and

$T^*(T^{-1})^* x = x$ $\forall x \in X$. This shows that
$(T^*)^{-1} = (T^{-1})^*$. ■

Example 6.12: $\ell_2(\mathbb{N})$

$T(x_1, x_2, \dots) = (x_1, 2x_2, 3x_3, \dots)$ with

$D(T) = \{x \in \ell_2(\mathbb{N}) : \sum_{n=1}^{\infty} n|x_n|^2 < \infty\}$. $\|T\| = \infty$,

$T^* = T$, and the inverse is bounded

$$T^{-1}(x_1, x_2, \dots) = (x_1, \tfrac{1}{2}x_2, 1/3\, x_3, \dots)$$

and $(T^{-1})^* = T^{-1}$. ■

6.3 Laplace operator

The study of partial differential equations (PDE's) is a major field of mathematics, pure and applied. Among the relevant equations, many are linear, and several involve the Laplace operator

$$\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i}$$

- Laplace equation $\Delta u = 0$
- Helmholtz $-\Delta u = \lambda u$
- Heat or diffusion equation: $\partial_t = u_t = \Delta u$
- Schrödinger's equation $-i\,\partial_t u = \Delta u$
- Wave equation $\partial_t^2 u = \Delta u$
- Poisson $-\Delta u = f(u)$

We briefly review the space of distributions and the notion of distributional (weak) derivative.

Dirac's "function" $\delta(x)$

$$\delta(x) = \begin{cases} 0 & \text{if } x \neq 0 \\ \infty & \text{if } x = 0 \end{cases} \qquad \text{s.t. } \int_{\mathbb{R}} \delta(x)\,dx = 1$$

This is motivated by physics. Mathematically, this does not make much sense.
The mathematician Laurent Schwarz eventually gave the delta / Dirac's function a proper meaning as a distribution.

Definition 6.13:

(a) $\Omega \subset \mathbb{R}^d$ open. The support of $f \in \mathcal{C}(\Omega)$ is

$$\operatorname{supp} f = \overline{\{x \in \Omega : f(x) \neq 0\}}$$

(b) The space of test functions is the space $\mathcal{C}_c^\infty(\Omega)$ of smooth functions on Ω with compact support. It is equipped with the following notion of convergence: $\varphi_n \to \varphi$ as $n \to \infty$ iff

- $\exists$ compact $K \subset \Omega$ s.t. $\operatorname{supp}(\varphi_n - \varphi) \subset K \ \forall n \in \mathbb{N}$
- For any $\alpha_1, \dots, \alpha_d \in \mathbb{N}$:

$$\frac{\partial^{\alpha_1 + \dots + \alpha_d}}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} \varphi_n(x) \longrightarrow \frac{\partial^{\alpha_1 + \dots + \alpha_d}}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} \varphi(x),$$

as $n \to \infty$, uniformly in x. The space is denoted by $\mathcal{D}(\Omega)$.

(c) A distribution is a continuous linear functional on the space of test functions $\mathcal{D}(\Omega)$

Remarks ① Given a locally integrable function f we define the corresponding distribution by

$$T_f(\varphi) = \int_\Omega f(x)\varphi(x)\,dx$$

② Dirac's / Delta - function

$$T_\delta(\varphi) = \int_\Omega \delta(x-a)\varphi(x)\,dx = \varphi(a)$$

③ Suppose $f \in \mathcal{C}^1(\mathbb{R})$. Then f is locally integrable and T_f exists.

$$T_f'(\varphi) = -T_f(\varphi') = -\int_{-\infty}^{\infty} f(x)\varphi'(x)\,dx$$

$$= \int_{-\infty}^{+\infty} f'(x)\varphi(x)\,dx = T_{f'}(\varphi) \qquad \varphi \in \mathcal{D}(\mathbb{R})$$

Definition 6.14: (a) Let $\alpha_1, \dots, \alpha_d \in \mathbb{N}$

The distributional derivative or weak derivative of a distribution T is defined by

$$\left(\frac{\partial^{\alpha_1+\dots+\alpha_d}}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}} T\right)(\varphi) = (-1)^{\alpha_1+\dots+\alpha_d}\, T\left(\frac{\partial^{\alpha_1+\dots+\alpha_d}}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}}\varphi\right)$$

(b) Let $k \in \mathbb{N}$. $H^k(\Omega)$ where $\Omega \subset \mathbb{R}^d$ open consists of functions $f \in L^2(\Omega)$ whose weak derivative of order less than k are also L^2-functions.

The space $H^k(\Omega)$ is called Sobolev space, and its norm, called Sobolev norm, is

$$\|f\|_{k,2} = \Big(\sum_{\substack{\alpha_1,\dots,\alpha_d \geq 0 \\ \alpha_1+\dots+\alpha_d = k}} \int_\Omega \Big| \frac{\partial^{\alpha_1+\dots+\alpha_d}}{\partial x_1^{\alpha_1}\dots\partial x_d^{\alpha_d}} f(x) \Big|^2 dx \Big)^{1/2};$$

$$(f,g)_{k,2} = \sum_{\substack{\alpha_1,\dots,\alpha_d \geq 0 \\ \alpha_1+\dots+\alpha_d = k}} \int_\Omega \overline{\frac{\partial^{\alpha_1+\dots+\alpha_d}}{\partial x_1^{\alpha_1}\dots\partial x_d^{\alpha_d}} f(x)} \frac{\partial^{\alpha_1+\dots+\alpha_d}}{\partial x_1^{\alpha_1}\dots\partial x_d^{\alpha_d}} g(x)\, dx$$

It is intuitively clear that a function that possesses derivatives, even weak derivatives, is smoother than a function that has none. Sobolev embedding theorem makes this intuition explicit.

<u>Theorem 6.15</u>: $\Omega \subset \mathbb{R}^d$ open. If $s > k + \frac{1}{2} d$,

then $H^s(\Omega) \subset \mathcal{C}_0^k(\Omega)$,

where $\mathcal{C}_0^k(\Omega)$ is the space of k-times differentiable functions that vanish at the boundary of Ω.

Proof. Any textbook, e.g. Folland or Adams – Sobolev spaces.

<u>Remarks:</u>

① $d=1$

$$\mathcal{C}_0^2(\Omega) \subset H^2(\Omega) \subset \mathcal{C}_0^1(\Omega) \subset H^1(\Omega) \subset \mathcal{C}_0(\Omega)$$

② $H^s(\Omega)$ complete, and $\mathcal{C}^\infty(\Omega)$ functions are dense.

$\Delta = \frac{d^2}{dx^2}$, $\Omega = (0,1)$

$\mathcal{D}(\Delta) = \mathcal{C}_c^2((0,1))$

(Why not choose $\mathcal{D}(\Delta) = \mathcal{C}^2((0,1))$? - Because $\mathcal{C}^2((0,1)) \not\subset L^2((0,1))$)

Δ with $\mathcal{D}(\Delta) = \mathcal{C}_c^2((0,1))$ symmetric

The closure of Δ exists, what is it?

$G(\Delta) = \{(f, f'') : f \in \mathcal{D}(\Delta)\}$

The completion of $\mathcal{C}_c^2((0,1))$ with respect to $\|\cdot\|_{H^2}$ is H_0^2,

we get $\mathcal{D}(\bar{\Delta}) = H_0^2((0,1))$ (see exercise)

Let us now find the adjoint of $\bar{\Delta}$

$\mathcal{D}(\bar{\Delta}^*) = \{g \in L^2((0,1)) : \exists h \text{ s.t. } (g, \Delta f) = (h, f)\ \forall f \in \mathcal{D}(\bar{\Delta})\}$

The condition is

$\int_0^1 g f'' \overset{\text{Fourier}}{=} \int_{-\infty}^{+\infty} k^2 \hat{g} \hat{f} = \int_{-\infty}^{\infty} \hat{h} \hat{f}$, it follows that

$\hat{h}(k) = k^2 \hat{g}(k)$ a.e. Then h exists as an L^2 function iff $\hat{h} \in L^2(\mathbb{R})$, i.e. iff $\int k^2 |\hat{g}|^2 < \infty$, i.e. iff $g \in H^2((0,1))$.

Finally, we check that g needs to belong to H_0^2.

Since g has weak first + second derivatives,

$$\int_0^1 g f'' = g f'\Big|_0^1 - \int g' f' = \underbrace{g f'\Big|_0^1}_{\text{must be zero}} - \underbrace{g' f\Big|_0^1}_{=0} + \int g'' f$$

Then, necessarily, $g(0) = g(1) = 0$, hence $g \in H_0^2((0,1))$.

<u>Summary:</u>

Δ with $\mathcal{D}(\Delta) = \mathcal{C}_c^2((0,1))$ is symmetric, but it is not self-adjoint.

It is essentially self-adjoint, since its closure $\overline{\Delta}$, with $\mathcal{D}(\overline{\Delta}) = H_0^2((0,1))$, is self-adjoint.

The End