Functional Analysis I

held by Prof. Dr. Moritz Weber in winter term ’17/’18

UNIVERSITÄT DES SAARLANDES

General information on organisation

Tutorials and admission for the final exam To take part in the final exam of this course, 50 % of the total points on all exercise sheets have to be achieved. Furthermore there is compulsory attendence for the tutorials associated to this course, one shouldn't miss more than two tutorials without good reason.

Advise regarding this transcript This transcript is in no way more valuable than a handwritten transcript of the lecture and can by no means replace attendance of the lecture or personal review of the content. Computer composition is no warrant for accuracy!

This transcript is written by a student, typos or general mistakes can not be ruled out. Hints towards mistakes in the transcript are explicitly asked for:

s9fhguen@stud.uni-saarland.de

Contents

0 Motivation

Analysis is the study of functions $f\colon \mathbb{C} \to \mathbb{C}$ (or more generally $f\colon A \to B$). We can ask ourselves for transformations of functions, i. e., mappings

$$T\colon \{f\colon A \to B\} \to \{f\colon A \to B\}.$$

If we do this, several questions emerge:

(i) Which structure do the spaces $\{f\colon A \to B\}$ have? For example

$$C([0,1]) = \{f\colon [0,1] \to \mathbb{C} \text{ continuous}\}$$

is a vector space via $f+g$ and λf (pointwise) for $\lambda \in \mathbb{C}, f, g \in C([0,1])$, furthermore it is normed via $\|\cdot\|_\infty$ and it is complete.

We need to understand topological concepts on vector spaces like the convergence of sequences of functions and generalisations to $\{f\colon A \to B\}$, which leads to the concepts of Banach spaces and Hilbert spaces.

(ii) In the setting $T\colon X \to Y$ with vector spaces X and Y, we are particularly interested in linear functions, a concept we know from linear algebra, e. g.

$$(Tf)(s) := \int_Z k(s,t) f(t)\, dt$$

for a suitable $k\colon Z \times Z \to \mathbb{C}$. In this sense, we will study "matrices of infinite size" (namely if X and Y are infinite dimensional). What are "eigenvalues", what is "diagonalizing", ...? If $\dim X, \dim Y < \infty$, then T being linear implies T being continuous, hence linear algebra. If $\dim X, \dim Y = \infty$, this implication doesn't hold. So, in a way, functional analysis is linear algebra and analysis coming together. The study of those maps will lead to the concept of operators on Hilbert— and Banach spaces.

(iii) For matrices (in linear algebra) we have $AB \neq BA$. This phenomenon of "non-commutativity" is important for operators on Hilbert spaces and also in quantum mechanics. If we want to study the space

$$\{T\colon X \to Y \text{ linear and continuous}\}$$

we will have to deal with non-commutativity. And we will meet algebraic structures, since $\{T\colon X \to Y\}$ is again a vector space and even an algebra (for instance if $X = Y$ via $ST := S \circ T$). We will deal with operator algebraic structures like Banach algebras, C^*-algebras and von Neumann algebras.

History of Functional Analysis Some of the main protagonists of functional analysis are:

- *Fredholm* ($\sim$ 1900) was interested in integral operators,
- *Hilbert* ($\sim$ 1910) was interested in spectral theory of general continuous operators,
- *Riesz* ($\sim$ 1910/1920) was interested in linear maps on normed spaces,
- *Banach* ($\sim$ 1920/1930) was the "founder of modern functional analysis",
- *von Neumann* ($\sim$ 1930/1940) worked on the mathematical foundation of quantum mechanics, introduced von Neumann algebras,
- *Gelfand* ($\sim$ 1940) introduced in C^*-algebras.

Literature suggestions Common literature on the topic are:

- John Conway, *A course in Functional analysis*, Springer, 1990,
- Friedrich Hirzebruch and Winfried Scharlau, *Einführung in die Funktionalanalysis*, Spektrum, 1996,
- Reinhold Meise und Dietmar Vogt, *Einführung in die Funktionalanalysis*, vieweg, 1992,
- Gert Pedersen, *Analysis Now*, Springer, 1989.

1 Topological vector spaces

There is a hierarchy of information regarding "place" and "convergence":

- topology – shape of boundaries, minimal requirement for continuity
- metric – distances
- normed – scale for distances from an "origin"

Definition 1.1 (Topological space): Let X be a set. A subset $\mathfrak{T} \subseteq \mathfrak{P}(X)$ is a *topology* on X, if

(i) $\varnothing, X \in \mathfrak{T}$,

(ii) $U, V \in \mathfrak{T} \Rightarrow U \cap V \in \mathfrak{T}$,

(iii) $M \subseteq \mathfrak{T} \Rightarrow \bigcup_{V \in M} V \in \mathfrak{T}$

hold. Elements $U \in \mathfrak{T}$ are called *open*, $A \subseteq X$ is called *closed*, if $A^{\mathsf{c}} := X \setminus A$ is open.

A topology $\mathfrak{T}$ is called *Hausdorff*, if $\forall\, x, y \in X, x \neq y$ there are $U, V \in \mathfrak{T}$ such that $x \in U, y \in V$ and $U \cap V = \varnothing$. If $Y \subseteq X$, then

$$\overline{Y} := \bigcap_{\substack{A \text{ closed} \\ Y \subseteq A}} A$$

is the *closure* of Y.

$N \subseteq X$ is a *neighbourhood* of $x \in X$, if there is an open set $U \subseteq X$ such that $x \in U \subset N$.

Remark 1.2: (i) If (X, d) is a *metric space* (i. e., there exists $d\colon X \times X \to [0, \infty)$ for which $d(x, y) = d(y, x)$, $d(x, z) \leq d(x, y) + d(y, z)$ and $d(x, y) = 0 \Leftrightarrow x = y$ hold), then the sets $M \subseteq X$ with

$$\forall\, a \in M \,\exists\, \varepsilon > 0 : B(a, \varepsilon) := \{x \in X \mid d(a, x) < \varepsilon\} \subseteq M$$

form a Hausdorff topology.

(ii) $\mathfrak{T} = \{\varnothing, X\}$ is the *trivial topology.* The trivial topology is not Hausdorff.

(iii) $V \subseteq X$ is open if and only if V is a neighbourhood for all $x \in V$.

Proof: "$\Rightarrow$" is trivial. For "$\Leftarrow$": Let $x \in V$ and choose $x \in U_x \subseteq V$ open, then $V = \bigcup_{x \in V} U_x$. ∎

(iv) Let $Y \subseteq X$. Then it holds: $z \in \overline{Y}$ if and only if $N \cap Y \neq \varnothing$ for all neighbourhoods N of z.

Proof: We have the following equivalences:

$$\begin{aligned} z \in \overline{Y} &\Leftrightarrow z \in A \text{ for all } A \supseteq Y \text{ closed} \\ &\Leftrightarrow \text{If } U \text{ is open, } U \cap Y = \varnothing, \text{ then } z \notin U \\ &\Leftrightarrow \text{For all open sets } U : z \in U \Rightarrow U \cap Y \neq \varnothing \end{aligned}$$ ■

In topological spaces we may define continuity, matching our notion of continuity in metric spaces.

Definition 1.3 (Continuity): A map $f\colon X \to Y$ between topological spaces is called *continuous* in $x \in X$, if for all neighbourhoods $N \subseteq Y$ of $f(x)$: $f^{-1}(N) \subseteq X$ is a neighbourhood of x. f is called *continuous*, if it is continuous in all $x \in X$.

Proposition 1.4: (i) *Let X, Y be topological spaces. $f\colon X \to Y$ is continuous if and only if $f^{-1}(U)$ is open for every $U \subset Y$ open.*

(ii) *Let X, Y be metric spaces. $f\colon X \to Y$ is continuous if and only if $\forall\, x \in X$ $\forall\, \varepsilon > 0\, \exists\, \delta > 0 : f(B(x, \delta)) \subset B(f(x), \varepsilon)$.*

Proof: (i) "$\Leftarrow$": Let $x \in X$ and $N \subseteq Y$ be a neighbourhood of $f(x)$. Without loss of generality let N be open. Then $f^{-1}(N)$ is open and $x \in f^{-1}(N)$.

"$\Rightarrow$": Let $U \subseteq Y$ be open and $x \in f^{-1}(U)$, then U is a neighbourhood of $f(x)$. Thus $f^{-1}(U)$ is a neighbourhood of x and by Remark 1.2 $f^{-1}(U)$ is open.

(ii) "$\Rightarrow$": Let $x \in X$ and $\varepsilon > 0$. Then $f^{-1}(B(f(x), \varepsilon))$ is a neighbourhood of x. Hence, there is a $\delta > 0$ such that $B(x, \delta) \subseteq f^{-1}(B(f(x), \varepsilon))$.

"$\Leftarrow$": If $N \subseteq Y$ is a neighbourhood of $f(x)$, then there is an $\varepsilon > 0$ with $B(f(x), \varepsilon) \subseteq N$. Thus, there is a $\delta > 0$ such that we have

$$B(x, \delta) \subseteq f^{-1}(B(f(x), \varepsilon)) \subseteq f^{-1}(N),$$

therefore $f^{-1}(N)$ is a neighbourhood of x. ■

In metric spaces we may also express continuity using sequences. In topological spaces, we need *nets*.

Definition 1.5: A set Λ is *ordered*, if there is a relation "$\leq$" such that

(i) $\lambda \leq \lambda$,

(ii) If $\lambda \leq \mu$, $\mu \leq \lambda$, then $\lambda = \mu$,

(iii) If $\lambda \leq \mu$ and $\mu \leq \nu$, then $\lambda \leq \nu$.

Λ is a *filtration*, if in addition $\forall\, \lambda, \mu \in \Lambda : \exists\, \nu \in \Lambda : \lambda \leq \nu, \mu \leq \nu$.

Remark 1.6: In general two arbitrary elements of a filtration are not comparable.

Definition 1.7: Let X be a topological space, $(x_\lambda)_{\lambda \in \Lambda} \subseteq X$ a family and Λ a filtration. Then $(x_\lambda)_{\lambda \in \Lambda}$ is called a *net*. The net *converges* to $x \in X$, if for every neighbourhood N of x there is a λ_0 such that $x_\lambda \in N$ for all $\lambda \geq \lambda_0$.

Example 1.8: (i) $(\Lambda, \leq) = (\mathbb{N}, \leq)$ gives the known concept of sequences.

(ii) Let Λ be the set of all partitions $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$ of the interval $[a, b] \subseteq \mathbb{R}$ and let $f\colon [a, b] \to \mathbb{R}$ be continuous. Then Λ is a filtration via $\lambda \geq \mu :\Leftrightarrow \lambda$ is a partition finer than μ. If we define

$$s_\lambda := \sum_{t_i \in \lambda} f(t_i)(t_i - t_{i-1}),$$

then $(s_\lambda)_{\lambda \in \Lambda}$ forms a net converging to $\int_a^b f(t)\, dt$.

Remark 1.9: (i) It is possible, that a net converges to two points. For instance in $\mathfrak{T} = \{\varnothing, X\}$ every net converges to every point.

(ii) Let X be a topological space and $Y \subseteq X$. Then $\overline{Y}$ is the set of all limit points of nets in Y.

Proof: "$\supseteq$": Let $(x_\lambda)_{\lambda \in \Lambda} \subseteq Y$ be a net with $x_\lambda \to z$ and let N be a neighbourhood of z. Thus, there is a $\lambda_0 \in \Lambda$ such that $x_\lambda \in N$ for all $\lambda \geq \lambda_0$. Then, $N \cap Y \neq \varnothing$ for all neighbourhoods N of z, so $z \in \overline{Y}$ by Remark 1.2.

"$\subseteq$": Let $z \in \overline{Y}$. Consider the set $\mathfrak{U}$ of all neighbourhoods of z. This is a filtration via $U \geq U' :\Leftrightarrow U \subseteq U'$. For any neighbourhood $U \in \mathfrak{U}$ choose $x_u \in U \cap Y$ ($U \cap Y \neq \varnothing$ via Remark 1.2), then we have $x_u \to z$. ■

(iii) $A \subseteq X$ is closed if and only if $(x_\lambda \to x,\, x_\lambda \in A \Rightarrow x \in A)$. This is a direct consequence of (ii).

(iv) Let $f\colon X \to Y$ be a mapping between topological spaces. Then f is continuous if and only if for all nets $(x_\lambda) \subseteq X$ with $x_\lambda \to x$ we have $f(x_\lambda) \to f(x)$.

Proof: See exercise 1 on sheet 1. ■

Definition 1.10: Let (X, d) be a metric space. A sequence $(x_n)_{n \in \mathbb{N}}$ is called *Cauchy sequence*, if $\forall\, \varepsilon > 0\, \exists\, N \in \mathbb{N} : \forall\, n, m \geq N : d(x_n, x_m) < \varepsilon$.

X is called *complete*, if all Cauchy sequences in X converge.

Example 1.11: (i) Consider $X = (0, 1)$ with the usual metric $d(x, y) := |x - y|$. This space is not complete.

(ii) Consider $X = (0, 1)$ with a metric mapping X in a bijective way to $\mathbb{R}$, then (X, d) is complete.

In the sequel, we need to understand how to complete a space and how to extend functions to these completions.

Theorem 1.12: *Let (X, d) be a metric space. Then there exists a unique (up to isometry) complete metric space $(\hat{X}, \hat{d})$ with isometric embedding $i\colon X \to \hat{X}$ (i. e., $\hat{d}(i(x), i(y)) = d(x, y)$) and $\overline{i(X)} = \hat{X}$.*

A map $\varphi\colon X \to Y$ between metric spaces with $d_Y(\varphi(x), \varphi(y)) \leq C d_X(x, y)$ for all $x, y \in X$ and fixed constant $C \geq 0$ may be extended in a unique way to a continuous map $\hat{\varphi}\colon \hat{X} \to Y$ with $\hat{\varphi} \circ i = \varphi$ in case Y is complete.

Proof: (i) We call two sequences $(x_n)_{n\in\mathbb{N}}$, $(x'_n)_{n\in\mathbb{N}}$ equivalent if and only if $d(x_n, x'_n) \to 0$, in this case we write $(x_n)_{n\in\mathbb{N}} \sim (x'_n)_{n\in\mathbb{N}}$. We then set

$$\hat{X} := \{[(x_n)_{n\in\mathbb{N}}] \mid (x_n)_{n\in\mathbb{N}} \text{ is a Cauchy sequence in } X\},$$

$\hat{d}([(x_n)], [(y_n)]) := \lim_{n\to\infty} d(x_n, y_n)$ and

$$\begin{aligned} i\colon X &\longrightarrow \hat{X} \\ x &\longmapsto [(x, x, x, \dots)]. \end{aligned}$$

Check, that "$\sim$" is indeed an equivalence relation on the set of Cauchy sequences with elements in X and that the limit $(d(x_n, y_n))_{n\in\mathbb{N}}$ exists and that $\hat{d}$ is a metric.

Proof (that the limit of $(d(x_n, y_n))_{n\in\mathbb{N}}$ exists): Via the triangular inequality we have

$$\begin{aligned} d(x_n, y_n) - d(x_m, y_m) &\le d(x_n, x_m) + d(x_m, y_m) + d(y_m, y_n) - d(x_m, y_m) \\ &< 2\varepsilon \end{aligned}$$

and similarly $-2\varepsilon < d(x_n, y_n) - d(x_m, y_m)$, so $(d(x_n, y_n))_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$, hence the limit exists. ■

The map $i\colon X \to \hat{X}$ is isometric, since

$$\hat{d}(i(x), i(y)) = \lim_{n\to\infty} d(x, y) = d(x, y)$$

and we have $\overline{i(X)} = \hat{X}$, since for $[(x_n)] \in \hat{X}$ we have $i(x_n) \to [(x_m)_{m\in\mathbb{N}}]$, because for $\varepsilon > 0$:

$$\hat{d}(i(x_n), [(x_m)]) = \lim_{k\to\infty} (x_n, x_k) < \varepsilon$$

for $n \ge N$.

(ii) $\hat{X}$ is complete: We construct a suitable diagonal sequence. Let $(\alpha_n)_{n\in\mathbb{N}}$ be Cauchy in $\hat{X}$, hence $\alpha_n = [(x_k^n)_{k\in\mathbb{N}}]$. Since "If $(a_n)_{n\in\mathbb{N}}$ is Cauchy and $(b_n)_{n\in\mathbb{N}}$ is a subsequence of (a_n), then $d(a_n, b_n) \to 0$", we may assume without loss of generality that $d(x_k^n, x_{k+l}^n) < \frac{1}{2^{k+1}}$. Also, we may assume

$$\hat{d}(\alpha_n, \alpha_{n+l}) < \frac{1}{2^{n+1}}$$

Choose $y_n := x_{k(n)}^n$ for $k(n) \ge n$ and $d(y_n, y_{n+1}) < \frac{1}{2^n}$.

Proof (that there is such (y_n)): Set $y_1 := x_1^1$. If y_n has been constructed, then from

$$\hat{d}(\alpha_n, \alpha_{n+1}) < \frac{1}{2^{n+1}},$$

we know, that there exists an $l \ge k(n) + 1$ such that $d(x_l^n, x_l^{n+1}) < \frac{1}{2^{n+1}}$. Put $y_{n+1} := x_l^{n+1}$. Then

$$d(y_n, y_{n+1}) \le d(x_{k(n)}^n, x_l^n) + d(x_l^n, x_l^{n+1}) \le \frac{1}{2^{k(n)+1}} + \frac{1}{2^{n+1}} < \frac{1}{2^n}.$$ ■

Then (y_n) is Cauchy in X, hence $[(y_n)_{n\in\mathbb{N}}] \in \hat{X}$. And $\alpha_m \to [(y_n)]$ for $m \to \infty$.

Proof (that $\alpha_m \to (y_n)$): Let $\varepsilon > 0$. Let N be such that $\frac{2}{2^{N-1}} < \varepsilon$, let $m \geq N$. Then $\hat{d}(\alpha_m, [(y_n)]) < \varepsilon$, because for $l \geq m$

$$\begin{aligned} d(x_l^m, y_l) &\leq d(x_l^m, x_{k(m)}^m) + d(y_m, y_l) \\ &< \frac{1}{2^{\min\{l,k(m)\}}} + \frac{1}{2^{m-1}} \leq \frac{2}{2^{N-1}} < \varepsilon. \end{aligned}$$

■

(iii) We set

$$\hat{\varphi}([(x_n)]) := \lim_{n\to\infty} \varphi(x_n),$$

this is welldefined, since $(\varphi(x_n))_{n\in\mathbb{N}}$ is a Cauchy sequence in Y (because we have

$$d_Y(\varphi(x_n), \varphi(x_m)) \leq C d_X(x_n, x_m)$$

and $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence). $\hat{\varphi}$ is continuous and unique. For the continuity: If we have $[(x_n^k)_{n\in\mathbb{N}} \to [(x_n)_{n\in\mathbb{N}}]$ for $k \to \infty$,

$$\begin{aligned} &d(\hat{\varphi}([(x_n^k)_{n\in\mathbb{N}}]), \hat{\varphi}([(x_n)_{n\in\mathbb{N}}])) \\ &\quad \leq d(\hat{\varphi}([(x_n^k)_{n\in\mathbb{N}}]), \varphi(x_n^k)) + d(\varphi(x_n^k), \varphi(x_n)) + d(\varphi(x_n), \hat{\varphi}([(x_n)_{n\in\mathbb{N}}])) \\ &\quad < 3\varepsilon \end{aligned}$$

for k, n large. For the uniqueness: Let φ' be another continuous map such that $\varphi' \circ i = \varphi$, then

$$\hat{\varphi}([(x_n)_{n\in\mathbb{N}}]) \leftarrow \varphi(x_k) = \varphi'([(x_k, x_k, x_k, \dots]) \to \varphi'([(x_n)]),$$

so $\hat{\varphi} = \varphi'$.

As a special case of the construction, we may prove that $\hat{X}$ and $\overline{X}$ are isomorphic, if $X \subseteq Y$ with Y complete (Exercise sheet 1).

For the uniqueness of $\hat{X}$: Let (Y, d_Y) be another complete metric space and $i\colon X \to Y$ be an isometric embedding, $\overline{i(X)} = Y$. Then, by (b), we have $\hat{i}\colon \hat{X} \to Y$ and

$$Y = \overline{i(X)} \cong \widehat{i(X)} \cong \hat{X}.$$

■

Example 1.13: (i) Let $X = \mathbb{Q}$, then $\hat{\mathbb{Q}} = \mathbb{R}$.

(ii) Let $X = C([0,1]) := \{f\colon [0,1] \to \mathbb{C} \text{ continuous}\}$,

$$\|f\|_\infty := \sup_{x\in[0,1]} |f(x)| \quad \text{and} \quad \|f\|_1 := \int_0^1 |f(x)|\, dx$$

and denote $d_\infty(f,g) := \|f-g\|_\infty$, $d_1(f,g) := \|f-g\|_1$. Then $(C([0,1]), d_\infty)$ is complete, but $(C([0,1]), d_1)$ isn't. We have

$$(\widehat{C([0,1]}), d_1) = (L^1([0,1]), \lambda).$$

In the following, with $\mathbb{K}$ we denote either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$.

Definition 1.14: Let X be a $\mathbb{K}$-vector space together with a topology. X is a *topological vector space*, if

$$+\colon X \times X \longrightarrow X \qquad \mu\colon \mathbb{K} \times X \longrightarrow X$$
$$(x, y) \longmapsto x + y \qquad (\lambda, x) \longmapsto \lambda x$$

are continuous.

Remark 1.15: (i) The continuity of the addition means: If $x_\lambda \to x$, $y_\mu \to y$, then $x_\lambda + y_\mu \to x + y$ with a filtration $\Lambda \times M$ with

$$(\lambda', \mu') \geq (\lambda, \mu) :\Leftrightarrow (\lambda' \geq \lambda \wedge \mu' \geq \mu).$$

(ii) If X is a topological vector space, $Y \subseteq X$ a subspace, then $\overline{Y} \subseteq X$ is also a vector space (If $x, y \in \overline{Y}$, then $\exists\, x_\lambda \to x, y_\mu \to y$, such that $x_\lambda + y_\mu \to x + y$, thus $x + y \in \overline{Y}$).

Theorem 1.16: *Let X, Y be topological vector spaces, $T\colon X \to Y$ a linear map. The following are equivalent:*

(i) *T is continuous (in all points),*

(ii) *T is continuous in some point,*

(iii) *T is continuous in 0.*

Proof: "(i) $\Rightarrow$ (iii)" and "(iii) $\Rightarrow$ (ii)" are trivial.
"(ii) $\Rightarrow$ (i)": Let T be continuous in $z \in X$. Let $x_\lambda \to x$, then $x_\lambda + (z - x) \to z$, thus $Tx_\lambda = T(x_\lambda + z - x) - T(z - x) \to T(x)$. ■

Definition 1.17: Let X be a $\mathbb{K}$-vector space. A map $\|\cdot\|\colon X \to \mathbb{R}$ is called a *norm*, if

(i) $\|x\| \geq 0$,

(ii) $\|\lambda x\| = |\lambda|\|x\|$,

(iii) $\|x + y\| \leq \|x\| + \|y\|$,

(iv) $\|x\| = 0 \Rightarrow x = 0$

hold for all $\lambda \in \mathbb{K}$, $x, y \in X$. Without (iv), $\|\cdot\|$ is called a *seminorm.*

Remark 1.18: Every normed vector space is a topological vector space via the topology induced by the metric $d(x, y) := \|x - y\|$.
Let $x_n \to x$ and $y_n \to y$, then

$$d(x_n + y_n, x + y) = \|(x_n + y_n) - (x + y)\| \leq \|x_n - x\| + \|y_n - y\| \to 0,$$

so the addition is continuous. The induced metric is translation invariant (i. e., $d(x + y, y + z) = d(x, y)$) and the norm is continuous (seen as a map $\|\cdot\|\colon X \to \mathbb{R}$).

Example 1.19: (i) Let $X = \mathbb{R}^n$ or $X = \mathbb{C}^n$ and $x = (x_1, \dots, x_n) \in X$. Then

$$\|x\|_1 := \sum_{i=1}^{n} |x_i| \quad , \quad \|x\|_2 := \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{\frac{1}{2}} \quad , \quad \|x\|_\infty := \max_{i=1,\dots,n} |x_i|.$$

declare norms on X. $\|\cdot\|_2$ is called the *Euclidean norm* and matches our geometric idea of distance.

(ii) Let K be a compact topological space and

$$X := C(K) := \{f \colon K \to \mathbb{C} \text{ continuous}\}.$$

Then

$$\|f\|_\infty := \sup_{x \in K} |f(x)|$$

declares a norm on X.

(iii) Let μ be a measure on $\mathbb{R}$ and $1 \leq p < \infty$. Then

$$L^p(\mu) := L^p(\mathbb{R}, \mu) := \{f \colon \mathbb{R} \to \mathbb{C} \mid \|f\|_p < \infty\}$$

where

$$\|f\|_p := \Big(\int_{\mathbb{R}} |f(x)|^p \, d\mu(x)\Big)^{\frac{1}{p}}$$

is a normed vector space. For $p = \infty$, set

$$L^\infty(\mu) := \{f \colon \mathbb{R} \to \mathbb{C} \mid \exists\, c \geq 0 : \{t \mid |f(t)| > c\} \text{ is a zero set}\}$$

and

$$\|f\|_\infty := \inf\{c \mid \{t \mid |f(t)| > c\} \text{ is a zero set}\}.$$

Definition 1.20: Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are called equivalent, if there are constants $C_1, C_2 > 0$ such that

$$C_1 \|x\|_1 \leq \|x\|_2 \leq C_2 \|x\|_1.$$

In this case, $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ are topologically isomorphic, since

$$\mathrm{id} \colon (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$$

is bijective and continuous in both directions. The norms generate the same topologies, since $x_n \to x$ in $(X, \|\cdot\|_1)$ implies $x_n \to x$ in $(X, \|\cdot\|_2)$ via

$$\|x_n - x\|_2 \leq C_2 \|x_n - x\|_1 \to 0.$$

One can show, that all the norms in Example 1.19 (i) are equivalent, in fact *all* norms on $\mathbb{R}^n$ (or $\mathbb{C}^n$ respectively) are equivalent. Hence, there is only one normed n-dimensional $\mathbb{R}$-vector space (or $\mathbb{C}$-vector space respectively). This is *not true* in the infinite dimensional case. Consider for instance the spaces $(C([0,1]), \|\cdot\|_\infty)$ and $(C([0,1]), \|\cdot\|_1)$ as in Example 1.13. We have

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

but there is no constant $C > 0$ such that $\|f\|_\infty \leq C\|f\|_1$: Take a function f_n as shown in Figure 1.1.

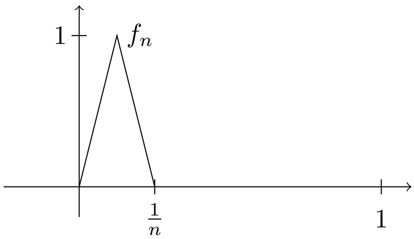


Figure 1.1: Function $f_n\colon [0,1] \to \mathbb{R} \hookrightarrow \mathbb{C}$

Then $\|f_n\|_1 = \frac{1}{2n}$ and $\|f_n\|_\infty = 1$. But suppose it exists $C > 0$, such that $1 = \|f_n\|_\infty \leq C\|f_n\|_1$, then

$$1 = \|f_n\|_\infty \leq \frac{C}{2n}$$

for all $n \in \mathbb{N}$, which is a contradiction.

Theorem 1.21: *Let X, Y be normed spaces, $T\colon X \to Y$ linear. Then the following are equivalent:*

(i) *T is continuous,*

(ii) $\exists\, C \geq 0\, \forall\, x \in X : \|Tx\| \leq C\|x\|$.

Proof: "(ii) $\Rightarrow$ (i)": It holds

$$\|Tx_n - Tx\| = \|T(x_n - x)\| \leq C\|x_n - x\| \to 0.$$

"(i) $\Rightarrow$ (ii)": Since T is continuous in 0, for $\varepsilon = 1$ there is a $\delta > 0$ such that $\|Ty\| \leq 1$ when $\|y\| \leq \delta$. Put $C := \frac{1}{\delta} > 0$, then it holds for any $x \in X$:

$$\|Tx\| = \frac{\|x\|}{\delta} \left\| T\left(\delta \frac{x}{\|x\|}\right) \right\| \leq \frac{\|x\|}{\delta} = C\|x\|.$$

■

Definition 1.22: Let $T\colon X \to Y$ be a linear map between normed spaces. Put

$$\|T\| := \inf\{C \geq 0 \mid \|Tx\| \leq C\|x\|\, \forall\, x \in X\} \in [0, \infty],$$

the *operator norm*. T is *bounded*, if $\|T\| < \infty$. Write

$$B(X,Y) := \{T\colon X \to Y \text{ linear, bounded}\},$$

we abbreviate $B(X) := B(X,X)$ and $X' := B(X,\mathbb{C})$.

We conclude that Theorem 1.21 tells us: "continious" is the same as "bounded".

Proposition 1.23: *Let X, Y be normed vector spaces and $T \in B(X,Y)$. Then:*

(i) $\|Tx\| \leq \|T\|\|x\|$ *for all* $x \in X$,

(ii) $\|T\| = \sup\{\|Tx\| \mid \|x\| = 1\}$,

(iii) $\|T\| = \sup\{\|Tx\| \mid \|x\| \leq 1\}$,

(iv) $\|T\| = \sup\{\frac{\|Tx\|}{\|x\|} \mid x \neq 0\}$.

Proof: (i) Choose $C_n \downarrow \|T\|$, then we have $\|Tx\| \leq C_n\|x\|$ for all $x \in X$.

(ii) Let $\alpha := \sup\{\|Tx\| \mid \|x\| = 1\} \leq \|T\|$, but we also have

$$\|Ty\| = \|y\| \left\| T\left(\frac{y}{\|y\|}\right)\right\| \leq \|y\|\alpha,$$

so $\|T\| < \alpha$.
(iii) and (iv) work similarly. ∎

Definition 1.24: A *Banach space* is a normed vector space which is complete (with respect to the metric induced by the norm).

Example 1.25: (i) Let K be compact, then $(C(K), \|\cdot\|_\infty)$ as in Example 1.19 is complete.

(ii) $L^p(\mu)$ is a Banach space for $1 \leq p \leq \infty$ as in Example 1.19.

(iii) Every finitedimensional normed vector space is complete (check that the normed vector spaces $(\mathbb{R}^n, \|\cdot\|_2)$ and $(\mathbb{C}^n, \|\cdot\|_2)$ are complete and then use the equivalence of norms mentioned in the remark following Definition 1.20).

Theorem 1.26: *Let X, Y be normed vector spaces. Then $B(X,Y)$ is a normed vector space (i. e., $\|\cdot\|$ from Definition 1.22 is indeed a norm). If Y is a Banach space, then $B(X,Y)$ is even a Banach space.*

Proof: (i) First, we want to show that $\|\cdot\|$ is indeed a norm. Let $S, T \in B(X,Y)$, then

$$\|(S+T)x\| \leq \|Sx\| + \|Tx\| \leq \|S\| + \|T|$$

for all $x \in X$ with $\|x\| = 1$, so we have the triangular inequality for $\|\cdot\|$. Furthermore, we have

$$\|(\lambda T)x\| = |\lambda|\|Tx\| \quad , \quad \|T\| = 0 \Leftrightarrow T = 0.$$

(ii) Let now Y be complete and let $(T_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $B(X,Y)$. Then $(T_n x)_{n\in\mathbb{N}}$ is a Cauchy sequence for all $x \in X$, so there is a limit $\lim_{n\to\infty} T_n x \in Y$. Put $S\colon X \to Y, Sx := \lim_{n\to\infty} T_n x$. S is linear, because

$$S(\lambda x + \mu y) \longleftarrow T_n(\lambda x + \mu y) = \lambda T_n x + \mu T_n y \longrightarrow \lambda Sx + \mu Sy,$$

and $T_n \to S$ (with respect to the operator norm on $B(X,Y)$!): For a given $\varepsilon > 0$ choose $N \in \mathbb{N}$ such that $\|T_n - T_m\| < \varepsilon \,\forall\, n,m \geq N$. Now let $x \in X$ so that $\|x\| = 1$ and choose $m(x) \geq N$, such that $\|(S - T_{m(x)})(x)\| < \varepsilon$, then

$$\|(S - T_n)x\| \leq \|(S - T_{m(x)})x\| + \|T_{m(x)} - T_n\|\|x\| < \varepsilon,$$

so $\|S - T_n\| < 2\varepsilon$. Eventually, S is continuous as a sum $S = T_N + (S - T_N)$ for some N with $\|S - T_N\| < \varepsilon$ (since $S - T_N$ is bounded and thus continuous and T_N is continuous by precondition). ■

Corollary 1.27: *$X' = B(X, \mathbb{C})$ is a Banach space.*

Proposition 1.28: *If X is a normed vector space, then the completion $\hat{X}$ is a Banach space.*

If T is linear and continuous and Y is a Banach space, then there is a unique linear continuous extension

$$\hat{T}\colon \hat{X} \to Y$$

such that $\|\hat{T}\| = \|T\|$.

Proof: $\hat{X}$ is a vector space via $\lambda[(x_n)_{n\in\mathbb{N}}] + \mu[(y_n)_{n\in\mathbb{N}}] := [(\lambda x_n + \mu y_n)_{n\in\mathbb{N}}]$ and it has a norm via

$$\|[(x_n)_{n\in\mathbb{N}}]\| := \lim_{n\to\infty} \|x_n\|.$$

Since

$$\hat{d}([(x_n)_{n\in\mathbb{N}}], [(y_n)_{n\in\mathbb{N}}]) = \lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} \|x_n - y_n\| = \|[(x_n + y_n)_{n\in\mathbb{N}}]\|$$

and $\hat{X}$ is complete with respect to $\hat{d}$ by Theorem 1.12, $\hat{X}$ is a Banach space. Since $d(Tx, Ty) = \|Tx - Ty\| \leq \|T\| d(x,y)$, $\hat{T}$ exists by Theorem 1.12 and one can check, that $\hat{T}$ is linear and $\|\hat{T}\| = \|T\|$ holds true. ■

Proposition 1.29: *Let $T \in B(X,Y)$ and $S \in B(Y,Z)$ for normed vector spaces X, Y, Z, then $\|ST\| \leq \|S\|\|T\|$.*

Proof: We have

$$\|(ST)x\| = \|S(Tx)\| \leq \|S\|\|Tx\| \leq \|S\|\|T\|\|x\|,$$

taking the supremum proves the claim. ■

Lemma 1.30: *In a Banach space, every absolute convergent series converges.*

Proof: Let $s_n := \sum_{k=1}^n x_k \in X$ and assume $\sum_{k=1}^\infty \|x_k\| < \infty$. Then $(s_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, since for $n \geq m$

$$\|s_n - s_m\| = \Big\| \sum_{k=m+1}^{n} s_k \Big\| \leq \sum_{k=m+1}^{n} \|x_k\| < \varepsilon,$$

so $(s_n)_{n\in\mathbb{N}}$ converges. ■

Definition 1.31: Let E be a normed vector space and let $F \subseteq E$ be a linear subspace. We put $x \sim y :\Leftrightarrow x - y \in F$, then

$$E/F := \{\dot{x} \mid x \in E\}$$

is called the *quotient space*, where $\dot{x}$ is the equivalence class with respect to $\sim$. E/F furthermore has a (semi) norm via

$$\|\dot{x}\| := \inf\{\|y\| \mid x \sim y\} = \inf\{\|x + z\| \mid z \in F\}.$$

Theorem 1.32: *Let E be a normed vector space, $F \subseteq E$ a linear subspace.*

(i) *E/F is a vector space via $\dot{x} + \dot{y} := (x + y)^{\cdot}$, $\lambda\dot{x} := (\lambda x)^{\cdot}$,*

(ii) *$\|\cdot\|$ is a semi norm on E/F and it is a norm if and only if F is closed.*

(iii) *If F is closed, then the canonical quotient map*

$$\begin{aligned} E &\longrightarrow E/F \\ x &\longmapsto \dot{x} \end{aligned}$$

is continuous, linear, has norm less or equal 1 and maps open sets to open sets.

(iv) *If F is closed and E is a banach space, then E/F is a Banach space.*

Proof: (i) The operations are well-defined, because

$$(x + F) + (y + F) = (x + y) + F,$$

for the scalar multiplication the prove works similarly.

(ii) Let $z_1, z_2 \in F$ such that $\|\dot{x}\| + \varepsilon = \|x + z_1\|$ and $\|\dot{y}\| + \varepsilon = \|y + z_2\|$. Then

$$\|\dot{x} + \dot{y}\| \leq \|(x + y) + z_1 + z_2\| \leq \|x + z_1\| + \|y + z_1\| \leq \|\dot{x}\| + \|\dot{y}\| + 2\varepsilon,$$

likewise $\|\lambda\dot{x}\| = |\lambda|\|\dot{x}\|$, so $\|\cdot\|$ is a semi norm.

Moreover $\|\dot{x}\| = 0 \Leftrightarrow \exists\, (z_n)_{n\in\mathbb{N}} \subseteq F : \|x + z_n\| \to 0 \Leftrightarrow x \in \overline{F}$. Now assume that F is closed, then $\|\dot{x}\| = 0$ implies $x \in \overline{F} = F$, hence $\dot{x} = 0$. Conversely, assume that $\|\cdot\|$ is a norm. Then $x \in \overline{F}$ implies, that $\|\dot{x}\| = 0$, therefore $\dot{x} = 0$, so $x \in F$. Thus $F \subseteq \overline{F} \subseteq F$, hence F is closed.

(iii) Since $\|\dot{x}\| \leq \|x\|$, the quotient map $x \mapsto \dot{x}$ is continuous with norm less or equal to 1. It remains to be shown, that $x \mapsto \dot{x}$ indeed maps open sets to open sets: Let $V \subseteq F$ be open and let $\dot{x} \in \dot{V}$. We need to show, that there is $\varepsilon > 0$, such that $B(\dot{x}, \varepsilon) \subseteq \dot{V}$. Without loss of generality, assume $x \in V$ (otherwise: $x + w \in V$ with $w \in F$); then there is $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq V$, because V is open. Let $\dot{z} \in B(\dot{x}, \varepsilon)$. Then $\|(z - x)^{\cdot}\| < \varepsilon$, hence we find $w \in F$, such that $\|z - x + w\| < \varepsilon$. But then $z + w \in B(x, \varepsilon) \subseteq V$ and therefore $\dot{z} = (z + w)^{\cdot} \in \dot{V}$.

(iv) It remains to be shown, that E/F is complete. Let $(\dot{x}_n)_{n \in \mathbb{N}} \subseteq E/F$ be a Cauchy sequence. Without loss of generality, $\|\dot{x}_n - \dot{x}_{n+1}\| < 2^{-(n+1)}$ (otherwise, pass to a subsequence). Hence, there is an $a_n \in E$, such that $\|a_n\| < 2^{-n}$ and $\dot{a}_n = \dot{x}_{n+1} - \dot{x}_n$. Then $s_n := \sum_{k=1}^n a_k$ converges absolutely, thus it converges in E by Lemma 1.30.

Put $z_{n+1} := x_n + \sum_{k=1}^n a_k$ with $z_1 := x_1$. Then $\dot{z}_n = \dot{x}_n$ and $(z_n)_{n \in \mathbb{N}}$ converges to some $z \in E$, so $\dot{z}_n \to \dot{z}$ due to the continuity of the quotient map. ■

Definition 1.33: Let X be a $\mathbb{K}$-vector space and let P be a family of seminorms on X. X is a *locally convex vector space*, if it is a topological vector space whose topology is generated by P: $U \subseteq X$ is open if and only if

$$\forall\, x \in U \,\exists\, n \in \mathbb{N} \,\exists\, p_1, \dots, p_n \,\exists\, \varepsilon_1, \dots, \varepsilon_n > 0 : B_{p_1}(x, \varepsilon_1) \cap \dots \cap B_{p_n}(x, \varepsilon_n) \subseteq U$$

holds. Here $B_{p_i}(x, \varepsilon) := \{y \in X \mid p_i(x - y) < \varepsilon\}$ mean the ε-balls with respect to p_i. Those are in particular open.

Remark 1.34: (i) If X is a normed vector space, then X is a locally convex vector space and thus even a topological vector space.

(ii) If $(x_\lambda)_{\lambda \in \Lambda} \subseteq X$ is a net, then $(x_\lambda) \to x$ if and only if $p(x_\lambda - x) \to 0 \,\forall\, p \in P$.

(iii) All seminorms $p \in P$ are continuous, i. e., if $x_\lambda \to x$, then $p(x_\lambda) \to p(x)$.

(iv) We may add arbitrary *continuous* seminorms without changing the topology. In particular, there is a maximal set of seminorms, namely the set of the continuous seminorms on X, defining a given locally convex topology.

(v) X is Hausdorff if and only if $\forall\, x \neq 0 \,\exists\, p \in P : p(x) > 0$.

Proof (of (v)): "$\Rightarrow$": The net $x_\lambda :\equiv x$ converges to x. Suppose $p(x) = 0 \,\forall\, p \in P$, then $x_\lambda \to 0$ by (ii). Because we have $x_\lambda \to x$ and $x_\lambda \to 0$, by Sheet 1, Exercise 2, $x = 0$ must hold.

"$\Leftarrow$": Assume $x_\lambda \to x$ and $x_\lambda \to y$ with $x \neq y$. By assumption we may find a $p \in P$ such that $p(x-y) > 0$. But then $p(x_\lambda - y) \to 0$ and $p(x_\lambda - y) \to p(x-y) > 0$, which is a contradiction to the continuity of p. ■

Example 1.35: (i) Let

$$X = C^\infty([0,1]) = \{f \colon [0,1] \to \mathbb{C} \text{ infinitely many times differentiable}\}$$

and $p_k(f) := \|f^{(k)}\|_\infty$, where $f^{(k)}$ is the k-th derivative. In this locally convex vector space, we have

$$f_n \to f \Leftrightarrow f_n^{(k)} \to f^{(k)}$$

for all k (uniform convergence of all derivatives).

(ii) Let $\Omega \subseteq \mathbb{C}$ be open, $(f_n)_{n\in\mathbb{N}} \subseteq C(\Omega)$. We say, that the sequence (f_n) converges to $f\colon \Omega \to \mathbb{C}$, if for all $K \subseteq \Omega$ compact the following statement holds:

$$\forall\, \varepsilon > 0\, \exists\, N \in \mathbb{N}\, \forall\, n \geq N : \|f_n - f\|_K < \varepsilon.$$

Here $\|f\|_K := \sup_{x\in K}|f(x)|$. Since $(C(K), \|\cdot\|_K)$ is complete, we obtain that $f|_K$ is continuous, in fact f is even continuous on the whole of Ω. We may prove: If f_n is holomorphic for all n, then f is holomorphic, too. Hence, this convergence is a useful one and it comes from seminorms $(\|\cdot\|_K)_{K\subseteq\Omega}$.

One can characterise all relatively compact subsets of $\mathcal{O}(\Omega)$ (in analogy to the theorem of Arzela-Ascoli) via the theorem of Montel.

Definition 1.36: Let X be a $\mathbb{K}$-vector space and $M \subseteq X$ a subset.

(i) M is *convex* if and only if $\forall\, x, y \in M\, \forall\, t \in [0, 1] : tx + (1 - t)y \in M$.

(ii) M is called *circled* if and only if $\forall\, x \in M\, \forall\, \lambda \in \mathbb{K}, |\lambda| \leq 1 : \lambda x \in M$.

(iii) M is called *absorbant* if and only if $\forall\, x \in X\, \exists\, \lambda > 0 : \lambda x \in M$.

Example 1.37 (in $\mathbb{R}^2$): We can illustrate the types of sets defined in Definition 1.36 in $\mathbb{R}^2$ with the following example sets:

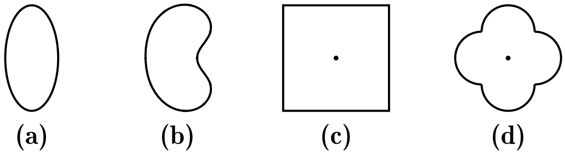


Figure 1.2: Examples for a convex set **(a)**, a non-convex set **(b)**, a circled convex set **(c)** and a circled, but non-convex set **(d)**.

Remark 1.38: In a locally convex vector space, $\bigcap_{k=1}^n B_{p_i}(0, \varepsilon_k)$ is convex, circled and absorbing. Such sets correspond to our idea of a ball arround zero.

Proposition 1.39: *Let X be a $\mathbb{K}$-vector space and let $M \subseteq X$ be a subset.*

(i) *If M is convex and absorbing, then the* "Minkowski functional"

$$p_M(x) := \inf\{\lambda > 0 \mid x \in \lambda M\}$$

is a sublinear functional, i. e., it holds (a) $\forall\, \alpha \geq 0\, \forall\, x \in X : p(\alpha x) = \alpha p(x)$ *and* (b) $p(x + y) \leq p(x) + p(y)\, \forall\, x, y \in X$.

(ii) *If in addition M is circled, then p_M is also a seminorm. Since*

$$\{x \mid p_M(x) < 1\} \subseteq M \subseteq \{x \mid p_M(x) \leq 1\},$$

the set M is "almost" a unit ball with respect to p_M.

Conversely: If p is a seminorm on X, we put $B_p := \{x \in X \mid p(x) < 1\}$ and $D_p := \{x \in X \mid p(x) \leq 1\}$. Then B_p and D_p are convex, circled and absorbing and $p_{B_p} = p_{D_p} = p$.

Proof: (i) (a) Let $x \in M$ and $\alpha \geq 0$ with $\alpha \neq 0$. Then

$$p_M(\alpha x) = \inf\{\lambda \mid \alpha x \in \lambda M\} = \inf\{\alpha\lambda \mid \alpha x \in \alpha\lambda M\} = \alpha p_M(x).$$

(b) Let $\varepsilon > 0$ and $x, y \in M$. Then we have $(p_M(x) + \varepsilon)^{-1} x \in M$ aswell as $(p_M(y) + \varepsilon)^{-1} y \in M$. Because M is convex, their convex combination

$$\frac{p_M(x) + \varepsilon}{p_M(x) + p_M(y) + 2\varepsilon}\left(\frac{1}{p_M(x) + \varepsilon} x\right) + \frac{p_M(y) + \varepsilon}{p_M(x) + p_M(y) + 2\varepsilon}\left(\frac{1}{p_M(y) + \varepsilon} y\right),$$

is an element of M, so

$$\frac{1}{p_M(x) + p_M(y) + 2\varepsilon}(x + y) \in M,$$

hence $p_M(x + y) \leq p_M(x) + p_M(y) + 2\varepsilon$. $\varepsilon \to 0$ proves the claim.

(ii) Let $x \in M$ and $\alpha \in \mathbb{K}, \alpha \neq 0$. Then

$$p_M(\alpha x) = \inf\{|\alpha|\lambda \mid \alpha x \in |\alpha|\lambda M\} = |\alpha| p_M(x).$$ ■

Theorem 1.40 (Characterisation of locally convex vector spaces): *Let X be a topological $\mathbb{K}$ vector space. X is locally convex if and only if every neighbourhood of zero contains an open, convex, circled, absorbant set.*

Proof: "⇒": This is Remark 1.38.
"⇐": Set $P := \{p_M \mid M \in \mathfrak{T} \text{ convex, circled, absorbant}, M \subseteq V, V \in \mathfrak{U}(0)\}$.
Then the given topology $\mathfrak{T}$ coincides with σ, the one induced by P.

Proof (That the induced topology coincides with the given one): Let V be open with resepect to $\mathfrak{T}$. We have to show, that V then is open with respect to σ. Let $z \in V$. Then $V' := V - z$ is a neighbourhood of zero with respect to $\mathfrak{T}$. By assumption, we find an open convex, circled, absorbant set $M \subseteq V'$. Then $0 \in \{x \mid p_M(x) < 1\} \subseteq M \subseteq V'$, thus V' is also a neighbourhood of zero with respect to σ. Hence V is a neighbourhood of z with respect to σ, then we use Remark 1.2.

The other inclusion is technical and left as exercise for the reader. ■

Remark 1.41: (i) The crucial point in Theorem 1.40 is: "There exists a *convex* set". In fact: In *any* topological vector space, any neighbourhood of zero contains an open, circled and absorbing set (If $U \in \mathfrak{U}(0)$, then $\mu^{-1}(U)$ is a neighbourhood of (0,0), where $\mu\colon \mathbb{K} \times X \to X$ is the multiplication map, i. e., $\exists\, \varepsilon > 0\, \exists\, V \in \mathfrak{U}(0)$ open such that $B(0, \varepsilon) \times V \subseteq \mu^{-1}(U)$, then put $W := \{\lambda v \mid |\lambda| < \varepsilon, v \in V\} \subseteq U$).

(ii) One can characterise metric vector spaces using seminorms (see Sheet 3, Exercise 1).

(iii) If $p_1, \dots, p_n$ are seminorms on X, then also $\sum_{i=1}^{n} p_i$ and $\max\{p_1, \dots, p_n\}$ are seminorms and

$$P' := \Big\{ \sum_{i=1}^{n} p_i : p_i \in P \Big\}, \qquad P'' := \Big\{ \max\{p_1, \dots, p_n\} : p_i \in P \Big\}$$

define the same topology.

But P', P'' are ordered ($p(x) \le q(x)\, \forall\, x$) and filtrations, too (If $p, q \in P'$, then $p + q \ge p, q$ and if $p, q \in P''$, then $\max\{p, q\} \ge p, q$).

Theorem 1.42: *Let (X_1, P_1), (X_2, P_2) be locally convex spaces, $T\colon X_1 \to X_2$ a linear map. Then T is continuous if and only if*

$$\forall\, q \in P_2\, \exists\, p \in P_1\, \exists\, C \ge 0 : q(Tx) \le Cp(x)\, \forall\, x \in X.$$

Proof: "$\Leftarrow$": Let $x_\lambda \to x$ and let $q \in P_2$. Then $q(Tx_\lambda - Tx) \le Cp(x_\lambda - x) \to 0$ (by Remark 1.34) for some $p \in P_1$.

"$\Rightarrow$": Let $q \in P_2$. Then $V := \{y \in X_2 \mid q(y) < 1\} \subseteq X_2$ is open. Because T is continuous, we know that $T^{-1}V \subseteq X_1$ is open, i. e., $\exists\, p_1, \dots, p_n \in P_1$ and $\exists\, \varepsilon_1, \dots, \varepsilon_n > 0$, such that

$$T(\{x \mid p_i(x) < \varepsilon_i, 1 \le i \le n\}) \subseteq V.$$

Now choose $\varepsilon < \min\{\varepsilon_1, \dots, \varepsilon_n\}$, $p \ge p_1, \dots, p_n$ (without loss of generality, P_1 is a filtration) and put $C := \varepsilon^{-1}$. Then we have: If $p(x) \le \varepsilon$, then $q(Tx) < 1$, thus

$$q(Tx) = q\left(T\left(x\frac{p(x)\varepsilon}{p(x)\varepsilon}\right)\right) = Cp(x)q\left(T\left(x\frac{\varepsilon}{p(x)}\right)\right) < Cp(x).$$

In case $p(x) = 0$ we have $q(Tx) < 1$, but also $p(Nx) = 0\, \forall\, N \in \mathbb{N}$, and we get that $Nq(Tx) = q(T(Nx)) < 1\, \forall\, N \in \mathbb{N}$, hence $q(Tx) = 0$ must hold. ■

2 Hahn-Banach Theorems

In order to understand spaces $\{f\colon \mathbb{C}\to\mathbb{C}\}$, evaluation maps are very useful, e. g.

$$\begin{aligned}\mathrm{ev}_x\colon \{f\colon \mathbb{C}\to\mathbb{C}\} &\longrightarrow \mathbb{C}\\ f &\longmapsto f(x).\end{aligned}$$

More general, we want to understand continuous functionals $f\colon X\to\mathbb{K}$ for some normed vector space X.

Definition 2.1: Let E be a $\mathbb{K}$ vector space. A *sublinear functional* on E is a map $p\colon E\to\mathbb{R}$ such that

(i) $p(\lambda x)=\lambda p(x)\,\forall\,\lambda\geq 0\,\forall\,x\in E$,

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

Example 2.2: (i) If p is a seminorm on a $\mathbb{K}$ vector space E, then p is a sublinear functional.

(ii) If $p\colon E\to\mathbb{R}$ is a linear function, then p is a sublinear functional as well.

In this chapter, we are interested in two kinds of questions:
(1) *Extensions*: Given a linear function $f\colon F\to\mathbb{R}$ with $f(x)\leq p(x)$ for all $x\in F$ for a subspace $F\subseteq E$. Can f be extended to a linear function $\tilde{f}\colon E\to\mathbb{R}$? Does $\tilde{f}(x)\leq p(x)$ still hold?
(2) *Separations*: Is there a continuous linear function $f\colon X\to\mathbb{R}$ such that $f(x)\leq 1$ on $M\subseteq X$ and $f(x_0)>1$ for a fixed $x_0\notin M$?

Theorem 2.3 (Hahn-Banach, root version): *Let E be an $\mathbb{R}$ vector space, $p\colon E\to\mathbb{R}$ be a sublinear functional, $F\subseteq E$ a linear subspace and let $f\colon F\to\mathbb{R}$ be linear with $f(x)\leq p(x)\,\forall\,x\in F$. Then there is a linear extension $\tilde{f}\colon E\to\mathbb{R}$ of f, such that $\tilde{f}(x)\leq p(x)\,\forall\,x\in E$.*

Proof: Let

$$\begin{aligned}M:=\{(G,g)\mid F\subseteq G\subseteq E &\text{ linear subspace},\\ &g\colon G\to\mathbb{R}\text{ linear}, g|_F=f, g(x)\leq p(x)\,\forall\,x\in G\}\end{aligned}$$

Then $M\neq\varnothing$, because $(F,f)\in M$, and M is ordered via

$$(G,g_1)\prec(G_2,g_2):\Leftrightarrow G_1\subseteq G_2, g_2|_{G_1}=g_1.$$

This order is reflexive (because $(G,g)\prec(G,g)$), it is antisymmetric (for (G,g), $(H,h)\in M$ holds: if $(G,g)\prec(H,h)\prec(G,g)$, then $G=H$ and $g=h$), transitive and inductive (if $(G_\alpha,g_\alpha)_{\alpha\in\Lambda}\subseteq M$ such that mutually two elements are comparable, then $G := \bigcup_{\alpha\in\Lambda} G_\alpha$ with $g(x) := g_\alpha(x)$ for $x \in G_\alpha \subseteq G$ is an upper bound). Via the Lemma of Zorn, it exists a maximal $(H, h) \in M$.

Assume $H \subsetneq E$, hence we have $z \in E \setminus H$. We define

$$\tilde{h} \colon \langle H, z\rangle \longrightarrow \mathbb{R}$$

such that $(H, h) \prec (\langle H, z\rangle, \tilde{h}) \in M$. For the definition: Let $x, y \in H$. Then $h(x) + h(y) = h(x + y) \leq p(x + y) \leq p(x + z) + p(y - z)$, so we can put

$$m := \sup_{y\in H}\{h(y) - p(y - z)\} \leq \inf_{x\in H}\{p(x + z) - h(x)\} =: M.$$

Now choose $a \in [m, M]$ and put $\tilde{h}(x + \lambda z) := h(x) + \lambda a$ for $x \in H$, $\lambda \in \mathbb{R}$. So, $\tilde{h}$ is well defined and linear, since H and $\{z\}$ are linearly independent and we have $(H, h) \prec (\langle H, z\rangle, \tilde{h})$.

It remains to be shown, that $\tilde{h}(x) \leq p(x)\,\forall\, x \in \langle H, x\rangle$. Let $\lambda > 0$, then

$$\tilde{h}(x + \lambda z) = h(x) + \lambda a \overset{a\leq M}{\leq} h(x) + \lambda\left(p\left(\frac{x}{\lambda} + z\right) - h\left(\frac{x}{\lambda}\right)\right) = p(x + \lambda z). \qquad \blacksquare$$

Let now $\lambda < 0$, then

$$\begin{aligned}\tilde{h}(x + \lambda z) = h(x) + \lambda a &\overset{a\geq m}{\leq} h(x) + \lambda m \\ &\leq h(x) + \lambda\left(h\left(-\frac{x}{\lambda}\right) - p\left(-\frac{x}{\lambda} - z\right)\right) = p(x + \lambda z).\end{aligned}$$

Example 2.4: On $\ell^\infty_{\mathbb{R}} = \{(a_n)_{n\in\mathbb{N}} \mid (a_n)_{n\in\mathbb{N}} \text{ bounded}, a_n \in \mathbb{R}\}$ there is a bounded linear functional $\mathrm{LIM}\colon \ell^\infty(\mathbb{R}) \to \mathbb{R}$ with

$$\liminf_{n\to\infty} a_n \leq \mathrm{LIM}((a_n)) \leq \limsup_{n\to\infty} a_n.$$

For convergent sequences $(a_n)_{n\in\mathbb{N}}$, we have that $\mathrm{LIM}((a_n)) = \lim_{n\to\infty} a_n$. For non-convergent sequences, it is this "Banach limit".

Theorem 2.5 (Hahn-Banach, seminorm version): *Let E be a $\mathbb{K}$ vector space, p a seminorm on E, $F \subseteq E$ a linear subspace, $f\colon F \to \mathbb{K}$ a linear functional and $|f(x)| \leq p(x)\,\forall\, x \in F$. Then there is a linear extension $\tilde{f}\colon E \to \mathbb{K}$ such that $|\tilde{f}(x)| \leq p(x)\,\forall\, x \in E$.*

Proof: (i) Let $\mathbb{K} = \mathbb{R}$: Since $f(x) \leq |f(x)|$, we obtain $\tilde{f}$ by Theorem 2.3 with $\tilde{f}(x) \leq p(x)$, but $-\tilde{f}(x) = \tilde{f}(-x) \leq p(-x) = p(x)$, so we have $|f(x)| \leq p(x)$.

(ii) Let $\mathbb{K} = \mathbb{C}$: Consider $u\colon F \to \mathbb{R}$ with $u(x) := \mathrm{Re}(f(x))$, then the inequality $|u(x)| \leq |f(x)| \leq p(x)$ holds. Via (i), we get an extension $\tilde{u}\colon E \to \mathbb{R}$, with $|\tilde{u}(x)| \leq p(x)$. Now put $\tilde{f}(x) := \tilde{u}(x) - \mathrm{i}\tilde{u}(\mathrm{i}x)$. Then $\tilde{f}(x) = f(x)$ for $x \in F$, because $u(\mathrm{i}x) = \mathrm{Re}(f(\mathrm{i}x)) = \mathrm{Re}(\mathrm{i}f(x)) = -\mathrm{Im}(f(x))$, moreover we have

$$\tilde{f}(x + y) = \tilde{f}(x) + \tilde{f}(y) \quad , \quad \tilde{f}(\lambda x) = \lambda\tilde{f}(x)$$

for all $x, y \in E$ and $\lambda \in \mathbb{R}$ as well as $\tilde{f}(\mathrm{i}x) = \tilde{u}(\mathrm{i}x) - \mathrm{i}\tilde{u}(-x) = \mathrm{i}\tilde{f}(x)$, thus $\tilde{f}(\lambda x) = \lambda \tilde{f}(x) \,\forall\, \lambda \in \mathbb{C}$.

Finally, for $x \in E$ there is a $\mu \in \mathbb{C}$ with $|\mu| = 1$ such that $|\tilde{f}(x)| = \mu \tilde{f}(x)$, so

$$|\tilde{f}(x)| = \mu \tilde{f}(x) = \tilde{f}(\mu x) = \mathrm{Re}(\tilde{f}(\mu x)) = \tilde{u}(\mu x) \leq p(\mu x) = p(x). \qquad \blacksquare$$

Corollary 2.6 ((ii) : Hahn-Banach, norm version): *Let E be a normed $\mathbb{K}$ vector space.*

(i) *For all $x \in E \setminus \{0\}$ there exists a continuous, linear $f\colon E \to \mathbb{K}$ with $f(x) = 1$,*

(ii) *If $F \subseteq E$ is a linear subspace and $f\colon F \to \mathbb{K}$ is continuous and linear, then there is a continuous, linear extension $\tilde{f}\colon E \to \mathbb{K}$ such that $\|\tilde{f}\| = \|f\|$.*

Proof: (i) Define $g\colon \langle x \rangle \to \mathbb{K}$ on the onedimensional linear subspace $\langle x \rangle \subseteq E$ via $g(\lambda x) := \lambda$ and a seminorm $p\colon E \to \mathbb{K}$ by $p(y) := \frac{1}{\|x\|}\|y\|$. Then we have

$$|g(\lambda x)| = |\lambda| \frac{\|x\|}{\|x\|} = p(\lambda x).$$

Theorem 2.5 ensures the existence of a linear extension $\tilde{f}\colon E \to \mathbb{K}$ such that $|\tilde{f}(y)| \leq p(y) = \frac{1}{\|x\|}\|y\|$, i. e., $\tilde{f}$ is continuous.

(ii) With $p(x) := \|f\|\|x\|$, we have $|f(x)| \leq p(x)$. By Theorem 2.5 we know, that there exists a $\tilde{f}\colon E \to \mathbb{K}$ such that $|\tilde{f}(x)| \leq p(x) = \|f\|\|x\|$, i. e., $\tilde{f}$ is continuous and $\|\tilde{f}\| \leq \|f\|$.

For $x \in F$, we have $f(x) = \tilde{f}(x)$, hence $\|f\| \leq \|\tilde{f}\|$ by Proposition 1.23. $\blacksquare$

Theorem 2.7 (Hahn-Banach, locally convex version): *Let E be a locally convex vector space and $F \subseteq E$ a linear subspace, $f\colon F \to \mathbb{K}$ continuous and linear. Then there is a continuous linear extension $\tilde{f}\colon E \to \mathbb{K}$.*

Proof: Since f is continuous, we find a seminorm p on E and $C > 0$ such that $|f(x)| \leq Cp(x)$ by Theorem 1.42. By Theorem 2.5 we find an extension $\tilde{f}$, such that $|\tilde{f}(x)| \leq Cp(x)$, i. e., $\tilde{f}$ is continuous by Theorem 1.42. $\blacksquare$

Theorem 2.8 (Hahn-Banach separation theorem): *Let E be a locally convex $\mathbb{R}$ vector space and let $M \subseteq E$ be closed and convex such that $0 \in M$. Let $x_0 \notin M$. Then there is a linear, continuous function $f\colon E \to \mathbb{R}$ with $f(x_0) > 1$ and $f(x) \leq 1$ for all $x \in M$.*

Proof: Let V be an open, convex circled, absorbant neighbourhood of zero such that $(x_0 + V) \cap M \neq \varnothing$. It exists by Theorem 1.40. Then $(x_0 + \frac{V}{2}) \cap (M + \frac{V}{2}) = \varnothing$ – we see this equality because if we have $x_0 + \frac{z_1}{2} = y + \frac{z_2}{2}$ with $z_1, z_2 \in V$ and $y \in M$, then $x_0 + \frac{1}{2}(z_1 - z_2) = y \in M$ holds, which is a contradiction.

As $M' := M + \frac{V}{2}$ is convex and absorbing, Proposition 1.39 ensures that $p_{M'}$ is a sublinear functional.

Now put

$$f_0 : \langle x_0 \rangle \longrightarrow \mathbb{R}$$
$$\lambda x_0 \longmapsto \lambda p_{M'}(x_0),$$

then $f_0(x) \leq p_{M'}(x)$ holds. By Theorem 2.3, there exists a function $f : E \to \mathbb{R}$ with $f(x_0) = f_0(x_0) = p_{M'}(x_0) > 1$, because $x_0 + \frac{V}{2} \notin M'$, and we also have that $f(x) \leq p_{M'}(x)$, hence f is continuous (in Proposition 1.39 we may prove that $p_{M'}$ is continuous as soon as out vector space is locally convex and $|f(x)| \leq C p_{M'}(x)$).

And: $\forall\, x \in M : x \in M + \frac{V}{2} = M'$, hence $p_{M'}(x) \leq 1$, i. e., $f(x) \leq 1$. ■

Corollary 2.9: *Let $M \subseteq E$ be a closed linear subspace of a locally convex $\mathbb{K}$ vector space E, $x_0 \in E \setminus M$. Then there is a continuous, linear function $f : E \to \mathbb{K}$ with $f(x_0) = 1$ and $f \equiv 0$ on M.*

Proof: (i) Let $\mathbb{K} = \mathbb{R}$. Theorem 2.8 ensures the existence of a linear map $f : E \to \mathbb{R}$ with $f(x_0) > 1$ and $f(x) \leq 1\, \forall\, x \in M$, hence $f(x) = 0\, \forall\, x \in M$ (because M is a subspace, $f(\lambda x) = \lambda f(x) \leq 1\, \forall\, \lambda \in \mathbb{R}$), then put $f' := \frac{f}{f(x_0)}$.

(ii) Let $\mathbb{K} = \mathbb{C}$. Then work with the real and imaginary parts. ■

Remark 2.10: Corollary 2.6 (ii) is also true for locally convex vector spaces: Choosing $M = \{0\}$ proves this.

A consequence of the Theorem of Hahn-Banach is the Theorem of Krein-Milman about the geometry of locally convex vector spaces. Let's state and prove it!

Definition 2.11: Let X be a vector space, $C \subseteq X$ be convex.

(i) A subset $M \subset C$ is called *extremal* in C, if it holds:

$$\forall\, x, y \in C\, \forall\, t \in (0,1) : tx + (1-t)y \in M \Rightarrow x, y \in M.$$

(ii) A point $z \in C$ is called an *extremal point* of C, if it holds:

$$\forall\, x, y \in C\, \forall\, t \in (0,1) : tx + (1-t)y = z \Rightarrow x = y = z.$$

Example 2.12 (Extremal sets and extremal points in $\mathbb{R}^2$): Consider the following example sets:

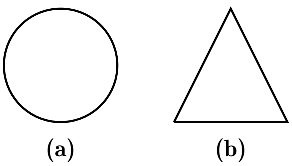


Figure 2.1: Examples for an extremal set **(a)** and a set with a boundary, that is an extremal set **(b)**.

Theorem 2.13 (Krein-Milman): *Let X be a locally convex vector space and let $\varnothing \neq C \subseteq X$ be a compact, convex subset.*

(i) *C contains at least one extremal point.*

(ii) *If $\varnothing \neq K \subseteq C$ is a compact, extremal subset, then it contains at least one extremal point of C.*

Proof: Consider $M := \{\varnothing \neq K \subseteq C \text{ extremal, compact}\}$. Then, M is ordered inductively by "$\supseteq$" ($K_1 \leq K_2 :\Leftrightarrow K_1 \supseteq K_2$ and $\varnothing \neq \bigcup_{i \in I} K_i$ is an upper bound for $(K_i) \subseteq M$, $K_i \leq K_{i+1}$). Via the Lemma of Zorn, there is a maximal element $K_0 \in M$. We now need to show, that K_0 consists of only one single point. Assume, there are $x_0 \neq y_0$ in K_0. By Corollary 2.9, there exists an $f\colon X \to \mathbb{K}$ that is continuous and linear, such that $f(x_0 - y_0) = 1$. Without loss of generality, we may assume $\operatorname{Re}(f(x_0)) \neq \operatorname{Re}(f(y_0))$ (otherwise work with $\operatorname{Im}(f)$). Since K_0 is compact, we have a minimum m of $\operatorname{Re}(f)$ on K_0. Now let $K^* := \{y \in K_0 \mid \operatorname{Re}(f(y)) = m\}$. Then

(1) $\varnothing \neq K^* \subsetneq K_0$, because m is a minimum and $\operatorname{Re}(f(x_0)) \neq \operatorname{Re}(f(y_0))$,

(2) K^* is compact, since $K^* = (\operatorname{Re}(f))^{-1}(\{m\})$ is a closed subset of a compact set,

(3) K^* is extremal in C, because it holds: If $tx + (1-t)y \in K^*$ for $t \in (0,1)$, $x, y \in K_0$, then $x, y \in K^*$. To see this, suppose $tx + (1-t)y \in K^*$, $t \in (0,1)$, $x, y \in K_0$. Then

$$t\operatorname{Re}(f(x)) + (1-t)\operatorname{Re}(f(y)) = \operatorname{Re}(f(tx + (1-t)y)) = m$$

and $\operatorname{Re}(f(x)), \operatorname{Re}(f(y)) \geq m$, hence $\operatorname{Re}(f(x)) = \operatorname{Re}(f(y)) = m$ and thereby $x, y \in K^*$.

Now let $tx + (1-t)y \in K^*$, $t \in (0,1)$, $x, y \in C$. Since K_0 is extremal in C, $x, y \in K_0$ and by (a), we have $x, y \in K^*$. By (i), (ii), (iii), K_0 is not maximal in M which is a contradiction, hence $K_0 = \{z\}$.

For (ii): If $K \in M$, then $\{z\} \subseteq K_0 \subseteq K$. ■

Lemma 2.14: *If $M \subseteq X$ is a subset, then its* convex hull

$$\operatorname{Konv}(M) := \bigcap_{M \subseteq C, C \text{ convex}} C$$

may be written as

$$\operatorname{Konv}(M) = \Big\{ \sum_{i=1}^{n} t_i x_i : x_1, \ldots, x_n \in M, t_i \geq 0, \sum_{i=1}^{n} t_i = 1, n \in \mathbb{N} \Big\}.$$

Proof: "$\supseteq$": $\operatorname{Konv}(M)$ is convex, so $\sum_{i=1}^n t_i x_i \in \operatorname{Konv}(M)$ for all $x_i \in \operatorname{Konv}(M)$: Via induction one sees that

$$\sum_{i=1}^{n+1} t_i x_i = t_1 x_1 + (1-t_1) \sum_{i=2}^{n+1} \frac{t_i}{(1-t_1)} x_i \in \operatorname{Konv}(M).$$

"$\subseteq$": The set $\{\sum_{i=1}^n t_i x_i : x_1, \ldots, x_n \in M, t_i \geq 0, \sum_{i=1}^n t_i = 1, n \in \mathbb{N}\}$ is convex, because

$$\begin{aligned} &t \sum_{i=1}^{n} t_i x_i + (1-t) \sum_{j=1}^{m} s_j y_i \\ &= \sum_i t_i' x_i' \in \left\{ \sum_{i=1}^n t_i x_i : x_1, \ldots, x_n \in M, t_i \geq 0, \sum_{i=1}^n t_i = 1, n \in \mathbb{N} \right\} \end{aligned}$$ ■

and it contains M.

Corollary 2.15 (Krein-Milman): *Let X be a locally convex vector space, $C \subseteq X$ a compact, convex subset. Then*

$$C = \overline{\operatorname{Konv}(\operatorname{Ext}(C))}$$

where $\operatorname{Ext}(C) := \{z \in C \mid z$ *is an extremal point of* $C\}$.

Proof: Let $A := \overline{\operatorname{Konv}(\operatorname{Ext}(C))}$, then $A \subseteq C$. Now assume $\exists\, x_0 \in C \setminus A$, without loss of generality it holds $0 \in A$ (otherwise translate the set) and $\mathbb{K} = \mathbb{R}$ (otherwise use Re, ...). A is closed and convex, by Theorem 2.7, we may find a continuous linear functional $f\colon X \to \mathbb{R}$, such that $f(x_0) > 1$ and $f(x) \leq 1 \,\forall\, x \in A$. The set

$$K := \{x \in C \mid f(x) = \max_{y \in C} f(y)\}$$

is non-empty (since C is compact), compact (since it is closed) and extremal in C (as in the proof of Theorem 2.13), therefore K contains an extremal point z of C. Hence, $z \in A$ and $f(z) \leq 1$. On the other hand, $z \in K$, hence $f(z) = \max_{y \in C} f(y) > 1$ which is a contradiction. ■

3 Dual spaces

In the first chapter, we saw that it holds: If E is a normed space, then its dual space

$$E' = B(E, \mathbb{K}) = \{f\colon E \to \mathbb{K} \text{ linear, continuous}\}$$

is a Banach space.

Remark 3.1: There is a bilinear (i. e., linear in both components) map

$$\begin{aligned} \langle \cdot, \cdot \rangle \colon E' \times E &\longrightarrow \mathbb{K} \\ \langle f, x \rangle &:= f(x) \end{aligned}$$

If E is a Hilbert space, then $E \cong E'$ and $\langle \cdot, \cdot \rangle$ is the inner product. This map is continuous: Let $(f_n)_{n\in\mathbb{N}} \to f$ and $(x_n)_{n\in\mathbb{N}} \to x$, then

$$|f_n(x_n) - f(x)| \leq \|f_n - f\|\|x_n\| + \|f\|\|x_n - x\| \to 0,$$

since $(x_n)_{n\in\mathbb{N}}$ is bounded as a converget sequence.

Theorem 3.2: *Let E be a normed vector space. There is a natural map $i\colon E \to E''$ such that $i(x)(f) := f(x)$ for $x \in E, f \in E'$. This map is linear, continuous and isometric (i. e., $\|i(x)\| = \|x\|$).*

Proof: i is well-defined, since $i(x)$ is continuous and linear (i. e., $i(x) \in E''$):

$$|i(x)(f)| = |f(x)| \leq \|f\|\|x\| \quad \Rightarrow \quad \|i(x)\| \leq \|x\|$$

which proves, that $i(x)$ is continuous. Moreover, $\|i\| \leq 1$ i. e., i is continuous. By Hahn-Banach (Corollary 2.6), we find $f \in E'$ for a given $x \in E$, such that $f(x) = \|x\|$, $\|f\| = 1$ (as an extension of $\langle x \rangle \to \mathbb{K}, \lambda x \to \lambda\|x\|$), therefore

$$|i(x)(f)| = |f(x)| = \|x\| \quad \Rightarrow \quad \|i(x)\| = \|x\|. \qquad \blacksquare$$

Remark 3.3: Since E'' is complete, we know that the completion of E is isometrically isomorphic to $\overline{i(E)} \subseteq E''$ (see Sheet 1).

Definition 3.4: E is called *reflexive*, if and only if $i\colon E \to E''$ is an isomorphism.

Example 3.5: (i) We have $c_0' = \ell^1$ and $\ell^{1\prime} = \ell^\infty$ by Sheet 2, so c_0 is not reflexive.

(ii) Put $c := \{(a_n)_{n\in\mathbb{N}} \mid a_n \in \mathbb{C}, (a_n) \text{ converges}\}$, then $c \supsetneq c_0$, but $c' = c_0'$ (see Sheet 4).

(iii) For $1 \leq p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, then $\ell^{p\prime} = \ell^q$ and $\ell^{q\prime} = \ell^p$, hence ℓ^p is reflexive for $1 < p < \infty$.

Theorem 3.6: *Let E, F be normed spaces, let $T\colon E \to F$ be a linear, continuous map.*

(i) *There is exactly one map $T'\colon F' \to E'$, such that $\forall\, x \in F', y \in E$*

$$\langle x, Ty\rangle_{F' \times F} = \langle T'x, y\rangle_{E' \times E}.$$

(ii) *The diagram*

$$\begin{array}{ccc} E & \xrightarrow{\quad T \quad} & F \\ {\scriptstyle i}\downarrow & & \downarrow{\scriptstyle i} \\ E'' & \xrightarrow[\quad T'' \quad]{} & F'' \end{array}$$

commutes, i. e., T'' extends T,

(iii) *T' is continuous, linear and it holds $\|T'\| = \|T\|$.*

Proof: (i) Put $T'(f) := f \circ T \in E'$. Then

$$\langle x, Ty\rangle = x(Ty) = (x \circ T)(y) = (T'x)(y) = \langle T'x, y\rangle$$

(ii) Let $x \in E$, $f \in F'$. Then we have

$$\begin{aligned} \big(T''(ix)\big)(f) &\overset{3.1}{=} \langle T''(ix), f\rangle_{F' \times F} \overset{(a)}{=} \langle ix, T'f\rangle_{E' \times E} \overset{3.1}{=} (ix)(T'f) \\ &\overset{3.2}{=} (T'f)(x) \overset{3.1}{=} \langle T'f, x\rangle_{E' \times E} \overset{(a)}{=} \langle f, Tx\rangle_{F' \times F} \overset{3.1}{=} f(Tx) \overset{3.2}{=} (i(Tx))(f) \end{aligned}$$

therefore, for all $f \in E'$, it holds that $T''(ix) = i(Tx)$ and as this holds for any $x \in E$, it holds that $T'' \circ i = i \circ T$.

(iii) That T' is linear is clear. Let $x \in F'$ and $y \in E$. Then it holds

$$|(T'x)(y)| = |x(Ty)| \le \|x\|\|Ty\| \le \|x\|\|T\|\|y\| \,\forall\, y \in E$$

hence $\|T'x\| \le \|x\|\|T\| \,\forall\, x \in F'$ which implies $\|T'\| \le \|T\|$. In particular, T' is continuous.

We also have $\|T''\| \le \|T'\| \le \|T\|$ and $\|T\| \le \|T''\|$, because

$$\begin{aligned} \|T''\| &= \sup\{\|T''y\| \mid y \in E'', \|y\| = 1\} \\ &\ge \sup\{\|T''ix\| \mid x \in E, \|x\| = 1\} \\ &= \sup\{\|Tx\| \mid x \in E, \|x\| = 1\} = \|T\| \end{aligned}$$

since $i\colon E \to E''$ with $\|i(x)\| = \|x\|$. ■

Remark 3.7: The norm on E is given by $\|x\| := \sup\{|f(x)| \mid f \in E', \|f\| = 1\}$ (see Sheet 3).

4 Theorem of Baire (and some consequences)

Consider $c_0 = \{(a_n)_{n\in\mathbb{N}} \mid a_n \in \mathbb{C}, \lim_{n\to\infty} a_n = 0\}$. This is a vector space. What is a basis of this vector space? One could try $e_n := (\delta_{i,n})_{n\in\mathbb{N}} \in c_0$, but this is not a basis of c_0; it only spans $f = \{(a_n)_{n\in\mathbb{N}} \mid a_m = 0 \text{ for } m \geq N \text{ and some } N \in \mathbb{N}\}$. We will see, that c_0 does not have a countable vector space basis. In fact: there is no Banach space of countable infinity vector space dimension. Why?

Theorem 4.1 (Baire): *Let X be a complete metric space and let $M_n \subseteq X$ be closed subsets, such that $X = \bigcup_{n\in I} M_n$, I countable. Then there is a $n_0 \in I$ and an open set $U \neq \varnothing$, such that $U \subseteq M_{n_0}$.*

Proof: Assume: $\forall x \in X \,\forall \varepsilon > 0 \,\forall n \in \mathbb{N} : B(x,\varepsilon) \cap X \setminus M_n \neq \varnothing$. Let $x_0 \in X$, $\varepsilon_0 > 0$. Then $B(x_0,\varepsilon_0) \cap (X \setminus M_1) \neq \varnothing$ is open, hence we can find $x_1 \in X$ and $0 < \varepsilon_1 < \frac{\varepsilon_0}{2}$ with $B(x_1, 2\varepsilon_1) \subseteq B(x_0,\varepsilon_0) \cap (X \setminus M_1)$. Chose inductively $x_n \in X$ and $0 < \varepsilon_n < \frac{1}{2^n}\varepsilon_0$, such that

$$B(x_n, 2\varepsilon_n) \subseteq B(x_{n-1}, 2\varepsilon_{n-1}) \cap (X \setminus M_n) \neq \varnothing$$

Hence, the sequence $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and thus converges to some $x \in X$. For $n \in \mathbb{N}$, we have $d(x_n, x) = \lim_{m\to\infty} d(x_n, x_m) \leq \varepsilon_n$, which implies $x \in B(x_n, 2\varepsilon_n) \subseteq X \setminus M_n$, so for all $n \in \mathbb{N}$: $x \notin \bigcup_{n\in\mathbb{N}} M_n = X$, which is a contradiction. ■

Remark 4.2: (i) Another formulation of Theorem 4.1 is: *Let X be a complete metric space, $U_n \subseteq X$ dense and open subsets of X for $n \in \mathbb{N}$, then also $\bigcap_{n\in\mathbb{N}} U_n$ is dense in X.*

(ii) For a subset $M \subseteq X$ of a topological space, M is said to be

(1) *nowhere dense*, if $\overline{M}$ has no (nonempty) open subsets,

(2) *of first category*, if $M = \bigcup_{n\in\mathbb{N}} M_n$, with M_n nowhere dense for all n,

(3) *of second category*, in any other case.

Another way to formulate Baire's theorem then is: *Every complete metric space X is of second category in itself.*

Corollary 4.3: *There is no Banach space of countable (infinite) vector space dimension.*

This Corollary will be proven on an exercise sheet to come.

Theorem 4.4 (Principle of uniform boundedness):

(i) *nonlinear: Let X be a complete metric space and*

$$F \subseteq C_{\mathbb{R}}(X) = \{f \colon X \to \mathbb{R} \mid f \text{ continuous}\}$$

such that $\sup_{f \in F} f(x) < \infty \,\forall\, x \in X$. Then there exists an open ball $U \subseteq X$ such that

$$\sup_{x \in U} \sup_{f \in F} f(x) < \infty.$$

(ii) *linear: Let X be a Banach space and Y be a normed space and $A \subseteq B(X, Y)$, such that $\sup_{T \in A} \|Tx\| < \infty \,\forall\, x \in X$. Then we also have*

$$\sup_{T \in A} \|T\| < \infty.$$

Proof: (i) Let $X = \bigcup_{n=1}^{\infty} M_n$ for $M_n := \{x \in X \mid f(x) \leq n \,\forall\, f \in F\}$. Those M_n can be written as

$$M_n = \bigcap_{f \in F} \{x \in X \mid f(x) \leq n\} = \bigcap_{f \in F} f^{-1}([-\infty, n]),$$

and thus are closed. Then use Theorem 4.1.

(ii) Put $F := \{f_T \colon X \to \mathbb{R}, x \mapsto \|Tx\| \mid T \in A\} \subseteq C_{\mathbb{R}}(X)$. Via (i) we have: There exists $U(x_0, R)$ and a $K \geq 0$, such that $\|Tx\| \leq K \,\forall\, x \in U, T \in A$. Let now $x \in X$ with $\|x\| = 1$. Thenn

$$\|Tx\| = \frac{2}{R}\left(T\left(\frac{R}{2}x\right)\right) \leq \frac{2}{R}\left\|T\left(\frac{R}{2}x + x_0\right)\right\| + \frac{2}{R}\|Tx_0\| \leq \frac{4}{R}K \,\forall\, T \in A. \qquad \blacksquare$$

Corollary 4.5: *Let E be a normed space, $M \subseteq E$ a subset, such that $f(M)$ is bounded for all $f \in E'$. Then M is bounded (If M is bounded in every one-dimensional direction, then M itself is bounded).*

This Corollary again will be proven on an exercise sheet to come.

Corollary 4.6 (Theorem of Banach-Steinhaus): *Let X be a Banach space and Y be a normed space, $(T_n)_{n \in \mathbb{N}} \subseteq B(X, Y)$. If $(T_n)_{n \in \mathbb{N}}$ converges pointwise, then the limit T is in $B(X, Y)$. Therefore we have $Tx = \lim_{n \to \infty} T_n x$ and we have $\|T\| \leq \sup_{n \in \mathbb{N}} \|T_n\| < \infty$.*

Proof: That T is linear, is clear. It remains to be shown, that T is also bounded. For $x \in X$, we have hat $\|T_n x\| \leq M_x \,\forall\, n \in \mathbb{N}$. By Theorem 4.4, we therefore have $s = \sup_{n \in \mathbb{N}} \|T_n\| < \infty$. Let now $x \in X$ with $\|x\| = 1$ and $\varepsilon > 0$. Then

$$\|Tx\| \leq \|Tx - T_n x\| + \|T_n x\| < \varepsilon + s \quad \forall\, \varepsilon > 0,$$

so $\|Tx\| \leq s \,\forall\, x : \|x\| = 1$, hence $\|T\| \leq s$. $\blacksquare$

Another important consequence of Baire's theorem (Theorem 4.1) is the so called "Open mapping theorem": Let E, F be complete metric vector spaces and let $T \in B(E,F)$. If T is bijective, then there exists $T^{-1}\colon F \to E$ and it is linear – but is T^{-1} continuous? Do we have $T^{-1} \in B(F,E)$? This is especially intersting for $E = F$: Is $B(X)$ closed under taking inverses (with respect to the multiplication $TS := T \circ S$)? We have to show, that $(T^{-1})^{-1}(U) \subseteq F$ is open for all open $U \subseteq E$.

Definition 4.7: A mapping $T\colon E \to F$ between topological vector spaces is called *open*, if $TU \subseteq F$ is open for all open $U \subseteq E$.

Lemma 4.8: *Let E, F be complete metric vector spaces, $T\colon E \to F$ continuous, linear. If T is open, then T is surjective.*

Proof: As $E \subseteq E$ is open, we have that $TE \subseteq F$ is an open subvectorspace. So there exists $r > 0$, such that $B(0,r) \subseteq TE$. If now $x \in F$ is given, we have $\frac{1}{n}x \in B(0,r)$ for suitable $n \in \mathbb{N}$, which implies

$$x = n\left(\frac{1}{n}x\right) \in TE.$$

■

Theorem 4.9: *Let E, F be complete normed vector spaces (or a translation invariant metric, i. e., $d(x+z, y+z) = d(x,y)$) and $T\colon E \to F$ continuous, linear. If T is surjective, then T is open.*

Proof: For $\varepsilon > 0$, we define $E_\varepsilon := B(0,\varepsilon) \subseteq E$. Then

$$\bigcup_{n\in\mathbb{N}} nTE_\varepsilon = \bigcup_{n\in\mathbb{N}} T(nE_\varepsilon) = T\Big(\bigcup_{n\in\mathbb{N}} nE_\varepsilon\Big) = TE = F.$$

Especially, it holds: $F = \bigcup_{n\in\mathbb{N}} \overline{nTE_\varepsilon}$, by Theorem 4.1 there exists $n_0 \in \mathbb{N}$, such that $\overline{n_0 TE_\varepsilon}$ contains an open set U. This implies $\overline{T(\frac{1}{2}E_\varepsilon)}$ contains $\frac{1}{2n_0}U$. This implies

$$\overline{T\left(\frac{1}{2}E_\varepsilon - \frac{1}{2}E_\varepsilon\right)} = \overline{T\left(\frac{1}{2}E_\varepsilon\right)} - \overline{T\left(\frac{1}{2}E_\varepsilon\right)}$$

contains $\frac{1}{2n_0}U - \frac{1}{2n_0}U \ni 0$. Results so far are: It exists an open E_ε, $E_\varepsilon \mapsto TE_\varepsilon$, such that $\overline{T(E_\varepsilon)} \supset V \ni 0$ for an open $V \subseteq F$. It remains to be shown, that $V \subseteq T(E_{3\varepsilon})$.

In this case: Let $U \subseteq E$, $y \in T(E)$ and let $x \in E$ such that $T(x) = y$. This implies:

$$\exists\, \delta > 0 : x + \delta E_{3\varepsilon} \subseteq E \Rightarrow T(x + \delta E_{3\varepsilon}) \supseteq y + \delta V.$$

Let $V_i = 2^{-i}V$, i. e., $V_i \subseteq \overline{TE_{2^{-i}\varepsilon}}$. Let $y \in V$, i. e., $y \in \overline{T(E_\varepsilon)}$. Then there exists $x_0 \in E_\varepsilon$, such that $Tx_0 \in y + V_1$. So there exists $x_1 \in E_{2^{-1}\varepsilon}$ with $Tx_0 + Tx_1 \in y \in V_2$ – this holds, because $Tx_0 - y \in V_1 \subseteq \overline{TE_{2^{-1}\varepsilon}}$. Inductively, we get: $\exists\, x_n \in E_{2^{-n}\varepsilon}$

with $s_n = \sum_{i=0}^{n} Tx_i \in y + V_{n+1}$. This implies $s_n \to y$ for $n \to \infty$ and $t_n := \sum_{i=0}^{n} x_i$ is Cauchy, because for $n \leq m$:

$$d(t_n, t_m) \leq 2^{-(n+1)}\varepsilon + 2^{-(n+2)}\varepsilon + 2^{-n}\varepsilon \leq 2^{-n}\varepsilon.$$

Because E is complete, $t_n \to z \in \overline{E_{2\varepsilon}}$. Because T is continuous, this implies

$$y \overset{n\to\infty}{\longleftarrow} s_n = Tt_n \overset{n\to\infty}{\longrightarrow} Tz \in T(E_{3\varepsilon}).$$ ■

Corollary 4.10: *Let E, F be Banach spaces, $T\colon E \to F$ linear and continuous. Then T is surjective if and only if T is open.*

Proof: "⇐" is Lemma 4.8, "⇒" is Theorem 4.9. ■

Corollary 4.11: *Let E, F be Banach spaces, $T\colon E \to F$ be linear and continuous. If T is bijective, then T^{-1} is continuous. Hence: If $T \in B(E, F)$ is bijective, then $T^{-1} \in B(F, E)$.*

Proof: For all $U \subseteq E$ open, it holds $(T^{-1})^{-1}(U) = TU \subseteq F$ is open. ■

Corollary 4.12: *Let E, F be Banach spaces, $T\colon E \to F$ be continuous and linear and bijective. Then E and F are isomorphic as Banach spaces.*

Proof: Because T and T^{-1} are continuous, we have the estimation

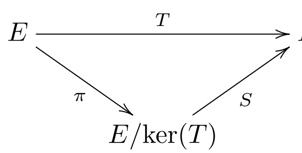
$$c_1\|x\|_E \leq \|Tx\|_F \leq c_2\|x\|_E.$$ ■

Example 4.13: Let E be a Banach space and let $M \subseteq E$ be a closed subspace. Then $E \to E/M$ is surjective, hence also open (see also Theorem 1.32). Let F be another Banach space, $T\colon E \to F$ be linear and continuous. Then

$$\ker(T) := \{x \in E \mid Tx = 0\} \subseteq E$$

is a closed subspace. Assume now, that T is surjective. Then, $E/\ker(T) \cong F$ as Banach spaces.

Proof: We have the diagram:



Set $S\dot{x} := Tx$. S is welldefined, because $\dot{x} = \dot{y}$, if $x - y \in \ker(T)$ which is the case, if $Tx = Ty$. That S is linear, is clear. For the continuity, we have

$$\|S\dot{x}\| = \|Tx\| = \inf\{\|Tz\| \mid z \sim x\} \leq \|T\|\|\dot{x}\|.$$

For the injectivity: If there are $\dot{x}, \dot{y} \in E/\ker(T)$ with $S\dot{x} = S\dot{y}$, then $Tx = Ty$, which implies that $x - y \in \ker(T)$, hence $\dot{x} = \dot{y}$. S is surjective, because T is surjective. Then use Corollary 4.12. ■

Definition 4.14: Let $T\colon E \to F$ be a map. Put

$$\mathrm{Graph}(T) := \{(x, Tx) \in E \times F\} \subseteq E \times F,$$

$\mathrm{Graph}(T)$ is called the *graph of* T.

Remark 4.15: $\mathrm{Graph}(T) \subseteq E \times F$ is closed (in the product topology) if and only if "$(x_n, Tx_n) \to (x, y) \Rightarrow Tx = y$" holds (i. e., $x_n \to x, Tx_n \to y$).

Theorem 4.16 (of the closed graph): *Let E, F be Banach spaces, $T\colon E \to F$ be linear. If $\mathrm{Graph}(T) \subseteq E \times F$ is closed, then T is continuous.*

Proof: The space $E \times F$ is again a normed space via $\|(x,y)\|_\infty := \max\{\|x\|, \|y\|\}$ or $\|(x,y)\|_1 := \|x\| + \|y\|$ or $\|(x,y)\|_2 := (\|x\|^2 + \|y\|^2)^{\frac{1}{2}}$. These norms are equivalent and describe the product topology on $E \times F$:

$$(x_n, y_n) \to (x, y) \Leftrightarrow x_n \to x, y_n \to y.$$

Hence $E \times F$ is a Banach space and $\mathrm{Graph}(T)$ is also a Banach space.

The maps

$$\pi_E\colon \mathrm{Graph}(T) \longrightarrow E \qquad \pi_F\colon \mathrm{Graph}(T) \longrightarrow F$$
$$(x, y) \longmapsto x \qquad\qquad (x, y) \longmapsto y$$

are linear and continuous. Moreover, π_E is even bijective, via Corollary 4.11 then also π_E^{-1} is also continuous. Then $T = \pi_F \circ \pi_E^{-1}$ is also continuous. ■

5 Hilbert spaces

In order to study linear maps $A\colon \mathbb{C}^n \to \mathbb{C}^n$ (i. e., $A \in M_n(\mathbb{C})$). The values $a_{i,j}$ are useful, with $A = (a_{i,j})$. We have an inner product on $\mathbb{C}^n$ via

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

for $x, y \in \mathbb{C}^n$ and using the canonical basis $\{e_1, \ldots, e_n\} \subseteq \mathbb{C}^n$ of $\mathbb{C}^n$, where $e_i = (\delta_{i,j})_{1\leq j\leq n} \in \mathbb{C}$, we have

$$\langle Ae_j, e_i\rangle = \left\langle \sum_{k=1}^{n} a_{k,j} e_k, e_i \right\rangle = \sum_{k=1}^{n} a_{k,j}\langle e_k, e_i\rangle = \sum_{k=1}^{n} a_{k,j}\delta_{k,j} = a_{i,j}.$$

Hence, for general linear maps A, it woud be nice to have an inner product and a “bases” (e_i) such that $\langle e_i, e_j\rangle = \delta_{i,j}$.

Definition 5.1: A map $\langle \cdot, \cdot\rangle\colon H \times H \to \mathbb{K}$ on a $\mathbb{K}$ vector space H is an *inner product* (*scalar product* appears in literature aswell), if it holds for all $x, y, z \in H$ and $\lambda \in \mathbb{K}$

(i) $\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle$,

(ii) $\overline{\langle x, y\rangle} = \langle y, x\rangle$,

(iii) $\langle x, x\rangle \geq 0$,

(iv) $\langle x, x\rangle = 0 \Rightarrow x = 0$.

A map, that satisfies (i)-(iii) (without (iv)), is called a *positive hermitian form.* A vector space H equipped with an inner product, is called *pre Hilbert space.*

Remark 5.2: We have

$$\langle z, \lambda x + \mu y\rangle \overset{\text{(ii)}}{=} \overline{\langle \lambda x + \mu y, z\rangle} \overset{\text{(i), (ii)}}{=} \overline{\lambda}\langle z, x\rangle + \overline{\mu}\langle z, y\rangle.$$

Hence, $\langle \cdot, \cdot\rangle$ is linear in the first component and anti-linear in the second (math convention as opposed to physicist convention).

Example 5.3: (i) $\mathbb{C}^n$ and $\mathbb{R}^n$ are pre Hilbert spaces via

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

(ii) $C([0, 1])$ is a pre Hilbert space via $\langle f, g\rangle := \int_0^1 f(t)\overline{g(t)}\, dt$,

(iii) $(L^2([0,1]), \lambda)$, where λ is the Lebesgue measure, or more general $(L^2(X), \mu)$ for a set X and a measure μ on a σ-Algebra $\mathfrak{A}$ on X, are pre Hilbert spaces via

$$\langle f, g\rangle = \int_X f(t)\overline{g(t)}\, d_\mu(t)\,.$$

For $X = \mathbb{N}$, $\mathfrak{A} = \mathfrak{P}(\mathbb{N})$ and $\mu = \zeta$, where ζ is the counting measure, $(L^2(\mathbb{N}), \zeta) = \ell^2$ with the scalar product

$$\langle (a_n), (b_n)\rangle = \sum_{n\in\mathbb{N}} a_n \bar{b}_n$$

Proposition 5.4: *Let H be a pre Hilbert space. Put $\|x\| := \sqrt{\langle x, x\rangle}$ for $x \in H$. Then it holds:*

(i) $\|x+y\|^2 = \|x\|^2 + 2\mathrm{Re}(\langle x, y\rangle) + \|y\|^2$,

(ii) *We have the* Cauchy-Schwarz-Inequality $|\langle x, y\rangle| \leq \|x\|\|y\|$. *We have equality if and only if x and y are linearly dependent,*

(iii) $\|\cdot\|$ *as defined in* (i) *is a norm,*

(iv) *For $y \in H$, $f_y(x) := \langle x, y\rangle$ is an element of the dual space H' such that $\|f_y\| = \|y\|$.*

Proof: (i) We calculate

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

which proves the claim.

(ii) Without loss of generality, we assume $y \neq 0$ (in case $y = 0$, $\langle x, y\rangle = \|y\| = 0$). For $\lambda \in \mathbb{K}$, we have

$$\langle x+\lambda y, x+\lambda y\rangle \overset{\text{(i)}}{=} \langle x, x\rangle + 2\mathrm{Re}\overline{\lambda}\langle x, y\rangle + |\lambda^2|\langle y, y\rangle.$$

With $\lambda := (-\langle x, y\rangle)(\langle y, y\rangle)^{-1}$, we have

$$\begin{aligned} 0 \leq \langle x+\lambda y, x+\lambda y\rangle &= \langle x, x\rangle - 2\mathrm{Re}\frac{\overline{\langle x, y\rangle}\langle x, y\rangle}{\langle y, y\rangle} + \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle} \\ &= \langle x, x\rangle - \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle} \end{aligned}$$

Moreover, $|\langle x, y\rangle|^2 = \langle x, x\rangle\langle y, y\rangle \Leftrightarrow 0 = \langle x+\lambda y, x+\lambda y\rangle \Leftrightarrow x = -\lambda y$.

(iii) It holds $\|x\| \geq 0$ by definition of the inner product, moreover it holds $\|\lambda x\| = |\lambda|\|x\|$ by definition aswell as $\|x\| = 0 \Rightarrow x = 0$.

For the triangular inequality, we compute

$$\|x+y\|^2 \overset{\text{(i)}}{=} \|x\|^2 + 2\mathrm{Re}\langle x, y\rangle + \|y\|^2 \leq (\|x\| + \|y\|)^2$$

(iv) f_y is linear. For the norm:

$$|f_y(x)| = |\langle x, y\rangle| \overset{\text{(ii)}}{\leq} \|x\|\|y\| \leq \|y\|,$$

so $f_y \in H'$. Since $f_y(y) = \langle y, y\rangle = \|y\|^2$, we have $\|f_y\| \geq \|y\|$. ■

Remark 5.5: (i) If $\langle\cdot,\cdot\rangle$ is just a positive hermetian form (rather than an inner procut), the Cauchy-Schwarz inequality still holds, $\|\cdot\|$ then only is a seminorm.

(ii) The map $x \mapsto \|x\|$ is continuous (see Remark 1.18) and $x \mapsto \langle x, y\rangle$ for fixed $y \in H$ is continuous as well (Proposition 5.4 (iv)), likewise $y \mapsto \langle x, y\rangle$ for fixed x.

Definition 5.6: A Hilbert space is a complete (with respect to $\|\cdot\|$ from Proposition 5.4 (iii)) pre Hilbert space.

Example 5.7: (i) $\mathbb{C}^n$ with $\langle x, y\rangle = \sum_{i=1}^n x_i \overline{y}_i$ is a Hilbert space (in fact, every finitedimensional pre Hilbert space is a Hilbert space),

(ii) $C([0,1])$ with $\langle f, g\rangle := \int_0^1 f(t)\overline{g(t)}\,dt$ is no Hilbert space.

(iii) $L^2(X, \mu)$ with $\langle f, g\rangle = \int_X f(t)\overline{g(t)}\,d_\mu(t)$ is a Hilbert space, in particular ℓ^2 is a Hilbert space. More generally

$$\ell^2(I) := \left\{(a_i)_{i\in I} : a_i \in \mathbb{C}\,\forall\, i \in I, \sum_{i\in I}|a_i|^2 < \infty\right\}$$

is a Hilbert space with $\langle (a_n), (b_n)\rangle = \sum_{i\in I} a_i \overline{b}_i$ for any index set I.

(iv) If H is a pre Hilbert space, then its completion (in the sense of chapter 1, using $\|\cdot\|$) $\widehat{H}$ is a Hilbert space with

$$\langle [(x_n)], [(y_n)]\rangle := \lim_{n\to\infty} \langle x_n, y_n\rangle.$$

Since $(\langle x_n, y_n\rangle)_{n\in\mathbb{N}}$ is a Cauchy sequence:

$$|\langle x_n y_n\rangle - \langle x_m, y_m\rangle| \leq \|x_n - x_m\|\|y_n\| + \|x_n\|\|y_n - y\| \to 0,$$

this inner product is well-defined and we have

$$\|[(x_n)]\| = \sqrt{\langle [(x_n)], [(x_n)]\rangle} = \lim_{n\to\infty} \sqrt{\langle x_n, x_n\rangle} = \lim_{n\to\infty} \|x_n\|.$$

The completion of $(C([0,1], \langle\cdot,\cdot\rangle)$ from Example 5.7 (ii) is $(L^2([0,1]), \lambda, \langle\cdot,\cdot\rangle)$.

(v) If $K \subseteq H$ is a closed subspace of a Hilbert space, then K is a Hilbert space, too.

Remark 5.8: If $x, y \neq 0$, then

$$\measuredangle(x,y) := \cos(\alpha) := \frac{|\langle x,y\rangle|}{\|x\|\|y\|}$$

defines an angle $\alpha \in [0, 2\pi]$ between x and y, with x is *orthogonal* to y, if $\langle x, y\rangle = 0$ We then write $x \perp y$.

Proposition 5.9: (i) *Let H be a pre Hilbert space. Then the* parallelogram identity *holds:*

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

(ii) *If H is a $\mathbb{C}$ vector space with a sesquiliniear form (for instance with an inner product), then*

$$\langle x,y\rangle = \frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^k \langle x + \mathrm{i}^k y, x + \mathrm{i}^k y\rangle.$$

If $\langle\cdot,\cdot\rangle$ is an inner product, we have $\langle x,y\rangle = \frac{1}{4}\sum_{k=0}^{3}\mathrm{i}^k\|x+\mathrm{i}^k y\|^2$.
If H is an $\mathbb{R}$ vector space with inner product, then

$$\langle x,y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2).$$

This identity is called the polarisation identity*.*

(iii) *If H is a normed space, it is a pre Hilbert space if and only if the parallelogram identity holds.*

Proof: (i) This is an exercise on Sheet 6.

(ii) It holds

$$\begin{aligned}\frac{1}{4}\sum_{k=0}^{3} \mathrm{i}^k \langle x + \mathrm{i}^k y, x + \mathrm{i}^k y\rangle &= \frac{1}{4}\sum_{k=0}^{3} \left(\mathrm{i}^k\langle x,x\rangle + \mathrm{i}^2\langle y,x\rangle + \langle x,y\rangle + \mathrm{i}^k\langle y,y\rangle\right)\\ &= \langle x,y\rangle.\end{aligned}$$

(iii) "$\Rightarrow$" is (i). "$\Leftarrow$": For $\mathbb{K} = \mathbb{C}$, we define

$$\langle x,y\rangle := \frac{1}{4}\sum_{k=0}^{3}\mathrm{i}^k\|x+\mathrm{i}^k y\|^2$$

and check, that this indeed is an inner product. ■

We wonder, if balls in Hilbert spaces round? What's the shape of balls in Hilbert spaces?

Let x be outside of such a ball B. Is there a unique element $x_0 \in B$, such that $\|x - x_0\| = \inf\{\|x-y\| \mid y \in B\}$?

If we take for example $H = \mathbb{R}^2$, then the unit balls with respect to the norms $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$ look like this:

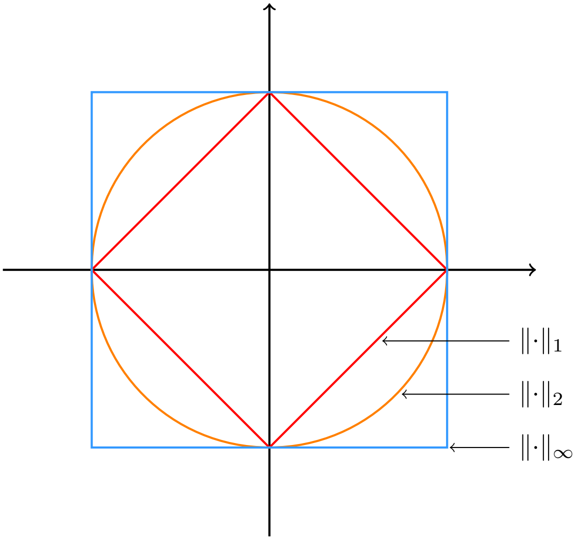


Figure 5.1: Unit balls in $\mathbb{R}^2$.

The answer to the above question is yes for $\|\cdot\|_2$, and no for $\|\cdot\|_1$, $\|\cdot\|_\infty$, e. g. for $x = (2, 0)$, $\|x, x_0\|_\infty \equiv 1 \,\forall\, x_0 \in \{(1, t) \mid t \in [-1, 1]\}$.

Theorem 5.10: *Let H be a Hilbert space, $A \subseteq H$ convex, closed and $x \in H \setminus A$. Then there is a unique element $x_0 \in A$, such that*

$$\|x - x_0\| = \inf\{\|x - y\| \mid y \in A\} =: \operatorname{dist}(x, A).$$

Proof: Put $d := \operatorname{dist}(x, A)$. Let $(y_n) \subseteq A$ be such that $\|y_n - x\| \to d$. Then (y_n) is Cauchy: If we put $z_n := y_n - x$, then

$$\begin{aligned} \|y_n - y_m\|^2 = \|z_n - z_m\|^2 &= 2(\|z_n\|^2 + \|z_m\|^2) - \|z_n + z_m\|^2 \\ &= 2(\|y_n - x\|^2 + \|z_m\|) - 4\|\frac{1}{2}(y_n + y_m) - x\|^2 \\ &\leq 4\varepsilon, \end{aligned}$$

for some $N \in \mathbb{N}$, because $\|y_n - x\| \leq d^2 + \varepsilon$ for $n, m \geq N$. Because A is closed, (y_n) converges to a point $x_0 \in A$, with $\|x - x_0\| = d$.

The uniqueness of x_0 remains to be shown. Let $x_0' \in A$ with $\|x - x_0'\| = d$. Then $(y_n) := (x_0, x_0', x_0, x_0', \dots)$ is a Cauchy sequence. ■

Remark 5.11: In a Hilbert space, the following holds:

$$\forall\, x_1, x_2 \in H \,\forall\, r_1, r_2 : r_1 + r_2 = d(x_1, x_2) : \exists!\, x_0 \in \overline{B(x_1, r_1)} \cap \overline{B(x_2, r_2)}.$$

Definition 5.12: Let H be a Hilbert space.

(i) $x, y \in H$ are said to be *orthogonal* (in signs: $x \perp y$), if $\langle x, y\rangle = 0$.

(ii) $M_1, M_2 \subseteq H$ are said to be *orthogonal* (in signs: $M_1 \perp M_2$), if $\langle x, y\rangle = 0 \,\forall\, x \in M_1, y \in M_2$.

(iii) Let $M \subseteq H$. The *orthogonal complement of* M is defined as

$$M^{\perp} := \{x \in H \mid x \perp y \,\forall\, y \in M\}$$

Remark 5.13: If $M \subseteq H$ is a subset of H, then $M^{\perp} \subseteq H$ is a closed linear subspace. Moreover: If $M \subseteq N \subseteq H$, then $M^{\perp} \supseteq N^{\perp}$. It holds $M^{\perp} = \overline{M}^{\perp}$. If M is a linear subspace, then $M^{\perp\perp} := (M^{\perp})^{\perp} = \overline{M}$.

Lemma 5.14 (Theorem of Pythagoras): *If H is a Pre Hilbert space, $x, y \in H$ with $x \perp y$. Then it holds:*

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

Proof: We easily calculate

$$\|x + y\|^2 = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle = \langle y, y\rangle = \|x\|^2 + \|y\|^2. \qquad \blacksquare$$

Theorem 5.15: *Let H be a Hilbert space and $K \subseteq H$ be a closed linear subspace. Furthermore let $x \in H$, $x_0 \in K$. Then $\|x - x_0\| = \operatorname{dist}(x, K)$ holds if and only if $x - x_0 \in K^{\perp}$.*

Proof: "$\Rightarrow$": Let $y \in K$, $\|y\| = 1$ and $z := x_0 - x$. We need to show, that $\langle z, y\rangle = 0$. Let $\alpha \in \mathbb{C}$. Then

$$\begin{aligned}\|z\|^2 = \operatorname{dist}(x, K)^2 \le \|x - (x_0 + \alpha y)\|^2 &= \|z - \alpha y\|^2 \\ &= \langle z, z\rangle - \alpha\langle y, z\rangle - \overline{\alpha}\langle z, y\rangle = \|\alpha\|^2\langle y, y\rangle,\end{aligned}$$

if we put $\alpha := \langle z, y\rangle$, we get $0 \le -\|\alpha\|^2$, hence $\alpha = 0$.

"$\Leftarrow$": Let $y \in K$. Then, with the theorem of Pythagoras, we get

$$\|x - y\|^2 = \|(x - x_0) + (x_0 - y)\|^2 = \|x - x_0\| + \|x_0 - y\|^2 \ge \|x - x_0\|^2. \qquad \blacksquare$$

Definition 5.16: Let $K_1, K_2 \subseteq H$ be two closed subspaces of a Hilbert space H, such that $K_1 \perp K_2$. We denote then

$$K_1 \oplus K_2 := \{x + y \mid x \in K_1, y \in K_2\} \subseteq H.$$

and call $K_1 \oplus K_2$ the *direct sum of K_1 and K_2*.

Lemma 5.17: *If $K_1, K_2 \subseteq H$ are closed subspaces of a Hilbert space, $K_1 \perp K_2$. Then $K_1 \cap K_2 = \{0\}$ and every element $x \in K_1 \oplus K_2$ has a unique decomposition of the form $x = x_1 + x_2$ with $x_1 \in K_1$, $x_2 \in K_2$.*

Proof: If $x \in K_1 \cap K_2$, then $x \perp x$, thus $\langle x, x\rangle = \|x\| = 0$, which implies $x = 0$.

Let $x = x_1' + x_2'$ be another such decomposition of x with $x_1' \in K_1$, $x_2' \in K_2$, then $x_1 - x_1' = x_2 - x_2' \in K_1 \cap K_2 = \{0\}$, thus the decomposition is unique. ■

Theorem 5.18: *Let $K \subseteq H$ be a closed linear subspace of a Hilbert space H. Then H decomposes as $H = K \oplus K^\perp$.*

Proof: $K^\perp \subseteq H$ is a closed subspace by Remark 5.13, and $K \perp K^\perp$ holds by definition. It remains to be shown, that $H \subseteq K \oplus K^\perp$. Let therefore $x \in H$. By Theorem 5.10, there exists $x_0 \in K$ with $\|x - x_0\| = \operatorname{dist}(x, K)$ and by Theorem 5.15, $x - x_0 \in K^\perp$. If we put $x_1 := x_0$, $x_2 := (x - x_0)$, we have $x = x_1 + x_2 \in K \oplus K^\perp$. ■

Remark 5.19: Such a decomposition theorem does not hold for general Banach spaces, since such "best" approximations Theorem 5.10 might not be unique. Decompositions $X = M \oplus N$ with $M \cap N = \{0\}$ and $M \times N = X$ of a Banach space X into Banach subspaces $M, N \subseteq X$ might neither exist, nor be unique. For the uniqueness: In a finitedimensional Banach space X, linear subspaces are automatically complete, linear hulls of subsets of bases of X give such subspaces. The Steinitz exchange lemma gives the non-uniqueness. For the non-existence: We may not find an N, such that $\ell^\infty = c_0 \oplus N$.

In fact: If X is a Banach space such that for each closed subspace $M \subseteq X$ we find a closed subspace $N \subseteq X$ with $X = M \oplus N$, then X is a Hilbert space.

Theorem 5.20 (Riesz representation theorem): *Let H be a Hilbert space. Then the map*

$$\begin{aligned} j\colon H &\longrightarrow H' \\ y &\longmapsto f_y = \langle \cdot, y\rangle \end{aligned}$$

is an antilinear, isometric isomorphism. Hence $H \cong H'$, in particular: H is reflexive. Thus for $f \in H'$ we find a $y \in H$, such that $f = f_y$.

Example 5.21: Let $g \in L^2(X, \mu)$. Put $f(h) := \int_X h\overline{g}\, d\mu$. Then f

$$\begin{aligned} f\colon L^2(X, \mu) &\longrightarrow \mathbb{C} \\ h &\longmapsto \int_X h\overline{g}\, d\mu \end{aligned}$$

is a linear functional. Are there more kinds of linear functinals? No! Let $f\colon L^2(X, \mu) \to \mathbb{C}$ be a linear functional. Then $f \in L^2(X, \mu)'$. By Theorem 5.20, there is a $g \in L^2$, such that $f = f_g$.

Proof (Theorem 5.20): First, we want to show the anti-linearity of j:

$$j(\lambda y_1 + \mu y_2)(x) = \langle x, \lambda y_1 + \mu y_2\rangle = \overline{\lambda}\langle x, y_1\rangle + \overline{\mu}\langle x, y_2\rangle = \overline{\lambda} j(y_1)(x) + \overline{\mu} j(y_2)(x),$$

i. e., j is anti-linear. Via Proposition 5.4, we have that

$$\|j(y)\| = \|f_y\| = \|y\|,$$

hence j is an isometry, and in particular injective. It remains to be shown, that j is indeed surjective. Let $0 \neq f \in H'$. Then $\ker(f) \subsetneq H$ is a closed subspace. By Theorem 5.18, $H = \ker(f) \oplus (\ker(f))^\perp$. Hence, we find $y \in (\ker(f))^\perp$, $y \neq 0$ and without loss of generality, we may assume $f(y) = 1$ (normalisation). Then $f(x)y - x \in \ker(f)$, hence

$$0 = \langle f(x)y - x, y\rangle = f(x)\|y\|^2 - \langle x, y\rangle \,\forall\, x \in H,$$

Put $z = \frac{y}{\|y\|^2} \in H$, then $f = f_z$:

$$f_z(x) = \frac{\langle x, y\rangle}{\|y\|^2} = f(x) \,\forall\, x \in H.$$

■

Remark: In the proof of Theorem 5.20, $\dim(\ker(f)^\perp) = 1$ holds.

Definition 5.22: A family $(x_i)_{i\in I}$ in a normed space X is called *summable* with value $s := \sum_{i\in I} x_i \in X$, if

$$\forall\, \varepsilon > 0 \,\exists\, J_0 \subseteq_{\text{fin}} I, |J_0| < \infty : \forall\, I \supseteq_{\text{fin}} J \supseteq J_0 : \Big\| \sum_{i\in J} x_i - s \Big\| < \varepsilon.$$

If I is countable, this is the usual notion of series.

Remark 5.23: (i) If $(x_i)_{i\in I}$ is summable, then only countably many elements may be non-zero.

Proof: For $\varepsilon = \frac{1}{n}$, we find $F_n \subseteq_{\text{fin}} I$, such that $\forall\, I \supseteq_{\text{fin}} J \supseteq F_n$:

$$\Big\| \sum_{i\in J} x_i - s \Big\| < \frac{1}{n}.$$

The union $F := \bigcup_{n\in\mathbb{N}} F_n$ then is countable, hence for $i \in I \setminus F$, we have

$$\|x_i\| = \Big\| \sum_{j\in F_n\cup\{i\}} x_j - \sum_{j\in F_n} x_j \Big\| \leq \Big\| \sum_{j\in F_n\cup\{i\}} x_j - s \Big\| + \Big\| s - \sum_{j\in F_n} x_j \Big\| < \frac{2}{n} \,\forall\, n,$$

i. e., $\|x_i\| = 0$. ■

(ii) We may formulate summability in the following way: A family $(x_i)_{i\in I}$ is summable to s if and only if the net $(s_F)_{F\subseteq_{\text{fin}} I}$ converges to s, where we put $s_F := \sum_{i\in F} x_i$.

(iii) We have the following rules: For summable families $(x_i)_{i\in I}$, $(y_i)_{i\in I}$, it holds:

$$\alpha\Big(\sum_{i\in I} x_i\Big) = \sum_{i\in I} \alpha x_i \quad \sum_{i\in I} x_i + \sum_{i\in I} y_i = \sum_{i\in I} x_i + y_i \quad \Big\langle \sum_{i\in I} x_i, y\Big\rangle = \sum_{i\in I} \langle x_i, y_i\rangle.$$

Proof: For example, $\alpha s \leftarrow \alpha \sum_{i\in I} x_i = \sum_{i\in I} \alpha x_i$, furthermore the inner product is continuous. ■

(iv) We have seen in Lemma 1.30, that in a Banach space every sequence $(x_i)_{i\in I}$ is summable if and only if $(\|x_i\|)_{i\in I}$ is summable. This can be shown for arbitrary families $(x_i)_{i\in I}$.

Lemma 5.24: *Let H be a Hilbert space and let $(x_i)_{i\in I}$ be a family of pairwise orthogonal elements. Then $(x_i)_{i\in I}$ is summable if and only if $(\|x_i\|^2)_{i\in I}$ is summable.*

Proof: Put $s_F := \sum_{i\in F} x_i$ and $t_F := \sum_{i\in F} \|x_i\|^2$ for $F \subseteq_{\text{fin}} I$. Hence, $\|s_F\|^2 = t_F$ due to Lemma 5.14.

"$\Rightarrow$": If $s_F \to s$, then $\|s_F\|^2 \to \|s\|^2$ by the continuity of $\langle\cdot,\cdot\rangle$.

"$\Leftarrow$": $(s_F)_{F\subseteq_{\text{fin}} I}$ is a Cauchy-net, i. e.,

$$\forall\, \varepsilon > 0\, \exists\, F_0 \subseteq_{\text{fin}} I\, \forall\, F, G \subseteq_{\text{fin}} I : F_0 \subseteq F, G : \|s_F - s_G\| < \varepsilon.$$

Indeed: Let $\varepsilon > 0$ and $F, G \subseteq_{\text{fin}} I$ such that $F_0 \subseteq F, G$. Then

$$\|s_F - s_G\|^2 = \Big\| \sum_{i\in F\cup G} \Big\|^2 = \|t_{F\cup G} - t_{F\cap G}\| < \varepsilon.$$

Now, choose $F_n \subseteq_{\text{fin}} I$ such that $\|s_F - s_G\| < \frac{1}{n}$ for $F, G \supseteq F_n$ and $F_n \subseteq F_{n+1}$. Then $(s_{F_n})_{n\in\mathbb{N}}$ is a Cauchy sequence. Hence, there is an $s \in H$ such that $s_{F_n} \to s$, i. e., for $F \supseteq F_n$:

$$\|s_F - s\| \leq \|s_F - s_{F_n}\| + \|s_{F_n} - s\| < \varepsilon + \varepsilon = 2\varepsilon$$ ■

Definition 5.25: A family $(e_i)_{i\in I}$ in a Hilbert space is called *orthonormal system*, if $\langle e_i, e_j\rangle = \delta_{i,j}$.

Lemma 5.26: *Let $(e_i)_{i\in I}$ be an orthonormal system in a Hilbert space H and let $x = \sum_{i\in I} \alpha_i e_i$ with $\alpha_i \in \mathbb{C}$. Then $\alpha_i = \langle x, e_i\rangle\ \forall\, i \in I$.*

Proof: Let $s_F := \sum_{i\in F} \alpha_i e_i$ for $F \subseteq_{\text{fin}} I$. Then $\langle s_F, e_k\rangle = \sum_{i\in F} \alpha_i \langle e_i, e_k\rangle = \alpha_k$, if $k \in F$. Hence

$$\langle x, e_k\rangle \longleftarrow \langle s_F, e_k\rangle = \alpha_k$$

for some $F \subseteq I$. ■

Theorem 5.27 (Bessel's inequality): *Let $(e_i)_{i\in I}$ be an orthonormal system in H. Then we have*

$$\sum_{i\in I}|\langle x, e_i\rangle|^2 \leq \|x\|^2 \quad \forall\, x \in H. \tag{5.1}$$

In particular, $(\sum_{i\in F}\langle x, e_i\rangle e_i)_{F\subseteq_{\text{fin}} I}$ converges. We have equality in Eq. (5.1) if and only if $x = \sum_{i\in I}\langle x, e_i\rangle e_i$.

Proof: We prove Theorem 5.27 only for $I = \mathbb{N}$ (for arbitrary index sets, work with techniques as in Lemma 5.26). Put $s_n := \sum_{i=1}^{n}\langle x, e_i\rangle e_i$. Then for $1 \leq k \leq n$, it holds

$$\langle s_n, e_k\rangle = \sum_{i=1}^{n}\langle x, e_i\rangle\delta_{i,k} = \langle x, e_k\rangle.$$

Hence $\langle x - s_n, e_k\rangle = 0\,\forall\, k \leq n$, i.e., $\langle x - s_n, s_n\rangle = 0$. By (Lemma 5.14), we then have

$$\|x\|^2 = \|x - s_n\|^2 + \|s_n\|^2 \geq \|s_n\| = \left\|\sum_{n=1}^{n}\langle x, e_i\rangle\right\| = \sum_{i=1}^{n}|\langle x, e_i\rangle|^2.$$

Therefore, $(\sum_{i=1}^{n}|\langle x, e_i\rangle|^2)_{n\in\mathbb{N}}$ converges and $\sum_{i\in\mathbb{N}}|\langle x, e_i\rangle|^2 \leq \|x\|^2$. Finally, $s_n \to x$ holds if and only if $\|s_n\|^2 \to \|x\|^2$. "$\Rightarrow$" is true due to the continuity of $\|\cdot\|$ and "$\Leftarrow$" is shown via $\|x\|^2 = \|x - s_n\|^2 + \|s_n\|^2 \to \|x\|^2$. ■

Theorem 5.28 (Parseval): *Let $(e_i)_{i\in I}$ be an othonormal system in H. Then the following are equivalent:*

(i) *$(e_i)_{i\in I}$ is a maximal orthonormal system (i.e., it is not included in a larger orthonormal system),*

(ii) *$x \perp e_i\,\forall\, i \in I \Leftrightarrow x = 0$,*

(iii) *$\forall\, x \in H : x = \sum_{i\in I}\langle x, e_i\rangle e_i$,*

(iv) *$\forall\, x \in H : \|x\|^2 = \sum_{i\in I}|\langle x, e_i\rangle|^2$,*

(v) *The linear span $\langle\sum_{i=1}^{N}\alpha_i e_i \mid \alpha_i \in \mathbb{C}, N \in \mathbb{N}\rangle$ is dense in H.*

If one (and thus all) of these conditions is satisfied by $(e_i)_{i\in I}$ is called an orthonormal basis.

Proof: "(i) $\Rightarrow$ (ii)": Let $x \neq 0$, $x \perp e_i\,\forall\, i \in I$, then $(e_i)_{i\in I} \cup \{\frac{x}{\|x\|}\}$ is a larger orthonormal system.

"(ii) $\Rightarrow$ (i)": If $(e_i)_{i\in I} \subsetneq (e_i)_{i\in I'}$, then there is $i_i \in I' \setminus I$: $e_{i_0} \perp e_i\,\forall\, i \in I$.

"(iii) $\Leftrightarrow$ (iv)": This is Theorem 5.27.

"(iii) $\Rightarrow$ (ii)": If $\langle x, e_i\rangle = 0\,\forall\, i \in I$, then $x = \sum_{i\in I}\langle x, e_i\rangle e_i = 0$.

"(ii) $\Rightarrow$ (iii)": By Theorem 5.27, the series $\sum_{i\in I}\langle x, e_i\rangle e_i$ converges. Put

$$z := x - \sum_{i\in I}\langle x, e_i\rangle e_i.$$

Then

$$\langle z, e_j \rangle = \langle x, e_j \rangle - \sum_{i \in I} \langle x, e_j \rangle \langle e_i, e_j \rangle = 0 \quad \forall j \in I,$$

but then $z = 0$.

"(iii) $\Rightarrow$ (v)": This is obvious.

"(v) $\Rightarrow$ (iii)": Let $x \perp e_i \, \forall i \in I$. Choose $x_n = \sum_{j=1}^N \alpha_{j,n} e_j \to X$ as $n \to \infty$. Then

$$\|x\|^2 \leftarrow \|x_n\|^2 = \langle x_n, x_n \rangle = |\langle x_n, x_n - x \rangle| \leq \|x_n\| \|x_n - x\| \to 0. \qquad \blacksquare$$

Remark 5.29: If $(e_i)_{i \in I}$ is an orthonormal basis, then the elements e_i are linear independent: For a finite linear combination of these e_i we have:

$$\sum_{i=1}^N \alpha_i e_i = 0 \Rightarrow 0 = \left\langle \sum_{i=1}^N \alpha_i e_i, e_j \right\rangle = \alpha_j \quad \forall j.$$

But an orthonormal base is no vector space base, since we might need infinite linear combination as opposed to finite linear combinations.

Theorem 5.30: *Let H be a Hilbert space, $(e_i)_{i \in I}$ and $(f_j)_{j \in J}$ two orthonormal bases. Then $|I| = |J|$ (i. e., an orthonormal base is not unique, but its cardinaility).*

Proof: If $|I| < \infty$, via Theorem 5.28 $|J| < \infty$ and as we know from linear algebra, we then know $|I| = |J|$.

If $|I|, |J| = \infty$, then $\varnothing \neq I_j := \{i \in I \mid \langle e_i, f_j \rangle \neq 0\}$ is countable via Remark 5.23 (i). Furthermore, $I = \bigcup_{j \in J} I_j$, hence $|I| \leq |J|$. Because the same argument works for J, then $|J| \leq |I|$. The Theorem of Schroeder-Bernstein-Cantor now ensures $|I| = |J|$. $\blacksquare$

Theorem 5.31: *Every Hilbert space H admits an orthonormal base.*

Proof: Let $(e_i)_{i \in I}$ be an orthonormal system in H. The set of orthonormal systems containing $(e_i)_{i \in I}$ is inductively ordered, hence by the Lemma of Zorn, there is a maximal orthonormal system in H. $\blacksquare$

Definition 5.32: Let H be a Hilbert space. The *(Hilbert space) dimension* $\dim H$ is defined as the cardinality of an orthonormal base of H. If $\dim H$ is countable, we call H *separable*.

Remark 5.33: Using the algorithm of *Gram-Schmidt*, we may show that a Hilbert space is separable if and only if it is separable as a Banach space (i. e., it contains a countable dense subset).

Proof: "⇒": Let $(e_i)_{i\in I}$ be a countable orthonormal base of H. Then

$$\Big\{\sum_{j=1}^{N} x_i e_i : N\in\mathbb{N}, x_j \in \mathbb{Q}+\mathrm{i}\mathbb{Q}\Big\} \subseteq H$$

is dense and countable.

"⇐": Let $E\subseteq H$ be dense. Choose a sequence $x_1, x_2, \ldots$ of linearly independent vectors, such that $\{x_i : i\in I\}$ is maximal linearly independent. Then the vectors defined by

$$e_1 := \frac{1}{\|x_1\|}x_1 \quad , \quad e_{n+1} := \frac{1}{\|x_{n+1}-\sum_{i=1}^{n}\langle x_i, e_i\rangle e_i\|}\Big(x_{n+1}-\sum_{i=1}^{n}\langle x_i, e_i\rangle e_i\Big)$$

where $e_{n+1}\neq 0$, because the x_i were assumed to be linearly independent, form an orthonormal base of H. ■

Example 5.34: (i) Let $H=\mathbb{C}^n$. Then $\dim\mathbb{C}^n = n$. Note that in this case, the vector space dimension and the Hilbert space dimension coincide.

(ii) Let $H=\ell^2$. Then $\dim\ell^2=\infty$ and ℓ^2 is separable with orthonormal base $(e_i)_{i\in\mathbb{N}}$, where $e_i := (\delta_{i,j})_{j\in\mathbb{N}}$.

Definition 5.35: Let H, K be Hilbert spaces. An *isomorphism between H and K* is a linear map

$$U\colon H\longrightarrow K$$

which is surjective and that satisfies

$$\langle Ux, Uy\rangle = \langle x, y\rangle \quad \forall\, x, y\in H.$$

Remark 5.36: An isomorphism $U\colon H\to K$ is injective, more precisely it is an isometry since for all $x\in H$:

$$\|Ux\| = \langle Ux, Ux\rangle = \langle x, x\rangle = \|x\|.$$

Therefore, U preserves the whole Hilbert space structure.

If $\dim H=\dim K<\infty$, surjectivity is granted because linear maps between vector spaces of the same finite dimension are surjective if and only if they are injective. However if $\dim H = \dim K = \infty$, we really need the surjectivity.

Theorem 5.37: *Two Hilbert spaces H, K are isomorphic if and only they have the same dimension.*

Remark 5.38: If H is a separable complex Hilbert space, then H is isomorphic to $\mathbb{C}^n$ or ℓ^2. Hence up to isomorphisms, there is only one infinitedimensional complex separable Hilbert space. In fact $L^2[0,1]\cong\ell^2 = L^2(\mathbb{N})$.

More general, if H has an orthonormal base $(e_i)_{i\in I}$ with an arbitrary index set I then $H\cong\ell^2(I)$.

Proof: "$\Rightarrow$": Let $U\colon H \to K$ be an isomorphism and let $(e_i)_{i\in I}$ be an orthonormal base of H. Then $(Ue_i)_{i\in I}$ forms an orthonormal system in K. If now $y \perp Ue_i$ for all $i \in I$, then there is $x \in H$, such that $Ux = y$, hence

$$\langle x, e_i\rangle = \langle Ux, Ue_i\rangle = 0 \,\forall\, i \in I,$$

so via the Parseval equality (Theorem 5.27), $x = 0$. Then, because U is linear, $y = 0$. By Theorem 5.28, $(Ue_i)_{i\in I}$ is an orthonormal base of K.

"$\Leftarrow$": Let $(e_i)_{i\in I}$ be an orthonormal base of H and $(f_i)_{i\in I}$ be an orthonormal base of K. Put $Ue_i := f_i$. This defines a linear isometry $U\colon H \to K$. ■

Example 5.39: (i) The Hilbert space $\mathbb{C}^n$ (with the standard inner product) has the orthonormal base $\{e_1, \ldots, e_n\}$, where $e_i = (\delta_{i,j})_{1\le j\le n}$.

(ii) The Hilbert space ℓ^2 has the orthonormal base $(e_i)_{i\in I}$ with $e_i = (\delta_{i,j})_{j\in\mathbb{N}}$. This is *no* vector space base of ℓ^2!

(iii) For an orthonormal base of $L^2([0,1])$ we will need the Theorem of Stone-Weierstraß.

6 Theorem of Stone-Weierstraß

We can approximate continuous functions $f \in C([0,1])$ with simpler functions, namely polynomials, with respect to $\|\cdot\|_\infty$.

For the proof of Stone-Weierstraß' theorem, we only need little very algebraic properties of $\{\text{Polynomials with complex coefficients}\} = \mathbb{C}[X] \subseteq C([0,1])$. Let K be a compact metric space. Consider

$$C(K) := \{f\colon K \to \mathbb{C} \mid f \text{ continuous}\}$$

endowed with $\|f\|_\infty := \sup_{x\in K}|f(x)|$. Then $(C(K), \|\cdot\|_\infty)$ is a Banach space.

Definition 6.1: A subset $A \subseteq C(K)$ is called a *-*subalgebra with unit* (or *unital* *-*subalgebra*), if

(i) $f, g \in A \Rightarrow fg \in A$,

(ii) $f, g \in A, \mu, \lambda \in \mathbb{C} \Rightarrow \mu f + \lambda g \in A$,

(iii) $f \in A \Rightarrow \overline{f} \in A$,

(iv) $1 \in A$.

A is said to *separate points*, if for all $s, t \in K, s \neq t$, there is $f \in A$, such that $f(s) \neq f(t)$.

Example 6.2: The set

$$P := \{\text{Polynomials in } X \text{ and } \overline{X}\} \subseteq C([0,1])$$

is a unital *-subalgebra seperating points. The same holds for

$$P_{\mathbb{R}} := \{\text{Polynomials in } \mathbb{R}\} \subseteq C_{\mathbb{R}}([0,1]) = \{f\colon [0,1] \to \mathbb{R} \mid f \text{ continuous}\}.$$

Theorem 6.3 (Stone-Weierstraß): *Let K be a compact space and $A \subseteq C(K)$ be a unital *-subalgebra, separating points. Then $A \subseteq C(K)$ dense (with respect to $\|\cdot\|_\infty$). In particular: If A is closed, then $A = C(K)$ holds.*

The same holds true for the real case, i. e., for a unital *-subalgebra separating points $A \subseteq C_{\mathbb{R}}(K)$.

Proof: Let $A \subseteq C([0,1])$ be a closed unital *-subalgebra separating points.

① *If $f \in A, f \geq 0$ then $\sqrt{f} \in A$.*

Proof (of ①): Without loss of generality let $0 \leq f \leq 1$ (normalisation!). Put $g := 1 - f$, then $0 \leq g \leq 1$ aswell. It holds

$$\sqrt{f(t)} = \sqrt{1 - g(t)} = 1 - \sum_{n=1}^{\infty} a_n g^n(t) \quad \forall\, t \in K$$

via the Taylor series expansion of g. This taylor series expansion converges uniformly on $[-1,1]$, hence

$$A \ni h_m := \sum_{n=1}^{m} a_n g^n(x) \to \sqrt{f} \quad (\text{with respect to } \|\cdot\|_\infty).$$

Because A is closed via assumption, we then have that $\sqrt{f} \in A$. ■

② *Let $f, g \in A$ be real-valued. Then* $\min\{f,g\}, \max\{f,g\} \in A$.

Proof (of ②): For $f, g \in A$ we have the representations

$$\min\{f,g\} = \frac{f+g-|f-g|}{2} \quad , \quad \max\{f,g\} = \frac{f+g+|f-g|}{2}.$$

Both of those are elements of A, because $|f| = \sqrt{f\overline{f}}$ for any $f \in A$. ■

③ *For a real-valued $f \in C(K)$ and $\varepsilon > 0$, there is $g \in A : \|f-g\|_\infty < \varepsilon$.*

Proof (of ③): For $s, t \in K$, $s \neq t$, there is $f_{s,t} \in A$ such that $f_{s,t}(s) = f(s)$, $f_{s,t}(t) = f(t)$: Since A separates points, there is $h \in A$ with $h(s) \neq h(t)$. Put

$$f_{s,t}(x) := f(t) + (f(s) - f(t))\frac{h(x)-h(t)}{h(s)-h(t)}.$$

Now put

$$U_t := \{x \in K \mid f_{s,t}(x) < f(x) + \varepsilon\}.$$

Due to the openness of $(f_{s,t} - f)^{-1}(\varepsilon, \infty)$, those U_t are open, furthermore $t \in U_t$ holds. Thus the family $(U_t)_{t\in K}$ is an open cover of K. Because K is compact, there are $t_1, \dots, t_n$, such that $K = \bigcup_{i=1}^n U_{t_i}$. Put $h_s := \min_{1\leq i\leq n} f_{s,t_i} \in A$ and put

$$V_s := \{x \in K \mid h_s(x) > f(x) - \varepsilon\}.$$

Again, the V_s are open and $(V_s)_{s\in K}$ forms an open cover of K, thus there are $s_1, \dots, s_m \in K$ with $K = \bigcup_{j=1}^m V_{s_j}$. Now put $g := \max_{1\leq j\leq m} h_{s_j} \in A$. It holds $h_{s_j} < f + \varepsilon$ for all j, so $g < f + \varepsilon$ and $g > f - \varepsilon$, i. e., $\|f-g\|_\infty < \varepsilon$. ■

Let now $f \in C(K)$ be an arbitrary continuous function. Via ③, there exist sequences $(g_n)_{n\in\mathbb{N}}, (h_n)_{n\in\mathbb{N}} \subseteq A$ such that $g_n \to \mathrm{Re}(f)$, $h_n \to \mathrm{Im}(f)$, i. e.,

$$g_n + \mathrm{i}h_n \to f \quad (\text{with respect to } \|\cdot\|_\infty)$$

and therefore, $A \subseteq C(K)$ is dense. ■

Corollary 6.4 (Theorem of Weierstraß): *The algebra P of polynomials is dense in $C_\mathbb{R}([0,1])$.*

Remark 6.5: The space $(P, \|\cdot\|_\infty)$ is no Banach space (refer to Sheet 5, Exercise 2). The completion is $(C([0,1]), \|\cdot\|_\infty)$.

Corollary 6.6: (i) *The set of polynomials* $\sum_{n=-N}^{N} a_n z^n$ *for* $a_n \in \mathbb{C}$ *is dense in* $C(\mathbb{S}^1)$*, where* $z^{-n} := \overline{z}^n$ *for* $n > 0$ *and*

$$\mathbb{S}^1 := \{z \in \mathbb{C} \mid |z| = 1\}.$$

(ii) *The set of functions* $e_n(t) := (\sqrt{2\pi})^{-1} e^{\mathrm{i}nt}$ *for* $t \in [0, 2\pi]$ *and* $n \in \mathbb{Z}$ *is an orthonormal base of* $(L^2([0,2\pi]), \lambda)$*.*

(iii) *It holds* $L^2([0,2\pi]) \cong \ell^2$ *as a Hilbert space, likewise* $L^2([0,1]) \cong \ell^2$*.*

Proof: (i) The set

$$\left\{ \sum_{n=-N}^{N} a_n z^n : a_n \in \mathbb{C}, N \in \mathbb{N} \right\} \subseteq C(\mathbb{S}^1)$$

is a unital *-subalgebra, seperating points ($z^n \cdot z^m = z^{n+m}$ for $n, m \in \mathbb{Z}$). Via Theorem 6.3, this *-subalgebra is dense.

(ii) Put

$$C_{\text{per}} := \{f \colon [0, 2\pi] \to \mathbb{C} \mid f(0) = f(2\pi)\}$$

and consider the mapping

$$\Phi \colon C_{\text{per}}[0, 2\pi] \longrightarrow C(\mathbb{S}^1)$$
$$\Phi(f)(e^{\mathrm{i}t}) := f(t).$$

Then Φ is isometric, surjective and $\Phi(\sqrt{2\pi} e_n)(z) = z^n$, as

$$\Phi(\sqrt{2\pi} e_n)(e^{\mathrm{i}t}) = \sqrt{2\pi} e_n(t) = e^{\mathrm{i}nt} = (e^{\mathrm{i}t})^n$$

with $z = e^{\mathrm{i}t}$. Since Φ is linear, linear combinations of e_n are mapped to polynomials $\sum_n a_n z^n$, therefore the linear combinations of e_n are dense in $C_{\text{per}}[0, 2\pi]$ with respect to $\|\cdot\|_\infty$. Since

$$\|f\|_2^2 = \int_0^{2\pi} |f(t)|^2 \, dt \leq \int_0^{2\pi} \|f\|_\infty \, dt \leq 2\pi \|f\|_\infty,$$

thus the linear combinations of the e_n are dense in $C_{\text{per}}[0, 2\pi]$ with respect to $\|\cdot\|_2$. We have $C_{\text{per}}[0, 2\pi] \subseteq C[0, 2\pi] \subseteq L^2[0, 2\pi]$ dense with respect to $\|\cdot\|_2$. It remains to be shown, that $(e_n)_{n\in\mathbb{Z}}$ are indeed an orthonormal system with respect to $\langle \cdot, \cdot \rangle$ in $L^2[0, 2]$, as it then is an orthonormal base via Theorem 5.28 (v).

It holds

$$\langle e_n, e_m \rangle = \int_0^{2\pi} e_n(t) \overline{e_n(t)} \, dt = \frac{1}{2\pi} \int_0^{2\pi} e^{\mathrm{i}(n-m)t} \, dt = \delta_{n,m},$$

so $(e_n)_{n\in\mathbb{Z}}$ is indeed an orthonormal base.

(iii) $(e_n)_{n\in\mathbb{Z}}$ is countable. ■

Remark 6.7: If $f \in C[0, 2\pi]$, then $\sum_{n\in\mathbb{Z}} c_n e^{int}$ is nothing but its Fourier series. Hence the statement about $(e_n)_{n\in\mathbb{Z}}$ being an orthonormal base is just Fourier Analysis. But note: The approximation is only with respect to $\|\cdot\|_2$ rather than pointwise or uniformal approximations.

7 Bounded operators on Hilbert spaces

Reminder 7.1: Let H be a Hilbert space. Then

$$B(H) := \{A\colon H \to H \text{ linear, bounded}\}$$

is a normed vector space with the operator norm (refer to Definition 1.22, Proposition 1.23 and Theorem 1.26). This vector space is even a Banach space, because H is a Banach space and furthermore an algebra via $ST := S \circ T$. Finally, we have $\|ST\| \leq \|S\|\|T\|$.

Example 7.2: (i) Let $H = \mathbb{C}^n$ and $e_1, \dots, e_n$ be the canonical base of H, then $A \in B(\mathbb{C}^n)$ is uniquely determined by $(a_{i,j})_{1\leq i,j\leq n}$ with $Ae_i = \sum_{j=1}^n a_{j,i}e_j$. Hence $B(\mathbb{C}^n) = M_n(\mathbb{C})$, the complex valued $n \times n$-matrices.

(ii) Let $H = L^2([0,1], \lambda)$ and $k\colon [0,1] \times [0,1] \to \mathbb{C}$ continuous (or more general: $k \in L^2([0,1] \times [0,1], \lambda^2)$, i. e.,

$$\int_0^1 \int_0^1 |k(s,t)|^2 \, ds \, dt < \infty.$$

Then $K\colon L^2[0,1] \to L^2[0,1]$, $(Kf)(s) := \int_0^1 k(s,t)f(t)\, dt$ is a bounded linear operator $K \in B(L^2[0,1])$, the *integral operator with kernel k*. K is bounded, as

$$\begin{aligned} \|Kf\|^2 = \int_0^1 |(Kf)(s)|^2 \, ds &= \int_0^1 \left| \int_0^1 k(s,t)f(t)\, dt \right|^2 ds \\ &= |\langle k(s,\cdot), \overline{f}\rangle|^2 \\ &\leq \int_0^1 |k(s,t)|^2 \, dt \|f\|^2, \end{aligned}$$

therefore $\|Kf\|^2 \leq \int_0^1 \int_0^1 |k(s,t)|^2 \, ds \, dt \|f\|_2^2 = \|k\|_2^2 \|f\|_2^2$, therefore $\|K\| \leq \|k\|_2^2$ or, if k is continuous: $\|K\| \leq \|k\|_\infty$.

Our idea is: K has a "continuous matrix" $(k(s,t))_{s,t\in[0,1]}$. Indeed, the above calculation works for any measure space $L^2(X,\mu)$, too, hence with $X = \{1, \dots, n\}$, $\mu(\{t\}) = 1$ and

$$e_i := \begin{cases} 1 & t = i, \\ 0 & \text{else,} \end{cases}$$

we have

$$Ke_i(s) = \sum_{t=1}^n k(s,t)e_i(t) = k(s,i) = \sum_{j=1}^n k(j,i)e_j(s).$$

Reminder 7.3: Let E, F be normed spaces and $T \in B(E,F)$. Then there is exactly one map $T' \in B(F', E')$ such that $\langle x, Ty\rangle_{F' \times F} = \langle T'x, y\rangle_{E'\times F}\, \forall\, x \in F', y \in E$ and $\|T'\| = \|T\|$ (3.6).

If E, F are Hilbert spaces, then $E' \cong E$, $F' \cong F$ (Theorem 5.20). So, how does T' look like, seen as an operator in $B(F,E)$?

Proposition 7.4: *Let H, K be Hilbert spaces and $A \in B(H,K)$. Then there is a unique operator $A^* \in B(K,H)$ with $\langle Ax, y\rangle = \langle x, A^*y\rangle\, \forall\, x \in H, y \in K$.*

Proof: Consider

$$\begin{aligned} f\colon H &\longrightarrow \mathbb{C} \\ x &\longmapsto \langle Ax, y\rangle. \end{aligned}$$

Then $f \in H'$, since f is linear and bounded:

$$|f(x)| = |\langle Ax, y\rangle| \leq \|Ax\|\|y\| \leq \|A\|\|x\|\|y\|,$$

so $\|f\| \leq \|A\|\|y\|$. By Theorem 5.20, we find $z_y \in H$ such that $f = f_{z_y}$, i. e., $\langle Ax, y\rangle = \langle x, z_y\rangle$.

Put $A^*y := z_y$, then A^* is linear, because

$$\langle x, A^*(\lambda y_1 + \mu y_2)\rangle = \langle Ax, \lambda y_1 + \mu y_2\rangle = \overline{\lambda}\langle Ax, y_1\rangle + \overline{\mu}\langle Ax, y_2\rangle = \langle x, \lambda A^* y_1 + \mu A^* y_2\rangle,$$

and A^* is bounded, because $\|A^*y\| = \|z_y\| = \|f_{z_y}\| = \|f\| \leq \|A\|\|y\|$, therefore $\|A^*\| \leq \|A\|$. ■

Remark 7.5: A^* is basically $A'\colon K' \to H'$ and $\|A^*\| = \|A\|$.

Proof: The diagramm

$$\begin{array}{ccc} K & \xrightarrow{\quad A^* \quad} & H \\ {\scriptstyle j_K}\downarrow & & \downarrow{\scriptstyle j_H} \\ K' & \xrightarrow{\quad A' \quad} & H' \end{array}$$

commutes, because

$$(A' j_K x)(y) = (A' f_x)(y) = f_x(Ay) = \langle Ay, x\rangle = \langle y, A^*x\rangle = f_{A^*x}(y) = (j_H A^* x)(y),$$

so $A' \circ j_K = j_H \circ A^*$ and $\|A*\| = \|j_H^{-1} \circ A' \circ j_K\| = \|A'\| = \|A\|$, using the results Theorem 5.20 and Theorem 3.6. ■

Proposition 7.6: *The map* $^*\colon B(H) \to B(H)$ *is*

(i) *antilinear, i. e.,* $(\mu A + \nu B)^* = \overline{\mu} A^* + \overline{\nu} B^*$, $\mu, \nu \in \mathbb{C}$,

(ii) *isometric, i. e.,* $\|A^*\| = \|A\|$,

(iii) *involutoric, i. e.,* $A^{**} = A$,

(iv) *it satisfies* $(AB)^* = B^*A^*$,

(v) *and the* C^**-condition:* $\|A^*A\| = \|A\|^2$,

(vi) *If* $A \in B(H)$ *is invertible, then also* A^* *is invertible and* $(A^*)^{-1} = (A^{-1})^*$.

Proof: We notice, that "$\langle Ax, y\rangle = 0\,\forall\, x, y \in H \Rightarrow A = 0$" (indeed: "$\langle Ax, Ax\rangle = 0 \Rightarrow Ax = 0\,\forall\, x \in H$"), hence: If

$$\langle Ax, y\rangle = \langle Bx, y\rangle\,\forall\, x, y \in H,$$

then $A = B$ (because then $\langle (A - B)x, y\rangle = 0\,\forall\, x, y$). We may thus check (i), (iii), (iv) directly:

$$\langle A^{**}x, y\rangle = \langle x, A^*y\rangle = \langle Ax, y\rangle\,\forall\, x, y \in H \Rightarrow A = A^*,$$
$$\langle (AB)^*x, y\rangle = \langle x, ABy\rangle = \langle A^*x, By\rangle = \langle B^*A^*x, y\rangle.$$

Ad (v): It holds $\|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^*Ax, x\rangle = \|A^*A\|\|x\|$, hence $\|A\|^2 \leq \|A^*A\|$, but $\|A^*A\| \leq \|A^*\|\|A\| = \|A\|^2$. (ii) was already shown in Remark 7.5. ■

Example 7.7: (i) Let $H = \mathbb{C}^n$ and $A = (a_{i,j})_{1\leq i,j\leq n} \in M_n(\mathbb{C}) = B(H)$. Then $A^* = (\overline{a}_{i,j})_{1\leq i,j\leq n}$, since

$$\langle A^*e_j, e_i\rangle = \langle e_j, Ae_i\rangle = \left\langle e_j, \sum_{k=1}^{n} a_{k,i}e_k \right\rangle = \sum_{k=1}^{n} \overline{a}_{k,i}\delta_{j,k} = \overline{a}_{j,i}.$$

(ii) Let $H = L^2[0, 1]$ and K as in Example 7.2. Then K^* is the integral operator with kernel $k^*(s, t) := \overline{k(s, t)}$, since

$$\begin{aligned}\langle K^*f, g\rangle = \langle f, Kg\rangle &= \int_0^1 f(t)\overline{K_g(t)}\, dt \\ &= \int_0^1 \int_0^1 f(t)\overline{k(t, s)}\, dt\, ds = \int_0^1 (K_{k^*}f)(s)\overline{g(s)}\, ds = \langle K_{k^*}f, g\rangle\end{aligned}$$

Proposition 7.8: *If* $A \in B(H)$, *then* $\ker(A) = \operatorname{im}(A^*)^\perp$, $\ker(A)^\perp = \overline{\operatorname{im}(A^*)}$.

Proof: If $x \in \operatorname{im}(A^*)^\perp$, then for all $y \in H : \langle x, A^*y\rangle = 0 = \langle Ax, y\rangle$ which holds if and only if $Ax = 0$, i. e., $x \in \ker(A)$. Furthermore $\ker(A)^\perp = \operatorname{im}(A^*)^{\perp\perp} = \overline{\operatorname{im}(A^*)}$ via Sheet 6. ■

Definition 7.9: (i) $A \in B(H)$ is called *selfadjoint* (or *hermitian*) if $A = A^*$,

(ii) $A \in B(H)$ is called *normal* if $A^*A = AA^*$,

(iii) $U \in B(H)$ is called *unitary*, if $U^*U = UU^* = 1 := \mathrm{id}_H \in B(H)$,

(iv) $V \in B(H)$ is called *isometry*, if $V^*V = 1$,

(v) $P \in B(H)$ is called *(orthogonal) projection*, if $P = P^* = P^2$.

Remark 7.10: (i) It holds $V^*V = 1$ if and only if

$$\langle Vx, Vy \rangle = \langle V^*Vx, y \rangle = \langle x, y \rangle \,\forall\, x, y \in H,$$

i. e., V is isometric in the previous sense.

(ii) U is unitary if and only if U is isometric and surjective, i. e., an isomorphism of Hilbert spaces as defined in Definition 5.35.

(iii) If P is a projection in the sense of Definition 7.9 (v), there is a closed subspace $K \subseteq H$, such that $H = K \oplus K^\perp$ and $P(x + y) = x$ for $x + y \in K \oplus K^\perp$. Conversely, if $K \subseteq H$ is a closed subspace with $P(x + y) = x$, then $P = P^* = P^2$ (see Sheet 7).

Remark 7.11: If $A = A^*$, then $\langle Ax, x \rangle \in \mathbb{R} \,\forall\, x \in H$:

$$\overline{\langle Ax, x \rangle} = \langle x, Ax \rangle = \langle A^*x, x \rangle = \langle Ax, x \rangle$$

and $\|A\| = \sup_{\|x\|=1} \{|\langle Ax, x \rangle|\}$.

8 Spectral values of bounded operators

As motivation for this chapter: Let $A \in M_n(\mathbb{C})$ with $A = A^*$. Then A is equivalent to the diagonal matrix of eigenvalues: $A \sim \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, more precisely: There is a unitary matrix $U \in M_n(\mathbb{C})$, such that $UAU^* = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.These matrices are determined by their eigenvalues. How about general bounded operators?

Now consider for $k \in L^2([0,1] \times [0,1])$ the "Fredholm integral equation"

$$\int_0^1 k(s,t) f(t)\, dt - \lambda f(s) = g(s)$$

for a given g and $\lambda \in \mathbb{C}$. We want to find such an (or this? Is it unique?) f. Abstractly speaking (in the situation of Example 7.2): What are solutions f of $Kf - \lambda f = g$? If $(K - \lambda 1)$ is invertible, $f = (K - \lambda 1)^{-1} g$ is a solution to our equation.

In finite dimensions $A - \lambda 1$ is non-injective if and only if λ is an eigenvalue (i. e., $\exists\, x \neq 0 : Ax = \lambda x$). $A - \lambda 1$ also is non-injective if and only if $A - \lambda 1$ is non-invertible as shown in any lecture on linear algebra. But if the involved spaces are infinite dimensional, "T injective if and only if T surjective if and inly if T invertible" do not hold.

We therefore have more possibilites to define a generalisation of eigenvalues. This will lead to the notion of the *spectrum*.

Definition 8.1: Let A be a $\mathbb{C}$-vector space.

(i) A is an *algebra* if there is a bilinear, associative multiplication on A that satisfies $\lambda(xy) = (\lambda x)y = x(\lambda y)\, \forall\, x, y \in A, \lambda \in \mathbb{C}$.

(ii) A is a *normed algebra* if A is an algebra, that is a normed vector space with $\|xy\| \leq \|x\|\|y\|$.

(iii) A is a *Banach algebra*, if A is a complete, normed algebra.

Example 8.2: (i) $(B(H), \|\cdot\|)$ is a Banach algebra, where H is a Hilbert space (or a Banach space),

(ii) $(C(K), \|\cdot\|_\infty)$ is Banach algebra, where K is a compact space.

Remark 8.3: The multiplication in a normed algebra is continuous (as in Remark 1.18).

Definition 8.4: Let A be a unital Banach algebra and let $x \in A$. Then

(i) $\operatorname{Sp}(x) := \{\lambda \in \mathbb{C} \mid \lambda 1 - x \text{ is not invertible}\} \subseteq \mathbb{C}$ is called the *spectrum of* x,

(ii) $\rho(x) := \mathbb{C} \setminus \operatorname{Sp}(x)$ is called the *resolvent set*.

$\mathrm{Sp}(x)$ is sometimes denoted $\sigma(x)$ in literature on the topic.

Remark 8.5: If now $A = B(H)$, then we may also define eigenvalues: Given an operator $T \in B(H)$, a complex number $\lambda \in \mathbb{C}$ is an *eigenvalue* of T, if there is a $0 \neq x \in H$ such that $Tx = \lambda x$.

The set $\sigma_p(T) := \{\text{Eigenvalues of } T\} \subseteq \mathbb{C}$ is called the *point spectrum* and we observe that if $\lambda \in \sigma_p(T)$, then $(\lambda 1 - T)$ is not injective and in particular $(\lambda 1 - T)$ is not invertible, hence we have $\sigma_p(T) \subseteq \mathrm{Sp}(T)$. Spectral values are generalised eigenvalues.

There are examples for $\sigma_p(T) = \mathrm{Sp}(T)$, $\sigma_p(T) \subsetneq \mathrm{Sp}(T)$, $\sigma_p(T) = \varnothing$ (but it always holds $\mathrm{Sp}(T) \neq \varnothing$).

Example 8.6: Let $X = C([1,2])$. Then $B(X)$ is a Banach algebra with unit (refer to Theorem 1.26 and Proposition 1.29). Consider

$$\begin{aligned} T\colon X &\longrightarrow X \\ (Tf)(t) &\longmapsto tf(t), \end{aligned}$$

then $T \in B(X)$ with $\|T\| = 2$. We have that $\sigma_p(T) = \varnothing$: If $Tf = \lambda f$, then $tf(t) = \lambda f(t) \,\forall t \in [1,2]$, hence $f(t) = 0 \,\forall t \neq \lambda$ and due to the continuity of f it holds $f \equiv 0$. In particular $\lambda 1 - T$ is injective for all $\lambda \in [1,2]$. For the surjectivity: First let $\lambda \in [1,2]$ and assume $\lambda 1 - T$ was surjective. Consider $g(t) \equiv 1 \,\forall t \in [1,2]$. Due to the surjectivity of T, there was $f \in C(X)$ such that $1 = g(t) = (\lambda 1 - T)f(t) = (\lambda 1 - t)f(t)$, but $(\lambda - t)f(t) = 0$ for $t = \lambda$.

Now let $\lambda \notin [1,2]$. Then $(\lambda 1 - T)^{-1}$ is given by $((\lambda 1 - T)^{-1} f)(t) = \frac{1}{\lambda - t} f(t)$, hence in this case $\mathrm{Sp}(T) = [1,2]$.

So our T has no eigenvalues, but many spectral values.

Lemma 8.7: *Let A be a unital Banach algebra.*

(i) *If $x \in A$ with $\|1 - x\| < 1$, then x is invertible and $x^{-1} = \sum_{n=0}^{\infty} (1-x)^n$.*

(ii) *If x is invertible and $y \in A$ with $\|x - y\| < (\|x^{-1}\|)^{-1}$, then y is invertible.*

(iii) $\mathrm{GL}(A) := \{x \in A \mid x$ *is invertible*$\}$ *is open, and*

$$\begin{aligned} \mathrm{GL}(A) &\longrightarrow \mathrm{GL}(A) \\ x &\longmapsto x^{-1} \end{aligned}$$

is continuous.

Proof: (i) For $z := 1 - x$, we have that $\|z\| < 1$, hence $\sum_{n=0}^{\infty} \|z\|^n$ is abosolute convergent ($\|z^n\| \leq \|z\|^n$ via the submultiplicativity), hence it is also convergent via Lemma 1.30. Moreover:

$$x \sum_{n=0}^{\infty} z^n \leftarrow (1-z) \sum_{n=0}^{N} z^n = \sum_{n=0}^{N} z^n - \sum_{n=1}^{N+1} z^n = 1 - z^{N+1} \to 1$$

since $\|z\| \leq 1$.

(ii) It holds that

$$\|1 - yx^{-1}\| = \|(x-y)x^{-1}\| \leq \|x-y\|\|x^{-1}\| < 1,$$

i. e., yx^{-1} is invertible and thus y is invertible.

(iii) For $\|x-y\| < \varepsilon < (\|x^{-1}\|)^{-1}$, we have $B(x,\varepsilon) \subseteq \mathrm{GL}(A)$ for $x \in \mathrm{GL}(A)$ by (ii). Now let $x_\lambda \to x$ with $x_\lambda, x \in \mathrm{GL}(A)$. Then $\|x_\lambda - x\| < (2\|x^{-1}\|)^{-1}\varepsilon$ for λ large and $0 < \varepsilon < 1$. Hence

$$\|1 - x_\lambda x^{-1}\| = \|(x - x_\lambda x^{-1}\| \leq \frac{\varepsilon}{2} < 1,$$

thus via (i), $x_\lambda x^{-1}$ is invertible with

$$xx_\lambda^{-1} = (x_\lambda x^{-1})^{-1} = \sum_{n=0}^{\infty}(1 - x_\lambda x^{-1})^n = 1 + \sum_{n=1}^{\infty}(1 - x_\lambda x^{-1})^n,$$

hence

$$\begin{aligned}\|x_\lambda^{-1} - x^{-1}\| = \|x^{-1}(xx_\lambda^{-1} - 1)\| &\leq \|x^{-1}\| \sum_{n=1}^{\infty}\|1 - x_\lambda x^{-1}\|^n \\ &\leq \|x^{-1}\| \sum_{n=1}^{\infty} \varepsilon \frac{1}{2^n} < \varepsilon\|x^{-1}\|.\end{aligned}$$

■

Proposition 8.8: *Let A be a unital Banach algebra and let $x \in A$. Then $\mathrm{Sp}(x)$ is compact and*

$$\mathrm{Sp}(x) \subseteq \{\lambda \in \mathbb{C} \mid |\lambda| \leq \|x\|\}.$$

Proof: Firstly, we notice that the resolvent set $\rho(x)$ can be written as $\rho(x) = f_x^{-1}(\mathrm{GL}(A))$, where $f_x \colon \mathbb{C} \to A, \lambda \mapsto \lambda 1 - x$ is continuous. Via Lemma 8.7, $\rho(x)$ is open and $\mathrm{Sp}(x)$ is closed.

Secondly, if $|\lambda| \geq \|x\|$, then $\lambda - x = \lambda(1 - \frac{x}{\lambda})$ is invertible via Lemma 8.7, since $|\lambda|^{-1}\|x\| < 1$, hence $\lambda \notin \mathrm{Sp}(x)$ and therefore $\mathrm{Sp}(x) \subseteq \{\lambda \in \mathbb{C} \mid |\lambda| \leq \|x\|\}$ is bounded. ■

Theorem 8.9: *If A is a unital Banach algebra, then $\mathrm{Sp}(x) \neq \varnothing$ for all $x \in A$.*

Proof: Let $x \in A$. For $\lambda \in \rho(x)$, put $R_\lambda(x) := (\lambda - x)^{-1}$.

① *We have $R_\lambda(x) - R_\mu(x) = (\mu - \lambda)R_\lambda(x)R_\mu(x) \,\forall\, \lambda, \mu \in \mathbb{C}$.*

Proof (of ①): It holds that

$$\begin{aligned}R_\lambda(x) - R_\mu(x) &= R_\lambda(x)R_\mu(x)(\mu - x) - (\lambda - x)R_\lambda(x)R_\mu(x) \\ &= (\mu - \lambda)R_\lambda(x)R_\mu(x),\end{aligned}$$

where we used $R_\lambda(x)R_\mu(x)(\mu - x) = (\mu - x)R_\lambda(x)R_\mu(x)$. In principle $ab \neq ba$ for $a, b \in A$, but here this doesn't cause issues because

$$(\lambda - x)(\mu - x) = (\mu - x)(\lambda - x) \quad \Longrightarrow \quad (\mu - x)R_\lambda(x).$$

■

② *Assume that x is invertible and let $f \in A'$ such that $f(x^{-1}) \neq 0$. Then $g\colon \rho(x) \to \mathbb{C}, g(\lambda) := f(R_\lambda(x))$ is holomorphic and $g(0) \neq 0$.*

Proof (of ②): Consider

$$\frac{g(\lambda) - g(\mu)}{\lambda - \mu} = f\left(\frac{R_\lambda(x) - R_\mu(x)}{\lambda - \mu}\right) \overset{①}{=} -f(R_\lambda(x)R_\mu(x)) \to -f(R_\lambda^2(x)) \text{ as } \mu \to \lambda$$

where we used, that $x \mapsto x^{-1}$ is continuous and therefore $\mu \mapsto R_\mu(x)$ is continuous aswell. Thus g is holomorphic with $g(0) = f(R_0(x)) = -f(x^{-1}) \neq 0$. ■

Assume, that $\mathrm{Sp}(x) = \varnothing$. Then $0 \notin \mathrm{Sp}(x)$, i. e., $0 - x$ is invertible and x is invertible. By the Theorem of Hahn-Banach Corollary 2.6, we find a functional $f \in A'$ with $f(x^{-1}) \neq 0$. Thus the function g from ② is a whole function.

③ *g is bounded, because $g(\lambda) \to 0$ for $\lambda \to \infty$.*

Proof (of ③): Put $z := 1 - \lambda^{-1}x$. Then $\|1 - z\| = |\lambda|^{-1}\|x\| < 1$ for $|\lambda|$ large. Via Lemma 8.7 (i), z is invertible with $z^{-1} = \sum_{n=0}^{\infty}(1 - z)^n$, hence

$$\|z^{-1}\| \leq \sum_{n=0}^{\infty}\|1 - z\|^n = (1 - \|1 - z\|)^{-1}$$

and thus

$$\|(1 - \lambda^{-1}x)^{-1}\| \leq \frac{1}{1 - \frac{\|x\|}{|\lambda|}}.$$

Finally, it holds

$$\begin{aligned} \|R_\lambda(x)\| &= \|(\lambda - x)^{-1}\| = |\lambda|^{-1}\|(1 - \lambda^{-1}x)^{-1}\| \\ &\leq \frac{1}{|\lambda|(1 - \frac{\|x\|}{|\lambda|})} = \frac{1}{|\lambda| - \|x\|} \to 0 \text{ as } |\lambda| \to \infty. \end{aligned}$$ ■

We conclude that by Liouville's Theorem g is constant and from $g(\lambda) \to 0$ as $|\lambda| \to 0$, we infer that $g \equiv 0$, which contradicts $g(0) \neq 0$. ■

Definition 8.10: Let A be a unital Banach algebra and let $x \in A$. Then

$$r(x) := \sup\{|\lambda| \mid \lambda \in \mathrm{Sp}(x)\}$$

is called the *spectral radius of x*.

Remark 8.11: From Theorem 8.9 we already know that $r(x) \leq \|x\|$.

Example 8.12: We may have $r(x) < \|x\|$: For instance with $A := M_2(\mathbb{C})$ and $x := \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. Then $\lambda - x = \left(\begin{smallmatrix} \lambda & -1 \\ 0 & \lambda \end{smallmatrix}\right)$ is invertible for all $\lambda \neq 0$, thus $\mathrm{Sp}(x) = \{0\}$. Hence $r(x) = 0$, but $\|x\| \neq 0$.

Theorem 8.13: *Let A be a unital Banach algebra and $x \in A$. Then*

$$r(x) = \lim_{n\to\infty} \sqrt[n]{\|x^n\|}.$$

Proof: If $\lambda \in \operatorname{Sp}(x)$, then $\lambda^n \in \operatorname{Sp}(x^n)$: From Analysis I, we know of the handy formula

$$\lambda^n - x^n = (\lambda - x)(\lambda^{n-1} + \lambda^{n-2}x + \cdots + \lambda x^{n-2} + x^{n-1}).$$

If $|\lambda^n| \leq \|x^n\|$, then $|\lambda| \leq \sqrt[n]{\|x^n\|}$ and thus $r(x) \leq \liminf_{n\to\infty} \sqrt[n]{\|x^n\|}$. We need to show, that $r(x) \geq \limsup_{n\to\infty} \sqrt[n]{\|x^n\|}$. Consider

$$R_z(x) := (z - x)^{-1} = \sum_{n=0}^{\infty} \frac{x^n}{z^{n+1}}$$

for $\|x\| \leq |z|$ (in particular $z \in \rho(x)$).

If this was a series in $\mathbb{C}$ (rather than A), its radius of convergence of this power series was $\limsup_{n\to\infty} \sqrt[n]{\|x^n\|}$. Since it is a series in A, we have to use the same trick as in the proof of Theorem 8.9. Let $f \in A'$. Like in the proof of Theorem 8.9, the function $g\colon \rho(x) \to \mathbb{C}, z \mapsto f(R_z(x))$ is holomorphic and $g(z) = \sum_{n=0}^{\infty} \frac{f(x^n)}{z^{n+1}}$ for $|z| > \|x\|$, in fact even for $|z| > r(x)$. Hence

$$\limsup_{n\to\infty} |f(x)|^{\frac{1}{n}} \leq r(x)$$

by the formula of Cauchy Hadamard for convergence radii of power series. For $r > r(x)$, we thus find an $N \in \mathbb{N}$ such that $|f(x^n)|^{\frac{1}{n}} < r \,\forall\, n \geq N$ and hence

$$\sup_{n\in\mathbb{N}} \left| \frac{f(x^n)}{r^n} \right| < \infty$$

for all $f \in A'$. By the principle of uniform boundedness Theorem 4.4, we see that $\{\frac{x^n}{r^n} \mid n \in \mathbb{N}\}$ is bounded. Hence there is $C > 0$ such that $\|x^n\| \leq Cr^n$ hence $\|x^n\|^{\frac{1}{n}} \leq C^{\frac{1}{n}} r$, which implies $\limsup_{n\to\infty} \|x^n\|^{\frac{1}{n}} \leq r \,\forall\, r \geq r(x)$ and thus

$$\limsup_{n\to\infty} \|x^n\|^{\frac{1}{n}} \leq r(x).$$

■

9 Compact operators and their spectral theorem

Remark 9.1: Let E be a normed vector space. Then the following are equivalent:

(i) E is finite-dimensional,

(ii) $\{x \in E \mid \|x\| \leq 1\}$ is compact.

If now $\dim H = \infty$, then $A(\{x \in H \mid \|x\| \leq 1\}) \subseteq H$ might be noncompact for some $A \in B(H)$ (for instance for $A = \mathrm{id}_H$). This makes spectral theory for $A \in B(H)$ much more complicated. Hence we first consider operators, which are in a way "close" to the finite setting.

We want a spectral theorem similar to the one in Linear Algebra:

Theorem: *Let H be a Hilbert space with* $\dim(H) < \infty$ *and let $A \in M_n(\mathbb{C})$ be* normal. *Then there is a unitary $U \in M_n(\mathbb{C})$ such that*

$$UAU^* = \mathrm{diag}(\lambda_1, \dots, \lambda_n),$$

where $\mathrm{Sp}(A) = \sigma_P(A) = \{\lambda_i \mid 1 \leq i \leq n\}$.

We will show such a theorem for *compact operators* $A \in K(H)$ first.

Definition 9.2: Let X, Y be Banach spaces. A linear operator $T\colon X \to Y$ is called *compact*, if $\overline{\{Tx \mid \|x\| \leq 1\}}$ is compact. We write

$$K(X,Y) := \{T\colon X \to Y \text{ linear, compact}\} \quad K(X) := K(X,X).$$

Remark 9.3: (i) We have $K(X,Y) \subseteq B(X,Y)$. Indeed, since $\overline{\{Tx \mid \|x\| \leq 1\}}$ is bounded, we find a constant C such that $\|Tx\| \leq C$ for all $\|x\| \leq 1$.

(ii) T is compact if and only if $\overline{TM}$ is compact for all bounded sets $M \subseteq X$.

Proof: "$\Leftarrow$": $M = \{x \mid \|x\| \leq 1\}$ is bounded.

"$\Rightarrow$": If M is bounded, then there is a constant C such that $\frac{1}{C}M \subseteq \{x \mid \|x\| \leq 1\}$ and thus $\overline{TM} \subseteq C\overline{\{Tx \mid \|x\| \leq 1\}}$. ∎

(iii) T is compact if and only if for any bounded sequence $(x_n)_{n\in\mathbb{N}}$ the sequence $(Tx_n)_{n\in\mathbb{N}}$ possesses a convergent subsequence.

Proof: "$\Rightarrow$": Let $(x_n)_{n\in\mathbb{N}}$ be bounded and without loss of generality let $\|x_n\| \leq 1$ for all $n \in \mathbb{N}$. Then

$$\overline{\{Tx_n \mid n \in \mathbb{N}\}} \subseteq \overline{\{Tx \mid \|x\| \leq 1\}}$$

is a compact subset. Therefore, $(Tx_n)_{n\in\mathbb{N}}$ has a convergent subsequence.

"$\Leftarrow$": Let $(Tx_n)_{n\in\mathbb{N}} \subseteq \overline{\{Tx \mid \|x\| \leq 1\}}$. Then $(Tx_n)_{n\in\mathbb{N}}$ has a convergent subsequence by assumption. ∎

Definition 9.4: A (twosided) *ideal* in an algebra A is a linear subspace $I \subseteq A$, such that $AI, IA \subseteq I$. If A is a Banach algebra and I is closed, we write $I \triangleleft A$.

Theorem 9.5: *Let X be a Banach space. Then $K(X) \triangleleft B(X)$.*

Proof: First, we want to show that $K(X)$ is a linear subspace. In order to show this, let $S, T \in K(X)$ and $M \subseteq X$ be bounded. Then

$$\overline{(S+T)M} \subseteq \overline{SM} + \overline{ST}$$

is compact, because "+" is continuous.

Let now $S \in K(X)$, $T \in B(X)$ and $M \subseteq X$ be bounded.

"$K(X)B(X) \subseteq K(X)$": TM is bounded (since T is bounded) and therefore, $\overline{S(TM)}$ is compact, thus $ST \in K(T)$.

"$B(X)K(X) \subseteq K(X)$": It holds that $TS(M) \subseteq T(\overline{SM})$, thus

$$\overline{TS(M)} \subseteq \overline{T(\overline{SM})} = T(\overline{SM}),$$

because $\overline{SM}$ is compact and T is continuous. Hence, $\overline{TS(M)}$ is compact.

It remains to be shown, that $K(X)$ is closed. Let $(T_n)_{n\in\mathbb{N}} \subseteq K(X)$ be a sequence with $T_n \to T \in B(X)$ (with respect to the operator norm). Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be bounded. If $(Tx_n)_{n\in\mathbb{N}}$ admits a convergent subsequence, we know that T is compact.

(i) *Construction of a subsequence $(y_k)_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$*: Since T_1 is compact, we find a subsequence $(x_k^{(1)})_{k\in\mathbb{N}}$, such that $(T_1 x_k^{(1)})_{k\in\mathbb{N}}$ converges. Inductively, choose a subsequence $(x_k^{(n+1)})_{k\in\mathbb{N}}$ of $(x_k^{(n)})_{k\in\mathbb{N}}$ such that $(T_{n+1}x_k^{(n+1)})_{k\in\mathbb{N}}$ converges. Since $(x_k^{(n)})_{k\in\mathbb{N}}$ is a subsequence of $(x_k^{(m)})_{k\in\mathbb{N}}$ when $m < n$, we know that $(T_m x_k^{(n)})_{k\in\mathbb{N}}$ converges, too. We therefore put $y_k := x_k^{(k)}$.

(ii) *The constructed sequence $(y_k)_{k\in\mathbb{N}}$ converges*: Let $\varepsilon > 0$, $n_0 \in \mathbb{N}$ such that $\|T - T_{n_0}\| < \varepsilon$ (which is possible since by assumption $T_n \to T$). Put

$$M := \sup_{k\in\mathbb{N}} \|y_k\| < \infty;$$

this is well-defined, since $(y_k) \subseteq (x_n)$ is bounded. Let $N \in \mathbb{N}$ such that $\|T_{n_0}y_k - T_{n_0}y_l\| < \varepsilon$ for all $k, l \geq N$ ($(y_k)_{k\in\mathbb{N}}$ is a subsequence of $(x_k^{(n_0)})_{k\in\mathbb{N}}$ up to finitely many entries, hence $(T_{n_0}y_k)_{k\in\mathbb{N}}$ converges). Then, for $k, l \geq N$:

$$\begin{aligned} \|Ty_k - Ty_l\| &\leq \|Ty_k - T_{n_0}y_k\| + \|T_{n_0}y_k - T_{n_0}y_l\| + \|T_{n_0}y_l - T_{y_l}\| \\ &< \varepsilon M + \varepsilon + \varepsilon M < 3M\varepsilon \end{aligned}$$

Hence (Ty_k) is Cauchy and thus converges. ■

Definition 9.6: An operator $T\colon X \to X$ is of *finite rank*, if TX is finite-dimensional. We put

$$E(X) := \{T \in B(X) \mid T \text{ of finite rank}\}$$

Remark 9.7: (i) If $T \in E(X)$, then $T \in K(X)$ since

$$\overline{\{Tx \mid \|x\| \leq 1\}} \subseteq \overline{\{y \in TX \mid \|y\| \leq C\}}$$

is compact via Remark 9.1 and Definition 9.2.

(ii) $\mathrm{id}\colon X \to X$ is compact if and only if X is finite-dimensional ("$\Leftarrow$": (i), "$\Rightarrow$": Definition 9.2).

(iii) As in Theorem 9.5, one can show that $E(X) \subseteq B(X)$ is a two-sided ideal, in general not closed. Also, $\overline{E(X)} \subseteq K(X)$.

Theorem 9.8: *Let H be a separable Hilbert space. Then $\overline{E(H)} = K(H)$.*

This theorem is wrong for general Banach spaces, which was shown by Enflo in 1973.

Proof: "$\subseteq$" holds in general. We now need to show "$\supseteq$": Let $T \in K(H)$ and let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal base of H – without loss of generality, $\dim(H) = \infty$. Consider $H_n := \langle e_1, \ldots, e_n \rangle$ and let $P_n \in B(H)$ be the orthogonal projection onto H_n (refer to sheet 8). Put $T_n := P_n T \in E(H)$. We need to show, that $\|T_n - T\| \to 0$. Let $x \in H$. Then we have

$$T_n(x) = P_n(Tx) = \sum_{k=1}^{n} \langle Tx, e_k \rangle e_k \to \sum_{i=1}^{\infty} \langle Tx, e_k \rangle e_k$$

via Theorem 5.28. This shows that $\|T_n x - Tx\| \to 0$ for all $x \in H$. Let $\varepsilon > 0$. Since T is compact, we have that $\{Tx \mid \|x\| \leq 1\} \subseteq \bigcup_{j=1}^{m} B(Tx_j, \varepsilon)$ for finitely many $x_1, \ldots, x_m$ with $\|x_i\| \leq 1$. Choose $N \in \mathbb{N}$ such that $\|Tx_j - T_n x_j\| < \varepsilon$ for $n \geq N$ and for all $1 \leq j \leq m$. Then, for arbitrary $x \in H$ with $\|x\| \leq 1$ it holds that $\|Tx - Tx_j\| < \varepsilon$.

Let $x \in H$ with $\|x\| \leq 1$. Then

$$\begin{aligned}\|Tx - T_n x\| &\leq \|Tx - Tx_j\| + \|Tx_j - T_n x_j\| + \|T_n x_j - T_n x\| \\ &\leq \varepsilon + \varepsilon + \|P_n\| \|Tx_j - Tx\| < 3\varepsilon,\end{aligned}$$

thus $\|T - T_n\| < \varepsilon$. ■

Remark 9.9: We proved that any $T \in B(H)$ may be approximated by a sequence $(T_n)_{n\in\mathbb{N}} \subseteq E(H)$ in the strong operator topology: $\forall x \in H$ with $\|x\| \leq 1$: $\|T_n x - Tx\| \to 0$. But for the norm topology, this approximation only holds for $T \in K(H)$: $\|T_n - T\| \to 0$.

Theorem 9.10: *Let H be a Hilbert space and $T \in B(H)$. Then T is compact if and only if T^* is compact.*

Proof: "$\Rightarrow$": Let $T \in K(H)$ be given. By Theorem 9.8, we find a sequence $(T_n)_{n\in\mathbb{N}} \subseteq E(H)$ with $T_n \to T$. Then

$$\|T_n^* - T^*\| = \|T_n - T\| \to 0,$$

hence $T_n^* \to T^*$ and $T_n^* = (P_n T)^* = T^* P_n = T^* P_n \in E(H)$. Theorem 9.8 does the job for us now.

"$\Leftarrow$": It holds $T^{**} = T$, then use the first part. ■

Remark 9.11: Let X, Y be Banach spaces, $T \in K(X, Y)$. Then $T' \in K(Y', X')$ is compact as well. This proves Theorem 9.10 for Banach spaces, since we can express $T^* = j^{-1} \circ T' \circ j$.

Example 9.12: (i) If $\dim H < \infty$, then $K(H) = B(H)$. In particular it holds $E(\mathbb{C}^n) = K(\mathbb{C}^n) = B(\mathbb{C}^n) = M_n(\mathbb{C})$.

(ii) If $H = L^2([0,1])$ with $k \in L^2([0,1] \times [0,1])$ and $K \in B(L^2([0,1]))$ as in Example 7.2, then $K \in K(L^2([0,1]))$.

Proof: Let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal base of $L^2([0,1])$. Check that

$$e_{n,m}(s,t) := e_n(s)\overline{e_m(t)}$$

is an orthonormal base of $L^2([0,1] \times [0,1])$. For

$$k = \sum_{n,m=1}^{\infty} \alpha_{n,m} e_{n,m},$$

put $k_N := \sum_{n,m=1}^{N} \alpha_{n,m} e_{n,m}$. Hence $\|k - k_N\|_2 \to 0$ as $N \to \infty$. For K_N, the integral operator with respect to k_N, we have that $K - K_N$ the integral operator with respect to $k - k_N$. Hence

$$\|K - K_N\| \leq \|k - k_N\|_2 \longrightarrow 0,$$

i. e., $K_N \to K$. K_N has finite rank as

$$\begin{aligned}(K_N f)(s) = \int_0^1 k_N(s,t) f(t)\, dt &= \sum_{n,m=1}^{N} \alpha_{n,m} \int_0^1 e_n(s)\overline{e_m(t)} f(t)\, dt \\ &= \sum_{n=1}^{N} e_n(s) \Big(\sum_{m=1}^{N} \alpha_{n,m} \langle f, e_m \rangle \Big),\end{aligned}$$

thus $K_N f \in \langle e_1, \ldots, e_n \rangle$ for all f. ■

Theorem 9.13 (Spectral theorem for selfadjoint compact operators): *Let H be a separable Hilbert space and let $T \in K(T)$ with $T = T^*$.*

(i) *If $\lambda \in \sigma_p(T)$, $\lambda \neq 0$, then the eigenspace* $\ker(\lambda - T)$ *is finite-dimensional.*[1]

(ii) *If $\lambda \notin \sigma_p(T)$, $\lambda \neq 0$, then $\lambda \notin \mathrm{Sp}(T)$ and $\sigma_p(T) \subseteq \mathbb{R}$.*

(iii) *The operator T has only countably many mutually different eigenvalues*

$$\{\lambda_1, \lambda_2, \dots\}$$

and the corresponding eigenspaces for $\lambda_i \neq 0$ are orthogonal to each other and are finite-dimensional. All eigenvalues are real and

$$\mathrm{Sp}(T) \subseteq \{0\} \cup \{\lambda_1, \lambda_2, \dots\}$$

We may decompose

$$T = \sum_{n=1}^{\infty} \lambda_n P_n$$

and call T a diagonal operator*.*

Proof: (i) Let $(e_i)_{i \in I}$ be an orthonormal base of $\ker(\lambda - T)$. Then for $i \neq j$ it holds

$$\|Te_i - Te_j\|^2 = \|\lambda e_i - \lambda e_j\|^2 = 2|\lambda|^2.$$

If $|I| = \infty$, then $(Te_i)_{i \in I}$ has no convergent subsequence which is a contradiction to the compactness.

(ii) We need to show, that $\mathrm{im}(\lambda - T) = H$. In this case, $\lambda - T$ is surjective and injective, hence invertible and $\lambda \notin \mathrm{Sp}(T)$.
(ii.1) *It holds $\sigma_p(T) \subseteq \mathbb{R}$.*

Proof (of (ii.1)): For $\lambda \in \sigma_p(T)$ and $x \neq 0$ with $Tx = \lambda x$ we have

$$\overline{\lambda}\langle x, x\rangle = \langle x, \lambda x\rangle = \langle x, Tx\rangle = \langle Tx, x\rangle = \langle \lambda x, x\rangle = \lambda\langle x, x\rangle. \qquad \blacksquare$$

(ii.2) *For a sequence $(x_n)_{n \in \mathbb{N}} \subseteq H$ with $\|x_n\|$ and $\|Tx_n - \xi x_n\| \to 0$, $\xi \neq 0$, it holds that $\xi \in \sigma_p(T)$.*

Proof (of (ii.2)): As T is compact, we can find a subsequence $(x_{n_k})_{k \in \mathbb{N}}$ such that $Tx_{n_k} \to y \in H$ as $k \to \infty$. Hence: $\xi x_{n_k} = Tx_{n_k} - (Tx_{n_k} - \xi x_{n_k}) \to y$ and therefore $x_{n_k} \to \xi^{-1} y$. But then $\xi y \leftarrow \xi(Tx_{n_k}) = T(\xi x_{n_k}) \to Ty$ and $y \neq 0$ since $1 \equiv \|x_{n_k}\| \to |\xi|^{-1}\|y\|$ implies that $\|y\| = |\xi| \neq 0$. $\blacksquare$

(ii.3) *It holds* $\mathrm{im}(\lambda - T) = \overline{\mathrm{im}(\lambda - T)}$.

[1] This holds for all compact operators on separable Hilbert spaces.

Proof (of (ii.3)): By (ii.2) there is $c > 0$ such that $\|Tx - \lambda x\| \geq c\|x\|$ for all $x \in H$; use $\|x\|^{-1}x$. Let $y \in \overline{\mathrm{im}(\lambda - T)}$ i.e., we have $x_n \in H$ with $\lambda x_n - Tx_n \to y$. Then

$$\|x_n - x_m\| \overset{\text{(ii.2)}}{\leq} \frac{1}{c}\|T(x_n - x_m) - \lambda(x_n - x_m)\| \to 0$$

Thus $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, we thus have $x_n \to x$ for some $x \in H$ and $y \leftarrow \lambda x_n - Tx_n \to (\lambda - T)x \in \mathrm{im}(\lambda - T)$. ■

Finally, we have

$$\mathrm{im}(\lambda - T) = \overline{\mathrm{im}(\lambda - T)} = \ker((\lambda - T)^*)^\perp = \ker(\overline{\lambda} - T)^\perp = \ker(\lambda - T)^\perp = H$$

as $\lambda \notin \sigma_p(T)$.

(iii) (iii.1) *There is $\lambda_1 \in \sigma_p(T)$ with $\lambda_1 \in \mathbb{R}$ such that $|\lambda_1| = \|T\|$.*

Proof (of (iii.1)): By sheet 8, we have that $\|T\| = \sup\{|\langle Tx, x\rangle| \mid \|x\| = 1\}$, hence we find $(x_n) \subseteq H$, $\|x_n\| = 1$ with $|\langle Tx_n, x_n\rangle| \to \|T\|$. We know $\langle Tx_n, x_n\rangle \in \mathbb{R}$ by (Remark 7.11), hence without loss of generality $\langle Tx_n, x_n\rangle \to \lambda_1 := \pm\|T\|$. Then

$$\begin{aligned}\|(T - \lambda_1)x_n\|^2 = \|Tx_n\|^2 - 2\lambda_1\langle Tx_n, x_n\rangle + \lambda_1^2 &\leq \lambda_1^2 - 2\lambda_1\langle Tx_n, x_n\rangle + \lambda_1^2 \\ &\leq 2\lambda_1(\lambda_1 - \langle Tx_n, x_n\rangle) \to 0\end{aligned}$$

By (ii.2) it holds $\lambda_1 \in \sigma_p(T)$. ■

(iii.2) Put $H_1 := \ker(\lambda - T)$ and P_1 the projection onto H_1. Decompose $H = H_1 \oplus H_1^\perp$. Then $TH_1 \subseteq H_1$, $TH_1^\perp \subseteq H_1^\perp$, because if $x \in H_1$, then $Tx = \lambda_1 x \in H_1$ and if $x \in H_1^\perp$, then for $y \in H_1$ it holds $\langle Tx, y\rangle = \lambda_1\langle x, y\rangle = 0$, thus $x \in H_1^\perp$.

Hence $T = \lambda_1 1 \oplus T_2 = \begin{pmatrix} \lambda_1 1 & \\ & T_2 \end{pmatrix} \in B(H_1 \oplus H_q^\perp)$. Now put $T_2 := T|_{H_1^\perp} \in K(H_1^\perp)$, then $T_2^* = T_2$. By (iii.1) we find $\lambda_2 \neq \lambda_1$ such that $|\lambda_2| \leq |\lambda_1|$ with $H_2 := \ker(\lambda_2 - T_2) = \ker(\lambda_2 - T)$.

(iii.3) Inductively we find a sequence $(\lambda_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ with $\lambda_n \in \sigma_p(T)$ for all n with $H_n := \ker(\lambda_n - T)$. The H_n are mutually orthogonal as we have for $x \in \ker(\lambda_n - T)$, $y \in \ker(\lambda_m - T)$:

$$\lambda_n\langle x, y\rangle = \langle Tx, y\rangle = \langle x, Ty\rangle = \lambda_m\langle x, y\rangle$$

and thus $\langle x, y\rangle = 0$ must hold.

(iii.4) *It holds $\lambda_n \to 0$.*

Proof (of (iii.4)): As $|\lambda_1| \geq |\lambda_2| \geq \dots$, there is $\alpha \geq 0$ with $|\lambda_n| \to \alpha$. For $x_n \in H_n$ with $\|x_n\|$, we find a convergent subsequence of $(Tx_n)_{n\in\mathbb{N}}$ as T is compact. Now it holds

$$\|Tx_n - Tx_m\|^2 = \|\lambda_n x_n - \lambda_m x_m\|^2 = |\lambda_n|^2 + |\lambda_m|^2 \geq 2\alpha^2$$

and thus $\alpha = 0$. ■

(iii.5) *It holds* $\sum_{n=1}^{N} \lambda_n P_n \to T$ *where* P_n *denotes the projection onto* $H_n = \ker(\lambda_n - T)$.

Proof (of (iii.5)): Let $x \in H$. Then we can decompose

$$x = x_0 + x_1 \in (H_1 \oplus \cdots \oplus H_N) \oplus (H_1 \oplus \cdots \oplus H_N)^\perp$$

and denote $H' := (H_1 \oplus \cdots \oplus H_N)^\perp$

$$\left\| \Big(T - \sum_{n=1}^{N} \lambda_n P_n\Big) x \right\| = \|Tx_1\| \leq \|T|_{H'}\| \leq |\lambda_{n+1}| \|x\| \to 0.$$

■

Of course, for Banach spaces we can't hope for a perfect analogon of the theorem for Hilbert spaces, as we don't have the orthogonal decomposition at hand. However, there is a generalisation of the spectral theorem of compact operators on Hilbert spaces:

Theorem 9.14 (Spectral theorem for compact operators on Banach spaces): *Let* X *be a Banach space,* $T \in K(X)$. *Then the following hold:*

(i) $\mathrm{Sp}(T)$ *has at most countable many elements and 0 is the only cluster point. If* X *is infinite-dimensional, then* $0 \in \mathrm{Sp}(T)$.

(ii) *If* $0 \neq \lambda \in \mathrm{Sp}(T)$, *then* $\lambda \in \sigma_p(T)$ *and* $\dim(\ker(\lambda - T)) < \infty$.

(iii) *For* $0 \neq \lambda \in \mathrm{Sp}(T)$, *there is a composition* $X = N_\lambda \oplus F_\lambda$ *(as described in (Remark 5.19)) such that* $\ker(\lambda - T) \subseteq N_\lambda$, $(\lambda - T)|_{N_\lambda}$ *is nilpotent and if* $0 \neq \mu \in \mathrm{Sp}(T)$ *with* $\lambda \neq \mu$, *then* $N_\lambda \subseteq F_\mu$.

10 Banach- and C^*-algebras and the Gelfand transformation

Definition 10.1: Let A be a normed $\mathbb{C}$-vector space.

(i) A is called a *normed algebra*, if A is an algebra such that $\|xy\| \leq \|x\|\|y\|$ for all $x, y \in A$.

(ii) A is called a *Banach algebra*, if A is a complete normed algebra.

(iii) An *involution* on an algebra is a map ${}^*\colon A \to A$ with

  (1) $(a+b)^* = a^* + b^*$, $(\lambda a)^* = \overline{\lambda} a^*$ for $a, b \in A$ and $\lambda \in \mathbb{C}$,

  (2) $(ab)^* = b^* a^*$,

  (3) $a^{**} = a$.

(iv) A **-Banach algebra* is a Banach algebra with an involution.

(v) A C^**-algeba*[1] is a *-Banach algebra such that $\|x^*x\| = \|x\|^2$ for all $x \in A$.

(vi) A(n) algebra / Banach algebra / C^*-algebra is called *commutative*, if it holds $xy = yx$ for all $x, y \in A$. It is called *unital*, if $1 \in A$.

(vii) Let A, B be Banach algebras. A map $\varphi\colon A \to B$ is called *(algebra) homomorphism*, if it is linear and multiplicative: $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in A$.

If A and B are *-Banach algebras and if we have $\varphi(a^*) = \varphi(a)^*$ for all $a \in A$, we call φ a *-homomorphism.

If $\|\varphi(a)\| = \|a\|$, we call φ *isometric*.

Example 10.2: (i) Let H be a Hilbert space. Then $B(H)$ is a unital C^*-algebra (see Proposition 7.6). If H is finite-dimensional, then $M_n(\mathbb{C})$ is a unital C^*-algebra via $(a_{i,j})^* = (\overline{a}_{j,i})$.

(ii) Let X be a compact topological space, then $(C(X), \|\cdot\|_\infty)$ is a unital commutative C^*-algebra via $f^*(t) := \overline{f(t)}$.

Remark 10.3: (i) An involution is bijective: "${}^{*^{-1}} = {}^*$"

(ii) If A is a unital *-Banach algebra, then $1^* = 1$, because it holds

$$1^* x = (x^* 1)^* = x^{**} = x.$$

If A is a unital C^*-algebra, then $\|1\| = 1$, because it holds

$$\|1\|^2 = \|1^* 1\| = \|1 \cdot 1\| = \|1\|.$$

Because $A \neq \{0\}$, it holds $\|1\| = 1$.

[1] First introduced by Gelfand in 1943.

(iii) It holds $(x^{-1})^* = (x^*)^{-1}$, because $(x^{-1})^* x^* = (xx^{-1})^* = 1$.

(iv) It holds $\mathrm{Sp}(x^*) = \{\overline{\lambda} \mid \lambda \in \mathrm{Sp}(x)\}$, because $\lambda - x$ is invertible if and only if $(\lambda - x)^*$ is invertible.

(v) If A is a C^*-algebra, then the involution is isometric, i. e., $\|x^*\| = \|x\|$, because

$$\|x\|^2 = \|x^*x\| \leq \|x^*\|\|x\| \Rightarrow \|x\| \leq \|x^*\| \Rightarrow \|x^*\| \leq \|x^{**}\| = \|x\|.$$

(vi) If A is a unital Banach algebra and $x \in A$, then

$$r(x) := \sup\{|\lambda| \mid \lambda \in \mathrm{Sp}(x)\} \leq \|x\|$$

is the spectral radius of x. We have the formula $r(x) = \lim_{n\to\infty} \sqrt[n]{\|x^n\|}$ (see Theorem 8.13). If now A is a unital C^*-algebra and if x is normal (i. e., $x^*x = xx^*$), then $r(x) = \|x\|$.

Proof: We have

$$\|x^2\|^2 = \|(x^2)^* x^2\| = \|x^*x^*xx\| = \|x^*xx^*x\| = \|(x^*x)^*(x^*x)\| = \|x^*x\|^2 = \|x\|^4,$$

thus $\|x\|^2 = \|x^2\|$. Inductively, we see that $\|x^{2^n}\| = \|x\|^{2^n}$, hence

$$r(x) = \lim_{n\to\infty} \sqrt[2^n]{\|x^{2^n}\|} = \|x\|.$$

■

Definition 10.4: (i) A *non-commutative monomial* in $x_1, \dots, x_n$ is an expression of the form

$$x_{i_1}^{k_1} x_{i_2}^{k_2} \cdots x_{i_m}^{k_m}$$

with $k_j \in \mathbb{N}, i_j \in \{1, \dots, n\}$. Note that in general $x_1 x_2 \neq x_2 x_1$! A *non-commutative polynomial* is a $\mathbb{C}$-linear combination of non-commutative monomials.

(ii) A non-commutative monomial in x and x^* is of the form

$$x^{k_1} {x^*}^{k_2} x^{k_3} {x^*}^{k_4} \cdots x^{k_m},$$

a non-commutative polynomial in x and x^* is defined standing to reason.

Remark 10.5: If x is normal, then any non-commutative monomial is of the form $x^k {x^*}^l$ with $k, l \in \mathbb{N}_0$ (rather than $x^{k_1} {x^*}^{k_2} \dots$). Also, the algebra of polynomials in x and x^* is commutative in this case.

Theorem 10.6 (Gelfand-Mazur): *If A is a (not necessarily commutative) unital Banach algebra, which is also a skew field.*[2] *Then $A \cong \mathbb{C}$ (as an algebra).*

[2]The german term for skew field is “Schiefkörper”. A skew field is an algebra in which every element is invertible.

Proof: Let $x \in A$. Then $\mathrm{Sp}(x) \neq \varnothing$ (refer to Theorem 8.9), i. e., we find a $\lambda \in \mathrm{Sp}(x)$ and hence $\lambda 1 - x$ is not invertible. Then $\lambda 1 - x = 0$, i. e., $x = \lambda 1 \in \mathbb{C}1$. ■

Definition 10.7: Let A be an algebra. A (two-sided) ideal $I \neq A$ is called *maximal*, if for every ideal $I \subseteq J \subset A$ it holds, that $I = J$ or $J = A$.

Proposition 10.8: *Let A be a Banach algebra.*

(i) *If $I \triangleleft A$ is a closed (two-sided) ideal, then A/I is a Banach algebra.*

(ii) *If $I \subseteq A$ is a (not necessarily closed, two-sided) ideal, then also $\overline{I} \subseteq A$ is a (two-sided) ideal.*

(iii) *If $I \subseteq A$ is a(n) (two-sided) ideal and if A is a unital Banach algebra, then the following are equivalent:*

   (1) *$I = A$,*

   (2) *$1 \in I$,*

   (3) *$I \cap \mathrm{GL}(A) \neq \varnothing$.*

(iv) *If A is unital, then every (two-sided) maximal ideal is closed.*

(v) *If A is unital, then every non-trivial (two-sided) ideal is contained in a maximal ideal.*

Proof: (i) Because $I \subseteq A$ is a closed linear subspace, we know from Theorem 1.32 that A/I is a Banach space. It is an algebra via $\dot{x}\dot{y} = (xy)^\bullet$ (this is well-defined: If $a, b \in I$, then $((x + a)(y + b))^\bullet = (xy + xb + ay + ab)^\bullet = (xy)^\bullet$). And we have $\|\dot{x}\dot{y}\| \leq \|\dot{x}\|\|\dot{y}\|$: Indeed, for $\varepsilon > 0$, we find $a, b \in I$ with $\|x + a\| \leq \|\dot{x}\| + \varepsilon$, likewise for $\|\dot{y}\|$. Hence:

$$\begin{aligned}\|\dot{x}\dot{y}\| = \left\|\big((x + a)(y + b)\big)^\bullet\right\| \leq \|(x + a)(y + b)\| &\leq \|x + a\|\|y + b\| \\ &\leq (\|\dot{x}\| + \varepsilon)(\|\dot{y}\| + \varepsilon).\end{aligned}$$

(ii) If $x \in \overline{I}$, then there is a sequence $(x_n)_{n\in\mathbb{N}} \subseteq I$ such that $x_n \to x$. Then $ax_n \to ax$ for $a \in A$, thus $ax \in \overline{I}$

(iii) If $1 \in I$, then $a = a1 \in I \,\forall\, a \in A$, thus $A \subseteq I$; If $x \in I \cap \mathrm{GL}(A) \neq \varnothing$, then $1 = xx^{-1} \in I$.

(iv) Let $I \subseteq A$ be maximal. Then $I \subseteq \overline{I} \subseteq A$. Hence $I = \overline{I}$ or $\overline{I} = A$. But as $I \neq A$, we have via (iii), that $I \subseteq \mathrm{GL}(A)^\mathsf{c}$. Because $\mathrm{GL}(A)$ is open, $\overline{I} \subseteq \mathrm{GL}(A)^\mathsf{c}$, therefore $\overline{I} \neq A$.

(v) By Zorn's Lemma we find a maximal element with respect to the ordering $I_1 \subseteq I_2$ containing I. If J_α are ideals such that $I \subseteq J_\alpha$ and $1 \notin J_\alpha$, then $\bigcup_\alpha J_\alpha \subseteq A$ is again an ideal with $1 \notin \bigcup_\alpha J_\alpha$. ■

Definition 10.9: Let A be a unital Banach algebra. A homomorphism $\varphi\colon A \to \mathbb{C}$, $\varphi \neq 0$ is called a *character of A*. We put

$$\mathrm{Spec}(A) := \{\varphi\colon A \to \mathbb{C} \text{ character}\}$$

and call $\mathrm{Spec}(A)$ the *spectrum of A*.

Lemma 10.10: *Let A be a unital Banach algebra, $\varphi \in \mathrm{Spec}(A)$.*

(i) *We have $\varphi(1) = 1$.*

(ii) *It holds for all $a \in A$: $\varphi(a) \in \mathrm{Sp}(a)$, and $\varphi(x) \neq 0$ for all $x \in \mathrm{GL}(A)$.*

(iii) *φ is continuous, $\|\varphi\| \leq 1$ (with $\|\varphi\| = 1$ if $\|1\| = 1$).*

(iv) *If A is a unital C^*-algebra, then φ is a *-homomorphism and $\|\varphi\| = 1$.*

Proof: (i) Since $\varphi \neq 0$ we find $x \in A$ such that $\varphi(x) \neq 0$. Then it holds $\varphi(x) = \varphi(x1) = \varphi(x)\varphi(1)$, thus $\varphi(1)$ is the unit element (with respect to multiplication) in the Banach algebra $\mathbb{C}$, i. e., $\varphi(1) = 1$.

(ii) Let $x \in A$ be invertible. Then $1 = \varphi(1) = \varphi(xx^{-1}) = \varphi(x)\varphi(x^{-1})$, thus $\varphi(x) \neq 0$. Furthermore it holds $\varphi(\varphi(a)1 - a) = 0$, thus $\varphi(a)1 - a$ is not invertible, hence $\varphi(a) \in \mathrm{Sp}(a)$.

(iii) We have that

$$\mathrm{Sp}(A) \subseteq \{\lambda \in \mathbb{C} \mid |\lambda| \leq \|a\|\},$$

thus $|\varphi(a)| \leq \|a\|$ for all $a \in A$, i. e., $\|\varphi\| \leq 1$. If $\|1\| = 1$, then $|\varphi(1)| = \|1\| = 1$, hence $\|\varphi\| = 1$.

(iv) Since $\|1\| = 1$ holds in unital C^*-algebras, we have that $\|\varphi\| = 1$. Let now $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ with

$$\varphi(x) = \alpha + \mathrm{i}\beta, \qquad \varphi(x^2) = \gamma + \mathrm{i}\delta.$$

We now need to show, that the equalities $\alpha = \gamma$, $\beta = -\delta$ hold, because then it held that $\varphi(x^2) = \overline{\varphi(x)}$ and φ were a *-homomorphism.

Assume $\beta + \delta \neq 0$, then

$$c := \frac{x + x^* - (\alpha + \beta)^*}{\beta + \delta}$$

satisfied $\varphi(c) = \mathrm{i}$ and for arbitrary $\lambda \in \mathbb{R}$ we thus found $\varphi(c + \lambda\mathrm{i}) = (1 + \lambda)\mathrm{i}$, hence $|1 + \lambda| \leq \|c + \lambda\mathrm{i}\|$. Therefore it held that

$$\begin{aligned}1 + 2\lambda + \lambda^2 = |1 + \lambda|^2 &\leq \|c + \lambda\mathrm{i}\|^2 \\ &= \|(c + \lambda\mathrm{i})^*(c + \lambda\mathrm{i})\| \\ &= \|(c - \lambda\mathrm{i})(c + \lambda\mathrm{i})\| = \|c^2 + \lambda^2\| \leq \|c\|^2 + \lambda^2,\end{aligned}$$

thus $1 + 2\lambda \leq \|c^2\|$ for all $\lambda \in \mathbb{R}$, which is a contradiction.

If we assume $\alpha - \gamma \neq 0$ and put

$$d := \frac{\mathrm{i}x + (\mathrm{i}x)^* + 2\beta 1}{\alpha - \beta},$$

we may proceed similarly. ■

For $x, y \in A$ and $\varphi \in \mathrm{Spec}(A)$ it holds, that

$$\varphi(xy) = \varphi(x)\varphi(y) = \varphi(y)\varphi(x) = \varphi(yx),$$

characters do not "see" the commutativity or non-commutativity of A. Thus, $\mathrm{Spec}(A)$ might not contain much information about A if A is non-commutative – on Sheet 10 we saw, that $\mathrm{Spec}(M_n(\mathbb{C})) = \varnothing$.

However if A is commutative, $\mathrm{Spec}(A)$ contains a lot of information of A.

Proposition 10.11: *Let A be a commutative unital Banach algebra. Then the map*

$$\begin{aligned} \mathrm{Spec}(A) &\longrightarrow \{\textit{Maximal ideals in } A\} \\ \varphi &\longmapsto \ker(\varphi) \end{aligned}$$

is bijective.

Proof: For the surjectivity let $I \subsetneq A$ be a maximal ideal (i. e., a two-sided ideal). Then I is closed by Proposition 10.8 (iv) and A/I is a Banach algebra via Proposition 10.8 (i). Also A/I is a skew field: Let $\pi\colon A \to A/I$ be the quotient map and let $a \in A$ with $\pi(a) \neq 0$. Put

$$J := \{bx + x \mid b \in A, x \in I\} \subseteq A.$$

Then J is a two-sided ideal in A, as for $(ba + x), (b'a + x') \in J$ and $c \in A$ it holds

$$(ba + x) + (b'a + x') = (b + b')a + (x + x'), \qquad c(ba + x) = (cb)a + cx,$$

the commutativity of J is obvious as A is commutative by precondition. Furthermore it holds $I \subseteq J$ with $b = 0$ and $I \neq J$ with $b = 1$ and $x = 0$; so $a \in J$ but $a \notin I$ (since $\pi(a) \neq 0$). By the maximality of I, we infer that $J = A$, thus there are $b \in A$ and $x \in I$ such that $1 = ba + x$, which implies $\pi(b)\pi(a) = \pi(ba + x) = \pi(1) = 1$, hence $\pi(a)$ is invertible and A/I is indeed a skew field. By Theorem 10.6 it holds that $A/I \cong \mathbb{C}$ and thus

$$\pi\colon A \longrightarrow A/I \cong \mathbb{C}$$

is a charakter with $\ker(\pi) = I$.

For the injectivity let $\varphi_1, \varphi_2 \in \mathrm{Spec}(A)$ with $\ker(\varphi_1) = \ker(\varphi_2)$. Then for any $a \in A$ it holds that $\varphi_1(a)1 - a \in \ker(\varphi_1) = \ker(\varphi_2)$, thus

$$0 = \varphi_2(\varphi_1(a) - a) = \varphi_1(a) - \varphi_2(a),$$

which implies $\varphi_1 = \varphi_2$.

For the well-definedness we check, that for $\varphi \in \operatorname{Spec}(A)$, $\ker(\varphi) \subseteq A$ is indeed a maximal two-sided ideal: For any $a \in A$ it holds that

$$\varphi(ax) = \varphi(a)\varphi(x) = 0,$$

furthermore $\ker(\varphi) \neq A$ since $\varphi \neq 0$ by $\varphi(1) = 1 \neq 0$. By Proposition 10.8 (v) and the surjectivity of the map we find $\psi \in \operatorname{Spec}(A)$ with $\ker(\varphi) \subseteq \ker(\psi)$ such that $\ker(\psi)$ is maximal. Then for any $a \in A$ it holds that $\varphi(a)1 - a \in \ker(\varphi) \subseteq \ker(\psi)$, i. e.,

$$0 = \psi(\varphi(a)1 - a) = \varphi(a) - \psi(a)$$

and thus $\varphi = \psi$. ■

Corollary 10.12: *Let A be a commutative unital Banach algebra and $a \in A$.*

(i) *a is invertible if and only if $\varphi(a) \neq 0$ for all $\varphi \in \operatorname{Spec}(A)$,*

(ii) *It holds $\lambda \in \operatorname{Sp}(a)$ if and only if $\varphi(a) = \lambda$ for some $\varphi \in \operatorname{Spec}(A)$, hence*

$$\operatorname{Sp}(a) = \{\varphi(a) \mid \varphi \in \operatorname{Spec}(A)\}.$$

Proof: (i) "⇒" was shown in Lemma 10.10 (ii), for "⇐" let $a \in A \setminus \operatorname{GL}(A)$. Then

$$I := \{ba \mid b \in A\}$$

is a non-trivial (since $1 \notin I$) two-sided ideal in A and thus $I \subseteq \ker(\varphi)$ for some $\varphi \in \operatorname{Spec}(A)$, i. e., $\varphi(a) = 0$ (hint: use Proposition 10.8 (v)).

(ii) We have

$$\lambda \in \operatorname{Sp}(a) \Leftrightarrow \exists\, \varphi \in \operatorname{Spec}(A) : \varphi(\lambda - a) = 0,$$

i. e., $\varphi(a) = \lambda$. ■

Proposition 10.13: *Let A be a unital Banach algebra. Then* $\operatorname{Spec}(A)$ *is compact (with respect to pointwise convergence of characters).*

Proof: We just want to give the idea of the proof here.

- By Tychonov's theorem, every product of compact spaces is compact (Tychonovs theorem is equivalent to the axiom of choice).
- Let E be a normed space and

$$(E')_1 := \{x \in E \mid \|x\| \leq 1\} \subseteq E'$$

the closed unit ball in the dual space of E. Endow $(E')_1$ with the locally convex topology of pointwise convergence

$$\varphi_\lambda \to \varphi \Leftrightarrow \varphi_\lambda(x) \to \varphi(x)\, \forall\, x \in E.$$

Then $(E')_1$ is a closed subset of the product $\prod_{x \in E_1} \{\lambda \in \mathbb{C} \mid |\lambda \leq 1\}$ which is compact by Tychonov's theorem, hence $(E')_1$ is compact. For $x \in (E')_1$ and $y \in E_1$ we then have $|x(y)| \leq \|x\|\|y\| \leq 1$.

- If now suffices to show that $\mathrm{Spec}(A) \subseteq (A')_1$ is a closed subset: Let $(\varphi_\lambda) \subseteq \mathrm{Spec}(A)$ be a net with $\varphi_\lambda \to \varphi \in (A')_1$ pointwise. Since $\varphi_\lambda(1) = 1$ for all λ, we have that $\varphi(1) = 1$ and moreover

$$\varphi(xy) \longleftarrow \varphi_\lambda(xy) = \varphi_\lambda(x)\varphi_\lambda(y) \longrightarrow \varphi(x)\varphi(y),$$

thus $\varphi \in \mathrm{Spec}(A)$. ■

Example 10.14: Let X be a compact topological Hausdorff space. Then $C(X)$ is a commutative unital Banach algebra (even a C^*-algebra). What is $\mathrm{Spec}(C(X))$? Let $t \in X$. Put

$$\begin{aligned}\varphi_t \colon C(X) &\longrightarrow \mathbb{C}\\ f &\longmapsto f(t),\end{aligned}$$

then $\varphi_t \in \mathrm{Spec}(C(X))$ for all $t \in X$. We have:

$$\begin{aligned}\Psi \colon X &\longrightarrow \mathrm{Spec}(C(X))\\ t &\longmapsto \varphi_t\end{aligned}$$

is a homeomorphism, i. e., $X \cong \mathrm{Spec}(C(X))$ as topological spaces.

- Ψ is continuous: For a net $(t_\lambda)_{\lambda\in\Lambda}$ in X with $t_\lambda \to t$, we have

$$\varphi_{t_\lambda}(f) = f(t_\lambda) \to f(t) = \varphi_t(f)$$

for all $f \in C(X)$, hence $\varphi_{t_\lambda} \to \varphi$.

- Ψ is injective: Let $s, t \in X$ with $s \neq t$. Find $f \colon X \to \mathbb{R}$ continuous such that $f(s) \neq f(t)$ (if X is metric, put $f(y) := d(s, y)$, if X not metric, the Lemma of Urysohn grants the existence of such a function). Then

$$\varphi_s(f) = f(s) \neq f(t) = \varphi_t(f),$$

hence $\varphi_s \neq \varphi_t$.

- Ψ is surjective: We need to show that for every maximal ideal $I \subseteq C(X)$ there is $t \in X$ such that $I = \ker(\varphi_t)$. For $\varphi \in \mathrm{Spec}(C(X))$ we find $t \in X$ such that $\ker(\varphi) = \ker(\varphi_t)$ via (Theorem 10.11). As in the proof of (Theorem 10.11) we thus have $\varphi = \varphi_t$.

- Ψ^{-1} is continuous: Ψ is a continuous bijective map between compact Hausdorff spaces, hence for $A \subseteq X$ closed, A is compact. Then

$$(\Psi^{-1})^{-1}(A) = \Psi(A) \subseteq \mathrm{Spec}(C(X))$$

is compact and thus closed.

Theorem 10.15: *Let A be a unital Banach algebra. Then, the* Gelfand transformation

$$\begin{aligned} \chi\colon A &\longrightarrow C(\mathrm{Spec}(A)) \\ x &\longmapsto \hat{x}, \end{aligned}$$

where $\hat{x}(\varphi) := \varphi(x)$,

(i) *…is a continuous algebra homomorphism,*

(ii) *If A is commutative, then $\|\chi(x)\|_\infty = r(x)$ and $\hat{x}(\mathrm{Spec}(A)) = \mathrm{Sp}(x)$.*

Proof: (i) We have

$$\begin{aligned} \widehat{x+y}(\varphi) &= \varphi(x+y) = \varphi(x) + \varphi(y) = \hat{x}(\varphi) + \hat{y}(\varphi) \\ \widehat{xy}(\varphi) &= \varphi(xy) = \varphi(x)\varphi(y) = \hat{x}\hat{y}(\varphi) \end{aligned}$$

as well as

$$|\hat{x}(\varphi)| = |\varphi(x)| \leq \|x\| \,\forall\, \varphi,$$

from which we conclude $\|\chi(x)\|_\infty = \|\hat{x}\|_\infty \leq \|x\|$.

(ii) It holds $\hat{x}(\mathrm{Spec}(A)) = \{\varphi(x) \mid \varphi \in \mathrm{Spec}(A)\} = \mathrm{Sp}(x)$, hence

$$r(x) = \sup\{|\lambda| \mid \lambda \in \mathrm{Sp}(x)\} = \sup\{|\hat{x}(\varphi)| \mid \varphi \in \mathrm{Spec}(A)\} = \|\hat{x}\|_\infty \qquad \blacksquare$$

Theorem 10.16: *Let A be a commutative unital C^*-algebra. Then, the Gelfand transformation χ is even an isometric algebra isomorphism respecting the involution, i. e., it is an isometric *-isomorphism.*

Proof: First, we want to show that χ is indeed a *-homomorphism. By Lemma 10.10, φ is a *-homomorphism for all $\varphi \in \mathrm{Spec}(A)$, hence

$$\widehat{x^*}(\varphi) = \varphi(x^*) = \overline{\varphi(x)} = \overline{\hat{x}(\varphi)},$$

i. e., χ is a *-homomorphism.

Secondly, we want to show that χ is isometric: Since A is commutative, we have $x^*x = xx^*$ for all $x \in A$. Using the results from Remark 10.3 and Theorem 10.15 (ii), it then holds that $\|x\| = r(x) = \|\chi(x)\|_\infty$, i. e., χ is isometric.

Thirdly, we need to show that χ is surjective. Therefore consider $\chi(A) \subseteq C(\mathrm{Spec}(A))$. $\chi(A)$ is a *-subalgebra, as for $\hat{x}, \hat{y} \in \chi(A)$ it holds that $\hat{x}\hat{y} = \widehat{xy} \in \chi(A)$ etc.

Fourthly, $\chi(A)$ separates points, because for $\varphi, \psi \in \mathrm{Spec}(A)$ with $\varphi \neq \psi$, there is $x \in A$ such that $\varphi(x) \neq \psi(x)$, thus $\chi(x)(\varphi) = \varphi(x) \neq \psi(x) = \chi(x)(\psi)$.

Finally, $\chi(A)$ is closed, because if $(\hat{x}_n)_{n\in\mathbb{N}} \subseteq \chi(A)$ is a Cauchy sequence, then

$$\|x_n - x_m\|_\infty = \|\chi(x_n - x_m)\|_\infty = \|\hat{x}_n - \hat{x}_m\| < \varepsilon$$

for some $n, m \geq N$. Thus $(x_n)_{n\in\mathbb{N}} \subseteq A$ is a Cauchy sequence, i. e., there is $x \in A$ such that $x_n \to x$. Hence $\hat{x}_n \to \hat{x}$. Therefore, $\chi(A)$ is complete and hence closed.

Now by Stone-Weierstraß we know that $\chi(A) = C(\mathrm{Spec}(A))$. $\blacksquare$

Corollary 10.17 (First fundamental theorem of C^*-algebras[3]):

(i) *If X is a compact topological space, then $C(X)$ is a commutative unital C^*-algebra,*

(ii) *If A is a commutative unital C^*-algebra, then there is a compact, topological space X such that $A \cong C(X)$. In fact $X = \mathrm{Spec}(A)$.*

Remark 10.18: As stated in Corollary 10.17, every commutative unital C^*-algebra is of the form $C(X)$. In some sense topology corresponds to commutative C^*-algebras, hence "non-commutative topology" corresponds to C^*-algebras.

This is somehow on of the keys to view C^*-algebras. For instance: If X is not connected, then $C(X)$ admits a non-trivial projection $f \in C(X)$ ($f = f^2 = \overline{f}$, $f \neq 0$, $f \neq 1$). Hence, we might ask: *Does a given non-commutative C^*-algebra have non-trivial projections?* i. e., *Is the associated "non-commutative topological space" connected?* For more on this, see "Elements of non-commutative geometry" by Varilly, Gracia-Bondia, Figueroa.

Remark 10.19: (i) For Banach algebras, Theorem 10.15 is wrong, in the sense that χ is no isomorphism. On Sheet 10, we showed that $\ell^1(\mathbb{Z})$ is a commutative unital Banach algebra with $\mathrm{Spec}(\ell^1(\mathbb{Z})) = \mathbb{T}$, where $\mathbb{T} := \{z \in \mathbb{C} \mid |z| = 1\}$. The Gelfand transformation

$$\chi\colon \ell^1(\mathbb{Z}) \longrightarrow C(\mathbb{T})$$

is nothing but the Fourier transform:

$$\chi(x)(z) = \chi(x)(\varphi_z) = \varphi_z(x) = \sum_{n\in\mathbb{Z}} \alpha_n z^n$$

with $x = (\alpha_n)_{n\in\mathbb{Z}}$. Then χ is injective, but not surjective.

(ii) Using the Gelfand transformation, we can show the Theorem of Wiener: If $f \in \ell(\mathbb{T})$ has an absolutely convergent Fourier expansion and if $f(z) \neq 0$ for all $z \in \mathbb{T}$, then also $1/f$ has an absolutely convergent Fourier expansion (study the image of $\ell^1(\mathbb{Z})$ under χ).

If $x \in A$, we know how to form $p(x) \in A$ for a polynomial p. But what about $f(x) \in A$ for a continuous function f, e. g. $f = \sqrt{\cdot}$? Using $f \in C(\mathrm{Sp}(x)) \cong C^*(x, 1) \subseteq A$, this is a commutative if x is normal.

Definition 10.20: Let A be a unital C^*-algebra.

(i) If $M \subseteq A$ is a subset, then

$$C^*(M) := \bigcap_{B\subseteq A\ C^*-\text{subalgebra}, M\subseteq B} B$$

is the smallest C^*-subalgebra of A containing M.

[3] In literature on the topic, this Theorem is also called the "Theorem of Gelfand-Naimark" (where the spelling of Naimark may vary due to different transliterations of his russian name).

(ii) If $x \in A$, then $C^*(x,1) := C^*(\{x,1\}) \subseteq A$.

Remark 10.21: (i) If A is a unital C^*-algebra, then

$$C^*(x,1) = \overline{\{\text{non-commutative polynomials in } x \text{ and } x^*\}} \subseteq A$$

since $\overline{\{\text{non-commutative polynomials}\}} \subseteq A$ is a C^*-subalgebra. This implies that

$$C^*(x,1) \subseteq \overline{\{\text{non-commutative polynomials}\}}$$

but also

$$\overline{\{\text{non-commutative polynomials}\}} \subseteq C^*(x,1).$$

(ii) If x is normal, then $C^*(x,1)$ is commutative, since $\{\text{polynomials in } x \text{ and } x^*\}$ is commutative. If we use Corollary 10.17 in this case, we get $C^*(x,1) \cong C(\text{Spec}C^*(x,1))$, we now want to get a better understanding of $C(\text{Spec}C^*(x,1))$.

Lemma 10.22: *Let A be a unital C^*-algebra, $x \in A$ normal.*

(i) $\text{Sp}_A(y) = \text{Sp}_{C^*(x,1)}(y)$ *for all* $y \in C^*(x,1)$,

(ii) *The map*

$$\begin{aligned} \chi(x) \colon \text{Spec}(C^*(x,1)) &\longrightarrow \text{Sp}(x) \\ \varphi &\longmapsto \varphi(x) \end{aligned}$$

is bijective.

Proof: (i) If $\lambda - y$ is invertible in $C^*(x,1)$ (i. e., $(\lambda - x)^{-1} \in C^*(x,1)$), then $\lambda - y$ is invertible in A. Let now $\lambda - y$ be invertible in A and consider

$$B := C^*(x, (\lambda - y)^{-1}, 1) \subseteq A.$$

Then B is commutative and unital: Indeed $x(\lambda - y) = (\lambda - y)x$, since $C^*(x,1)$ is commutative and $x, (\lambda - y) \in C^*(x,1)$, thus $x(\lambda - y)^{-1} = (\lambda - y)^{-1}x$. Hence

$$\{\text{polynomials in } x, x^*, (\lambda - y)^{-1}, ((\lambda - y)^{-1})^*, 1\}$$

is commutative. By Corollary 10.17, the Gelfand transform $\chi_B \colon B \to C(\text{Spec}(B))$ is thus a *-isomorphism. We need to show that $\chi_B(C^*(x,1)) \subseteq C(\text{Spec}(B))$ is a closed *-subalgebra separating points. For the point-separating property: Let $\varphi, \psi \in \text{Spec}(B)$ with $\varphi|_{C^*(x,1)} = \psi|_{C^*(x,1)}$. Then it holds $\varphi(\lambda - y) = \psi(\lambda - y)$ which implies $\varphi((\lambda - y)^{-1}) = \psi((\lambda - y)^{-1})$ and thus $\varphi|_B = \varphi|_B$. Via the Theorem of Stone-Weierstraß, we see that $\chi_B(C^*(x,1)) = C(\text{Spec}(B)) = \chi_B(B)$. Because χ_B is injective, it holds that $C^*(x,1) = B \ni (\lambda - y)^{-1}$.

(ii) We know from Corollary 10.17 (ii), that $\chi(x)$ is surjective. For the injectivity: Let $\varphi, \psi \in \operatorname{Spec}(C^*(x,1))$ with $\varphi(x) = \psi(x)$. Then

$$C := \{y \in C^*(x,1) \mid \varphi(y) = \psi(y)\} \subseteq C^*(x,1)$$

is a closed *-subalgebra of A containing $\{x,1\}$. Thus $C^*(x,1) \subseteq C$, since $C^*(x,1)$ is the smallest such C^*-algebra. Hence $C = C^*(x,1)$. Therefore $\varphi = \psi$. $\chi(x)$ is continuous as $\chi(x) \in C(\operatorname{Spec}(C^*(x,1))$ and $\chi(x)^{-1}$ continuous since $\chi(x)$ is continuous between compact Hausdorff spaces. ■

Theorem 10.23 (Continuous functional calculus): *Let A be a unital C^*-algebra, $a \in A$ normal. There is a unique isometric *-isomorphism*

$$\Phi \colon C(\operatorname{Sp}(a)) \longrightarrow C^*(a,1) \subseteq A$$

such that $\Phi(\mathrm{id}) = a$ and $\Phi(1) = 1$. We write $f(a) := \Phi(f)$.

In particular, if f is a polynomial in X and X^*, then $f(a)$ is the polynomial applied to a and a^*.

Proof: *Existence*: Via Remark 10.21, we can identify $C(\operatorname{Sp}(a)) \cong C(\operatorname{Spec}(C^*(a,1))$ and via Theorem 10.15, we can identify $C(\operatorname{Spec}(C^*(a,1)) \cong C^*(a,1)$, thus there is such Φ. We have the diagram

$$\begin{array}{ccccc} C(\operatorname{Sp}(a)) & \xrightarrow{\cong} & C(\operatorname{Spec}(C^*(a,1)) & \xrightarrow{\cong} & C^*(a,1) \\ f \longmapsto f \circ \hat{a} & & & & \hat{a} \longleftarrow a \end{array}$$

thus $\mathrm{id} \mapsto \mathrm{id} \circ \hat{a} = \hat{a} \mapsto a$ or alternatively with $\varphi(a) = \lambda$: $\Phi^{-1}(a)(\lambda) = \hat{a}(\varphi) = \varphi(a) = \lambda$ and thus $\Phi^{-1}(a) = \mathrm{id}$.

Uniqueness: Let $\Psi \colon C^*(a,1) \to C(\operatorname{Sp}(a))$ be a *-isomorphism with $\Psi(a) = \mathrm{id}$, $\Psi(1) = 1$. Then

$$C = \{y \in C^*(a,1) \mid \Psi(y) = \Phi^{-1}(y)\} \subseteq A$$

is a C^*-algebra containing $\{a,1\}$ (as in Remark 10.21). Thus $C^*(a,1) = C$, i. e., $\Psi = \Phi^{-1}$. ■

Proposition 10.24: *The functional calculus for $x \in A$ normal in a C^*-algebra A has the following properties:*

(i) $(f+g)(x) = f(x)+g(x)$, $(fg)(x) = f(x)g(x)$, $\overline{f}(x) = f(x)^* \; \forall f,g \in C(\operatorname{Sp}(x))$,

(ii) $\operatorname{Sp}(f(x)) = f(\operatorname{Sp}(x)) \; \forall f \in C(\operatorname{Sp}(x))$,

(iii) *If $g \in C(f(\operatorname{Sp}(x)), f \in C(\operatorname{Sp}(x))$, then $(g \circ f)(x) = g(f(x))$,*

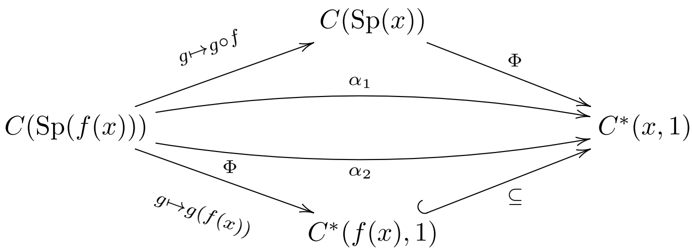
(iv) *If x is selfadjoint, then $\operatorname{Sp}(x) \subseteq \mathbb{R}$ and we may decompose $x = x_+ + x_-$ with $\operatorname{Sp}(x_+), \operatorname{Sp}(x_-) \subseteq [0,\infty)$ and $x_+x_- = 0$.*

Proof: (i) It holds $\Phi(f+g) = \Phi(f)+\Phi(g)$, $\Phi(fg) = \Phi(f)\Phi(g)$ and so on because Φ is a * homomorphism.

(ii) We have

$$\begin{aligned}\lambda \notin \mathrm{Sp}(f(x)) &\Leftrightarrow \lambda - f(x) = \Phi(\lambda 1 - f)\\ &\Leftrightarrow \lambda 1 - f \text{ is invertible}\\ &\Leftrightarrow f(\mu) \neq \lambda \,\forall\, \mu \in \mathrm{Sp}(x) \Leftrightarrow \lambda \notin f(\mathrm{Sp}(x))\end{aligned}$$

(iii) We have the diagram



and put $A := \{h \in C(\mathrm{Sp}(f(x)) \mid \alpha_1(h) = \alpha_2(h)\} \subseteq C(\mathrm{Sp}(f(x)))$. Then A is a closed *-subalgebra, separating points (as $\mathrm{id} \in A$!). Via the Theorem of Stone-Weierstraß, it holds that $A = C(\mathrm{Sp}(f(x)))$.

(iv) For $\mathrm{id} \in C(\mathrm{Sp}(x))$, we have $\Phi(\overline{\mathrm{id}}) = \Phi(\mathrm{id})^* = x^* = x = \Phi(\mathrm{id})$. As Φ is injective, it follows that $\overline{\mathrm{id}} = \mathrm{id}$ and thus $\mathrm{Sp}(x) \subseteq \mathbb{R}$. With

$$h_+(t) := \begin{cases} t & t \geq 0, \\ 0 & \text{otherwise}, \end{cases} \qquad h_-(t) := \begin{cases} -t & t \leq 0, \\ 0 & \text{otherwise}, \end{cases}$$

we have $\mathrm{id} = h_+ - h_-$. Put $x_+ := h_+(x)$ and $x_- := h_-(x)$. ■

Example 10.25: Let A be a unital C^*-algebra, $u \in A$ unitary (i. e., $u^*u = uu^* = 1$). Assume that there is $\lambda_0 \in \mathbb{S}^1$ with $\lambda_0 \notin \mathrm{Sp}(u) \subseteq \mathbb{S}^1$ (where the last inclusion is to be shown on sheet 11). Then $f(z) := \arg(z) := \theta$ for $z = e^{\mathrm{i}\theta}$ is continuous on $\mathrm{Sp}(u)$ and real valued. Thus, $x := f(u) \in A$ is selfadjoint (via Proposition 10.24) we infer: $x^* = (f(u))^* = \overline{f}(u) = f(u) = x$) and $e^{\mathrm{i}x} = u$ (since for $g(t) := e^{\mathrm{i}t}$, we have $g \circ f = \mathrm{id}$). Hence, in this case ($\mathrm{Sp}(u) \subsetneq \mathbb{S}^1$), we may write u in "polar coordinates".

Definition 10.26: Let A be a unital C^*-algebra, $x \in A$. x is called *positive*, if $x = x^*$ and $\mathrm{Sp}(x) \subseteq [0, \infty)$.

Proposition 10.27: *Every positive element in a unital C^*-algebra A admits a unique positive square root, i. e., if $x \in A$ is positive, there is exactly one $y \in A$ positive with $y^2 = x$. In particular, every positive operator in $B(H)$ has a positive square root.*

Proof: See Exercise sheet 12, exercise 2. ■

11 Spectral theorem for normal operators on Hilbert spaces

As Motivation for this chapter, we recall the spectral theorem from linear algebra:

Theorem: *Let H be a Hilbert space with $\dim(H) < \infty$ and let $A \in M_n(\mathbb{C})$ with $A = A^*$. Then A may be diagonalised, i. e.,*

$$A = \sum_{\lambda \in \mathrm{Sp}(A)} \lambda P_\lambda$$

where $\mathrm{Sp}(A) = \sigma_P(A) = \{\lambda_i \mid 1 \leq i \leq n\}$, P_λ is the projection onto the eigenspace $\ker(\lambda 1 - A)$. We have $P_\lambda \perp P_\mu$ (i. e., $P_\lambda H \perp P_\mu H$, or equivalently $P_\lambda P_\mu = 0$) if $\lambda \neq \mu$ and $H = \bigoplus_{\lambda \in \mathrm{Sp}(A)} P_\lambda H = \bigoplus_{\lambda \in \mathrm{Sp}(A)} \ker(\lambda 1 - A)$ and $\sum_{\lambda \in \mathrm{Sp}(A)} P_\lambda = 1$, hence

$$A \cong \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

If $\dim(H) = \infty$ and $A = A^* \in K(H)$, we know from Chapter 9, that we may decompose $A = \sum_{\lambda \in \mathrm{Sp}(A)} \lambda P_\lambda$, where P_λ are the projections onto $\ker(\lambda 1 - A)$ (which are finite-dimensional and non-trivial), $P_\lambda \perp P_\mu$ for $\mu \neq \lambda$ and $\mathrm{Sp}(A) \subseteq \{\text{eigenvalues of } A\} \cup \{0\}$.

Now, what happens if $A = A^* \in B(H) \setminus K(H)$? First of all we face the problem, that "$\lambda \in \mathrm{Sp}(A) \Rightarrow \lambda$ eigenvalue" does not hold, so we cannot simply take the projections P_λ onto the eigenspaces.

Our idea:

(i) Write

$$A = \int_{\mathrm{Sp}(A)} t \, dE(t),$$

where E is a "measure with values in $B(H)$", which leads to the notion of *Spectral measures*.

(ii) By Theorem 10.23, we know that $C(\mathrm{Sp}(A)) \cong C^*(A, 1) \subseteq B(H)$ for A normal, i. e., "the whole information on A is in its spectrum" (in fact "spectral theorem = diagonalise + whole info is in the spectrum"), which will lead to the extension of the continuous functional calculus to $C(\mathrm{Sp}(A)) \subseteq B_b(\mathrm{Sp}(A))$.

Example 11.1: Let $H = L^2([0, 1])$ and

$$\begin{aligned} A \colon H &\longrightarrow H \\ t &\longmapsto t f(t), f \in H \end{aligned}$$

Then $A = A^*$, as

$$\langle g, Af \rangle = \int_0^1 g(t)\overline{tf(t)}\,dt = \int_0^1 tg(t)\overline{f(t)}\,dt = \langle Ag, f \rangle,$$

$\|A\| \leq 1$, as

$$\|Af\|^2 = \int_0^1 |tf(t)|^2\,dt \leq \int_0^1 |f(t)|^2\,dt = \|f\|^2$$

and $\mathrm{Sp}(A) = [0,1]$ (as in (Example 8.6)), but $\sigma_p(A) = \varnothing$ (no eigenvalues). Hence "$A = \sum \lambda P_\lambda$" does not make sense.

If $\Delta \subseteq \mathbb{R}$ is a small interval around $\lambda \in \mathrm{Sp}(A)$, then we have $Af \approx \lambda f$, if $f|_{\Delta^c} = 0$. Consider the orthogonal projection

$$\begin{aligned} E(\Delta) \colon H &\longrightarrow H_\Delta \\ f &\longmapsto f|_\Delta \end{aligned}$$

and put $H_\Delta := \{f \in H \mid f|_{\Delta^c} = 0 \text{ almost surely}\} \subseteq H$. Then $E(\Delta)$ behaves like a measure: For example we have

- If $\Delta_1 \cap \Delta_2 = \varnothing$, then $E(\Delta_1)E(\Delta_2) = 0$,
- If $\Delta = \Delta_1 \mathbin{\dot\cup} \Delta_2$, then $E(\Delta) = E(\Delta_1) + E(\Delta_2)$,
- If $\dot{\bigcup}_{i=1}^n \Delta_i = [0,1]$, then $\sum_{i=1}^n E(\Delta_i) = 1$,
- If $\dot{\bigcup}_{i=1}^n \Delta_i = [0,1]$, then $Af = \sum_{i=1}^n AE(\Delta_i)f \approx \sum_{i=1}^n \lambda_i f$.

Hence $A \approx \int \lambda\, dE(\lambda)$. How do we get such a "spectral measure" E? Let $f = \chi_\Delta$ be the characteristic function of Δ, i. e.,

$$\chi_\Delta(t) = \begin{cases} 1 & t \in \Delta, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\chi_\Delta(A) \mathrel{\widehat{=}} \int \chi_\Delta(t)\, dE(\lambda) = E(\Delta)$. Hence, we may define $E(\Delta) := \chi_\Delta(A)$, as soon as we are allowed to use functional calculus with measurable functions (it is clear that $\chi_\Delta^2 = \chi_\Delta = \overline{\chi_\Delta}$, hence $E(\Delta)$ would then be a projection).

We want an extension

$$\begin{array}{ccc} B_b(\mathrm{Sp}(x)) & \longrightarrow & W^*(x,1) \subseteq B(H) \\ \uparrow \subseteq & & \uparrow \subseteq \\ C(\mathrm{Sp}(x)) & \longrightarrow & C^*(x,1) \end{array}$$

Definition 11.2: (i) Let X be a set and let $(f_n)_{n\in\mathbb{N}} \subseteq \{f \colon X \to \mathbb{C}\}$. $(f_n)_{n\in\mathbb{N}}$ *converges bounded pointwise* to $f \colon X \to \mathbb{C}$, if

(1) $f_n(x) \to f(x)$ for all $x \in X$,

(2) There is $C > 0$ such that for all $n \in \mathbb{N}$ it holds that $|f_n(x)| \leq C$ for all $x \in X$.

(ii) Let X be a compact metric space. Put

$$B_b(X) := \{f\colon X \to \mathbb{C} \text{ bounded, Borel measurable functions}\}.$$

(iii) Let H be a Hilbert space, $(x_\lambda)_{\lambda\in\Lambda} \subseteq B(H)$, $x \in B(H)$. We say $(x_\lambda)_{\lambda\in\Lambda}$ *converges weakly* to x ($x_\lambda \overset{w}{\to} x$) if it holds

$$\langle x_\lambda \xi, \nu\rangle \to \langle x\xi, \eta\rangle \,\forall\, \xi, \eta \in H.$$

(iv) We call $W^*(x,1) := \overline{C^*(x,1)}^w \subseteq B(H)$ for $x \in B(H)$ the *weak closure* (with respect to (iii)).

Remark 11.3: (i) The convergence $x_\lambda \overset{w}{\to} x$ is given by the locally convex toplogy $(\zeta_{\xi,\eta})_{\xi,\eta\in H}$, where $\zeta_{\xi,\eta}(x) := |\langle x\xi, \eta\rangle|$. It is called the "weak operator topology". We have

$$x_\lambda \xrightarrow{\|\cdot\|} x \Rightarrow x_\lambda \xrightarrow{w} x$$

as

$$|\langle (x_\lambda - x)\xi, \eta\rangle| \leq \|x_\lambda - x\|\|\xi\|\|\eta\| \to 0,$$

thus $C^*(x,1) \subseteq W^*(x,1) \subseteq B(H)$ is still a C^*-algebra ($\|\cdot\|$-closed *-subalgebra). In fact $W^*(x,1)$ is a *von Neumann algebra.* von Neumann algebras correspond to non-commutative measure theory as C^*-algebras correspond to non-commutative topology.

(ii) In the weak operator topology, the multiplication is not continuous, but we have "$x_\lambda \overset{w}{\to} x \Rightarrow x_\lambda y \overset{w}{\to} xy$". However, the involution is continuous.

Lemma 11.4: *Let X be a compact metric space. Then $B_b(X)$ form the smallest set $M \subseteq \{f\colon X \to \mathbb{C}\}$*

(i) *containing $C(X)$,*

(ii) *closed under bounded pointwise convergence.*

Proof: ① *It holds $C(X) \subseteq B_b(X)$, and $B_b(X)$ is closed under pointwise convergence.*

Proof (of ①): Any $f \in C(X)$ is measurable and bounded, since X is compact. For a sequence $(f_n)_{n\in\mathbb{N}}$, the pointwise limit $\lim_{n\to\infty} f_n(t) =: f(t)$ is measurable and bounded. ■

② *Let M be the smallest set with* (i) *and* (ii). *Then M is a vector space.*

Proof (of ②): Put $M_f := \{g\colon X \to \mathbb{C} \mid f + g \in M\}$ for $f\colon X \to \mathbb{C}$. Then M_f satisfies (i) and (ii), if $f \in C(X)$. Hence $M \subseteq M_f$, thus $f + M \subseteq M$ for all $f \in C(X)$. This implies that $C(X) + M \subseteq M$ and therefore M_f satisfies (i) and (ii) for all $f \in M$. It thus holds $M \subseteq M_f$ and hence $M + M \subseteq M$. Likewise $\lambda M \subseteq M$. ■

③ *For all* $f, g \in M$ *it holds that* $\max\{f, g\} \in M$.

Proof (of ③): The set $M' := \{h\colon X \to \mathbb{C} \mid |h| \in M\}$ satisfies (i) and (ii), thus $M \subseteq M'$. We now have

$$\max\{f, g\} = \frac{f + g + |f - g|}{2},$$

thus for $f, g \in M$ it holds that $\max\{f, g\} \in M$. ■

④ $\mathfrak{G} := \{A \subseteq X \mid \chi_A \in M\}$ *is a* σ*-algebra containing all open sets* $U \subseteq X$.

Proof (of ④): That $\varnothing, X \in \mathfrak{G}$ is clear. Furthermore we have that $\chi_{A^\mathsf{c}} = 1 - \chi_A \in M$, $\chi_{\cup_{i=1}^n A_i} = \max\{\chi_{A_1}, \dots \chi_{A_n}\} \in M$ and $\chi_{\cup_{i\in\mathbb{N}} A_i} \in M$ via (ii). ■

⑤ *It holds* $B_b(\lambda) \subseteq M$.

Proof (of ⑤): Let $f \in B_b(X)$. There are elementary functions $M \ni g = \sum_{i=1}^n \alpha_i \chi_{A_i}$ approximating f in the sense of (ii). ■

We thus can understand $B_b(X) = \overline{C(X)}^{(b)}$.

Theorem 11.5 ("Measurable functional calculus"): *Let* $x \in B(H)$ *be normal. The functional calculus*

$$\Phi\colon C(\mathrm{Sp}(x)) \longrightarrow C^*(x, 1) \subseteq B(H)$$

admits a unique extension

$$\Phi'\colon B_b(\mathrm{Sp}(x)) \longrightarrow W^*(x, 1) := \overline{C^*(x, 1)}^{(b)} \subseteq B(H)$$

such that Φ' *is a* **-homomorphism with* $\|\Phi'(f)\| \leq \|f\|_\infty$ *and*

$$\text{“} f_n \to f \text{ bounded pointwise} \Rightarrow \Phi'(f_n) \xrightarrow{w} \Phi'(f) \text{”.}$$

Again, we write $f(x) := \Phi'(f)$ *for all* $f \in B_b(\mathrm{Sp}(x))$.

The idea of the proof is to extend $\langle \Phi(f)\xi, \eta\rangle$, but use $\langle \Phi(f)\xi, \xi\rangle$ instead in order to apply Theorem of Fischer-Riesz: $\langle \Phi(f)\xi, \eta\rangle = \int f \, d\mu$. This integral we then want to extend to $f \in B_b$ and get back "$\langle \Phi'(f)\xi, \eta\rangle$" using the polarisation identity.

Proof: Let for the uniqueness Ψ be another *-homomorphism like Φ (so, it extends Φ, is a *-homomorphism, $\|\Psi\| \leq 1$ and is continuous with respect to the bounded / pointwise topology). Then by Lemma 11.4

$$B_b(\mathrm{Sp}(x)) \subseteq M := \{f \in B_b(\mathrm{Sp}(x)) \mid \langle \Phi'(f)\xi, \eta\rangle = \langle \Psi(f)\xi, \eta\rangle \, \forall \, \xi, \eta \in H\},$$

and as $M \subseteq B_b(\mathrm{Sp}(x))$, $\Phi = \Psi$ as desired.

For the existence let $\xi \in H$. Put $\Delta_\xi \colon C(\mathrm{Sp}(x)) \to \mathbb{C}$ via

$$\Delta_\xi(f) := \langle f(x)\xi, \xi\rangle,$$

then Δ_ξ is positive, linear and continuous. By the Theorem of Fischer-Riesz, there is a measure μ such that for all $f \in C(\mathrm{Sp}(x))$:

$$\Delta_\xi(f) = \int_{\mathrm{Sp}(x)} f(t)\, d\mu(t)\,.$$

Now put

$$\begin{aligned} \Delta'_\xi \colon B_b(\mathrm{Sp}(x)) &\longrightarrow \mathbb{C} \\ f &\longmapsto \int_{\mathrm{Sp}(x)} f(t)\, d\mu(t), \end{aligned}$$

then Δ'_ξ is a positive, linear, continuous extension of Δ_ξ with $|\Delta'_\xi(f)| \leq \|f\|_\infty \|\xi\|^2$. For $f \in B_b(\mathrm{Sp}(x))$ and $\xi, \eta \in H$, put

$$B_f(\xi, \eta) := \frac{1}{4} \sum_{k=0}^{3} \mathrm{i}^k \Delta'_{\xi + \mathrm{i}^k \eta}(f).$$

This B_f then is a sesquilinear form on H. For $f \in C(\mathrm{Sp}(x))$, B_f has the properties

$$B_f(\xi, \xi) = \langle f(x)\xi, \xi\rangle, \qquad |B_f(\xi, \eta)| \leq \|f\|_\infty \|\xi\|\|\eta\|.$$

Indeed, for the first property:

$$\begin{aligned} B_f(\xi, \xi) &= \frac{1}{4} \sum_{k=0}^{3} \mathrm{i}^k \langle f(x)(\xi + \mathrm{i}^k \xi), \xi + \mathrm{i}^k \xi\rangle \\ &= \frac{1}{4} \sum_{k=0}^{3} \mathrm{i}^k (1 + \mathrm{i}^k)\overline{(1 + \mathrm{i}^k)} \langle f(x)\xi, \xi\rangle = \langle f(x)\xi, \xi\rangle, \end{aligned}$$

and

$$|B_f(\xi, \eta)| = |\langle f(x)\xi, \xi\rangle| \leq \|f(x)\xi\|\|\xi\| \leq \|f(x)\|\|\xi\| \leq \|f\|_\infty \|\xi\|.$$

Then, for the second property, we get $|B_f(\xi, \eta)|^2 \leq |B_f(\xi, \xi)||B_f(\eta, \eta)|$.

As B_f is a sesquilinear form, $\xi \mapsto B_f(\xi, \eta)$ (for fixed $\eta \in H$) is a linear functional on H and thus by Theorem 5.20, there is $\zeta_\eta \in H$ such that $B_f(\xi, \eta) = \langle \xi, \zeta_\eta \rangle$. Now, put $T_\eta := \zeta_\eta$. Then $T \in B(H)$. Finally, we put $\Phi'(f) := T^*$. Then we have $\|\Phi'(f)\| = \|T^*\| = \|T\| \leq \|f\|_\infty$, as

$$\|T\eta\| = \|\zeta_\eta\| = \|f_{\zeta_\eta}\| = \sup_{\|\xi\|=1} |\langle \xi, \zeta_\eta \rangle| = \sup_{\|\xi\|=1} |\langle B_f(\xi, \eta) \rangle| \leq \|f\|_\infty \|\eta\|$$

Hence, for $f \in C(\mathrm{Sp}(x))$ it holds that

$$\langle f(x)\xi, \xi \rangle = \Delta_\xi(f) = \Delta'_\xi(f) = B_f(\xi, \xi) = \langle \xi, \zeta_\xi \rangle = \langle \xi, T\xi \rangle = \langle T^*\xi, \xi \rangle = \langle \Phi'(f)\xi, \xi \rangle$$

thus Φ' is an extension of Φ. We check that Φ' is a w-continuous *-homomorphism.■

Corollary 11.6 (Weak spectral theorem): *If $x = x^* \in B(H)$ (or normal), then x may be approximated in $\|\cdot\|$ by diagonal operators.*

Proof: There are numers $a, b \in \mathbb{R}$ such that $\mathrm{Sp}(x) \subseteq [a, b]$. Let $\varepsilon > 0$ and let $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$ be a partition of $[a, b]$, such that $\max_{1 \leq i \leq n} |t_i - t_{i-1}| < \varepsilon$. Then

$$\left\| \mathrm{id}_{[a,b]} - \sum_{i=1}^{n} t_{i-1} \chi|_{(t_{i-1}, t_i]} \right\|_\infty < \varepsilon$$

Put $E_i := \chi|_{(t_{i-1}, t_i]}(x)$. Then the E_i are projections and

$$\left\| x - \sum_{i=1}^{n} t_{i-1} E_i \right\| \leq \left\| \mathrm{id}_{[a,b] \cap \mathrm{Sp}(x)} - \sum_{i=1}^{n} t_{i-1} \chi_{(t_{i-1}, t_i] \cap \mathrm{Sp}(x)} \right\|_\infty < \varepsilon.$$

■

Definition 11.7: Let H be a Hilbert space and let (Y, M) be a measurable space. A map

$$E \colon M \longrightarrow \{\text{projections in } B(H)\}$$

is called a *spectral measure*, if

(i) $E(\varnothing) = 0$,

(ii) $E(Y) = 1$,

(iii) $E(\bigcup_{i \in \mathbb{N}} M_i) = \sum_{i \in \mathbb{N}} E(M_i)$ for all $M_i \in M$ mutually disjoint.

Lemma 11.8: *Let $\xi \in H$. Then*

$$\mu_\xi \colon M \longrightarrow [0, \infty)$$
$$A \longmapsto \langle E(A)\xi, \xi \rangle$$

is a measure.

Theorem 11.9: *Let $x \in B(H)$ be normal. Then $E\colon \{\text{Borel sets in } \mathrm{Sp}(x)\} \to B(H)$ defined by*

$$E(A) := \chi_A(x)$$

is a spectral measure with:

(i) *If $\xi \in H$ is an eigenvector to the eigenvalue $\lambda \in \mathbb{C}$, then for all $f \in B_b(\mathrm{Sp}(x))$: $f(x)\xi = f(\lambda)\xi$,*

(ii) *$E(\{\lambda\})$ is the orthogonal projection onto the eigenspace corresponding to $\lambda \in \mathbb{C}$. $\lambda \in \mathbb{C}$ is an eigenvalue if and only if $E(\{\lambda\}) \neq 0$.*

(iii) *If $\lambda \in \mathrm{Sp}(x)$ is isolated, then λ is an eigenvalue.*

Proof: It holds $\chi_\varnothing(x) = 0$, $\chi_{\mathrm{Sp}(x)}(x) = 1$ and $\sum_{i=1}^n \chi_{M_i} \to \chi_{\cup_{i\in\mathbb{N}} M_i}$ pointwise and bounded, thus

$$\sum_{i=1}^n E(M_i) \xrightarrow{w} \sum_{i=1}^\infty E(M_i), \qquad \sum_{i=1}^n \chi_{M_i}(x) \longrightarrow \chi_{\cup_{i\in\mathbb{N}} M_i}(x) = E\Big(\dot{\bigcup_{i\in\mathbb{N}}} M_i\Big)$$

(i) The formula is true for monomials in x and x^*, hence for polynomials, hence for $f \in C(\mathrm{Sp}(x))$, as

$$p_n \xrightarrow{\|\cdot\|} f \Rightarrow f(x)\xi \longleftarrow p_n(x)\xi = p_n(\lambda)\xi \longrightarrow f(\lambda)\xi$$

and then also for $f \in B_b(\mathrm{Sp}(x))$ by Lemma 11.4.

(ii) It holds

$$xE(\{\lambda\})\xi = \mathrm{id}_{\mathrm{Sp}(x)}(x)\chi_{\{\lambda\}}(x)\xi = (\mathrm{id}_{\mathrm{Sp}(x)}\chi_{\{\lambda\}})(x)\xi = \lambda\chi_{\{\lambda\}}(x)\xi = \lambda E(\{\lambda\})$$

thus $E(\{\lambda\})H \subseteq \mathrm{Eig}_\lambda$, if λ is an eigenvalue.

Conversely for all $\xi \in \mathrm{Eig}_\lambda$ it holds that $\chi_{\{\lambda\}}(x)\xi = \chi_{\{\lambda\}}(\lambda)\xi$, thus by (i) $\mathrm{Eig}_\lambda \subseteq E(\{\lambda\})H$.

(iii) If $\lambda \in \mathrm{Sp}(x)$ is isolated, then the function

$$f(t) := \begin{cases} 1 & t = \lambda \\ 0 & t \in \mathrm{Sp}(x) \setminus \{\lambda\} \end{cases}$$

is continuous, hence $\|E(\{\lambda\})\| = \|f(x)\| = \|f\|_\infty \neq 0$ and thus $E(\{\lambda\}) \neq 0$. ■

Corollary 11.10 (Spectral theorem): *Let $x \in B(H)$ be normal and E be the spectral measure from Theorem 11.9. Put*

$$\int_{\mathrm{Sp}(x)} f(t)\, dE(t) := z \in B(H)$$

where z is an operator given by $\langle z\xi,\xi\rangle := \int_{\mathrm{Sp}(x)} f(t)\,d\mu_\xi(t)$ and

$$\begin{aligned}\mu_\xi\colon \{\textit{Borel sets on } \mathrm{Sp}(x)\} &\longrightarrow [0,\infty)\\ A &\longmapsto \langle E(A)\xi,\xi\rangle\end{aligned}$$

is the measure from Lemma 11.8. Then $x = \int_{\mathrm{Sp}(x)} t\,dE(t)$, $f(x) = \int_{\mathrm{Sp}(x)} f(t)\,dE(t)$ for all $f \in B_b(\mathrm{Sp}(x))$ and

$$\begin{aligned}B_b(\mathrm{Sp}(x)) &\longrightarrow B(H)\\ f &\longmapsto \int f\,dE\end{aligned}$$

is a *-homomorphism.

Proof: For $f(x) = \Phi'(f)$ we have (refer to the proof of Theorem 11.5) for $f \in C(\mathrm{Sp}(x))$

$$\langle f(x)\xi,\xi\rangle = \Delta_\xi(f) = \int f\,d\mu$$

and

$$\langle E(A)\xi,\xi\rangle = \int \chi_A\,d\mu = \mu(A)$$

μ from the proof Theorem 11.5: $\mu = \mu_\xi$. ■

Example 11.11: Let $H = L^2([0,1])$ and

$$\begin{aligned}A\colon H &\longrightarrow H\\ f &\longmapsto \mathrm{id} f.\end{aligned}$$

Then

$$\begin{aligned}E\colon \{\text{Borelsets in } [0,1]\} = \mathfrak{B}([0,1]) &\longrightarrow B(H)\\ B &\longmapsto E(B)\end{aligned}$$

with $E(B)f = \chi_B(A)f = f|_B$ defines the spectral measure from Lemma 11.8. It holds

$$\mu_f(B) = \langle E(B), f\rangle = \int_0^1 \chi_B(t)f(t)\overline{f(t)}\,dt = \int_B |f(t)|^2\,dt,$$

i. e., μ_f is gives as the Lebesgue-measure with density $t \mapsto |f(t)|^2$. We thus have

$$\begin{aligned}\left\langle \left(\int_0^1 t\,dE(t)\right) f, f\right\rangle = \int_{\mathrm{Sp}(A)} t\,d\mu_f(t) &= \int_0^1 t|f(t)|^2\,dt\\ &= \int_0^1 (Af)(t)\overline{f(t)}\,dt = \langle Af, f\rangle,\end{aligned}$$

which shows $A = \int t\,dE(t)$.

12 Unbounded operators

In Chapter 1, we have seen that each linear operator $A\colon H \to H$ is bounded (i. e., continuous) if $\dim(H) < \infty$. If $\dim(H) = \infty$, this is not true anymore (although constructing "Examples" is a bit tricky). The theory of unbounded (linear) operators deals with such situations, is however not complementary to the theory of bounded operators but rather a vast generalisation thereof. In fact it also captures many interesting operators, such as differential operators and observables in quantum mechanics, that even fail to be defined on all of H. Developing this general frame is the goal of this chapter.

Example 12.1: Consider the complex Hilbert space

$$H = L^2(\mathbb{R}) = \left\{ f\colon \mathbb{R} \to \mathbb{C} \text{ measurable} : \int_{\mathbb{R}} |f(t)|^2 \, d\lambda < \infty \right\} /\mathcal{N}.$$

On the subspace $D(Q) := \{f \in L^2(\mathbb{R}) \mid \int_{\mathbb{R}} t^2 |f(t)|^2 \, d\lambda < \infty\}$ we may define the *position operator*

$$\begin{aligned} Q\colon H \supset D(Q) &\longrightarrow H \\ f &\longmapsto \mathrm{id} f. \end{aligned}$$

This linear operator fails to be bounded since $f_n := \chi_{[n,n+1]} \in D(Q)$ satisfies

$$\|f\|_2^2 = \int_{\mathbb{R}} |\chi_{[n,n+1]}(t)|^2 \, d\lambda = \int_n^{n+1} 1 \, dt = 1,$$

but

$$\|Qf\|_2^2 = \int_{\mathbb{R}} t^2 \chi_{[n,n+1]}(t) \, d\lambda = \int_n^{n+1} t^2 \, dt \geq n^2 \overset{n\to\infty}{\longrightarrow} \infty.$$

Similarly, the *momentum operator*

$$\begin{aligned} P\colon H \supset D(P) &\longrightarrow H \\ t &\longmapsto \mathrm{i} f'(t) \end{aligned}$$

is only defined on $D(P) := \{f \in C^1(\mathbb{R}) \cap L^2(\mathbb{R}) \mid f' \in L^2(\mathbb{R})\}$ and fails to be bounded since

$$\begin{aligned} f_n\colon \mathbb{R} &\longrightarrow \mathbb{C} \\ t &\longmapsto \left(\frac{n}{\pi}\right)^{\frac{1}{4}} \exp\left(-\frac{nt^2}{2}\right) \end{aligned}$$

satisfies $\|f_n\|_2 = 1$, but $\|Pf_n\|_2 = \frac{1}{2}\sqrt{2n} \to \infty$ as $n \to \infty$.

Definition 12.2: Let H be a Hilbert space.

(i) *An (unbounded) operator* T on H is given by a linear map $T\colon D(T) \to H$, where $D(T) \subseteq H$ is a linear subspace, called the *domain of* T.

(ii) If $\overline{D(T)} = H$, we say that T is *densely defined*.

(iii) $G(T) := \{(x, Tx) \mid x \in D(T)\} \subseteq H \times H$ is called the *graph* of T.

(iv) An operator S is called an *extension of* T ($T \subseteq S$) if $D(T) \subseteq D(S)$ and $Tx = Sx$ for all $x \in D(T)$ or equivalently, if $G(T) \subseteq G(S)$.

(v) T is called *closed*, if $G(T)$ is closed in $H \times H$.

(vi) T is called *closeable*, if T admits an extension S such that S is closed.

Remark 12.3: (i) If H, K are Hilbert spaces, then $H \times K$ becomes a Hilbert space ($H \oplus K$) with the inner product

$$\langle (x_1, y_1), (x_2, y_2) \rangle := \langle x_1, x_2 \rangle_H + \langle y_1, y_2 \rangle_K.$$

Clearly a sequence $(x_n, y_n)_{n \in \mathbb{N}}$ in $H \times K$ converges to (x, y), i. e., $(x_n, y_n) \to (x, y)$, if and only if $x_n \to x$ and $y_n \to y$. Thus, an operator T on H is closed if and only if it holds:

$$\text{“}(x_n \in D(T),\, x_n \to x \in H \text{ and } Tx_n \to y \in H) \Rightarrow (x \in D(T) \text{ and } Tx = y)\text{”},$$

see Remark 4.15. If T is closed and $D(T) = H$, then T is continuous by the closed graph theorem Theorem 4.16.

(ii) A linear subspace $G \subseteq H \oplus H$ is the graph of an operator if and only if $G \cap (\{0\} \times H) = \{(0, 0)\}$.

Proof: “$\Rightarrow$”: If $G = G(T)$ for some operator T, then $(x, y) \in G \cap (\{0\} \times H)$ satisfies $x = 0$ and thus $y = Tx = 0$.

“$\Leftarrow$”: Put $D(T) := \{x \in H \mid \exists y \in H : (x, y) \in G\}$ and define $T\colon D(T) \to H$ by $Tx := y$ for each $x \in D(T)$, where $y \in H$ is chosen such that $(x, y) \in G$. This is well-defined: If there are points $(x, y_1), (x, y_2) \in G$, then $(0, y_1 - y_2) \in G \cap (\{0\} \times H)$, thus $y_1 - y_2 = 0$, i. e., $y_1 = y_2$. ∎

(iii) For an operator $T\colon D(T) \to H$, the following statements are equivalent:

(1) T is closeable,

(2) The *separating space of* T, that is given by

$$\mathfrak{S}(T) := \{y \in H \mid \exists\, (x_n)_{n \in \mathbb{N}} \subseteq D(T) : x_n \to 0, Tx_n \to y\},$$

is $\{0\}$,

(3) $\overline{G(T)} \cap (\{0\} \times H) = \{(0, 0)\}$.

Proof: "(1) $\Rightarrow$ (2)": Let S with $T \subseteq S$ be closed. Take $y \in \mathfrak{S}(T)$ and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $D(T)$ with $x_n \to 0$ and $Tx_n \to y$. Since $T \subseteq S$, we have that $(x_n, Tx_n) \in G(T) \subseteq G(S)$ and thus, since $G(S)$ is closed, $(x_n, Tx_n) \to (0, y) \in G(S)$, i. e., $y = 0$.

"(2) $\Rightarrow$ (3)": If $(0, y) \in \overline{G(T)}$ is given, we find $(x_n)_{n\in\mathbb{N}}$ in $D(T)$ such that $(x_n, Tx_n) \to (0, y)$, i. e., $x_n \to 0$ and $Tx_n \to y$. Thus $y \in \mathfrak{S}(T) = \{0\}$ and therefore $y = 0$.

"(3) $\Rightarrow$ (1)": If (3) holds, then (ii) yields that $\overline{G(T)}$ is the graph of some operator $\overline{T}$ on H which is thus closed. Moreover $T \subseteq \overline{T}$ since $G(T) \subseteq \overline{G(T)} = G(\overline{T})$, hence T is closeable (Note, that $D(\overline{T}) = \overline{D(T)}^{\|\cdot\|_T}$, where the *graph norm* $\|\cdot\|_T$ is given by $\|x\|_T^2 := \|x\|_H^2 + \|Tx\|^2$). ■

(iv) If $T\colon D(T) \to H$ is closeable, there is a unique minimal closed extension $\overline{T}$ of T called the *closure of* T. We have that $G(\overline{T}) = \overline{G(T)}$.

Proof: The existence of a closed extension $\overline{T}$ of T with the property $G(\overline{T}) = \overline{G(T)}$ was established in the proof of "(3) $\Rightarrow$ (1)" in part (iii). Now, if $S \supseteq T$ is any other closed extension of T, we have $G(\overline{T}) = \overline{G(T)} \subseteq \overline{G(S)} = G(S)$ because S is an extension of T and S is closed, and hence $\overline{T} \subseteq S$. This shows the minimality of $\overline{T}$ and thus the uniqueness. ■

(v) Let $T\colon D(T) \to H$ be closeable. A subspace $D \subseteq D(T)$ is called a *core for* T, if $\overline{T|_D} = \overline{T}$.

(vi) If $T\colon D(T) \to H$ is injective we may consider $T^{-1}\colon D(T^{-1}) \to H$, where $D(T^{-1}) := \operatorname{im}(T)$. Then T is closed if and only if T^{-1} is closed (Note that $U\colon H \oplus H \to H \oplus H, (x, y) \mapsto (y, x)$ is an isometric isomorphism).

(vii) In general, the composition of closed operators (with suitable domains) is not closed. There are even examples of continuous linear operators S and closed linar operators T for which ST is not closed. Remarkably, TS is under these conditions, with suitable domain, always closed.

What about existence of adjoint operators? Even in the case of bounded operators, some work had to be done for that purpose; see (Proposition 7.4).

Theorem 12.4: *Let $T\colon D(T) \to H$ be densely defined.*

(i) *Put $D(T^*) := \{y \in H \mid x \mapsto \langle Tx, y\rangle$ is continuous on $D(T)\}$. Then $D(T^*)$ is a linear subspace of H.*

(ii) *For each $y \in D(T^*)$ there is a unique element $T^*y \in H$ such that for all $x \in D(T)$ holds:*

$$\langle Tx, y\rangle = \langle x, T^*y\rangle.$$

The induced operator $T^*\colon D(T^*) \to H, y \mapsto T^*y$ is linear.

(iii) *$G(T^*)$ is closed, i. e., T^* is a closed operator. More precisely: We have*

$$G(T^*) = V(G(T)^\perp) = \big(V(G(T))\big)^\perp$$

for a unitary operator $V\colon H \oplus H \to H \oplus H, (x, y) \mapsto (y, -x)$ which sataisfies $V^2 = -\mathrm{id}$. We have

$$H \oplus H = \overline{V(G(T))} \oplus G(T^*)$$

and $\ker(T^*) = \mathrm{im}(T)^\perp$.

(iv) *If T is closed, then $D(T^*) \subseteq H$ is dense. Then $T^{**} := (T^*)^* = T$.*

(v) *$D(T^*) \subseteq H$ is dense if and only if T is closeable. If T is closeable, then $\overline{T} = T^{**}$.*

(vi) *If $T \subseteq S$, then $S^* \subseteq T^*$.*

Proof: (i) Define $f_y\colon D(T) \to \mathbb{C}, x \mapsto \langle Tx, y\rangle$ for $y \in H$. Since for any $x \in D(T)$ and $y, z \in D(T^*)$ it holds that

$$f_{\lambda y + \mu z}(x) = \overline{\lambda} f_y(x) + \overline{\mu} f_z(x),$$

we know that $\lambda y + \mu z \in D(T^*)$ for all $\lambda, \mu \in \mathbb{C}$.

(ii) Take $y \in D(T^*)$. Since f_y is continuous and thus bounded, there is $C > 0$ so that $|f_y(x)| \leq C\|x\|$ for all $x \in D(T)$. By Theorem 1.12, f_y admits a continuous and linear extension $f_y'\colon H = \overline{D(T)} \to H$. By Theorem 5.20, we may thus find $z \in H$ such that for all $x \in H$ holds:

$$f_y'(x) = \langle x, z\rangle.$$

Then we have for all $x \in D(T)$

$$\langle Tx, y\rangle = f_y(x) = f_y'(x) = \langle x, z\rangle.$$

Suppose that there were $z_1, z_2 \in H$ satsifying $\langle x, z_1\rangle = \langle Tx, y\rangle = \langle x, z_2\rangle$ for all $x \in D(T)$. Then by the properties of the inner product, we had for all $x \in D(T)$

$$\langle x, z_1 - z_2\rangle = 0,$$

and since $\overline{D(T)} = H$, it followed that $z_1 - z_2 = 0$, i. e., $z_1 = z_2$. Thus, we may put $T^*y := z$; checking that T^* is linear is then straight forward.

(iii) We have that $V^*(x, y) = (-y, x)$ since

$$\begin{aligned}\langle V(x_1, y_2), (x_2, y_2)\rangle &= \langle (y_1, x_1), (x_2, y_2)\rangle \\ &= \langle y_1, x_2\rangle - \langle x_1, y_2\rangle = \langle (x_1, y_1), (-y_2, x_2)\rangle.\end{aligned}$$

Thus $V^*V = 1 = VV^*$, i. e., V is a unitary operator; $V^2 = -\mathrm{id}$ is clear. Furthermore:

$$\begin{aligned}(x, y) \in G(T^*) &\Leftrightarrow x \in D(T^*), y \in T^*x \\ &\Leftrightarrow \langle Tz, x\rangle = \langle z, y\rangle \, \forall\, z \in D(T) \\ &\Leftrightarrow 0 = \langle (z, Tz), (-y, x)\rangle \, \forall\, z \in D(T) \\ &\Leftrightarrow 0 = \langle V(z, Tz), (x, y)\rangle \, \forall\, z \in D(T) \\ &\Leftrightarrow (x, y) \perp V(G(T)), \text{ i. e., } (x, y) \in (V(G(T)))^\perp.\end{aligned}$$

Thus $G(T^*) = (V(G(T)))^\perp = V(G(T)^\perp)$, where the last equality holds because V is a unitary operator. Since $(V(G(T)))^\perp$ is closed (see Remark 5.13), we see that $G(T^*)$ is closed. By Theorem 5.18, we have

$$H \oplus H = \overline{V(G(T))} \oplus (V(G(T)))^\perp = \overline{V(G(T))} \oplus G(T^*).$$

As in the proof of Proposition 7.8 one shows that $\ker(T^*) = \mathrm{im}(T)^\perp$.

(iv) Let T be closed. Since V is a unitary operator that satisfies $V^2 = -\mathrm{id}$, (iii) yields that

$$H \oplus H = G(T) \oplus V(G(T^*)),$$

i. e., $G(T) = (V(G(T^*))^\perp$. Take $z \in D(T^*)^\perp$. We want to show that $z = 0$. Since for all $y \in D(T^*)$ holds that

$$\langle (0, z), V(y, T^*y)\rangle = \langle (0, z), (T^*y, -y)\rangle = -\langle z, y\rangle = 0,$$

we see that $(0, z) \in (V(G(T^*)))^\perp = G(T)$, hence $z = T(0) = 0$. Thus $D(T^*)^\perp = 0$, i. e., $\overline{D(T^*)} = H$. We may apply (iii) to T^*, which yields (since T^* is closed)

$$H \oplus H = V(G(T^*)) \oplus G(T^{**}),$$

thus $G(T^{**}) = (V(G(T^*)))^\perp = G(T)$, i. e., $T^{**} = T$.

The remaining statements (v) and (vi) are shown similarly. ∎

Remark 12.5: Let $T \in B(H)$ be normal and let E be its spectral measure. The Borel functional calculus

$$\begin{aligned} B_b(\mathrm{Sp}(T)) &\longrightarrow W^*(T, 1) \subseteq B(H) \\ f &\longmapsto \int_{\mathrm{Sp}(T)} f(t)\, dE(t)\end{aligned}$$

admits an extension to $B(\mathrm{Sp}(T))$ with values being unbounded operators: If $f \in B(\mathrm{Sp}(T))$ is given, then

$$D_f := \left\{ x \in H : \int_{\mathrm{Sp}(T)} |f(t)|^2 \, d\mu_x(t) < \infty \right\}$$

with the measure μ_x given by $\mu_x(A) = \langle E(A)x, x\rangle$ is a dense subspace of H; there is a unique operator $f(T)$ on H with $D(f(T)) = D_f$, that satisfies

$$\langle f(T)x, x\rangle = \int_{\mathrm{Sp}(T)} f(t)\, d\mu_x(t)$$

for all $x \in D_f$; see also (Theorem 11.9). We write

$$f(T) = \int_{\mathrm{Sp}(T)} f(t)\, dE(t)\,.$$

Example 12.6: The momentum operator (see Example 12.1)

$$\begin{aligned} Q\colon L^2(\mathbb{R}) \supseteq D(Q) &\longrightarrow L^2(\mathbb{R}) \\ f &\longmapsto \mathrm{id} f \end{aligned}$$

on $D(Q) := \{f \in L^2(\mathbb{R}) \mid \int_{\mathbb{R}} |tf(t)|^2\, dt < \infty\}$ is densely defined. Indeed: For all $f \in L^2(\mathbb{R})$ it holds that $f\chi_{[-N,N]} \in D(Q)$. By Theorem 12.4, the adjoint operator Q^* exists. How does it look like? We have

$$\begin{aligned} g \in D(Q^*) &\Leftrightarrow f \mapsto \langle Qf, g\rangle \text{ (continuous on } D(Q)) \\ &\Leftrightarrow \exists C > 0\, \forall f \in D(Q) : \left| \int_{\mathbb{R}} f(t)\overline{tg(t)}\, dt \right| = |\langle Qf, g\rangle| \le C\|f\|_2 \\ &\Leftrightarrow (t \mapsto tg(t)) \in L^2(\mathbb{R}) \Leftrightarrow g \in D(Q) \end{aligned}$$

and $Q^*g = Qg$, since for all $f \in D(Q)$ holds

$$\langle Qf, g\rangle \int_{\mathbb{R}} f(t)\overline{tg(t)}\, dt = \langle f, Qg\rangle,$$

thus $Q^* = Q$.

Definition 12.7: Let $T\colon D(T) \to H$ be a densely defined operator.

(i) T is called *symmetric*, if $T \subseteq T^*$,

(ii) T is called *self-adjoint*, if $T = T^*$,

(iii) T is called *maximally symmetric*, if T is symmetric and if there is a symmetric operator $T \subseteq S$, then $T = S$.

For such operators (self-adjoint or symmetric), we hope for a “nice” theory like in the bounded case.

Remark 12.8: (i) If T is symmetric, then for all $x, y \in D(T)$ it holds that

$$\langle Tx, y\rangle = \langle x, Ty\rangle.$$

(ii) Selfadjoint operators are closed by Theorem 12.4 (iii). Symmetric operators might not be closed, but they are always closeable ($T \subseteq T^*$, T^* is closed).

(iii) Selfadjoint operators are maximally symmetric (If there are operators $T = T^*$ and $S \subseteq S^*$, then $T \subseteq S \Rightarrow S^* \subseteq T^* = T \subseteq S \subseteq S^*$ holds and thus $T = S$). The converse is *not* true.

Definition 12.9: Let T be an operator on H. Then

$$\mathrm{Res}(T) := \{\lambda \in \mathbb{C} \mid \lambda - T\colon D(T) \to H \text{ invertible}\}$$

is called the *Resolvent set* and $\mathrm{Sp}(T) := \mathbb{C} \setminus \mathrm{Res}(T)$ is called the *spectrum of* T.

Lemma 12.10: *Let T be symmetric. Then we have the following statements:*

(i) $\|Tx + \mathrm{i}x\|^2 = \|Tx\|^2 + \|x\|^2 = \|Tx - \mathrm{i}x\|^2$ *for all* $x \in D(T)$,

(ii) T *is closed if and only if* $\mathrm{im}(T+\mathrm{i})$ *is closed which holds if and only if* $\mathrm{im}(T-\mathrm{i})$ *is closed,*

(iii) $T + \mathrm{i}$, $T - \mathrm{i}$ *are injective,*

(iv) *If T is also closed, then* $\mathrm{im}(\lambda - T)$ *is closed and* $\lambda - T$ *is injective for all* $\lambda \in \mathbb{C} \setminus \mathbb{R}$,

(v) *If* $\mathrm{im}(T + \mathrm{i}) = H$ *or* $\mathrm{im}(T - \mathrm{i}) = H$, *then T is maximally symmetric.*

Proof: (i) It holds that

$$\begin{aligned}\|Tx + \mathrm{i}x\|^2 &= \langle Tx, Tx\rangle + \langle Tx, \mathrm{i}x\rangle + \langle \mathrm{i}x, Tx\rangle + \langle \mathrm{i}x, \mathrm{i}x\rangle \\ &= \|Tx\|^2 - \mathrm{i}\langle Tx, x\rangle + \mathrm{i}\langle Tx, x\rangle + \|x\|^2 = \|Tx\|^2 + \|x\|^2.\end{aligned}$$

(ii) The map

$$\begin{aligned} G(T) &\longrightarrow \mathrm{im}(T + \mathrm{i}) \\ (x, Tx) &\longmapsto (T + \mathrm{i})x \end{aligned}$$

is surjective and it is isometric since

$$\|(x, Tx)\|^2 = \|x\|^2 + \|Tx\|^2 = \|(T + \mathrm{i})x\|^2.$$

Then $G(T)$ is closed if and only if $\mathrm{im}(T + \mathrm{i})$ is closed:

In general, let

$$H_1 \supseteq K \xrightarrow{\alpha} L \subseteq H_2$$

be surjective and isometric. If L is closed, then K is closed, since for a sequence $(x_n)_{n\in\mathbb{N}} \subseteq K$ with $x_n \to x \in H_1$, it holds that $L \ni \alpha(x_n) \to \alpha(x)$. As L is closed, $\alpha(x) \in L$ holds and thus $x = \alpha^{-1}(\alpha(x)) \in K$. If K is closed, then L is closed, as for a sequence $(y_n) \subseteq L$ with $y_n \to y \in H_2$, the sequence $(x_n) \subseteq K$ with $\alpha(x_n) = y_n$ is a Cauchy sequence due to the isometric property of α. Because K is closed, $(x_n) \to x \in K$, hence $\lim_{n\to\infty} \alpha(x_n) = \alpha(x) = y \in L$ and therefore, L is closed as well.

(iii) If $(T+\mathrm{i})x = 0$, then by (i), $x = 0$. Thus, $T + \mathrm{i}$ is injective.

(iv) Let T be closed and symmetric. Let $\lambda = a + \mathrm{i}b \in \mathbb{C}$ with $b \neq 0$. Let $x \in D(T)$. Then

$$\|(\lambda - T)x\|^2 \geq b^2\|x\|^2,$$

indeed:

$$\begin{aligned}\|(\lambda - T)x\|^2 &= |\lambda|^2\|x\|^2 - \lambda\langle x, Tx\rangle - \overline{\lambda}\langle Tx, x\rangle + \|Tx\|^2 \\ &= (a^2+b^2)\|x\|^2 - 2a\langle x, Tx\rangle + \|Tx\|^2 \\ &= b^2\|x\|^2 + \|(a-T)x\|^2 \geq b^2\|x\|^2.\end{aligned}$$

Hence $\lambda - T$ is injective: If $(\lambda - T)x = 0$, then $x = 0$, and $\mathrm{im}(\lambda - T)$ is closed: If $(\lambda - T)x_n \to y$, then $((\lambda - T)x_n)$ is a Cauchy-sequence and thus (x_n) is a Cauchy sequence. Therefore there is $x \in D(T)$ with $x_n \to x$ and $(\lambda - T)x_n \to (\lambda - T)x$.

(v) Let $T \subsetneq S$. Then $T + \mathrm{i} \subsetneq S + \mathrm{i}$, but then $S + \mathrm{i}$ is not injective (as $T + \mathrm{i}$ is surjective). By (iii), S cannot be symmetric. ■

Proposition 12.11: *Let T be closed and symmetric. Then the following are equivalent:*

(i) *T is selfadjoint,*

(ii) *i and $-\mathrm{i}$ are no eigenvalues of T^* (i. e., $\ker(T^* \pm \mathrm{i}) = \{0\}$),*

(iii) *$\mathrm{im}(T+\mathrm{i}) = \mathrm{im}(T - \mathrm{i}) = H$.*

Proof: "(i) ⇒ (ii)": Let $(T^* - \lambda)x = 0$ for some $\lambda \in \mathbb{C} \setminus \mathbb{R}$ (for instance, $\lambda = \pm\mathrm{i}$). Then

$$\lambda\langle x, x\rangle = \langle\lambda x, x\rangle = \langle T^*x, x\rangle = \langle x, Tx\rangle = \langle x, T^*x\rangle = \langle x, \lambda x\rangle = \overline{\lambda}\langle x, x\rangle,$$

Since $\lambda \neq 0$, we infer $\langle x, x\rangle = 0$ and thus $x = 0$.

"(ii) ⇔ (iii)": We have the equialities

$$\begin{aligned}\ker(T^* + \mathrm{i}) = \{0\} &\Leftrightarrow \mathrm{im}(T - \mathrm{i})^\perp = 0 && \text{(Theorem 12.4 (iii))} \\ &\Leftrightarrow \mathrm{im}(T - \mathrm{i}) = H && \text{(Lemma 12.10 (ii)).}\end{aligned}$$

"(iii) ⇒ (i)": We have $T \subseteq T^*$. We need to show, that $D(T^*) \subseteq D(T)$. Let $x \in D(T^*)$. Since $\mathrm{im}(T - \mathrm{i}) = H$, we find $y \in D(T) = D(T - \mathrm{i})$ such that $(T - \mathrm{i})y = (T^* - \mathrm{i})x$. Since $y \in D(T) \subseteq D(T^*) = D(T^* - \mathrm{i})$, we have $(T^* - \mathrm{i})y = (T - \mathrm{i})y = (T^* - i)x$ (because $T \subseteq T^*$). Since $T^* - \mathrm{i}$ is injective by (ii), this imlies $x = y \in D(T)$. ■

Lemma 12.12: *If T is closed, then $(\lambda - T)^{-1} \in B(H)$ for all $\lambda \in \mathrm{Res}(T)$. Moreover $\mathrm{Sp}(T) \subseteq \mathbb{C}$ is closed.*

Proof: $\lambda - T$ is closed, since for $D(\lambda - T) \supseteq (x_n)$ with $x_n \to x$ and $(\lambda - T)x_n \to y$, we have $x_n \in D(T)$ and hence $x \in D(T)$, since T is closed. Therefore we have $x \in D(\lambda - T)$ and $y \leftarrow (\lambda - T)x_n = \lambda x_n - Tx_n \to \lambda x - Tx = (\lambda - T)x$, now apply Remark 12.3 (iii).

Because $\lambda - T$ is closed, $G(\lambda - T)$ is closed and

$$G((\lambda - T)^{-1}) = \{(y, x) \mid (x, y) \in G(\lambda - T)\}$$

is closed as well. Hence, $(\lambda - T)^{-1}$ is closed.

$D((\lambda - T)^{-1}) = H$: For $\lambda \in \mathrm{Res}(T)$, we have that $(\lambda - T)\colon D(T) \to H$ is invertible, thus $(\lambda - T)^{-1}\colon H \to D(T)$ is bounded via Theorem 4.16.

Let $\lambda \in \mathrm{Res}(T)$ and $\mu \in \mathbb{C}$ with $|\lambda - \mu| < \|(\lambda - T)^{-1}\|^{-1}$. Then

$$\mu - T = ((\mu - \lambda)(\lambda - T)^{-1} + 1)(\lambda - T)$$

where $((\mu - \lambda)(\lambda - T)^{-1} + 1)$ is bounded and invertible by Lemma 8.7, as $\|1 - [(\mu - \lambda)(\lambda - T)^{-1} + 1]\| = |\lambda - \mu| \|(\lambda - T)^{-1}\| < 1$, thus $\mu \in \mathrm{Res}(T)$ and $B(\lambda, \|(\lambda - T)^{-1}\|) \subseteq \mathrm{Res}(T)$. Therefore, $\mathrm{Res}(T)$ is open and $\mathrm{Sp}(T) \subseteq \mathbb{C}$ is closed. ∎

Remark 12.13: $\mathrm{Sp}(T)$ is not compact in general. For instance $\mathrm{Sp}(Q) = \mathbb{R}$, where Q is the ... operator from Example 12.6.

Proposition 12.14: (i) *If T is selfadjoint, then* $\mathrm{Sp}(T) \subseteq \mathbb{R}$*.*

(ii) *If T is closed and symmetric and* $\mathrm{Sp}(T) \subseteq \mathbb{R}$*, then T is selfadjoint.*

Proof: (i) Let $\lambda \in \mathbb{C} \setminus \mathbb{R}$. As $T = T^*$, T is closed by Theorem 12.4 (iii). By Lemma 12.10 (iv), we know that $\lambda - T$, $\overline{\lambda} - T$ are injective, $\mathrm{im}(\lambda - T)$ is closed. We thus know that $\mathrm{im}(\lambda - T)^{\perp} = \ker(\overline{\lambda} - T) = \{0\}$ and therefore $\mathrm{im}(\lambda - T) = \overline{\mathrm{im}(\lambda - T)} = H$. We conclude, that $\lambda - T$ is injective and surjective, hence $\lambda \notin \mathrm{Sp}(T)$.

(ii) If $\mathrm{Sp}(T) \subseteq \mathbb{R}$, then $\pm\mathrm{i} \notin \mathrm{Sp}(T)$, thus $\mathrm{im}(T \pm \mathrm{i}) = H$. By Proposition 12.11 it now holds that $T = T^*$. ∎

Remark 12.15: Let T be closed and symmetric, but $T \neq T^*$. Then $\mathrm{Sp}(T) \subsetneq \mathbb{R}$ and $\mathrm{i} \in \mathrm{Sp}(T)$ or $-\mathrm{i} \in \mathrm{Sp}(T)$ (since if $\pm\mathrm{i} \notin \mathrm{Sp}(T)$, it held that $\mathrm{im}(T \pm \mathrm{i}) = H$ and thus T was selfadjoint in ths case). One can show that $\ker(\lambda - T^*)$ has constant dimension for all $\lambda \in \mathbb{C}_+ := \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}$, likewise for $\mathbb{C}_-$. Hence if $\mathrm{i} \in \mathrm{Sp}(T)$, then $\mathbb{C}_+ \subseteq \mathrm{Sp}(T)$ and if $-\mathrm{i} \in \mathrm{Sp}(T)$, then $\mathbb{C}_- \subseteq \mathrm{Sp}(T)$. Thus, there are only four possiblities for T closed and symmetric:

(i) $\mathrm{Sp}(T) = \mathbb{C}_+ \cup \mathbb{R}$ ($\mathrm{i} \in \mathrm{Sp}(T)$, $-\mathrm{i} \notin \mathrm{Sp}(T)$),

(ii) $\mathrm{Sp}(T) = \mathbb{C}_- \cup \mathbb{R}$ ($\mathrm{i} \notin \mathrm{Sp}(T)$, $-\mathrm{i} \in \mathrm{Sp}(T)$),

(iii) $\mathrm{Sp}(T) = \mathbb{C}$ ($\mathrm{i} \in \mathrm{Sp}(T)$, $-\mathrm{i} \in \mathrm{Sp}(T)$),

(iv) $\mathrm{Sp}(T) \subseteq \mathbb{R}$ (i. e., $T = T^*$).

Therefore whether or not T is selfadjoint depends on $T \pm \mathrm{i}$. In any case, $T \pm \mathrm{i}$ is injective (by Lemma 12.10) and $\mathrm{im}(T \pm \mathrm{i})$ is closed. The question whether or not $\pm\mathrm{i} \in \mathrm{Sp}(T)$ amounts to the question, whether $\mathrm{im}(T \pm \mathrm{i}) = H$.

Definition 12.16: Let T be closed and symmetric. We define

$$n_+(T) := \dim(\mathrm{im}(T+\mathrm{i})^\perp) \in [0,\infty], \qquad n_-(T) := \dim(\mathrm{im}(T-\mathrm{i})^\perp) \in [0,\infty]$$

and call $n_\pm(T)$ the *defect indices*.

Corollary 12.17: *Let T be closed and symmetric. Then T is selfadjoint if and only if $n_+ = n_- = 0$.*

We already showed this statement in Proposition 12.11.

Remark 12.18: For bounded selfadjoint operators, we have a spectral theorem. How about for unbounded selfadjoint operators? One of the occuring problems is, that $\mathrm{Sp}(T) \subseteq \mathbb{R}$ might be unbounded. Let's make the spectrum compact: Consider the mapping

$$\begin{aligned} \alpha\colon \mathbb{R} &\longrightarrow \mathbb{S}^1 \setminus \{1\} \\ t &\longmapsto \frac{t-\mathrm{i}}{t+\mathrm{i}}. \end{aligned}$$

α is bijective and has the inverse map

$$\begin{aligned} \beta\colon \mathbb{S}^{-1} &\longrightarrow \mathbb{R} \\ z &\longmapsto \mathrm{i}\frac{1+z}{1-z} \end{aligned}$$

(check that $|t-\mathrm{i}/t+\mathrm{i}| = |\overline{t+\mathrm{i}}/t+\mathrm{i}| = 1$ and that $\alpha(0) = -1$, $\alpha(1) = -\mathrm{i}$).

Definition 12.19: Let T be closed and symmetric. Define $D(U) := \mathrm{im}(T+\mathrm{i})$. The *Cayley-transform* of T then is

$$U := (T-\mathrm{i})(T+\mathrm{i})^{-1}\colon \mathrm{im}(T+\mathrm{i}) \longrightarrow \mathrm{im}(T-\mathrm{i}).$$

Remark 12.20: By Lemma 12.10, $\mathrm{im}(T \pm \mathrm{i})$ is closed and $T \pm \mathrm{i}$ is injective. Hence U is well-defined.

Theorem 12.21: *Let T be closed and symmetric and U be its Cayley-transform.*

(i) *U is an isometry, U is closed,*

(ii) $\mathrm{im}(1-U) = D(T)$,

(iii) *$1-U\colon D(U) \to D(U)$ is injective,*

(iv) $T = \mathrm{i}(1+U)(1-U)^{-1}$,

(v) *$U \in B(H)$ is unitary if and only if T is selfadjoint.*

Proof: (i) Let $y = (T+\mathrm{i})x \in \operatorname{im}(T+\mathrm{i})$ with $x \in D(T)$. Then $Uy = (T-\mathrm{i})x$. Now, by Lemma 12.10, we know that

$$\|Uy\|^2 = \|(T-\mathrm{i})x\|^2 = \|(T+\mathrm{i})x\|^2 = \|y\|^2.$$

Since $\operatorname{im}(T-\mathrm{i}) = \operatorname{im}(U)$ is closed, U is closed by Remark 12.3.

(ii) Let $y = (T+\mathrm{i})x \in D(U) = \operatorname{im}(T+\mathrm{i})$ with $x \in D(T)$. Then

$$(1-U)y = (T+\mathrm{i})x - (T-\mathrm{i})x = 2\mathrm{i}x \in D(T),$$

thus $\operatorname{im}(1-U) = D(T)$.

(iii) If $(1-U)y = 0$, then $2\mathrm{i}x = 0$, thus $x = 0$ which implies $y = 0$ (again: $y = (T+\mathrm{i})x$).

(iv) Similarly we see that $(1+U)((T+\mathrm{i})x) = 2Tx$. Hence for $x \in D(T)$ we have

$$\begin{aligned}
\mathrm{i}(1+U)(1-U)^{-1}\left(\frac{2\mathrm{i}}{2\mathrm{i}}x\right) &= \mathrm{i}(1+U)(1-U)^{-1}\left(\frac{1}{2\mathrm{i}}(1-U)(T+\mathrm{i})x\right)\\
&= \frac{1}{2}(1+U)(T+\mathrm{i})x = Tx.
\end{aligned}$$

(v) If T is selfadjoint, by Proposition 12.11 it holds that $D(U) = \operatorname{im}(T+\mathrm{i}) = H$ and $\operatorname{im}(U) = \operatorname{im}(T-\mathrm{i}) = H$. Hence $U\colon H \to H$ is isometric and surjective and therefore unitary.

Conversely, if U is unitary, then it holds $\operatorname{im}(T+\mathrm{i}) = D(U) = H$ respectively $\operatorname{im}(T-\mathrm{i}) = \operatorname{im}(U) = H$ and thus by Proposition 12.11, T is selfadjoint. ∎

Remark 12.22: Let V be a closed isometric operator on H such that $1-V$ is injective. Then V is the Cayley-transform of a closed symmetric operator defined as in Theorem 12.21 (iv).

Theorem 12.23: *Let T be closed and symmetric. Then we have*

(i) *T is selfadjoint if and only if $n_+ = n_- = 0$,*

(ii) *T is maximally symmetric if and only if $n_+ = 0$ or $n_- = 0$,*

(iii) *T has a selfadjoint extension if and only if $n_+ = n_-$.*

Proof: (i) is the statement from Corollary 12.17. As for (ii) and (iii): Let S be closed and symmetric and let $T \subseteq S$. Furthermore let V be the Cayley-transform of S. Then $U \subseteq V$. Hence, if $n_+ = 0$ or $n_- = 0$, U cannot be extended in the way which implies statement (ii).

If $n_+ = n_-$, we may find such a unitary extension V and Theorem 12.21 (v) gives statement (iii). ∎

Example 12.24: The momentum operator $(Tf)(t) = \mathrm{i} f'(t)$ on $L^2(0,1)$ is symmetric, but not selfadjoint. Using Theorem 12.23, we can see that there is a selfadjoint extension.

Theorem 12.25 (Spectral theorem for selfadoint unbounded operators): *Let T be selfadjoint. Let U be its Cayley-transform, then $U \in B(H)$ is unitary, i. e., it holds $\mathrm{Sp}(U) \subseteq \mathbb{S}^1$. Let $E\colon \{$Borel sets in $\mathrm{Sp}(U)\} \to B(H)$ be the spectral measure associated to U. Define*

$$\begin{aligned} F\colon \{\textit{Borel sets in } \mathbb{R}\} &\longrightarrow B(H) \\ A &\longmapsto E(\beta^{-1}(A)) \end{aligned}$$

with β from Remark 12.18. Then F is the spectral measure which is concentrated on $\mathrm{Sp}(T)$ (i. e., if $A \cap \mathrm{Sp}(T) = \varnothing$, then $F(A) = 0$). Put

$$S := \int_{\mathbb{R}} t \, dF_t \qquad \textit{on } D(S) := \Big\{ x \in H : \int_{\mathbb{R}} t^2 \, d\mu_x(t) < \infty \Big\}.$$

Then $S = T$, i. e., in this sense we may "diagonalise" T.