Vorlesung aus dem Sommersemester 2011

Functional Analysis

Prof. Thomas Østergaard Sørensen, Ph. D.

geTEXt von Viktor Kleen & Florian Stecker

Contents

0 Introduction **2**

1 Topological and metric spaces **2**

2 Banach and Hilbert spaces **16**

3 Lebesgue integration **32**
3.1 Measurable Functions . . . 35
3.2 p-integrable functions, $p \geq 1$ . . . 36

4 Cornerstones of functional analysis **39**
4.1 3 consequences of Baire's theorem . . . 41

5 Topics on operators **47**

0 Introduction

Example. Consider the ordinary differential equation $\frac{\mathrm{d}x}{\mathrm{d}t}(t) = f\big(t, x(t)\big)$ with boundary conditions $x(t = t_0) = x_0$ with $x\colon I \to \mathbb{R}^n$ for some interval $I \subseteq \mathbb{R}$ and $f\colon I \times \mathbb{R}^n \to \mathbb{R}^n$. By integrating both sides, we get an integral equation for x:

$$x(t) = x_0 + \int_{t_0}^{t} f\big(s, x(s)\big)\,\mathrm{d}s.$$

Define a map $K\colon C(I;\mathbb{R}^n) \to C(I;\mathbb{R}^n)$ by

$$\big(K(x)\big)(t) = x_0 + \int_{t_0}^{t} f\big(s, x(s)\big)\,\mathrm{d}s.$$

Now the integral equation becomes the fixed point equation $K(x) = x$. But K is *not* linear (consider for example $f(t, x) = t\langle x, x\rangle x$).

Example. For an example of a linear problem consider a map $k\colon [0,1]^2 \to \mathbb{R}$ and for $x\colon [0,1] \to \mathbb{R}$ let

$$(Kx)(t) = \int_0^1 k(t,s)x(s)\,\mathrm{d}s.$$

This defines a linear map $K\colon C(I;\mathbb{R}) \to C(I;\mathbb{R})$. The idea now is to study the linear map K and solutions to the equations $Kx = y$ and $Kx = \lambda x$.

1 Topological and metric spaces

We will start by generalising the concept of “continuous functions”, i.e. of “continuity”. We will first talk about the euclidean topology on $\mathbb{R}^n$. For $x, y \in \mathbb{R}^n$ let

$$\|x - y\|_{\mathrm{Eucl}} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

and let $d_{\mathrm{Eucl}}(x, y) = \|x - y\|_{\mathrm{Eucl}}$. A subset $U \subseteq \mathbb{R}^n$ is called *open* if and only if (iff) for all $x_0 \in U$ there exists $\varepsilon > 0$ such that (s.t.) if $\|y - x_0\|_{\mathrm{Eucl}} < \varepsilon$ then $y \in U$. Writing $B_\varepsilon(x_0) = \{y \in \mathbb{R}^n \colon \|x - y\|_{\mathrm{Eucl}} < \varepsilon\}$ we can say that U is open iff for all $x_0 \in U$ there exists $\varepsilon > 0$ s.t. $B_\varepsilon(x_0) \subseteq U$. In particular $B_\varepsilon(x)$ is open for all $x \in \mathbb{R}^n$ und all $\varepsilon > 0$. We denote $\mathscr{T}_{\mathrm{Eucl}}$ the family of all open subsets of $\mathbb{R}^n$:

$$\mathscr{T}_{\mathrm{Eucl}} = \{U \subseteq \mathbb{R}^n \colon U \text{ is open}\}.$$

Note that $\mathscr{T}_{\mathrm{Eucl}}$ is a *subfamily* of the powerset $2^{\mathbb{R}^n}$ of $\mathbb{R}^n$. The following should be known:

Proposition 1.1.

- *$\emptyset$ and $\mathbb{R}^n$ are open.*
- *If $U_1, U_2 \in \mathscr{T}_{\mathrm{Eucl}}$, then $U_1 \cap U_2 \in \mathscr{T}_{\mathrm{Eucl}}$.*
- *If I is some index set and $(U_i)_{i\in I}$ is a family of sets in $\mathscr{T}_{\mathrm{Eucl}}$, then $\bigcup_{i\in I} U_i \in \mathscr{T}_{\mathrm{Eucl}}$.*

We used all of this to study continuity of functions. Let $f\colon \mathbb{R}^n \to \mathbb{R}^m$ be any map. f is called *continuous at* $x_0 \in \mathbb{R}^n$ iff for all $\varepsilon > 0$ there exists $\delta > 0$ s.t. if $\|x - x_0\|_{\mathbb{R}^n} < \delta$ then $\|f(x) - f(x_0)\|_{\mathbb{R}^m} < \varepsilon$, i.e. $f\big(B_\delta(x_0)\big) \subseteq B_\varepsilon\big(f(x_0)\big)$. We say f is *continuous* iff it is continuous at all $x_0 \in \mathbb{R}^n$. Recall the following:

Proposition 1.2. *A map $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is continuous iff $f^{-1}(U)$ is open (in $\mathbb{R}^n$) for all open $U \subseteq \mathbb{R}^m$.*

We shall use Proposition 1.1 and 1.2 to generalise the concept of continuity of maps between other sets than $\mathbb{R}^n$ and $\mathbb{R}^m$.

Definition 1.3. A *topological space* $T = \{A, \mathscr{T}\}$ consist of a non-empty set A and a family $\mathscr{T}$ of subsets of A (i.e. $\mathscr{T} \subseteq 2^A$) satisfying

1. $\emptyset, A \in \mathscr{T}$.
2. If $U_1, U_2 \in \mathscr{T}$, then $U_1 \cap U_2 \in \mathscr{T}$.
3. If I is some index set and $(U_i)_{i\in I}$ is a family of sets in $\mathscr{T}$, then $\bigcup_{i\in I} U_i \in \mathscr{T}$.

Then the collection $\mathscr{T}$ of subsets of A is called a *topology* on/for A and the elements of A are called *points*. The elements of $\mathscr{T}$ are called *open sets*.

Remark.
1. In general, $\mathscr{T} \subseteq 2^A$, but $\mathscr{T} \neq 2^A$.
2. It follows, by induction, that the intersection of *finitely* many open sets is open.
3. Let $A \neq \emptyset$ and let $\mathscr{T} = \{\emptyset, A\}$. Then $\{A, \mathscr{T}\}$ is a topological space; it is called an *indiscrete space*.
4. Let $A \neq \emptyset$ and let $\mathscr{T} = 2^A$. Then $\{A, \mathscr{T}\}$ is also a topological space; it is called a *discrete space*. In particular any set (with at least two points) can be given *several* topologies.

Definition 1.6. Let $\mathscr{T}_1, \mathscr{T}_2$ be two topologies on the same set A, then we say that $\mathscr{T}_1$ is *stronger* or *finer* than $\mathscr{T}_2$ iff $\mathscr{T}_1 \supsetneq \mathscr{T}_2$, and that then $\mathscr{T}_2$ ist *weaker* or *coarser* than $\mathscr{T}_1$.

Remark.
1. Given two topologies $\mathscr{T}_1, \mathscr{T}_2$ on the same set these do *not* need to be comparable in the sense above.
2. The discrete topology (2^A) is stronger than any other topology and the indiscrete topology is ($\{\emptyset, A\}$) is weaker than any other topology.

Using the notion of topologies and inspired by Proposition 1.2, we can generalise continuity.

Definition 1.7. Given two topological spaces $T_i = \{A_i, \mathscr{T}_i\}$, $i = 1, 2$, a map $f\colon A_1 \to A_2$ is called *continuous* iff $f^{-1}(U) \in \mathscr{T}_1$ for all $U \in \mathscr{T}_2$. For emphasis, we say that f is $(\mathscr{T}_1, \mathscr{T}_2)$-continuous.

Definition 1.8. For T_1, T_2 as above, and $a \in A_1$, f is said to be *continuous at a* iff for any $U_2 \in \mathscr{T}_2$, with $f(a) \in U$, there exists a $U_1 \in \mathscr{T}_1$ s.t. $a \in U_1$ and $f(U_1) \subseteq U_2$.

Remark. f is continuous iff f is continuous at all $a \in A$.

Proposition 1.9.

1. *Let $T = \{A, \mathscr{T}\}$ be a topological space and let $\mathrm{id}\colon A \to A$ be the identity map. Then* id *is $(\mathscr{T}, \mathscr{T})$-continuous.*
2. *Any constant map $f\colon A_1 \to A_2$ is continuous.*

Proof.

1. Let $U \subseteq A$ be open. Then $\mathrm{id}^{-1}(U) = U \in \mathscr{T}$.
2. Let $U \subseteq A_2$ be open. Then if $f^{-1}(U) = \emptyset$ if $a \notin U$ and $f^{-1}(U) = A_1$ if $a \in U$. In either case $f^{-1}(U)$ is open. □

The result of Proposition 1.9 is reassuring (but not surprising); same for the next result, which however shows the strength of the definitions.

Proposition 1.10. *Let $T_i = \{A_i, \mathscr{T}_i\}$, $i = 1, 2, 3$, be three topological spaces and assume that $f\colon A_1 \to A_2$ and $g\colon A_2 \to A_3$ are continuous maps. Then $g \circ f$ is continuous.*

Proof. Let $U \in \mathscr{T}_3$. Then since g is $(\mathscr{T}_2, \mathscr{T}_3)$-continuous, $g^{-1}(U) \in \mathscr{T}_2$. Also $f^{-1}(V) \in \mathscr{T}_1$ for any $V \in \mathscr{T}_2$ since f is $(\mathscr{T}_1, \mathscr{T}_2)$-continuous. In particular $f^{-1}(g^{-1}(U)) \in \mathscr{T}_1$ — but $f^{-1}(g^{-1}(U)) = (g \circ f)^{-1}(U)$. □

Definition 1.11. Let $T = \{A, \mathscr{T}\}$ be a topological space and let $H \subseteq A$, $H \neq \emptyset$. Then the *induced topology* (or relative topology) on H is defined by

$$\mathscr{T}_H = \{V \subseteq H \colon \exists U \in \mathscr{T}.\ V = H \cap U\} = \{U \cap H \colon U \in \mathscr{T}\}.$$

Then $\{H, \mathscr{T}_H\}$ is called a *topological subspace* of $T = \{A, \mathscr{T}\}$.

Definition 1.12. Let $T_i = \{A_i, \mathscr{T}_i\}$, $i = 1, 2$, be two topological spaces. A map $f\colon A_1 \to A_2$ is called a *homeomorphism* of topological spaces iff f is a bijection and both f and f^{-1} are continuous. If such a map exists, T_1 and T_2 are called *homeomorphic.*

Definition 1.13. Let $T = \{A, \mathscr{T}\}$ be a topological space. We call $V \subseteq A$ *closed* iff $A \setminus V \in \mathscr{T}$. I.e. a set is closed if its complement is open.

Example 1.14.

1. $[a, b] \subseteq \mathbb{R}$ is closed in the euclidean topology since $\mathbb{R} \setminus [a, b] = (-\infty, a) \cup (b, \infty)$.
2. $[a, b) \subseteq \mathbb{R}$ is closed in the discrete topology. (Obviously, $[a, b) \subseteq \mathbb{R}$ is *not* closed in the euclidean topology.)
3. $[a, \infty) \subseteq \mathbb{R}$ is closed in the euclidean topology on $\mathbb{R}$ since $\mathbb{R} \setminus [a, \infty) = (-\infty, a)$ is open.

Proposition 1.15. *Let $T = \{A, \mathscr{T}\}$ be a topological space. Then*

1. *$\emptyset$ and A are closed.*
2. *The union of two (and, hence by induction any finite number) of closed sets is closed.*
3. *The intersection of any number of closed sets is closed.*

Proof. Use the definition of "topology" and de Morgan's laws. □

Definition 1.16. A *neighbourhood* of a point $x \in A$, where $T = \{A, \mathscr{T}\}$ is a topological space, is a set $V \subseteq A$ s.t. there exists $U \in \mathscr{T}$ with $x \in U \subseteq V$.

Definition 1.17. Let $T = \{A, \mathscr{T}\}$ be a topological space and let $x \in A$, and $H \subseteq A$. The point x is called a *limit point* of H iff every open set containing x, contains some point of H other than x ($U \in \mathscr{T}, U \ni x \implies U \cap (H \setminus \{x\}) \neq \emptyset$).

Example.

1. In $\{\mathbb{R}, \mathscr{T}_{\text{Eucl}}\}$ the point a is a limit point of both (a, b) and $[a, b]$ (i.e. limit points of a set may or may not belong to the set).
2. Let $H = \{0\} \cup (1, 2) \subseteq \mathbb{R}$ with the euclidean topology. Then 0 is not a limit point of H (the set of limit points of H is $[1, 2]$). Hence the points of the set may or may not be limits points of the set.

Definition 1.18. The *closure* $\overline{H}$ of $H \subseteq A$ is the union of H and its limit points.

Proposition 1.19. *$x \in \overline{H}$ iff for any open set U containing x the intersection $H \cap U$ is nonempty.*

Proposition 1.20. *Let $T = \{A, \mathscr{T}\}$ be a topological space. Then*

1. *H is closed in T iff $H = \overline{H}$.*
2. *If $H \subseteq K$ then $\overline{H} \subseteq \overline{K}$.*
3. *$\overline{\overline{H}} = \overline{H}$.*
4. *$\overline{H}$ is closed in T.*

Proof. 2 follows from the definitions. 4 follows from 1 and 3. Assume $H \subseteq A$ is closed. Since $H \subseteq \overline{H}$ by definition, we need to prove $\overline{H} \subseteq H$, or $A \setminus \overline{H} \supseteq A \setminus H$. Let $x \in A \setminus H$. Since H is closed, $A \setminus H$ is open, but $(A \setminus H) \cap H = \emptyset$. Hence, x is not a limit point of H. So $x \in A \setminus \overline{H}$. Conversely, assume $H = \overline{H}$. Let $x \in A \setminus H$. Then $x \notin H$ and x is not a limit point of H, since $H = \overline{H}$. So there exists $U_x \in \mathscr{T}$ with $U_x \ni x$ and $U_x \cap H = \emptyset$. Hence, $U_x \subseteq A \setminus H$. Then $\bigcup_{x \in A \setminus H} U_x = A \setminus H$. Since all U_x are open, $A \setminus H$ is open, hence H is closed.

Clearly $\overline{H} \subseteq \overline{\overline{H}}$. To prove $\overline{H} \supseteq \overline{\overline{H}}$, let $x \in \overline{\overline{H}}$. Then, for any open set U with $x \in U$, $\overline{H} \cap U \neq \emptyset$. Let $y \in \overline{H} \cap U$. Then U is an open set containing y, and $y \in \overline{H}$, hence $H \cap U \neq \emptyset$. Hence for any open set U containing x, $U \cap H \neq \emptyset$, hence $x \in \overline{H}$. So, $\overline{\overline{H}}$, so $\overline{H} = \overline{\overline{H}}$. □

Definition 1.21. Let $T = \{A, \mathscr{T}\}$ be a topological space, $H \subseteq A$. H is called *(everywhere) dense* in T (in A) iff $\overline{H} = A$. T (or A) is called a *separable* topological space iff it has a countable dense subset.

Example. $\mathbb{Q} \subseteq \mathbb{R}$ is dense in the euclidean topology. So is $\mathbb{R} \setminus \mathbb{Q}$. Also $\mathbb{Q}^n \subseteq \mathbb{R}^n$ is dense. Since $\mathbb{Q}^n$ is countable, $\mathbb{R}^n$ with the euclidean topology is separable. But $\mathbb{R}$ is not separable in the discrete topology.

Definition 1.22. The *interior* H° or $\operatorname{Int}(H)$ of a set $H \subseteq A$ is the union of all open subsets of H, i.e.

$$H^\circ = \bigcup \{U \subseteq H \colon U \text{ is open}\}.$$

Then $H^\circ \subseteq H$ and H° is open. It is the largest open subset of H.

Example.

1. $[a, b]^\circ = (a, b)$ in $\{\mathbb{R}, \mathscr{T}_{\text{Eucl}}\}$.
2. $[a, b]^\circ = [a, b]$ in $\{\mathbb{R}, 2^{\mathbb{R}}\}$.
3. Let $(a, b] \subseteq (-\infty, b] = H$ and give H the induced topology fro $\{\mathbb{R}, \mathscr{T}_{\text{Eucl}}\}$. Then $(a, b]^\circ = (a, b]$.
4. $\mathbb{Q}^\circ = \emptyset$ in $\{\mathbb{R}, \mathscr{T}_{\text{Eucl}}\}$.

Definition 1.23. Let $T = \{A, \mathscr{T}\}$ be a topological space. Then $H \subseteq A$ is called *nowhere dense* iff $\operatorname{Int}(\overline{H}) = \emptyset$.

Example. $\left\{\frac{1}{n} \colon n \in \mathbb{N}\right\} \subseteq \mathbb{R}$ is nowhere dense in $\{\mathbb{R}, \mathscr{T}_{\text{Eucl}}\}$.

Proposition 1.24. *$H \subseteq A$ is nowhere dense in T iff $A \setminus \overline{H}$ is (everywhere) dense in T.*

Proof. Use the fact $x \in \overline{H} \iff \forall U \in \mathscr{T} (x \in U \Longrightarrow U \cap H \neq \emptyset)$. □

Corollary 1.25. *A* closed *subset H of A is nowhere dense in T if and only if $A \setminus H$ is dense in T.*

Definition 1.26. The *boundary* ∂H of a set $H \subseteq A$ in a topological space $T = \{A, \mathscr{T}\}$ is defined by $\partial H = \overline{H} \cap (\overline{A \setminus H})$.

Definition 1.27. A sequence in a topological space $T = \{A, \mathscr{T}\}$ is a map $S \colon \mathbb{N} \to A$. We shall normally write $x_n = S(n)$ and $\{x_n\}_{n \in \mathbb{N}} \subseteq A$.

Definition 1.28. A sequence $\{x_n\}_{n \in \mathbb{N}} \subseteq A$ is said to *converge to* $x \in A$ ("$x_n \to x$ as $n \to \infty$") iff for every open set U containing x, there exists $N \in \mathbb{N}$ such that $n \geq N$ implies $x_n \in U$.

Example 1.29. Let T be any indiscrete space, let $\{x_n\}_{n \in \mathbb{N}}$ be any sequence in $T = \{A, \{\emptyset, A\}\}$, and let $x \in A$. Then $x_n \to x$ as $n \to \infty$: Let U be open in T, such that U contains x — so $U = A$. Hence, $x_n \in U$ for all $n \in \mathbb{N}$. There are not "enough" open sets in this space.

Definition 1.30. A topological space $T = \{A, \mathscr{T}\}$ is called a *Hausdorff-space* iff for all $x, y \in A$, $x \neq y$, there exist $U, V \in \mathscr{T}$, $U \cap V = \emptyset$, with $x \in U$ and $y \in V$.

Proposition 1.31. *In a Hausdorff space, limits of convergent sequences are unique, i.e. if $x_n \to x$ and $x_n \to y$ as $n \to \infty$, then $x = y$.*

Proof. Assume for contradiction that $\{x_n\}_{n\in\mathbb{N}} \subseteq A$, $T = \{A, \mathscr{T}\}$, and $x, y \in A$, $x \neq y$, with $x_n \to x$ and $x_n \to y$ as $n \to \infty$. Since T is Hausdorff and $x \neq y$ there exist $U, V \in \mathscr{T}$, $U \cap V = \emptyset$, $U \ni x$, $V \ni y$. Since $x_n \to x$ as $n \to \infty$ and $x \in U \in \mathscr{T}$, there exists $N_x \in \mathbb{N}$ such that $n \geq N_x$ implies $x_n \in U$. Also, $x_n \to y$ as $n \to \infty$ and $y \in V \in \mathscr{T}$, so there exists $N_y \in \mathbb{N}$ such that $n \geq N_y$ implies $x_n \in V$. Let $N = \max\{N_x, N_y\}$, then $n \geq N$ implies $x_n \in U \cap V = \emptyset$, a contradiction. □

Proposition 1.32.

1. *Any subspace of a Haudorff space is Hausdorff.*
2. *Let $T_i = \{A_i, \mathscr{T}_i\}$, $i = 1, 2$, be topological spaces and let $f\colon A_1 \to A_2$ be continuous. If T_2 is Hausdorff and f is injective, then T_1 is Hausdorff.*

Definition 1.33. A *metric space* $M = \{A, d\}$ consists of a nonempty set A, and a map $d\colon A \times A \to \mathbb{R}$ satisfying for $x, y, z \in A$

1. $d(x, y) \geq 0$ and $d(x, y) = 0 \iff x = y$.
2. $d(x, y) = d(y, x)$.
3. $d(x, y) \leq d(x, z) + d(z, y)$.

The map d is called a *metric* on A (or a *distance function*).

Example.

1. $\mathbb{R}^n$ with $d = d_{\text{Eucl}}$ is a metric space.
2. Let $A \neq \emptyset$ any set, and define

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

   for all $x, y \in A$. This is a metric, called the *discrete metric*.
3. Let $A = \mathbb{R}^2$, $x = (x_1, x_2) \in \mathbb{R}^2$, $y = (y_1, y_2) \in \mathbb{R}^2$ and let, for $p \geq 1$ (not necessarily $p \in \mathbb{N}$),

$$d_p(x, y) = \left(|x_1 - y_1|^p + |x_2 - y_2|^p\right)^{\frac{1}{p}}.$$

   Note, that for $p = 2$, $d_p = d_{\text{Eucl}}$. Then $\{\mathbb{R}^2, d_p\}$ is a metric space for any $p \geq 1$. Additionally let $d_\infty(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$. This is also a metric.
4. Let $A = \mathbb{C}$, $z_1, z_2 \in \mathbb{C}$, and define $d(z_1, z_2) = |z_1 - z_2|$. Then $\{\mathbb{C}, d\}$ is a metric space.

Proposition 1.34. *Let $M = \{A, d\}$ be a metric space. We will denote by $B_r(x; d) = \{y \in A\colon d(x, y) < r\}$ the open ball of radius r around x. Let*

$$\mathscr{T}_d = \{U \subseteq A\colon \forall x \in U\ \exists \varepsilon > 0.\ B_\varepsilon(x; d) \subseteq U\}.$$

Then $T = \{A, \mathscr{T}_d\}$ is a topological space.

Proof. Obviously $\emptyset, A \in \mathscr{T}_d$. Let $U_1, U_2 \in \mathscr{T}_d$, and let $x \in U_1 \cap U_2$. Then $x \in U_1 \in \mathscr{T}_d$, so there exists $\varepsilon_1 > 0$ such that $B_{\varepsilon_1}(x; d) \subseteq U_1$. Similarly, since $x \in U_2 \in \mathscr{T}_d$, there exists

$\varepsilon_2 > 0$ such that $B_{\varepsilon_2}(x; d) \subseteq U_2$. Let $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\} > 0$, then $B_\varepsilon(x; d) \subseteq U_1 \cap U_2$, so $U_1 \cap U_2 \in \mathscr{T}_d$.

Let I be some index set, and assume $U_i \in \mathscr{T}_d$, for all $i \in I$, and let $x \in \bigcup_{i\in I} U_i$, that is, there exists $i_0 \in I$ such that $x \in U_{i_0} \in \mathscr{T}_d$. Then there exists $\varepsilon > 0$ such that $B_\varepsilon(x; d) \subseteq U_{i_0} \subseteq \bigcup_{i\in I} U_i$. Hence, $\bigcup_{i\in I} U_i \in \mathscr{T}_d$. So $\mathscr{T}_d$ is a topology. □

Example. For $A = \mathbb{R}^n$, $\mathscr{T}_{\text{Eucl}} = \mathscr{T}_{d_{\text{Eucl}}}$.

In other words, any metric space is a topological space. Note however, that the converse is not true, i.e. there are topological spaces whose topology does not come from a metric:

Definition 1.35. A topological space $T = \{A, \mathscr{T}\}$ which comes from a metric space in this way (i.e. there exists a metric d such that $\mathscr{T} = \mathscr{T}_d$) is called a *metrizable space.* ($\mathscr{T}_d$ is called the topology arising from d).

Example.

1. The discrete topology on a set A comes from the discrete metric on A.
2. On the other hand, no indiscrete space with more than two points is metrizable.
3. There exist much more interesting (but also more complicated) examples of non-metrizable spaces. Sometimes it is also more useful to work directly with the topology.

Proposition 1.36. *Every metric space $\{A, d\}$ is a Hausdorff space.*

Proof. Let $x, y \in A$, $x \neq y$. Let $\varepsilon = d(x, y) > 0$. Then $B_{\varepsilon/2}(x; d) \cap B_{\varepsilon/2}(y; d) = \emptyset$, and $B_{\varepsilon/2}(x; d), B_{\varepsilon/2}(y; d) \in \mathscr{T}_d$ containing x and y respectively. □

Proposition 1.37. *Let $M = \{A, d\}$ be a metric space. A subset $H \subseteq A$ is dense iff for all $x \in A$ and $\varepsilon > 0$, $B_\varepsilon(x; d) \cap H \neq \emptyset$.*

Definition 1.38. A subset $K \subseteq A$ of a metric space $\{A, d\}$ is called bounded iff there exist $a \in A$, $R > 0$ s.t. $K \subseteq B_R(a; d)$.

Remark. If this holds for some $a \in A$, then it holds for any $\tilde{a} \in A$ with R replaced by $\tilde{R} = R + d(a, \tilde{a})$, since $d(x, \tilde{a}) \leq d(x, a) + d(a, \tilde{a}) < R + d(a, \tilde{a}) = \tilde{R}$ for all $x \in K$.

Remark. If K is bounded, and x_0, R as above, $x, y \in K$, then $d(x, y) \leq d(x, x_0) + d(x, y) < 2R < \infty$. So the following definition makes sense.

Definition 1.39. If $M = \{A, d\}$ is a metric space and $K \subseteq A$ is bounded, then the diameter $\text{diam}(K)$ of K is defined by

$$\text{diam}(K) = \sup\{d(x, y) \colon x, y \in K\}.$$

Proposition 1.40. *The union of any finite number of bounded sets is bounded.*

Proof. By induction, it is enough to proof this for 2 sets, which is left as an exercise.

Proposition 1.41. *Let $N_i = \{A_i, d_i\}$, $i = 1, 2$, be metric spaces, and let $f \colon A_1 \to A_2$ be a map. Let $T_i = \{A_i, \mathscr{T}_{d_i}\}$, $i = 1, 2$, be the corresponding topological spaces.*

1. *The map f is continuous iff*

$$\forall a \in A_1\ \forall \varepsilon > 0\ \exists \delta > 0 \big(d_1(x,a) < \delta \implies d_2(f(x), f(a))\big).$$

2. *The map f is continuous at a iff*

$$\forall \varepsilon > 0\ \exists \delta > 0 \big(d_1(x,a) < \delta \implies d_2(f(x), f(a))\big).$$

Proof. Assume f is continuous. Let $a \in A_1$, and $\varepsilon > 0$. Note that $B_\varepsilon(f(a); d_2) \subseteq A)2$ is an open set in A_2, so by assumption $f^{-1}(B_\varepsilon(f(a); d_2)) \in \mathscr{T}_{d_1}$. Since $a \in f^{-1}(B_\varepsilon(f(a); d_2))$, there exists $\delta > 0$ such that $B_\delta(a; d_1) \subseteq f^{-1}(B_\varepsilon(f(a); d_2))$, i.e. $f(B_\delta(a; d_1)) \subseteq B_\varepsilon(f(a); d_2)$.

Conversely, assume the "ε-δ-condition" holds, and let $U \in \mathscr{T}_{d_2}$. Then let $a \in f^{-1}(U)$, i.e. $f(a) \in U \in \mathscr{T}_{d_2}$, so there exists $\varepsilon > 0$ s.t. $B_\varepsilon(f(a); d_2) \subseteq U$. So, by assumption there exists $\delta > 0$ s.t. $f(B_\delta(a; d_1)) \subseteq B_\varepsilon(f(a); d_2) \subseteq U$. Hence, $B_\delta(a; d_1) \subseteq f^{-1}(U)$. The proof of 2. is left as an exercise. ☐

Definition 1.42. Let $M = \{A, d\}$ be a metric spaces, $X \neq \emptyset$, and let $f\colon X \to A$ be a map. Then f is called *bounded* iff $f(X) \subseteq A$ is bounded.

Example 1.43. Let $A = \mathbb{R}^n$, let $p \in (1, \infty)$ and let

$$d_p(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{\frac{1}{p}}.$$

Then $\{\mathbb{R}^n, d_p\}$ are metric spaces. Also, let $d_\infty(x,y) = \max\{|x_i - y_i| \colon i = 1, \dots, n\}$. Then $\{\mathbb{R}^n, d_\infty\}$ is also a metric space. This will be proven in the tutorials. Also prove *Hölder's inequality*: For $p \in (1, \infty)$, $x, y \in \mathbb{R}^n$

$$\left|\sum_{i=1}^{n} x_i y_i\right| \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |x_i|^q\right)^{\frac{1}{q}}$$

where $\frac{1}{p} + \frac{1}{q} = 1$. This implies *Minkowski's inequality*: For $p \in (1, \infty)$, $x, y \in \mathbb{R}^n$

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}}$$

Generalising this to "infinite coordinates", let $\mathscr{M}(\mathbb{N}; \mathbb{R})$ be the set of real sequences, i.e. maps $\mathbb{N} \to \mathbb{R}$. We would like to define $d_p(x,y)$ for $x = \{x_n\}_{n\in\mathbb{N}}, y = \{y_n\}_{n\in\mathbb{N}} \in \mathscr{M}(\mathbb{N}; \mathbb{R})$. However "often" $d_p(x,y) = \infty$. The solution is to restrict to a subset of $\mathscr{M}(\mathbb{N}; \mathbb{R})$. Define

$$\ell_p = \left\{\{x_n\}_{n\in\mathbb{N}} \in \mathscr{M}(\mathbb{N}; \mathbb{R}) \colon \sum_{i=1}^{n} |x_i|^p < \infty\right\}$$

Note that $\ell_p \subsetneq \mathscr{M}(\mathbb{N}; \mathbb{R})$. Also let

$$\ell_\infty = \{\{x_n\}_{n\in\mathbb{N}} \in \mathscr{M}(\mathbb{N}; \mathbb{R}) \colon \exists K \in \mathbb{R}\ \forall n \in \mathbb{N}.\ |x_n| \leq K\}$$

be the set of bounded real sequences. Note that $\ell_\infty = \mathscr{B}(\mathbb{N};\mathbb{R})$, so (ℓ_∞, d_∞) is a metric space. For $1 < p < \infty$, $x, y \in \ell_p$, fix $N \in \mathbb{N}$, then

$$t_N := \left(\sum_{i=1}^{N} |x_i - y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{N} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{N} |y_i|^p\right)^{\frac{1}{p}} \leq$$
$$\leq \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{\frac{1}{p}} < \infty$$

since $x, y \in \ell_p$. Furthermore $\{t_N\}_{N\in\mathbb{N}}$ is increasing and bounded above, so $\{t_N\}$ is convergent and $d_p(x,y) < \infty$ is well-defined, i.e. $d_p\colon \ell_p \times \ell_p \to \mathbb{R}$. Note, that $d_p(x,y) \geq 0$, $d_p(x,x) = 0$ and $d_p(x,y) = 0$ implies $x = y$. For the triangle inequality, let $x, y, z \in \ell_p$, let $N \in \mathbb{N}$, then

$$s_N := \left(\sum_{i=1}^{N} |x_i - y_i|^p\right)^{\frac{1}{p}} = \left(\sum_{i=1}^{N} |x_i - z_i + z_i - y_i|^p\right)^{\frac{1}{p}} \leq$$
$$\leq \left(\sum_{i=1}^{N} |x_i - z_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{N} |z_i - y_i|^p\right)^{\frac{1}{p}} \leq d_p(x,z) + d_p(z,y)$$

Since $\{s_N\}_{N\in\mathbb{N}}$ is increasing and bounded above, it follows that $d_p(x,y) \leq d_p(x,z) + d_p(z,y)$. Hence, $\{\ell_p, d_p\}$, $1 < p \leq \infty$, are metric spaces. Note that $\ell_p \neq \ell_q$ for $p \neq q$.

Definition 1.44. Let $A \neq \emptyset$. A family $\mathscr{U} \subseteq 2^A$ is called a *cover* for A iff $A = \bigcup \mathscr{U}$. Let $T = \{A, \mathscr{T}\}$ be a topological space. A cover $\mathscr{U} \subseteq 2^A$ is called *open* iff $\mathscr{U} \subseteq \mathscr{T}$. A *subcover* $\mathscr{V}$ of a cover $\mathscr{U}$ is a subfamily $\mathscr{V} \subseteq \mathscr{U}$ such that $A = \bigcup \mathscr{V}$. A topological space is called *compact* ("is a compact space") iff every open cover has a finite subcover.

Definition 1.45. A subset $H \subseteq A$ where $T = \{A, \mathscr{T}\}$ is a topological space, is compact iff $\{H, \mathscr{T}_H\}$ is compact, where $\mathscr{T}_H$ is the induced topology.

Proposition 1.46. *A subset of $\mathbb{R}^n$ is compact in the euclidean topology iff it is both closed and bounded.*

Proposition 1.47. *Any compact set in a metric space is bounded.*

Proof. Exercise.

Proposition 1.48. *Let $T = \{A, \mathscr{T}\}$ be a Hausdorff space, and assume $C \subseteq A$ is compact. Then C is closed.*

Proof. Exercise.

Remark. In particular, in a metric space any compact set is closed and bounded.

Proposition 1.49. *Let $T = \{A, \mathscr{T}\}$ be a compact topological space, and assume $C \subseteq A$ is closed. Then C is compact.*

Proof. Assume $\mathscr{U} \subseteq \mathscr{T}$ covers C, i.e. $C \subseteq \bigcup \mathscr{U}$. Since C is closed, the set $U_0 = A \setminus C$ is open. So $\mathscr{V} = \mathscr{U} \cup \{U_0\}$ is an open cover of A. Since A is compact, there exists a finite subcover $\mathscr{W} \subseteq \mathscr{V}$. Then $\mathscr{W} \setminus \{U_0\} \subseteq \mathscr{U}$ is finite and covers C. □

Proposition 1.50. *Let $f\colon A_1 \to A_2$ be a continuous map between topological spaces $T_i = \{A_i, \mathscr{T}_i\}$, $i = 1, 2$. If T_1 is compact, then the image $\mathrm{R}\, f$ is compact.*

Proof. It is enough to look at the case when f is surjective. Let $\mathscr{U}$ be an open cover of A_2: $A_2 \subseteq \bigcup \mathscr{U}$ and $\mathscr{U} \subseteq \mathscr{T}_2$. Since f is continuous, $f^{-1}(V) \in \mathscr{T}_1$ for all $V \in \mathscr{U}$. Also, $\bigcup_{V \in \mathscr{U}} f^{-1}(V) = A_1$. Hence $\{f^{-1}(V)\colon V \in \mathscr{U}\}$ is an open cover of A_1. Since A_1 is compact, there exist $f^{-1}(V_1), \ldots, f^{-1}(V_k)$ such that $A_1 = \bigcup_{i=1}^{k} f^{-1}(V_i)$. Then $A_2 = \bigcup_{i=1}^{k} V_i$, and so $\{V_1, \ldots, V_k\}$ is a finite subcover of $\mathscr{U}$ for A_2. □

Corollary 1.51. *Let $\{C, \mathscr{T}\}$ be a compact space, $\{A, d\}$ a metric space, and assume $f\colon C \to A$ is continuous. Then f is bounded.*

Corollary 1.52. *Assume $\{C, \mathscr{T}\}$ is a compact space, and that $f\colon C \to \mathbb{R}$ is continuous. Then f attains its bounds, i.e there exist $x_N, x_M \in C$ such that $f(x_N) \leq f(x) \leq f(x_M)$ for all $x \in C$.*

Definition 1.53. Let $M = \{A, d\}$ be a metric space. A sequence $\{x_n\} \subseteq A$ is called a *Cauchy sequence* (is said “to be Cauchy”) iff

$$\forall \varepsilon > 0\, \exists N \in \mathbb{N}\, (n, m \geq N \implies d(x_n, x_m) < \varepsilon)\,.$$

Lemma 1.54. *Any convergent sequence is Cauchy.*

Proof. Assume $\{x_n\} \subseteq A$ is convergent in a metric space $M = \{A, d\}$, $x_n \to x$ as $n \to \infty$. Let $\varepsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $d(x_n, x) < \frac{\varepsilon}{2}$ for all $n \geq N$. So if $n, m \geq N$, then $d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \varepsilon$. □

Remark. Not all Cauchy sequences are convergent, for example $\left\{\frac{1}{n}\right\}_{n \in \mathbb{N}} \subseteq (0, 1]$ in $\{(0, 1], d_{\mathrm{Eucl}}\}$ is Cauchy but not convergent.

Definition 1.55. Let $M = \{A, d\}$ be a metric space. M is called *complete* iff every Cauchy sequence in M is convergent in M.

Proposition 1.56. $\{\mathbb{R}^n, d_{\mathrm{Eucl}}\}$ *is a complete metric space. So is* $\{\mathbb{C}, |\cdot|\}$.

Lemma 1.57. *Let $\{A, d\}$ be a metric space. Then $K \subseteq A$ is closed iff for any sequence $\{x_n\} \subseteq K$, $x_n \to x$ as $n \to \infty$ implies $x \in K$.*

Proof. Problem 5 of Sheet 2.

Remark. $\{x_{n_k}\}_{k \in \mathbb{N}}$ is a *subsequence* of $\{x_n\}_{n \in \mathbb{N}}$ — formally $S\colon \mathbb{N} \to A, n \mapsto s(n) \equiv x_n$ — is defined by an *injective*, *increasing* function $\varphi\colon \mathbb{N} \to \mathbb{N}$ so $S \circ \varphi\colon \mathbb{N} \to A, k \mapsto x_{n_k}$.

Lemma 1.58. *In a metric space $\{A, d\}$, if the Cauchy sequence $\{x_n\} \subseteq A$ has a convergent subsequence $\{x_{n_k}\}$, say, $x_{n_k} \to x$ as $k \to \infty$, then $\{x_n\}$ also converges to x.*

Proof. Let $\varepsilon > 0$, and choose $N \in \mathbb{N}$ s.t. $n, m \geq N$ implies $d(x_n, x_m) < \frac{\varepsilon}{2}$. Also, choose $K \in \mathbb{N}$ s.t. $k \geq K$ implies $d(x_{n_k}, x) < \frac{\varepsilon}{2}$. For any $n \geq N$, choose $k \geq K$ so large that $n_k \geq N$. Then, for $n \geq N$, $d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) < \varepsilon$. $\square$

Definition 1.59. A subspace C of a metric space $M = \{A, d\}$ is called *sequentially compact* in itself (in M) if and only if every sequence in C has a subsequence which converges in C (in M).

Theorem 1.60. *A subspace C of a metric space is compact iff it is sequentially compact in itself.*

Proof. Later.

Corollary 1.61. *Any bounded sequence in $\{\mathbb{R}^d, d_{\text{Eucl}}\}$ has a convergent subsequence.*

Proof. Let $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^d$ be a bounded sequence, and let $S = \{x_n \colon n \in \mathbb{N}\} \subseteq \mathbb{R}^d$. Then $\overline{S}$ is bounded. So, $\overline{S}$ is closed and bounded, hence compact, hence sequentially compact. So $\{x_n\} \subseteq \overline{S}$ has a convergent subsequence. $\square$

Proposition 1.62. *Any compact metric space is complete.*

Proof. Let $\{x_n\} \subseteq A$ be Cauchy in a metric space $M = \{A, d\}$. Since M is compact, it is sequentially compact in itself, hence $\{x_n\}$ has a convergent subsequence. So, by 1.58, $\{x_n\}$ is convergent. Hence M is complete. $\square$

Proposition 1.63. *Let $M = \{A, d\}$ be a metric space and $H \subseteq A$. Then*
1. *if $\widetilde{M} = \{H, d\}$ is complete, then H is closed in M.*
2. *if M is complete, and $H \subseteq A$ is closed, then $\widetilde{M}$ is complete.*

Proof.
1. Let $x \in \overline{H}$. Then there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq H$ such that $x_n \to x$ as $n \to \infty$. Since $\{x_n\}$ is convergent, it is Cauchy. Since $\{x_n\} \subseteq \widetilde{M}$ is Cauchy, it is convergent with limit in H. By uniqueness of limits, this limit is x. So, $x \in H$, hence $H = \overline{H}$.
2. By assumption, $H = \overline{H}$. Let $\{x_n\}_{n\in\mathbb{N}} \subseteq H$ be a Cauchy sequence in $\widetilde{M}$. But then $\{x_n\}_{n\in\mathbb{N}} \subseteq A$ is a Cauchy sequence in M. Since M is complete, there is an $x \in A$ such that $x_n \to x$ as $n \to \infty$. Since H is closed and $\{x_n\} \subseteq H$, it follows that $x \in H$. Hence $\widetilde{M}$ is complete. $\square$

Proposition 1.64. *Let X be any set, and let $M = \{A, d\}$ be a metric space. Denote by $\mathscr{B}(X, A)$ the set of bounded maps $X \to A$, and let*

$$d_\infty(f, g) = \sup_{x\in X} d\big(f(x), g(x)\big).$$

Then $\{\mathscr{B}(X, A), d_\infty\}$ is complete iff $\{A, d\}$ is complete.

Proof. Assume $M = \{A, d\}$ is not complete. Take any non-convergent Cauchy sequence $\{x_n\}_{n\in\mathbb{N}}$. Let, for $n \in \mathbb{N}$, $f_n \colon X \to A, t \mapsto x_n$. Then $d_\infty(f_n, f_m) = d(x_n, x_m)$, so $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathscr{B}(X, A)$ is Cauchy in d_∞. But $\{f_n\}_{n\in\mathbb{N}}$ is not convergent, since if $d_\infty(f_n, f) \to 0$ as $n \to \infty$ for some $f \in \mathscr{B}(X, A)$ then, since $d(f_n(t), f(t)) \le d_\infty(f_n, f)$ for all $t \in X$, $x_n = f_n(t) \to f(t)$ as $n \to \infty$. But $\{x_n\}$ is not convergent.

On the other hand, assume $M = \{A, d\}$ is complete. Let $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathscr{B}(X, A)$ be any Cauchy sequence. Let $\varepsilon > 0$. Since $\{f_n\}$ is Cauchy, there exists $N \in \mathbb{N}$ such that $n, m \ge N$ implies $d_\infty(f_n, f_m) < \varepsilon$. Hence, for any $x \in X$ fixed, $d(f_n(x), f_m(x)) \le d_\infty(f_n, f_m) < \varepsilon$. So $\{f_n(x)\}_{n\in\mathbb{N}}$ is Cauchy in d. Since $M = \{A, d\}$ is complete , $\{f_n(x)\}_{n\in\mathbb{N}}$ is convergent. Let $f(x) = \lim_{n\to\infty} f_n(x)$. Then $f \colon X \to A$. We need to prove that $f \in \mathscr{B}(X, A)$ and that $d_\infty(f_n, f) \to 0$ as $n \to \infty$. Since, for all $a \in A$ fixed, the map $A \to \mathbb{R}, x \mapsto d(a, x)$ is continuous, it follows that $\lim_{m\to\infty} d(f_n(x), f_m(x)) = d(f_n(x), f(x))$. Hence, $d(f_n(x), f(x)) \le \varepsilon$ for $n \ge N$ and all $x \in X$. Then $d_\infty(f_n, f) \le \varepsilon$. Hence $f \in \mathscr{B}(X, A)$ and $d_\infty(f_n, f) \to 0$ as $n \to \infty$. Hence, the Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ is convergent to an element in $\mathscr{B}(X, A)$. So $\{\mathscr{B}(X, A), d_\infty\}$ is complete. □

Example. ℓ_∞ is complete.

Definition 1.65. Let $M_i = \{A_i, d_i\}$, $i = 1, 2$, be two metric spaces, and define

$$\mathscr{C}(A_1, A_2) = \{f \colon A_1 \to A_2 \colon f \text{ is } (d_1, d_2)\text{-continuous}\}$$
$$\mathscr{C}_b(A_1, A_2) = \{f \in \mathscr{C}(A_1, A_2) \colon f \text{ is bounded}\}$$

Then $\mathscr{C}_b(A_1, A_2) \subseteq \mathscr{C}(A_1, A_2)$ and $\mathscr{C}_b(A_1, A_2) \subseteq \mathscr{B}(A_1, A_2)$. Also, if $\{A_1, d_1\}$ is compact, then $\mathscr{C}_b(A_1, A_2) = \mathscr{C}(A_1, A_2)$ (for example $\mathscr{C}([0, 1], \mathbb{R})$).

Theorem 1.66. *Let $M_i = \{A_i, d_i\}$, $i = 1, 2$, be two metric spaces. Then $\{\mathscr{C}_b(A_1, A_2), d_\infty\}$ is a complete metric space iff $\{A_2, d_2\}$ is complete.*

Proof. If $\{A_2, d_2\}$ is not complete, then neither is $\{\mathscr{C}_b(A_1, A_2), d_\infty\}$ (same proof as in 1.64). On the other hand, assume $\{A_2, d_2\}$ is complete, and let $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathscr{C}_b(A_1, A_2)$ be Cauchy in d_∞. Since then $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy in $\mathscr{B}(A_1, A_2)$ which is complete, there exists $f \in \mathscr{B}(A_1, A_2)$ such that $d_\infty(f_n, f) \to 0$ as $n \to \infty$. We shall prove that f is continuous at a for all $a \in A_1$. Let $a \in A_1$, $\varepsilon > 0$. Let $N \in \mathbb{N}$ such that $n \ge N$ implies $d_\infty(f_n, f) < \varepsilon$. Then $n \ge N$ implies $d_2(f_n(x), f(x)) < \varepsilon$ for all $x \in A_1$. Since f_N is continuous at A, so there exists $\delta > 0$ such that $d_1(x, a) < \delta$ implies $d_2(f_N(x), f_N(a)) < \varepsilon$. Hence, $d_1(x, a) < \delta$ implies $d_2(f(x), f(a)) \le d_2(f(x), f_N(x)) + d_2(f_N(x), f_N(a)) + d_2(f_N(a), f(a)) < 3\varepsilon$. □

Definition 1.67. A map $f \colon A_1 \to A_2$ for a metric space $M_i = \{A_i, d_i\}$, $i = 1, 2$, is *uniformly continuous* (on A_1) iff

$$\forall \varepsilon > 0\, \forall x \in A_1\, \exists \delta > 0 \big(d_1(x, y) < \delta \implies d_2(f(x), f(y)) < \varepsilon\big).$$

Proposition 1.68. *A continuous map on a compact space is uniformly continuous.*

Proof. Let $f \colon A_1 \to A_2$ be a continuous map between metric spaces $M_i = \{A_i, d_i\}$, $i = 1, 2$, and assume M_1 is compact. Then, for any $\varepsilon > 0$, there exists $\delta(x) > 0$

such that $d_1(x,y) < 2\delta(x)$ implies $d_2(f(x), f(y)) < \varepsilon$. Then $\mathscr{U} = \{B_{\delta(x)}(x; d_1) \colon x \in A_1\}$ is an open cover of A_1. Since M_1 is compact, there exist $x_1, \ldots, x_N$ such that $A_1 = \bigcup_{i=1}^N B_{\delta(x_i)}(x_i; d)$. Let $\delta = \min\{\delta(x_1), \ldots, \delta(x_N)\} > 0$ and let $x, y \in A_1$ such that $d_1(x,y) < \delta$. Then there is $i_0 \in \{1, \ldots, N\}$ such that $x \in B_{\delta(x_{i_0})}(x_{i_0}; d)$, so $d_1(x, x_{i_0}) < \delta(x_{i_0}) < 2\delta(x_{i_0})$, hence $d_2(f(x), f(x_{i_0})) < \varepsilon$. Also, $d(y, x_{i_0}) \leq d_1(y,x) + d_1(x, x_{i_0}) < \delta(x_{i_0}) + \delta(x_{i_0}) = 2\delta(x_{i_0})$, hence, $d_2(f(y), f(x_{i_0})) < \varepsilon$. So $d_2(f(x), f(y)) \leq d_2(f(x), f(x_{i_0})) + d_2(f(x_{i_0}), f(y)) < 2\varepsilon$. □

Definition 1.69. A metric space $M = \{A, d\}$ is called *totally bounded* or *pre-compact* iff for all $\varepsilon > 0$ there exist finitely many $x_1, \ldots, x_N \in A$ such that $A \subseteq \bigcup_{i=1}^N B_\varepsilon(x_i; d)$.

Theorem $1.\widetilde{60}$. *Let $M = \{A, d\}$ be a metric space, $C \subseteq A$. Then the following are equivalent:*
(a) C is compact.
(b) C is sequentially compact.
(c) C is complete and totally bounded.

Proof.

$(a) \Rightarrow (b)$ Let $\{x_n\} \subseteq C$ be any sequence. Let $S_k = \overline{\{x_n \colon n \geq k\}}$. Then S_k is closed and $\bigcap_{k=1}^\infty S_k \neq \emptyset$, for assume otherwise and let $U_k = (A \smallsetminus S_k) \cap C$. Then U_k is open in the relative topology on C, and $\bigcup_{k=1}^\infty U_k = C \cap \bigcup_{k=1}^\infty S_k^c = C \cap (\bigcap_{k=1}^\infty S_k)^c = C$. So $C = U_1 \cup \cdots \cup U_N$ for some N.Then $C \cap S_1 \cap \ldots S_N = \emptyset$ which is impossible. Then $\{x_n\}$ has a convergent subsequence.

$(b) \Rightarrow (c)$ Let $\{x_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence in C. Then $\{x_n\}_{n\in\mathbb{N}}$ has convergent subsequence, say, $x_{n_k} \to x$ as $k \to \infty$, with $x \in C$ since C is sequentially compact. So by 1.58, $\{x_n\}$ is also convergent, with the same limit. Hence, C is complete.
Assume that C is not totally bounded. Then there exists $\varepsilon_0 > 0$ such that for no choice of finitely many points $\{x_1, \ldots, x_N\}$ do we have $C \subseteq \bigcup_{i=1}^N B_{\varepsilon_0}(x_i; d)$. In particular for all $x \in A$, $C \smallsetminus B_{\varepsilon_0}(x; d) \neq \emptyset$. Let $y_1 \in C$ be arbitrary. Define inductively $y_n \in C$ such that $y_n \in C \smallsetminus \bigcup_{i=1}^{n-1} B_{\varepsilon_0}(y_i; d)$. Then $\{y_i\}_{i\in\mathbb{N}} \subseteq C$, and, for any $m \neq k$, $d(y_m, y_k) > \varepsilon_0$. Hence, no subsequence of $\{y_i\}_{i\in\mathbb{N}}$ will converge — a contradiction, since C is sequentially compact.

$(c) \Rightarrow (a)$ Assume that there exists an open cover $\mathscr{U} = \{U_i\}_{i\in I}$ of C with no finite subcover. We will construct, inductively, a sequence of open balls B_n, with radii 2^{-n} and centres x_n. Since C is totally bounded, there exists $\{y_1, \ldots, y_M\} \in C$ such that $C \subseteq \bigcup_{i=1}^M B_{1/2}(y_i; d)$. Then at least one of the $B_{1/2}(y_i; d)$'s cannot be covered by finitely many U_i's (otherwise, so could C). Let B_1 be one of these balls; $B_1 = B_{1/2}(x_1; d)$. Assume now $B_{n-1} = B_{2^{1-n}}(x_{n-1}; d)$ chosen, for some $n \geq 2$. Again, there exist $\{z_1, \ldots, z_K\}$ such that $C \subseteq \bigcup_{i=1}^K B_{2^{-n}}(z_i; d)$; of all the $B_{2^{-n}}(z_i; d)$ which have non-empty intersection with B_{n-1}, at least one cannot be covered by finitely many U_i's. So let B_n be such a ball, so $B_n = B_{2^{-n}}(x_n; d)$, $B_n \cap B_{n-1} \neq \emptyset$ and B_n cannot be covered by finitely many U_i's. This gives a sequence $\{x_n\}_{n\in\mathbb{N}}$ which is Cauchy: For $y \in B_n \cap B_{n-1}$, $d(x_{n-1}, x_n) \leq d(x_{n-1}, y) + d(y, x_n) < 2^{1-n} + 2^{-n} < 2^{2-n}$. So, for $m > n$, $d(x_n, x_m) \leq d(x_n, x_{n+1}) + \cdots + d(x_{m-1}, x_m) < 2^{2-n} + \cdots + 2^{2-m} < 8 \cdot \frac{1}{2^n}$. Hence, $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy, so convergent, i.e. there exists $x \in C$ such that $x_n \to x$

as $n \to \infty$. Since $\mathscr{U}$ is an open cover for C, there exists U_{i_0} such that $x \in U_{i_0}$, and some $r > 0$ such that $B_r(x; d) \subseteq U_{i_0}$. Since $x_n \to x$ as $n \to \infty$, there exists $N \in \mathbb{N}$ such that $m \geq N$ implies $d(x, x_m) < \frac{r}{2}$. Choose m such that $2^{-m} < \frac{r}{2}$. Then $B_m = B_{2^{-m}}(x_m; d) \subseteq B_r(x; d) \subseteq U_{i_0}$ — a contradiction to the construction of the B_n's: none of the B_n's can be covered by finitely many balls. □

Theorem 1.70 (Arzelà-Ascoli)**.** *Let $\{A_1, d_1\}$ be a compact metric space and $\{A_2, d_2\}$ a complete metric space. $M \subseteq \mathscr{C}(A_1, A_2)$ is compact iff the following holds:*
(a) For all $x \in A_1$, the set $M(x) = \{f(x) \colon f \in M\} \subseteq A_2$ is compact.
(b) M is equicontinuous*, i.e.*

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall x, y \in A_1\, \forall f \in M \big(d_1(x, y) < \delta \implies d_2(f(x), f(y)) < \varepsilon\big)$$

(c) M is closed.

Proof.
"⇒" Assume M is compact. Then it is closed. Note that for $f, g \in (A_1, A_2)$ and $x \in A_1$, $d_2\big(f(x), g(x)\big) \leq d_\infty(f, g)$. So $\phi_x \colon \mathscr{C}(A_1, A_2) \to A_2, f \mapsto f(x)$ is (d_∞, d_2)-continuous. Since $M \subseteq \mathscr{C}(A_1, A_2)$ is compact, the set $M(x) = \phi_x(M)$ is compact. Let $\varepsilon > 0$. Then there exists $\{B_{\varepsilon/3}(f_1; d_\infty), \ldots, B_{\varepsilon/3}(f_N; d_\infty)\}$ such that $M \subseteq \bigcup_{i=1}^N B_{\varepsilon/3}(f_i; d_\infty)$. Now each $f_i \colon A_1 \to A_2$ is uniformly continuous since $\{A_1, d_1\}$ is compact, so there exists $\delta > 0$ such that $d_1(x, y) < \delta$ implies $d_2(f_j(x), f_j(y)) < \varepsilon/3$ for $j = 1, \ldots, N$. Let $f \in M$, and $x, y \in A_1$ with $d_1(x, y) < \delta$. Then there exists $j_0 \in \{1, \ldots, N\}$ such that $f \in B_{\varepsilon/3}(f_{j_0}; d_\infty)$. So $d_2(f(x), f(y)) \leq d_2(f(x), f_{j_0}(x)) + d_2(f_{j_0}(x), f_{j_0}(y)) + d_2(f_{j_0}(y), f(y)) < \varepsilon$. Hence, M is equicontinuous.
"⇐" Since $M \subseteq \mathscr{C}(A_1, A_2)$ is closed, and $\{\mathscr{C}(A_1, A_2), d_\infty\}$ is complete, $\{M, d_\infty\}$ is complete. Let $\varepsilon > 0$, and choose $\delta > 0$ such that $d_1(x, y) < \delta$ implies $d_2\big(f(x), f(y)\big) < \varepsilon/4$ for all $f \in M$. Since A_1 is compact, there exist $x_1, \ldots, x_N \in A_1$ such that $A_1 \subseteq \bigcup_{j=1}^N B_\delta(x_j; d_1)$. Similarly, since all $M(x_i)$, $i = 1, \ldots, N$, are compact, there exists $y_1, \ldots, y_P \in A_2$ such that $B = \bigcup_{i=1}^N M(x_i) \subseteq \bigcup_{k=1}^P B_{\varepsilon/4}(y_k; d_2)$. Let $\Phi = \{\phi \colon \{1, \ldots, N\} \to \{1, \ldots, P\}\}$. Then for any $\phi \in \Phi$ define $M_\phi = \{f \in M \colon d_2(f(x_j), y_{\phi(j)}) < \varepsilon/4 \text{ for } j = 1, \ldots, N\}$. Then $M = \bigcup_{\phi \in \Phi} M_\phi$. Let $\phi \in \Phi$, and $f, g \in M_\phi$. For all $x \in A_1$, there exists $j \in \{1, \ldots, N\}$ such that $d_1(x, x_j) < \delta$. Then $d_2(f(x), f(x_j)) < \varepsilon/4$ and $d_2(g(x), g(x_j)) < \varepsilon/4$. So, $d_2(f(x), g(x)) \leq d_2\big(f(x), f(x_j)\big) + d_2\big(f(x_j), y_{\phi(j)}\big) + d_2\big(y_{\phi(j)}, g(x_j)\big) + d_2\big(g(x_j), g(x)\big) < \varepsilon$. Hence, $d_\infty(f, g) \leq \varepsilon$. So M_ϕ is contained in a ball of radius 2ε. Hence, since $|\Phi| < \infty$, M is contained in a union of finitely many balls of radius 2ε. Hence, M is totally bounded. Hence, by 1.60, M is compact. □

Theorem 1.71 (Baire's theorem)**.** *Let $M = \{A, d\}$ be a complete metric space, and let $\{V_n\}_{n \in \mathbb{N}}$ be a countable family of dense, open subsets $V_n \subseteq A$. Then $\bigcap_{n=1}^\infty V_n$ is dense.*

Proof. We need to prove that if $W \subseteq A$ is open and $W \neq \emptyset$, then $\bigcap_{n=1}^\infty V_n \cap W \neq \emptyset$. Let $W \subseteq A$ be open. Since V_1 is dense, $V_1 \cap W \neq \emptyset$. Since V_1 and W are open, there exists $x_1 \in A$, $r_1 > 0$ such that $\overline{B_{r_1}(x_1; d)} \subseteq V_1 \cap W$ and $0 < r_1 < 1$. Assume $n \geq 2$ and x_{n-1}, r_{n-1} have been chosen. Then, since V_n is dense, $V_n \cap B_{r_{n-1}}(x_{n-1}; d) \neq \emptyset$, and since V_n is open, there exists $x_n \in A$, $r_n > 0$ such that $\overline{B_{r_n}(x_n;d)} \subseteq V_n \cap B_{r_{n-1}}(x_{n-1};d)$ and $0 < r_n < \frac{1}{n}$. This gives sequences $\{x_n\}_{n\in\mathbb{N}} \subseteq A$, and $\{r_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$.

If $i, j > n$, then $x_i, x_j \in B_{r_n}(x_n; d)$. So

$$d(x_i, x_j) \leq d(x_i, x_n) + d(x_n, x_j) < 2r_n < \frac{2}{n}$$

So $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy. Since M is complete, there exists $x \in \overline{A}$ such that $x_n \to x$ as $n \to \infty$. Since $x_i \in \overline{B_{r_n}(x_n;d)}$ for all $i \geq n$, we get that $x \in \overline{B_{r_n}(x_n;d)}$ for all $n \in \mathbb{N}$. Hence, $x \in V_n$ for all $n \in \mathbb{N}$. Also, $x \in W$, hence $x \in \bigcap_{n=1}^{\infty} V_n \cap W$. □

2 Banach and Hilbert spaces

"All maths" is about solutions to equations (existence, uniqueness, properties). Linear Algebra is about the equation $Ax = b$ for a matrix A and vectors b and x. Some problems — for example diagonalisation of matrices — can be turned into such equations. All of this is assumed known. In particular, the axioms of a vectorspace are assumed known (all vectorspaces will be over $\mathbb{R}$ or $\mathbb{C}$ for which we will write $\mathbb{K}$). Also, all vectorspaces will be nontrivial, i.e. not $\{0\}$.

Definition 2.1. Let X be a $\mathbb{K}$-vectorspace.

1. A map $p\colon X \to [0, \infty)$ is called a *semi-norm* iff
   (a) $p(\lambda x) = |\lambda| p(x)$ for all $x \in X$ and $\lambda \in \mathbb{K}$.
   (b) $p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$.
2. A semi-norm p is called a *norm* iff $p(x) = 0$ implies $x = 0$. In this case we will write $\|x\| := p(x)$.

The pair $\{X, p\}$ is called a *semi-normed space* and $\{X, \|\cdot\|\}$ is called *normed space.*

Remark. (a) implies $p(0) = 0$.

Remark. A normed space is a metric space: Define $d(x, y) := \|x - y\|$. Then d is a metric. This is the canonical metric we will use when treating normed spaces.

Proposition 2.2. *Let $\{X, \|\cdot\|\}$ be a normed space. Then*

1. *If $x_n \to x$ as $n \to \infty$ and $y_n \to y$ as $n \to \infty$, then $x_n + y_n \to x + y$ as $n \to \infty$.*
2. *If $\lambda_n \to \lambda$ as $n \to \infty$ and $x_n \to x$ as $n \to \infty$, then $\lambda_n x_n \to \lambda x$ as $n \to \infty$.*
3. *If $x_n \to x$ as $n \to \infty$ then $\|x_n\| \to \|x\|$ as $n \to \infty$.*

I.e. the vectorspace-structure and the topological structure are compatible.

Proof.

1. $\|(x_n + y_n) - (x + y)\| \leq \|x_n - x\| + \|y_n - y\| \to 0$ as $n \to \infty$.
2. $\|\lambda_n x_n - \lambda x\| \leq \|\lambda_n x_n - \lambda_n x\| + \|\lambda_n x - \lambda x\| = |\lambda_n| \|x_n - x\| + |\lambda_n - \lambda| \|x\| \to 0$.
3. This follows from $|\|x\| - \|y\|| \leq \|x - y\|$, since $|\|x_n\| - \|x\|| \leq \|x_n - x\| \to 0$. □

Definition 2.3. A normed space $\{X, \|\cdot\|\}$ which is complete is called a *Banach space.*

Example 2.4.

(a) $\mathbb{R}^n$ with $\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$ — or, more generally, $\{\mathbb{R}^n, \|\cdot\|_p\}$, with $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ for $1 \le p < \infty$ and $\|x\|_\infty = \max_{i=1,\dots,n} |x_i|$.

(b) $\ell_\infty(\mathbb{K}) = \{x\colon \mathbb{N} \to \mathbb{K}, i \mapsto x_i \colon x \text{ is bounded}\}$ with $\|x\|_\infty = \sup_{i\in\mathbb{N}} |x_i|$. $\ell_\infty(\mathbb{K})$ is complete, since $\{\ell_\infty(\mathbb{K}), d_\infty\} = \{\mathscr{B}(\mathbb{N},\mathbb{K}), d_\infty\}$. In fact, let $\{Y, \|\cdot\|_Y\}$ be a Banach space, and $M \neq \emptyset$ any set. Then define $\ell_\infty(M,Y) = \mathscr{B}(M,Y)$ and $\|f\|_\infty = \sup_{t\in M} \|f(t)\|_Y$. Then $\{\ell_\infty(M,Y), \|\cdot\|_\infty\}$ is a Banach space.

(c) Let $M = \{A, d\}$ be a metric space, X a Banach space and $\mathscr{C}_b(A,X)$ the continuous and bounded maps from A to X. Write $\|f\|_\infty = \sup_{t\in M} \|f(t)\|_X$. Then $\mathscr{C}_b(A,X)$ with $\|\cdot\|_\infty$ is a Banach space.

(d) $\{C^\alpha, \|\cdot\|_\infty\}$ is a Banach space.

(e) $\{C^1[0,1], \|\cdot\|_{C^1}\}$ is a Banach space, where $\|f\|_{C^1} = \sup_{t\in[0,1]} |f(t)| + \sup_{t\in[0,1]} |f'(t)|$. Note that $\sup_{t\in[0,1]} |f'(t)|$ is a semi-norm but not a norm.

(f) $\ell_p = \{x\colon \mathbb{N} \to \mathbb{K}\colon \sum_{i=1}^\infty |x_i|^p < \infty\}$ with $\|x\|_p = (\sum_{i=1}^\infty |x_i|^p)^{1/p}$ is a normed vectorspace. This is a Banach space: Let $\{x_n\}_{n\in\mathbb{N}} \subseteq \ell_p$ be a Cauchy sequence, i.e. $x_n \in \ell_p$: $x_n \colon \mathbb{N} \to \mathbb{K}, i \mapsto x_n(i)$. Let $\varepsilon > 0$. Since $\{x_n\}$ is Cauchy, there exists $N \in \mathbb{N}$ such that $\|x_n - x_m\|_p < \varepsilon$ for $n, m \ge N$. Then $|x_n(i) - x_m(i)| \le \|x_n - x_m\|_p < \varepsilon$ for $n, m \ge N$ and all $i \in \mathbb{N}$, hence, $\{x_n(i)\}_{n\in\mathbb{N}} \subseteq \mathbb{K}$ is a Cauchy sequence, for all $i \in \mathbb{N}$. Since $\mathbb{K}$ is complete, there exists, for all $i \in \mathbb{N}$, $x(i) \in \mathbb{K}$ such that $x_n(i) \to x(i)$ as $n \to \infty$. This defines a sequence $x = \{x(i)\}_{i\in\mathbb{N}}$ in $\mathbb{K}$. For $n, m \ge N$ and for all $M \in \mathbb{N}$,

$$\left(\sum_{i=1}^M |x_n(i) - x_m(i)|^p\right)^{1/p} \le \left(\sum_{i=1}^\infty |x_n(i) - x_m(i)|^p\right)^{1/p} = \|x_n - x_m\|_p < \varepsilon$$

For n fixed, let $m \to \infty$ in the inequality $\left(\sum_{i=1}^M |x_n(i) - x_m(i)|^p\right)^{1/p} < \varepsilon$; we get that $\left(\sum_{i=1}^M |x_n(i) - x(i)|^p\right)^{1/p} \le \varepsilon$ for all $M \in \mathbb{N}$. Hence,

$$\|x_n - x\|_p = \left(\sum_{i=1}^\infty |x_n(i) - x(i)|^p\right)^{1/p} \le \varepsilon$$

for all $n \ge N$. Hence, $x - x_n \in \ell_p$ for all $n \ge N$. So $x = (x - x_n) + x_n \in \ell_p$ and $\|x_n - x\|_p \to 0$ as $n \to \infty$, hence $x_n \to x$ as $n \to \infty$.

Proposition 2.5. *Let $\{X, \|\cdot\|\}$ be a normed space. Then X is a Banach space iff every absolutely convergent series is convergent: If $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ is a sequence such that $\sum_{n=1}^\infty \|x_n\| < \infty$, then there exists $x \in X$ such that $\lim_{M\to\infty} \left\|x - \sum_{n=1}^M x_n\right\| = 0$, i.e. $x = \lim_{M\to\infty} \sum_{n=1}^M x_n =: \sum_{n=1}^\infty x_n$.*

Proof.

"$\Rightarrow$" The sequence $\left\{\sum_{i=1}^M x_n\right\}_{M\in\mathbb{N}}$ is Cauchy in $\{X, \|\cdot\|\}$ if $\sum_{n=1}^\infty \|x_n\| < \infty$.

“$\Leftarrow$” Assume $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy. For all $\varepsilon > 0$ there exists $N(\varepsilon) \in \mathbb{N}$ such that $n, m \geq N(\varepsilon)$ implies $\|x_n - x_m\| < \varepsilon$. Do this for $\varepsilon = \varepsilon_k = 2^{-k}$, $k \in \mathbb{N}$, i.e. there exists $N_k \in \mathbb{N}$ such that $n, m \geq N_k$ implies $\|x_n - x_m\| < 2^{-k}$. Using this, define inductively a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ such that $\|x_{n_{k+1}} - x_{n_k}\| < 2^{-k}$. Let $y_n = x_{n_{k+1}} - x_{n_k}$. Then $\sum_{k=1}^{\infty} \|y_k\| < \sum_{k=1}^{\infty} 2^{-k} < \infty$. So, $\sum y_k$ is absolutely convergent, hence, by assumption, there exists $y \in X$ such that $\lim_{M\to\infty} \left\| y - \sum_{k=1}^{M} y_k \right\| = 0$. So $\lim_{M\to\infty} \|y - (x_{n_{M+1}} - x_{n_1})\| = 0$. Hence, $\{x_n\}_{n\in\mathbb{N}}$ has a convergent subsequence, and is Cauchy. So, by 1.58, also $\{x_n\}_{n\in\mathbb{N}}$ is convergent. Hence, X is Banach. □

Definition 2.6. Let X be a vectorspace over $\mathbb{K}$.

(a) A subset $C \subseteq X$ is called *convex* iff $x, y \in C$, $\lambda \in [0,1]$ implies $\lambda x + (1-\lambda) y \in C$.

(b) The *convex hull* of a subset $A \subseteq X$ is

$$\operatorname{co}(A) = \left\{ \sum_{k=1}^{n} s_k x_k : n \in \mathbb{N}, x_k \in A, s_k \in [0,1], \sum_{k=1}^{n} s_k = 1 \right\}$$

the set of all linear convex combinations of elements in A.

(c) A subset $A \subseteq X$ is called *absolutely convex* iff $x, y \in A$, $s, t \in \mathbb{K}$, $|s| + |t| \leq 1$ implies $sx + ty \in A$. In particular, A is convex.

(d) The *absolutely convex hull* of a subset $A \subseteq X$ is

$$\Gamma(A) = \left\{ \sum_{i=1}^{n} s_k x_k : n \in \mathbb{N}, x_k \in A, s_k \in \mathbb{K}, \sum_{k=1}^{n} |s_k| \leq 1 \right\}$$

Now, let X be normed.

(e) X is called *strictly normed* (or *strictly convex*) iff for $\|x\| = \|y\| = 1$, $\left\| \frac{1}{2}(x+y) \right\| = 1$ implies $x = y$.

(f) X is called *uniformly convex* iff for sequences $\{x_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}} \subseteq X$, $\lim_{n\to\infty} \|x_n\| = 1$, $\lim_{n\to\infty} \|y_n\| = 1$ and $\lim_{n\to\infty} \left\| \frac{1}{2}(x+y) \right\| = 1$ implies $\|x_n - y_n\| \to 0$ as $n \to \infty$.

As usual we denote $B_1(0) = \{y : \|y\| < 1\}$ the *unit ball* in X. Also, $S_1(0) = \{y : \|y\| = 1\}$.

Remark.

1. $\overline{B_1(0)} = \{y : \|y\| \leq 1\} = B_1(0) \cup S_1(0)$.
2. $B_1(0)$ is convex (any open ball is convex).

Definition 2.7. Let X be a vectorspace. Then $U \subseteq X$ is called a *linear subspace* iff $x, y \in U$, $\lambda \in \mathbb{K}$ implies $x + \lambda y \in U$. Then $x \sim y \iff x - y \in U$ defines an equivalence relation and the quotient X/U is a vectorspace. We write $[x] = x + U \in X/U$.

Lemma 2.8. *Let $\{X, p\}$ be a semi-normed space.*

(a) $N = \{x \in X : p(x) = 0\}$ is a linear subspace of X.

(b) $\|[x]\| = p(x)$ defines a norm on X/N.

(c) If every Cauchy-sequence in $\{X, p\}$ converges, then $\{X/N, \|\cdot\|\}$ is Banach.

Proof.

(a) $0 \leq p(x+\lambda y) \leq p(x)+|\lambda|p(y) = 0$ if $x, y \in N$. So $x+\lambda y \in N$.

(b) $\|[x]+[y]\| = p(x+y) \leq p(x)+p(y) = \|[x]\|+\|[y]\|$, $\|\lambda[x]\| = \|[\lambda x]\| = p(\lambda x) = |\lambda|p(x) = |\lambda|\|[x]\|$. Note: if $y \sim x$, then $x-y \in N$, so $p(x) = p(x-y+y) \leq p(x-y)+p(y) = p(y)$ and $p(y) = p(y-x+x) \leq p(y-x)+p(x) = p(x)$. Hence, $\|[x]\|$ is well-defined. Also, $\|[x]\| = 0$ implies $p(x) = 0$, so $x \in N$, i.e. $[x] = 0$.

(c) Clearly, $\{[x_n]\}_{n\in\mathbb{N}}$ is Cauchy or convergent in $\{X/N, \|\cdot\|\}$ iff $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy or convergent in $\{X, p\}$. □

Lemma 2.9. *Let X be a normed space, and $U \subseteq X$ be a linear subspace. Then $\overline{U}$ is also a linear subspace.*

Proof. Let $x, y \in \overline{U}$, $\lambda \in \mathbb{K}$. Then there exist $\{x_n\}, \{y_n\} \subseteq U$ such that $x_n \to x$, $y_n \to y$ as $n \to \infty$. Then, since U is a linear subspace, $x_n + \lambda y_n \in U$. On the other hand, by 2.2, $x_n + \lambda y_n \to x + \lambda y$ as $n \to \infty$. Hence, $x + \lambda y \in \overline{U}$. So $\overline{U}$ is a linear subspace. □

Definition 2.10. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on the same vectorspace X are called equivalent iff there exist $c, C > 0$ such that $c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1$.

Remark 2.11.

- Two equivalent norms have exactly the same convergent sequences and give rise to the same topology.
- Any two norms on $\mathbb{R}^n$ are equivalent.
- If $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent norms on X, then $\{X, \|\cdot\|_1\}$ is Banach space iff $\{X, \|\cdot\|_2\}$ is Banach.
- Let $X = \mathscr{C}[0,1] = \mathscr{C}([0,1], \mathbb{R})$, and $\|f\|_\infty = \sup_{t\in[0,1]} |f(t)|$, and

$$\|f\|_1 = \int_0^1 |f(t)|\,\mathrm{d}t.$$

  Then $\|\cdot\|_\infty$ and $\|\cdot\|_1$ are norms. Note, that $\|f\|_1 \leq \|f\|_\infty$ for all $f \in \mathscr{C}[0,1]$. Assume there exists $C_0 > 0$ such that $\|f\|_\infty \leq C_0\|f\|_1$. Take

$$f(t) = \begin{cases} 1 - C_0 t & t \in [0, \frac{1}{C_0}] \\ 0 & t \in [\frac{1}{C_0}, 1] \end{cases}$$

  Then $\|f\|_\infty = 1$ and $\|f\|_1 = \frac{1}{2C_0}$. So $1 \leq \frac{1}{2}$ — a contradiction. So these two norms are not equivalent. Note, that $(\mathscr{C}[0,1], \|\cdot\|_\infty)$ is Banach but $(\mathscr{C}[0,1], \|\cdot\|_1)$ is not.

Proposition 2.12. *Let X, Y be normed spaces, with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$.*

1. *$\|(x,y)\|_p := (\|x\|_X^p + \|y\|_Y^p)^{1/p}$ defines a norm on $X \oplus Y$ for $1 \leq p < \infty$. We denote this normed space $X \oplus_p Y$. (Also $\|(x,y)\|_\infty = \max\{\|x\|_X, \|y\|_Y\}$)*
2. *Any $\|(\cdot,\cdot)\|_p$, $\|(\cdot,\cdot)\|_q$ are equivalent norms on $X \oplus Y$.*
3. *If X and Y are Banach spaces, then $X \oplus_p Y$ is Banach.*

Definition 2.13. Let $M = \{A, d\}$ be a metric space and $U \subseteq A$ a subset. The *distance* from $x \in A$ to U is defined as

$$d(x, U) := \inf_{a\in U} d(x, a).$$

A point $a \in U$ such that $d(x, a) = d(x, U)$ is called a *best approximation* to x in A.

Remark.
(1) Such a point is not necessarily unique.
(2) If U is compact, then there exists at least one best approximation (for all x) since the map $a \mapsto d(x,a)$ is continuous. We shall see more later on the existence and uniqueness of best approximations, especially for the case of linear subspaces of Banach spaces.

Proposition 2.14 (Riesz' Lemma)**.** *Let X be a normed space, $U \subseteq X$ a linear subspace such that $\overline{U} = U$ and $U \neq X$. Let $\delta \in (0,1)$. Then there exists $x_\delta \in X$ with $\|x_\delta\| = 1$ and $\|x_\delta - u\| \geq 1 - \delta$ for all $u \in U$.*

Proof. Let $x \in X \setminus U$. Since $U = \overline{U}$, $d(x,U) > 0$. Since $\delta \in (0,1)$, $d(x,U) < \frac{d(x,U)}{1-\delta}$. Since $d(x,U) = \inf_{a \in U} d(x,a)$, there exists $u_\delta \in U$ such that $d(x,u_\delta) < \frac{d(x,U)}{1-\delta}$. Let $x_\delta = \frac{x-u_\delta}{\|x-u_\delta\|}$. Then $\|x_\delta\| = 1$, and for all $u \in U$

$$\begin{aligned}\|x_\delta - u\| &= \left\| \frac{x-u_\delta}{\|x-u_\delta\|} - u \right\| = \left\| \frac{x}{\|x-u_\delta\|} - \frac{u_\delta}{\|x-u_\delta\|} - u \right\| \\ &= \frac{1}{\|x-u_\delta\|} \|x - (u_\delta + \|x-u_\delta\| u)\| \geq \frac{d(x,U)}{\|x-u_\delta\|} \geq 1 - \delta\end{aligned}$$ □

Definition 2.15. Let X, Y be two $\mathbb{K}$-vectorspaces. A map $T \colon X \to Y$ is called *linear* iff $T(\alpha x_1 + x_2) = \alpha T(x_1) + T(x_2)$ for all $x_1, x_2 \in X$ and $\alpha \in \mathbb{K}$. The *kernel* $\mathrm{N}(T) = T^{-1}(\{0\})$ of T is a linear subspace of X. The *image* (or *range*) $\mathrm{R}(T) = \{Tx \colon x \in X\}$ of T is a linear subspace of Y. We shall often (for linear T) write Tx instead of $T(x)$. We call T a *linear operator.*

Theorem 2.16. *For normed spaces X, Y and a linear operator $T \colon X \to Y$, the following are equivalent:*
(a) There exists $C > 0$ such that $\|Tx\|_Y \leq C\|x\|_X$.
(b) T is uniformly continuous on X.
(c) There exists $a \in X$ such that T is continuous at a.
(d) $\|T\| = \sup\limits_{\substack{x \in X \\ \|x\| \leq 1}} \|Tx\|_Y < \infty$.

Proof.
$(a) \Rightarrow (b)$ For $x, y \in X$, $\|T(x) - T(y)\|_Y = \|T(x-y)\|_Y \leq C\|x-y\|_X$, so, for $\varepsilon > 0$, $d(u,v) < \frac{\varepsilon}{C}$ implies $d(T(u), T(v)) < \varepsilon$.
$(b) \Rightarrow (c)$ Trivial.
$(c) \Rightarrow (d)$ For $\varepsilon = 1$, there exists $\delta > 0$ such that $\|x - a\|_X \leq \delta$ implies $\|Tx - Ta\|_Y \leq 1$. For all $x \in X$ with $\|x\| \leq 1$, $\|(a + \delta x) - a\|_X \leq \delta$. So $\|T(\delta x)\|_Y = \|T(a+\delta x) - T(a)\|_Y \leq 1$. So $\|Tx\|_Y \leq \frac{1}{\delta} < \infty$ for all $x \in X$ with $\|x\| \leq 1$. So $\|T\| \leq \frac{1}{\delta} < \infty$.
$(d) \Rightarrow (a)$ For $x \neq 0$, $\left\| \frac{x}{\|x\|_X} \right\|_X = 1$, so $\left\| T\left(\frac{x}{\|x\|_X}\right) \right\|_X \leq \sup_{\|x\|_X \leq 1} \|Tx\|_Y = \|T\|$. Hence, $\frac{1}{\|x\|_X} \|Tx\|_Y \leq \|T\|$, so $\|Tx\|_Y \leq \|T\| \|x\|_X$. □

Remark. In this case, the number $\|T\|$ in (d) is the smallest number such that (a) holds, i.e. $\|T\| = \sup_{x \neq 0} \frac{\|Tx\|_Y}{\|x\|_X}$. It is called the *operator norm* of T.

Definition 2.17. Let X, Y be normed spaces, and T a linear operator such that one (hence, all) of the conditions in 2.16 holds. Then T is called a *bounded linear operator.* The set of all such operators is denoted $B(X, Y)$. If $X = Y$, we write $B(X)$.

Remark 2.18.

1. $B(X, Y)$ is the set of all continuous and linear maps from X to Y. However, if $T \in B(X, Y)$, then it is *not* a bounded map as defined in 1.42: The range $\mathrm{R}(T)$ is not a bounded subset of Y. However, the image $T\big(B_1(0, \|\cdot\|_X)\big)$ of the unit ball in X is a bounded subset of Y.
2. Not all linear maps are bounded, i.e. there exist discontinuous linear maps; these are called *unbounded operators.* Let $X = \mathscr{C}^1[0, 2\pi]$, $Y = \mathscr{C}[0, 2\pi]$, $\|\cdot\|_X = \|\cdot\|_Y = \|\cdot\|_\infty$, and let $T = \frac{\mathrm{d}}{\mathrm{d}x} \colon X \to Y$. Then T is well-defined and linear. But T is not bounded: Let $f_n(t) = e^{int}$. Then $\|f_n\|_\infty = 1$ but $Tf_n = (in)f_n$, so $\|Tf_n\|_\infty = n$. Hence, there does not exist $C > 0$ such that $\|Tf\|_\infty \le C\|f\|_\infty$. (But try $\|f\|_{\mathscr{C}^1} = \|f\|_\infty + \|f'\|_\infty$ on X.)
3. If $T \colon X \to Y$ is bounded and one chooses an equivalent norm on X or Y, or on both, then T remains bounded. Note, however, that the number $\|T\| = \sup_{\|x\|_X \le 1} \|Tx\|_Y$ might very well change.
4. If $T \colon X \to Y$ is linear and $\dim X < \infty$, then T is bounded, in particular, any linear map $\mathbb{R}^n \to \mathbb{R}^m$ is bounded: Choose a basis $\{e_1, \dots, e_n\}$ of X and define, for $x = \sum_{i=1}^n x_i e_i$, $\|x\|_1 = \sum_{i=1}^n |x_i|$. Then $\|\cdot\|_1$ is a norm on X. Since T is linear, $Tx = \sum_{i=1}^n x_i T(e_i)$. So $\|Tx\|_1 \le \sum_{i=1}^n |x_i| \|Te_i\|_Y$. Let $C = \max_{i=1,\dots,n} \|Te_i\|_Y$. Then $\|Tx\|_Y \le \sum_{i=1}^n |x_i| C = C\|x\|_1$. So T is $(\|\cdot\|_1, \|\cdot\|_Y)$-bounded. Since $\dim X < \infty$, $\|\cdot\|_1$ is equivalent to $\|\cdot\|_X$. Hence, T is $(\|\cdot\|_X, \|\cdot\|_Y)$-bounded.
5. Let X, Y be normed spaces. Then $B(X, Y)$ is a vectorspace: $(\alpha T + S)(x) := \alpha T(x) + S(x)$ for $\alpha \in \mathbb{K}$ and $T, S \in B(X, Y)$. This defines a linear map $\alpha T + S \colon X \to Y$. Also, if $x \in X$, $\|x\|_X \le 1$, then
$$\|(\alpha T + S)x\|_Y \le |\alpha| \|Tx\|_Y + \|Sx\|_Y \le |\alpha| \|T\| + \|S\|,$$
hence $\|\alpha T + S\| \le |\alpha| \|T\| + \|S\| < \infty$. Hence, $\alpha T + S \in B(X, Y)$. Also, $\|T\| = \sup_{\|x\|_X \le 1} \|Tx\|_Y$ defines a norm on $B(X, Y)$: Clearly, $T = 0 \iff \|T\| = 0$. From above, $\|T + S\| \le \|T\| + \|S\|$ and $\|\lambda T\| = \sup_{\|x\|_X \le 1} \|(\lambda T)x\|_Y = |\lambda| \|T\|$. Hence, $(B(X, Y), \|\cdot\|)$ is a normed vectorspace. Note, that if $\dim X = m < \infty$ and $\dim Y = n < \infty$, then $B(X, Y)$ can be identified with $\mathbb{K}^{n \times m} \cong \mathbb{K}^{n \cdot m}$.
6. Let X, Y, Z be normed vectorspaces, and let $T \in B(X, Y)$ and $S \in B(Y, Z)$. Then $ST \in B(X, Z)$ and $\|ST\| \le \|S\| \|T\|$, since for $x \in X$, $\|x\|_X \le 1$:
$$\|S(Tx)\|_Z \le \|S\| \|Tx\|_Y \le \|S\| \|T\| \|x\|_X \le \|S\| \|T\|.$$
7. If X is normed an $Y = \mathbb{L}$, then $B(X, \mathbb{K}) =: X'$ is called the *dual space* of X. It is a normed linear space. Note, that $L(X, \mathbb{K})$ is the *algebraic dual* of X. An element of $B(X, \mathbb{K})$ is called a *bounded linear functional.* For example, $T \colon \mathscr{C}[0, 1] \to \mathbb{K}, x \mapsto x(0)$ is in $\mathscr{C}[0, 1]'$, $\|T\| = 1$; $T \colon : \mathscr{C}^1[0, 1] \to \mathbb{K}, x \mapsto x(0) + x'(1)$ is in $\mathscr{C}^1[0, 1]'$, $\|T\| = 1$; $T \colon \mathscr{C}[0, 1] \to \mathbb{K}, x \mapsto \int_0^1 x(t)\,\mathrm{d}t$ is in $\mathscr{C}[0, 1]'$, $\|T\| = 1$ and, for any $g \in \mathscr{C}[0, 1]$, $T \colon \mathscr{C}[0, 1] \to \mathbb{K}, x \mapsto \int_0^1 x(t) g(t)\,\mathrm{d}t$ is in $\mathscr{C}[0, 1]'$ with $\|T\| = \int_0^1 |g(t)|\,\mathrm{d}t$.

Proposition 2.19. *Let X, Y be normed spaces, and let $(B(X,Y), \|\cdot\|)$ be the normed space of bounded linear operators.*
(a) If Y is Banach, then so ist $B(X,Y)$.
(b) X' is a Banach space.

Remark. The result is independent of whether or not X is Banach.

Proof. (b) follows immediately from (a), since $\mathbb{K}$ is complete. Let $\{T_n\}_{n\in\mathbb{N}} \subseteq B(X,Y)$ be Cauchy. Since, for all $x \in X$, $\|T_n x - T_m x\|_Y = \|(T_n - T_m)x\|_Y \le \|T_n - T_m\|\|x\|_X$, $\{T_n x\}_{n\in\mathbb{N}} \subseteq Y$ is Cauchy for all $x \in X$. Since Y is Banach, $\{T_n x\}$ is convergent in Y. Let $Tx = \lim_{n\to\infty} T_n x$. So $T\colon X \to Y$ is linear, since for $x_1, x_2 \in X$, $x \in \mathbb{K}$,

$$T(\alpha x_1 + x_2) = \lim_{n\to\infty} T_n(\alpha x_1 + x_2) \overset{2.2}{=} \alpha \lim_{n\to\infty} T_n(x_1) + \lim_{n\to\infty} T_n(x_2) = \alpha T(x_1) + T(x_2).$$

Let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that $n, m \ge N$ implies $\|T_n - T_m\| < \varepsilon$. Let $x \in X$, $\|x\|_X \le 1$. Take an $m > N$ such that $\|T_m x - Tx\|_Y < \varepsilon$. Then, for all $n \ge N$,

$$\|Tx - T_n x\|_Y \le \|Tx - T_m x\|_Y + \|T_m x - T_n x\|_Y \le \varepsilon + \|T_m - T_n\|\|x\|_X.$$

Hence, $\|Tx - T_m x\|_Y \le 2\varepsilon$ for all $x \in X$ with $\|x\|_X \le 1$. So,

$$\|T - T_n\| = \sup_{\|x\|_X \le 1} \|Tx - T_n x\|_Y \le 2\varepsilon < \infty \quad \forall n \ge N$$

So, $T - T_n \in B(X,Y)$, hence $T = (T - T_n) + T_n \in B(X,Y)$ and $T_n \to T$ as $n \to \infty$ in $B(X,Y)$. ☐

Remark. $0 \in B(X,Y)$. If $X = Y$, then we denote the identity map by I. Clearly, $\|I\| = 1$ and $I \in B(X)$. Since $S, T \in B(X,Y)$ implies $ST \in B(X)$, $B(X)$ is a $\mathbb{K}$-algebra.

Definition 2.20. Let X, Y be normed spaces.

(a) A linear map $T\colon X \to Y$ is called an *isomorphism* iff T is bijective and both T and T^{-1} are bounded, i.e. an isomorphism is a linear homeomorphism.
(b) A surjective linear map $T\colon X \to Y$ is called an *isometry from X on Y* iff $\|Tx\|_Y = \|x\|_X$ for all $x \in X$, in particular, T is an isomorphism.
(c) A linear map $T\colon X \to Y$ is called an *isometry from X in Y* iff $T\colon X \to \mathrm{R}(T)$ is an isometry of X on $\mathrm{R}(T)$.
(d) X and Y are called *isomorphic* (written $X \simeq Y$) iff there exists an isomorphism $X \to Y$. They are called *isometric* (or *isometrically isomorphic*) iff there exists an isometry from X on Y (written $X \cong Y$).
(e) If a linear map $T\colon X \to Y$ is injective, it is called an *embedding* of X in Y (and if $T \in B(X,Y)$, then T is called a continuous/bounded embedding).
(f) If a linear map $P\colon X \to Y$ satisfies $P^2 = P$, it is called a *projection.*
(g) If $T \in B(X,Y)$ is bijective then $T^{-1} \in B(Y,X)$ (i.e. the inverse is automatically bounded). The proof of this is nontrivial, and we shall do this later. We call T *invertible.*

Remark.

1. Both "$\simeq$" and "$\cong$" are equivalence relations.
2. Normed spaces of the same finite dimension are always isomorphic. However $\{\mathbb{R}^2, \|\cdot\|_2\}$ and $\{\mathbb{R}^2, \|\cdot\|_1\}$ are not isometrically isomorphic.
3. The question which ("known") Banach spaces are isomorphic or isometrically isomorphic to which other spaces is/was an important one.
4. If $T\colon X \to Y$, and $\dim X = \dim Y = \infty$, then it is (in general) not enough that T is injective or surjective to conclude that T is a bijection.

Proposition 2.21. *Let X be a normed space, Y a Banach space, $V \subseteq X$ a linear subspace, and $T\colon V \to Y$ a continuous linear map (i.e. $T \in B(V, Y)$). Then there exists a unique extension $\overline{T}\colon \overline{V} \to Y$ (i.e. $\overline{T}|_V = T$), with $\overline{T} \in B(\overline{V}, Y)$ and $\|\overline{T}\| = \|T\|$.*

Proof. Assume $x \in \overline{V}$; then there exists $\{v_n\} \subseteq V$ such that $\|x - v_n\| \to 0$ as $n \to \infty$. So, if $\overline{T}$ exists, then $\overline{T}x = \lim_{n\to\infty} \overline{T}v_n = \lim_{n\to\infty} Tv_n$. This proves uniqueness. Let $x \in \overline{V}$ and take $\{v_n\} \subseteq V$ such that $v_n \to x$ as $n \to \infty$. Then $\|Tv_n - Tv_m\| = \|T(v_n - v_m)\| \leq \|T\|\|v_n - v_m\|$. Since $\{v_n\}$ is convergent, it is Cauchy, so this proves that $\{Tv_n\} \subseteq Y$ is Cauchy, hence, since Y is Banach, it is convergent. If also $\{u_n\} \subseteq V$ with $u_n \to x$ as $n \to \infty$, then

$$\|Tu_n - Tv_n\| = \|T(u_n - v_n)\| \leq \|T\|\|u_n - v_n\| \leq \|T\|\big(\|u_n - x\| + \|x - v_n\|\big) \xrightarrow{n\to\infty} 0,$$

hence, $\lim_{n\to\infty} Tu_n - Tv_n = 0$. Hence, $\lim_{n\to\infty} Tu_n = \lim_{n\to\infty} Tv_n$ by 2.2. So, $\overline{T}x = \lim_{n\to\infty} Tv_n$ is well-defined ($x \in \overline{V}$, $v_n \to x$ as $n \to \infty$). Clearly, $\overline{T}$ is an extension of T. Also, $\overline{T}\colon \overline{V} \to Y$ is linear (take $x, y \in \overline{V}$, $\lambda \in \mathbb{K}$, take $\{v_n\}, \{w_n\} \subseteq V$ s.t. $v_n \to x$ as $n \to \infty$, $w_n \to y$ as $n \to \infty$ and use the definition of $\overline{T}$, linearity of T, and 2.2). Since T is bounded (on V), we have $\|Tv_n\| \leq \|T\|\|v_n\|$. Taking $n \to \infty$, by 2.2 $\|\overline{T}x\| \leq \|T\|\|x\|$, hence $\|\overline{T}\| \leq \|T\|$. So $\overline{T} \in B(\overline{V}, Y)$, and since $\|T\| \leq \|\overline{T}\|$, we get $\|T\| = \|\overline{T}\|$. □

Remark.

1. In particular, if $V \subseteq X$ ($V \neq X$), $\overline{V} = X$, $T \in B(V, Y)$, Y Banach, then there exists a unique $\overline{T} \in B(X, Y)$ extending T with $\|\overline{T}\| = \|T\|$.
2. If $T\colon V \to Y$ is an isometry, then also $\overline{T}$ is an isometry. However, if T is injective, one cannot be sure that also $\overline{T}$ is injective.
3. Note, that the special case $Y = \mathbb{K}$, gives extensions of certain linear bounded functionals.

We shall now study a special class of normed spaces, namely those where the norm comes from a scalar product.

Definition 2.22. Let H be a $\mathbb{K}$-vectorspace. A map $\langle\cdot,\cdot\rangle\colon H \times H \to \mathbb{K}$ is called *scalar product* (or *inner product*) iff for all $x, y_1, y_2, y \in H$ and $\lambda \in \mathbb{K}$

(a) $\langle x, \lambda y_1 + y_2\rangle = \lambda\langle x, y_1\rangle + \langle x, y_2\rangle$.

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

(c) $\langle x, x\rangle \geq 0$ and $\langle x, x\rangle = 0$ iff $x = 0$.

Note that $\langle \lambda x_1 + x_2, y\rangle = \overline{\lambda}\langle x_1, y\rangle + \langle x_2, y\rangle$. If $\mathbb{K} = \mathbb{R}$, $\langle \cdot, \cdot\rangle$ is called *bilinear*, if $\mathbb{K} = \mathbb{C}$, $\langle \cdot, \cdot\rangle$ is called a *sesquilinear* form. Property (c) is called *positive definiteness*. Property (b) is called *symmetry*. Hence, $\langle x, x\rangle = \overline{\langle x, x\rangle} \in \mathbb{R}$. The space $(H, \langle \cdot, \cdot\rangle)$ is called a *pre-Hilbert space*.

Proposition 2.23. *Let $(H, \langle \cdot, \cdot\rangle)$ be a pre-Hilbert space, and let $\|x\| = \sqrt{\langle x, x\rangle}$ for $x \in H$. Then*

1. *$\|\cdot\|$ is a norm on H.*
2. *$|\langle x, y\rangle| \leq \|x\|\|y\|$ with equality if $x = \lambda y$ for $\lambda \in \mathbb{K}$ (Cauchy-Schwarz inequality).*
3. *$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$ (parallelogramme rule).*

Proof.

1. $\|\cdot\|$ is positive definite by definition. Also, for $x \in H$, $\alpha \in \mathbb{K}$, $\|\alpha x\|^2 = \langle \alpha x, \alpha x\rangle = \overline{\alpha}\alpha\langle x, x\rangle = |\alpha|^2\|x\|^2$. The triangle inequality follows from 2:
$$\begin{aligned}\|x + y\|^2 = \langle x + y, x + y\rangle = \|x\|^2 + \|y\|^2 + 2\operatorname{Re}\langle x, y\rangle \leq \\ \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2.\end{aligned}$$
2. Let $\lambda \in \mathbb{K}$ be arbitrary, $x, y \in H$, then
$$0 \leq \langle x + \lambda y, x + \lambda y\rangle = \|x\|^2 + \overline{\lambda}\langle y, x\rangle + \lambda\langle x, y\rangle + |\lambda|^2\|y\|^2$$
Taking $\lambda = -\overline{\langle x, y\rangle}/\|y\|^2$, Cauchy-Schwarz follows.
3. Follows from the first 2 lines in the computation in 1. □

Definition. Hence, a pre-Hilbert space $(H, \langle \cdot, \cdot\rangle)$ gives rise to a normed space $(H, \|\cdot\|)$, $\|x\| = \sqrt{\langle x, x\rangle}$. If this space is complete, $(H, \langle \cdot, \cdot\rangle)$ is called a *Hilbert space*.

Remark. For $\mathbb{K} = \mathbb{R}$
$$\langle x, y\rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$
and for $\mathbb{K} = \mathbb{C}$
$$\langle x, y\rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2).$$
This is called *polarization identity*. So, the scalar product defines the norm, on the other hand, the scalar product is uniquely determined by the norm.

Proposition 2.24. *A normed space X is a pre-Hilbert space iff*
$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad \forall x, y \in X \tag{$*$}$$

Proof. If X is a pre-Hilbert space, then $(*)$ holds. So assume $(*)$ holds, and set $(\mathbb{K} = \mathbb{R})$
$$\langle x, y\rangle := \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2).$$
Then (!) one proves that this does define a scalar product on X. (For $\mathbb{K} = \mathbb{C}$, use the polarization identity). □

In the proof of the above proposition one needs the following lemma.

Lemma 2.25. *The scalar product on a pre-Hilbert space is a continuous map* $H \times H \to \mathbb{K}$.

Proof. Form the Cauchy-Schwarz inequality, it follows that

$$|\langle x_1, y_1\rangle - \langle x_2, y_2\rangle| = |\langle x_1 - x_2, y_1\rangle + \langle x_2, y_1 - y_2\rangle| \leq \|x_1 - x_2\|\|y_1\| + \|x_2\|\|y_1 - y_2\|.$$

This proves continuity. □

Example 2.26.

1. $\mathbb{C}^n$ with $\langle x, y\rangle = \sum_{i=1}^n \overline{x_i} y_i$.
2. ℓ_2 with $\langle x, y\rangle = \sum_{i=1}^\infty \overline{x_i} y_i$, since for $x, y \in \ell_2(\mathbb{N})$, $N \in \mathbb{N}$,

$$\left|\sum_{i=1}^N \overline{x_i} y_i\right| \leq \left(\sum_{i=1}^N |x_i|^2\right)^{1/2} \left(\sum_{i=1}^N |y_i|^2\right)^{1/2} \leq \|x\|_2\|y\|_2$$

   and $\langle x, x\rangle = \|x\|_2$.
3. Let $H = \mathscr{C}([0,1], \mathbb{C})$ and define

$$\langle f, g\rangle = \int_0^1 \overline{f(t)} g(t)\,\mathrm{d}t$$

   This is a scalar product, so $(H, \langle\cdot,\cdot\rangle)$ is a pre-Hilbert space. However, this is not a Hilbert space. We shall "repair" this later, when studying Lebesgue-integration.
4. Let $H = \mathscr{C}^k([0,1], \mathbb{C})$, and let

$$\langle f, g\rangle_{\mathscr{C}^k} = \sum_{j=0}^k \left\langle f^{(j)}, g^{(j)}\right\rangle$$

   with $\langle\cdot,\cdot\rangle$ the scalar product in 3. This gives a pre-Hilbert space.

Definition 2.27. Let H be a pre-Hilbert space.

(a) If $\langle x, y\rangle = 0$ then we say that x and y are *orthogonal* and write $x \perp y$. In this case it follows that $\|x\|^2 + \|y\|^2 = \|x + y\|^2$.

(b) Let $Y, Z \subseteq H$ be two subsets of H. Then we call Y and Z *orthogonal* iff $\langle z, y\rangle = 0$ for all $z \in Z$ and $y \in Y$. If Y, Z are linear subspaces, then $Y \cap Z = \{0\}$ if Y and Z are orthogonal.

(c) For a subset $Y \subseteq H$ we define the *orthogonal complement* of Y by

$$Y^\perp = \{x \in H : \forall y \in Y.\ x \perp y\}$$

Then $Y \cap Y^\perp = \{0\}$ if Y is a linear subspace.

Remark.

1. $A^\perp$ is always a linear closed subspace of H.
2. $(\overline{A})^\perp = A^\perp$.

3. $A \subseteq (A^\perp)^\perp$.

Proposition 2.28. *Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space, and let $K \subseteq H$ be a closed and convex subset and let $x_0 \in H$. Then there exists a unique $x \in K$ such that $\|x_0 - x\| = d(x_0, K)$, i.e. there exists a unique best approximation to x_0 in K.*

Proof. This is trivial if $x_0 \in K$. So assume $x_0 \notin K$. Also, assume $x_0 = 0$ (otherwise, subtract x_0 everywhere). Since $d := d(x_0, K) = \inf_{y \in K} \|y\|$, there exists a sequence $\{y_n\} \subseteq K$ such that $\|y_n\| \to d$ as $n \to \infty$. We aim to prove that $\{y_n\}$ is Cauchy. Use the parallelogramme rule to get

$$\left\|\frac{y_n + y_m}{2}\right\|^2 + \left\|\frac{y_n - y_m}{2}\right\|^2 = \frac{1}{2}\left(\|y_n\|^2 + \|y_m\|^2\right)$$

Note that $\frac{y_n+y_m}{2} \in K$ since K is convex. So, $\left\|\frac{y_n+y_m}{2}\right\|^2 \geq d^2$. Also, $\frac{1}{2}\left(\|y_n\|^2 + \|y_m\|^2\right) \to d^2$ as $n, m \to \infty$. Hence, $\|y_n - y_m\|^2 \to 0$ as $n, m \to \infty$, hence $\{y_n\}$ is Cauchy. So, let $x = \lim_{n\to\infty} y_n \in H$. Then $x \in K$. Also (by 2.2), $\|x\| = \lim_{n\to\infty} \|y_n\| = d$. So x is a best approximation of x_0.

Assume $x, \tilde{x} \in K$, $\|x\| = \|\tilde{x}\| = \inf_{y\in K} \|y\| = d$, $x \neq \tilde{x}$. Then, by the parallelogramme rule

$$\left\|\frac{x + \tilde{x}}{2}\right\|^2 < \left\|\frac{x + \tilde{x}}{2}\right\|^2 + \left\|\frac{x - \tilde{x}}{2}\right\|^2 = \frac{1}{2}\left(\|x\|^2 + \|\tilde{x}\|^2\right) = d^2$$

Hence, $\left\|\frac{x+\tilde{x}}{2}\right\| < d$ and $\frac{x+\tilde{x}}{2} \in K$ — a contradiction. □

Remark. This gives a map $P\colon H \to H$ with $\|x - P(x)\| = \inf_{y\in K} \|x - y\| = d(x, K)$. Clearly, $P(x) \in K$ for $x \in H$, so $P(H) \subseteq K$. So $P(P(x)) = P(x)$ for all $x \in H$, i.e. $P^2 = P$, so P is a projection (not necessarily linear). This typically is used when K is a closed linear subspace of H.

Proposition 2.29. *Let H be a Hilbert space, $K \subseteq H$ be convex and closed and $x_0 \in H$. Then the following are equivalent for $x \in K$:*

(i) $\|x_0 - x\| = d(x_0, K)$.

(ii) $\operatorname{Re}\langle x_0 - x, y - x\rangle \leq 0$ *for all* $y \in K$.

Proof.

(ii) ⇒ (i) This follows from

$$\begin{aligned}\|x_0 - y\|^2 = \|(x_0 - x) + (x - y)\|^2 &= \|x_0 - x\|^2 + 2\operatorname{Re}\langle x_0 - x, x - y\rangle + \|x - y\|^2 \\ &\geq \|x_0 - x\|^2\end{aligned}$$

for all $y \in K$. So $\|x_0 - x\| = d(x_0, K)$.

(i) ⇒ (ii) Let $y \in K$, and for $\lambda \in [0, 1]$ let $y_\lambda = (1 - \lambda)x + \lambda y \in K$. So

$$\begin{aligned}\|x_0 - x\|^2 \leq \|x_0 - y_\lambda\|^2 &= \langle x_0 - x + \lambda(x - y), x_0 - x + \lambda(x - y)\rangle = \\ &= \|x_0 - x\|^2 + 2\operatorname{Re}\langle x_0 - x, \lambda(x - y)\rangle + \lambda^2\|x - y\|^2\end{aligned}$$

Hence, $\operatorname{Re}\langle x_0 - x, y - x\rangle \leq \frac{\lambda}{2}\|x - y\|^2$ for $\lambda \in (0, 1]$, so $\operatorname{Re}\langle x_0 - x, y - x\rangle \leq 0$. □

Theorem 2.30 (Orthogonal projections)**.** *Let $U \neq \{0\}$ be a closed linear subspace of a Hilbert space H. Then there exists a linear projection $P_U\colon H \to H$ with $P_U(H) = U$, $\|P_U\| = 1$, and $\mathrm{N}(P_U) = U^\perp$. Also, $I - P_U$ is a projection on $U^\perp$ with $\|I - P_U\| = 1$ (except if $U = H$), and $H = U \oplus_2 U^\perp$. P_U is called the orthogonal projection on U.*

Proof. Note that U is closed and convex, so $P_U\colon H \to H$ is defined (see above), with $P_U(x)$ the best approximation to x in U. We have seen above that $P_U^2 = P_U$ and $P_U(H) = U$. By 2.29 $\mathrm{Re}\,\langle x_0 - P_U(x_0), y - P_U(x_0)\rangle \leq 0$ for all $y \in U$. Since U is a linear subspace, $y - P_U(x_0) \in U$ for all $y \in U$. Hence, (put $y = \tilde{y} + P_U(x_0), \tilde{y} \in U$) $\mathrm{Re}\,\langle x_0 - P_U(x_0), \tilde{y}\rangle \leq 0$ for all $\tilde{y} \in U$. Now do the same for $-\tilde{y} \in U$ and $i\tilde{y} \in U$ ($\mathbb{K} = \mathbb{C}$). Then

$$\langle x_0 - P_U(x_0), \tilde{y}\rangle = 0 \quad \forall \tilde{y} \in U \tag{$*$}$$

Hence, $P_U(x_0)$ is the unique element in U such that

$$x_0 - P_U(x_0) \in U^\perp \tag{$**$}$$

Since $U^\perp$ is a linear subspace of H it follows that, if $x_1, x_2 \in H$, $\lambda \in \mathbb{K}$, $x_1 - P_U(x_1) \in U^\perp$, $x_2 - P_U(x_2) \in U^\perp$, so $(x_1 + \lambda x_2) - \big(P_U(x_1) + \lambda P_U(x_2)\big) \in U^\perp$. But $P_U(x_1 + \lambda x_2)$ is the unique w such that $(x_1 + \lambda x_2) - w \in U^\perp$. Hence, $P_U(x_1 + \lambda x_2) = P_U(x_1) + \lambda P_U(x_2)$, i.e. P_U is linear. By construction, $\mathrm{R}(P_U) = U$ and from $(**)$ it follows that $P_U(x_0) = 0$ iff $x_0 \in U^\perp$. So $\mathrm{N}(P_U) = U^\perp$. Then also $I - P_U$ is a projection:

$$(I - P_U)^2(x) = (I - P_U)(x - P_U(x)) = x - P_U(x) - P_U(x) + P_U^2(x) = x - P_U(x) = (I - P_U)(x).$$

Also $\mathrm{R}(I - P_U) = U^\perp$ and $\mathrm{N}(I - P_U) = U$. From Pythagoras it follows that $\|x\|^2 = \|(x - P_U(x)) + P_U(x)\|^2 = \|x - P_U(x)\|^2 + \|P_U(x)\|^2$, hence $\|P_U(x)\| \leq \|x\|$ for all x, so $\|P_U\| \leq 1$. On the other hand (for any projection) $\|P_U\| = \|P_U^2\| =\leq \|P_U\|^2$, hence $\|P_U\| = 0$ or $\|P_U\| \geq 1$. So $\|P_U\| = 1$ and similarly $\|I - P_U\| = 1$. By Pythagoras' theorem it is clear that $H = U \oplus_2 U^\perp$. □

Corollary 2.31. *Let U be a linear subspace of a Hilbert space H. Then $\overline{U} = (U^\perp)^\perp$.*

Proof. $U^\perp = \big(\overline{U}\big)^\perp$ is a closed subspace, and by 2.30 $P_{(U^\perp)^\perp} = I - P_{U^\perp} = I - (I - P_{\overline{U}}) = P_{\overline{U}}$. So $\overline{U} = \mathrm{R}(P_{\overline{U}}) = \mathrm{R}(P_{(U^\perp)^\perp}) = \big(U^\perp\big)^\perp$. □

Theorem 2.32 ((Fréchet-)Riesz representation theorem)**.** *For any Hilbert space H the map $\Phi\colon H \to H' = B(H, \mathbb{K}), y \mapsto \langle y, \cdot\rangle$ is bijective, isometric, and conjugate linear, i.e. $\Phi(\lambda y_1 + y_2) = \overline{\lambda}\Phi(y_1) + \Phi(y_2)$. In other words, for every $x' \in H'$ there exists a* unique *$y \in H$ such that $x'(x) = \langle y, x\rangle$ for all $x \in H$.*

Proof. Take any $y \in H$. Then $\Phi(y) \in B(H, \mathbb{K})$, because $H \ni x \mapsto \langle y, x\rangle$ is linear and bounded, since, for all $x \in H$, $|\Phi(y)x| = |\langle y, x\rangle| \leq \|y\|\|x\|$. So $\|\Phi(y)\|_{H'} \leq \|y\|$. Furthermore, Φ is isometric, since for any $y \in H \setminus \{0\}$

$$\|y\| = \left\langle y, \frac{y}{\|y\|}\right\rangle \leq \sup_{\|x\| \leq 1} |\langle y, x\rangle| = \|\Phi(y)\|_{H'},$$

hence $\|\Phi(y)\|_{H'} = \|y\|$.

To see that Φ is surjective, take $x' \in H'$, $x' \neq 0$. Then $U := \mathrm{N}(x') = (x')^{-1}(\{0\})$ is a closed subspace of H. Hence, $H = U \oplus_2 U^{\perp}$, where $U^{\perp} \neq \{0\}$ since $x' \neq 0$.Take any $y \in U^{\perp}$, $\|y\| = 1$, and set $a = x'(y) \in \mathbb{K}$. Then, for all $x \in H$, $x'(x)y - x'(y)x \in \mathrm{N}(x') \perp y$. Hence, $0 = \langle y, x'(x)y - x'(y)x\rangle = x'(x)\langle y, y\rangle - x'(y)\langle y, x\rangle$ for all $x \in H$. Hence, $x'(x) = \langle \overline{a}y, x\rangle$ for all $x \in H$, i.e. $x' = \Phi(\overline{a}y)$. Additionally, $\Phi(0) = 0$. □

Definition 2.33. Let X be a $\mathbb{K}$-vector space. $\mathscr{B} \subseteq X$ is called *algebraic basis*, or *Hamel basis*, iff $\mathscr{B}$ is linearly independent and $\mathrm{span}(\mathscr{B}) = X$. $|\mathscr{B}|$ is called the *algebraic dimension* of X.

Theorem 2.34. *Every vector space has an algebraic basis.*

Lemma 2.35 (Zorn)**.** *If $(M, \leq)$ is a nonempty, partially ordered set in which every nonempty totally ordered subset $\mathscr{C} \subseteq M$ has an upper bound in M, then M contains a maximal element.*

For purposes of functional analysis algebraic bases are de facto useless, because:

Proposition 2.36. *Let X be a Banach space and $\mathscr{B} \subseteq X$ an algebraic basis. If $|\mathscr{B}| = \infty$, then $\mathscr{B}$ is uncountable.*

Definition 2.37. Let X be a Banach space and I be any index set.

(a) A map $x\colon I \to X$ is called *family*, written $\{x(i)\}_{i\in I}$. We denote by $F(I)$ the set of all finite subsets of I.

(b) A family $\{x_i\}_{i\in I} \subseteq X$ is called *absolutely summable*, iff

$$\|x\|_1 = \sum_{i\in I} \|x_i\|_X := \sup\Big\{ \sum_{i\in \tilde{I}} \|x_i\|_X : \tilde{I} \in F(I)\Big\} < \infty.$$

We write

$$\ell_1(I, X) = \{x\colon I \to X\colon \|x\|_1 < \infty\}.$$

(c) For $x \in \ell_1(I, X)$ define the *support* of x by $\mathrm{supp}(x) = \{i \in I\colon x(i) \neq 0\}$.

(d) For $x \in \ell_1(I, X)$, $x = \{x_i\}_{i\in I}$, we can find a bijection $\varphi\colon \mathbb{N} \to J \supseteq \mathrm{supp}(x)$ (if necessary, take a countable $J \supseteq I$ and define $x(j) = 0$ for $j \in J \setminus I$). Then

$$\sum_{i\in I} x_i := \sum_{k=1}^{\infty} x_{\varphi(k)}$$

Here, $\sum_{k=1}^{\infty} x_{\varphi(k)} = \lim_{K\to\infty} \sum_{k=1}^{K} x_{\varphi(k)}$ converges absolutely, hence this is independent of the choice of φ.

(e) A family $\{x_i\}_{i\in I} \subseteq \mathbb{K}$, is called *square summable* iff

$$\|x\|^2 := \sum_{i\in I} |x_i|^2 = \sup\Big\{ \sum_{i\in \tilde{I}} |x_i|^2 \colon \tilde{I} \in F(I) \Big\} < \infty.$$

Again $\operatorname{supp}(x)$ is countable if $x \in \ell_2(I) = \{x\colon I \to \mathbb{K}\colon \|x\|^2 < \infty\}$. Define a scalar product by

$$\langle x, y\rangle = \sum_{i\in I} \overline{x_i} y_i := \sum_{k=1}^{\infty} \overline{x_{\varphi(k)}} y_{\varphi(k)}$$

for $x = \{x_i\}_{i\in I}, y = \{y_i\}_{i\in I} \in \ell_2(I)$ and some bijection $\varphi\colon \mathbb{N} \to J \supseteq \operatorname{supp}(x) \cap \operatorname{supp}(y)$. Completeness of $\ell_2(I)$ follows from the completeness of $\ell_2(\mathbb{N})$, hence $\ell_2(I)$ is a Hilbert space.

Remark. If $x \in \ell_1(I, X)$, then $\operatorname{supp}(x)$ is countable. In fact, for all $n \in \mathbb{N}$, $S_n = \{i \in I\colon \|x_i\|_X \geq \frac{1}{n}\}$ is finite. So, $\operatorname{supp}(x) = \bigcup_{n\in\mathbb{N}} S_n$ is countable.

Definition 2.38. A set $\{e_i\colon i \in I\} \subseteq H$ in a pre-Hilbert space H is called *orthonormal system* iff for all $i, j \in I$, $\langle e_i, e_j\rangle = \delta_{ij}$. An orthonormal system E is called *maximal* iff $E^\perp = \{0\}$. For $x \in H$, the numbers $\hat{x}(i) = \langle e_i, x\rangle$, $i \in I$, are called *Fourier coefficients* of x.

Example.

1. In $\ell_2(I)$ the canonical unit vectors $e_k\colon I \to X, i \mapsto \delta_{ik}$, $k \in I$, form a maximal orthonormal system.
2. $[0, 2\pi] \to \mathbb{C}, t \mapsto (2\pi)^{-1/2} e^{ikt}$, $k \in \mathbb{Z}$, form an orthonormal system in the pre-Hilbert space $\mathscr{C}[0, 2\pi]$ with the scalar product

$$\langle f, g\rangle = \int_0^{2\pi} \overline{f(t)} g(t)\, \mathrm{d}t.$$

Lemma 2.39. *Let $\{e_i\colon i \in I\}$ be an orthonormal system in a pre-Hilbert space H. For every finite subset $J \subseteq I$ we have, for any family $\{x_i\}_{i\in J} \subseteq \mathbb{K}$,* Pythagoras' identity

$$\Big\| \sum_{i\in J} x_i e_i \Big\|^2 = \sum_{i\in J} |x_i|^2$$

and, for any $x \in H$,

$$0 \leq \Big\| x - \sum_{i\in J} \hat{x}(i) e_i \Big\|^2 = \|x\|^2 - \sum_{i\in J} |\hat{x}(i)|^2 \tag{$*$}$$

Proof. Clearly,

$$\Big\| \sum_{i\in J} x_i e_i \Big\|^2 = \sum_{i,j\in J} \overline{x_i} x_j \langle e_i, e_j\rangle = \sum_{i\in J} |x_i|^2.$$

Using this and $\langle x + y, x + y\rangle = \|x\|^2 + \|y\|^2 + 2\operatorname{Re}\langle x, y\rangle$ for any $x, y \in H$, we get

$$\Big\| x - \sum_{i\in J} \hat{x}(i) e_i \Big\|^2 = \|x\|^2 + \sum_{i\in J} |\hat{x}(i)|^2 - 2\operatorname{Re} \sum_{i\in J} \hat{x}(i) \langle x, e_i\rangle = \|x\|^2 - \sum_{i\in J} |\hat{x}(i)|^2. \qquad \square$$

Corollary 2.40 (Bessel's inequality)**.** *Let $\{e_i\colon i \in I\}$ be an orthonormal system in a pre-Hilbert space H. Then, for any $x \in H$, only countably many Fourier coefficients $\hat{x}(i)$ are nonzero and*

$$\sum_{i\in I} |\hat{x}(i)|^2 \leq \|x\|^2.$$

In particular, $\{\hat{x}(i)\}_{i\in I} \in \ell_2(I)$.

Proof. By the lemma, for any $J \in F(I)$,

$$\sum_{i\in J} |\hat{x}(i)|^2 \leq \|x\|^2.$$

Hence,

$$\sum_{i\in I} |\hat{x}(i)|^2 = \sup_{J\in F(I)} \sum_{i\in J} |\hat{x}(i)|^2 \leq \|x\|^2. \qquad \square$$

Remark 2.41.

(a) Given some $x = \{x_i\}_{i\in I} \in \ell_2(I)$ and some orthonormal system $\{e_i\colon i \in I\}$ in a Hilbert space H, one can construct $\sum_{i\in I} x_i e_i$. In fact, pick some bijection $\varphi\colon \mathbb{N} \to J \supseteq \operatorname{supp}(x)$ as in definition 2.37 and observe, using Pythagoras, that

$$\Big\| \sum_{k=m}^{n} x_{\varphi(k)} e_{\varphi(k)} \Big\|^2 = \sum_{k=m}^{n} |x_{\varphi(k)}|^2 \xrightarrow{m,n\to\infty} 0.$$

So $\left\{\sum_{k=1}^{n} x_{\varphi(k)} e_{\varphi(k)}\right\}_{n\in\mathbb{N}}$ is Cauchy in the Hilbert space H, hence is convergent. We define

$$\sum_{i\in I} x_i e_i := \lim_{n\to\infty} \sum_{k=1}^{n} x_{\varphi(k)} e_{\varphi(k)}.$$

This is independent of the choice of φ, because $y = \sum_{i\in I} x_i e_i$ satisfies

$$\forall \varepsilon > 0\, \exists I_\varepsilon \in F(I)\, \forall J \in F(I).\ J \supseteq I_\varepsilon \implies \Big\| \sum_{i\in J} x_i e_i - y \Big\| < \varepsilon.$$

(b) Let $\{e_i\colon i \in I\}$ be an orthonormal system in a Hilbert space H. By Bessel's inequality

$$\mathscr{F}\colon H \to \ell_2(I), x \mapsto \{\hat{x}(i) = \langle e_i, x\rangle\}_{i\in I}$$

is linear and bounded, since

$$\|\mathscr{F}(x)\|^2 = \sum_{i\in I} |\hat{x}(i)|^2 \leq \|x\|^2$$

hence $\|\mathscr{F}\| \leq 1$. $\mathscr{F}\colon H \to \ell_2(I)$ is always surjective: given $x = \{x_i\}_{i\in I} \in \ell_2(I)$, define

$$y = \sum_{i\in I} x_i e_i.$$

Then $\mathscr{F}(y) = x$, since for any $j \in I$, (with notation as in (a))

$$\langle e_j, y\rangle = \lim_{n\to\infty} \sum_{k=1}^{n} \Big\langle e_j, x_{\varphi(k)} \Big\rangle e_{\varphi(k)} = x_j.$$

$\mathscr{F}$ is injective iff $\{e_i \colon i \in I\}$ is maximal.

Theorem 2.42. *Let $\{e_i \colon i \in I\}$ be an orthonormal system in a Hilbert space H. Then the following are equivalent*

(a) For all $x \in H$, $x = \sum_{i\in I} \hat{x}(i) e_i$.

*(b) For all $x \in H$, $\|x\|^2 = \sum_{i\in I} |\hat{x}(i)|^2$ (*Parseval's identity*)*

(c) $\mathscr{F} \colon H \to \ell_2(I)$ is isometric.

(d) $\operatorname{span}\{e_i \colon i \in I\}$ is dense in H.

(e) $\mathscr{F} \colon H \to \ell_2(I)$ is injective.

(f) $\{e_i \colon i \in I\}$ is maximal.

Proof.

(a)⇔(b) Follows from $(*)$.

(b)⇔(c) Follows from the definition of "isometry".

(a)⇒(d) Clear.

(d)⇒(f) $\{e_i \colon i \in I\}^\perp = \overline{\operatorname{span}\{e_i \colon i \in I\}}^\perp = \{0\}$.

(f)⇒(e) $\mathscr{F}(x) = 0$ is equivalent to $\langle e_i, x\rangle = 0$ for all $i \in I$., i.e $x \in \{e_i \colon i \in I\}^\perp = \{0\}$.

(e)⇒(a) Note $\mathscr{F}(x) = \{\hat{x}(i)\}_{i\in I}$ and

$$\mathscr{F}\Big(\sum_{i\in I} \hat{x}(i) e_i\Big) = \{\hat{x}(i)\}_{i\in I}.$$

Hence, by injectivity of $\mathscr{F}$, $x = \sum_{i\in I} \hat{x}(i) e_i$. □

Remark 2.43. A maximal orthonormal system is also called *complete* or *orthonormal basis*. If $\dim H = \infty$, then an orthonormal basis in general is *not* an algebraic basis, i.e. the expansion $x = \sum_{i\in I} \hat{x}(i) e_i$ in general has infinitely many summands.

Theorem 2.44.

(a) Every Hilbert space H has an orthonormal basis $\{e_i \colon i \in I\}$. In particular, H is isometrically isomorphic to $\ell_2(I)$.

(b) H has a countable orthonormal basis iff H is separable. In this case $H \cong \ell_2(\mathbb{N})$, if H is infinite dimensional.

Proof.

(a) Write $\mathfrak{M}$ for the set of all orthonormal systems in H. Then $\mathfrak{M}$ is partially ordered by $\subseteq$. Let $\mathfrak{C} \subseteq \mathfrak{M}$ be a totally ordered subset. Then $\hat{B} = \bigcup \mathfrak{C}$ is an orthonormal system and an upper bound of $\mathfrak{C}$ in $\mathfrak{M}$. Indeed $\hat{B} \in \mathfrak{M}$, since take $e_1, e_2 \in \hat{B}$, $e_1 \neq e_2$. Then $e_i \in B_i \in \mathfrak{C}$, $i = 1, 2$. Since $\mathfrak{C}$ is totally ordered by $\subseteq$, we have $B_1 \subseteq B_2$ or $B_2 \subseteq B_1$, say $B_1 \subseteq B_2$. Then $e_1, e_2 \in B_2$, hence $e_1 \perp e_2$. So by Zorn's lemma, there is a maximal element $M \in \mathfrak{M}$. M is an orthonormal basis, since if $\overline{\operatorname{span} M} \neq H$, there would exist some $e \in H$, $\|e\| = 1$ such that $e \perp M$ and $M \cup \{e\} \supsetneq M$ — a contradiction.

(b) If H has a countably infinite orthonormal basis $\{e_i : i \in \mathbb{N}\}$ then $\mathscr{F} : H \to \ell_2(\mathbb{N})$ is an isometric isomorphism. Since $\ell_2(\mathbb{N})$ is separable, H must be separable. Conversely, every orthonormal system $\{e_i : i \in I\}$ is discrete, since for all $i, j \in I$, $i \neq j$, $\|e_i - e_j\|^2 = 2$. Hence, if the orthonormal system $\{e_i : i \in I\}$ is uncountable, H cannot be separable. □

Remark. All orthonormal bases of a Hilbert space H have the same cardinality. This cardinality is then called the *Hilbert space dimension* of H.

3 Lebesgue integration

In Lebesgue integration the concept of measure is essential. But to make this concept useful one has to consider σ-algebras different from the power set. In fact Vitali proved in 1905 that there can be no measure $\mu : 2^{\mathbb{R}^d} \to [0, \infty]$ such that $\mu([0,1]^d) = 1$ and $\mu \circ \beta = \mu$ for every rigid motion β. Even worse, Banach and Tarski proved in 1924 that for any two bounded sets $A, B \subseteq \mathbb{R}^d$ such that $A^\circ \neq \emptyset \neq B^\circ$ there exist disjoint $C_1, \dots, C_n \subseteq \mathbb{R}^d$ and rigid motions $\beta_1, \dots, \beta_n : \mathbb{R}^d \to \mathbb{R}^d$ such that $\beta_1(C_1), \dots, \beta_n(C_n)$ are disjoint and

$$A = \bigcup_{\ell=1}^{n} C_\ell, \quad B = \bigcup_{\ell=1}^{n} \beta_\ell(C_\ell).$$

This shows that there have to exist sets for which the notion of volume does not make sense. Instead one has to consider σ-algebras:

Definition 3.1. Let X be a set. A system of sets $\mathfrak{A} \subseteq 2^X$ is called *σ-algebra* iff

(i) $\emptyset \in \mathfrak{A}$.
(ii) For any $A \in \mathfrak{A}$, $A^c = X \setminus A \in \mathfrak{A}$.
(iii) For any countable family $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathfrak{A}$, $\bigcup_{i \in \mathbb{N}} A_i \in \mathfrak{A}$.

Example.

(a) 2^X is a σ-Algebra.
(b) For any index set I and σ-algebras $\mathfrak{A}_i$, $i \in I$, $\bigcap_{i \in I} \mathfrak{A}_i$ is again a σ-algebra.
(c) Take any $\mathscr{E} \subseteq 2^X$. Then

$$\sigma(\mathscr{E}) := \bigcap \{\mathfrak{A} \subseteq 2^X : \mathscr{E} \subseteq \mathfrak{A} \text{ and } \mathfrak{A} \text{ is a } \sigma\text{-algebra}\}$$

is a σ-algebra. $\sigma(\mathscr{E})$ is called the σ-algebra generated by $\mathscr{E}$.

(d) Let $\{X, \mathscr{T}\}$ be a topological space. Then $\mathscr{B}(X) := \sigma(\mathscr{T})$ is called *Borel σ-algebra.* We have $\mathscr{B}(\mathbb{R}^d) \subsetneq 2^{\mathbb{R}^d}$.

Definition 3.2. Let $\mathfrak{A}$ be a σ-algebra. Then a map $\mu\colon \mathfrak{A} \to [0, \infty]$ is called *measure* iff $\mu(\emptyset) = 0$ and it is *σ-additive*, i.e. for any countable *disjoint* family $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathfrak{A}$,

$$\mu\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n).$$

Example.

(a) Let X be a set and define $\zeta\colon 2^X \to [0, \infty]$ by

$$\zeta(A) = \begin{cases} n & \text{if } |A| = n \in \mathbb{N} \\ \infty & \text{if } A \text{ is infinite} \end{cases}$$

ζ is called *counting measure* on X.

(b) For any $a \in X$, the measure δ_A defined on 2^X by

$$\delta_a(A) := \begin{cases} 1 & a \in A \\ 0 & a \notin A \end{cases}$$

is called *Dirac measure* at a.

Definition 3.3. A system $\mathfrak{h} \subseteq 2^X$ is called *semi-ring* iff

(i) $\emptyset \in \mathfrak{h}$.

(ii) For all $A, B \in \mathfrak{h}$, $A \cap B \in \mathfrak{h}$.

(iii) For all $A, B \in \mathfrak{h}$ there exist disjoint $C_1, \dots, C_n \in \mathfrak{h}$ such that $A \setminus B = C_1 \cup \dots \cup C_n$.

Example. For $a = (a_1, \dots, a_d), b = (b_1, \dots, b_d) \in \mathbb{R}^d$ we write $a \leq b$ iff $a_j \leq b_j$ for all $1 \leq j \leq n$, and $(a, b] = (a_1, b_1] \times \dots \times (a_d, b_d]$. Then

$$J^d = \{(a, b]\colon a, b \in \mathbb{R}^d, a \leq b\}$$

and

$$J^d_{\mathbb{Q}} = \{(a, b]\colon a, b \in \mathbb{Q}^d, a \leq b\}$$

are *semi-rings* and $\sigma(J^d_{\mathbb{Q}}) = \sigma(J^d) = \mathscr{B}(\mathbb{R}^d)$.

Definition 3.4. Let $\mathfrak{h}$ be a semi-ring. A map $\mu\colon \mathfrak{h} \to [0, \infty]$ is called *content* if $\mu(\emptyset) = 0$ and μ is finitely additive, i.e. for all disjoint $A_1, \dots, A_n \in \mathfrak{h}$ such that $\bigcup_{i=1}^n A_i \in \mathfrak{h}$, $\mu(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i)$. A content is called *premeasure* if it is σ-additive, i.e. for all disjoint $A_1, A_2, \dots \in \mathfrak{h}$ such that $\bigcup_{i=1}^\infty A_i \in \mathfrak{h}$, $\mu(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu(A_i)$.

Example.
(a) Define $\lambda^d\colon J^d \to [0,\infty)$ by

$$\lambda^d\big((a,b]\big) = \prod_{i=1}^{n}(b_i - a_i)$$

for all $a,b \in \mathbb{R}^d$, $a \le b$. λ^d is called *Lebesgue-content.* It can be shown that λ^d is a premeasure.

(b) Let $F\colon \mathbb{R} \to \mathbb{R}$ be a monotonically increasing function. Then $\mu_F\big((a,b]\big) = F(b) - F(a)$, $a \le b$, defines the *Lebesgue-Stieltjes content* associated with F. μ_F is a premeasure iff F is upper semicontinuous.

Definition 3.5. An *exterior measure* is a map $\eta\colon 2^X \to [0,\infty]$ such that

(i) $\eta(\emptyset) = 0$.

(ii) $A \subseteq B$ implies $\eta(A) \le \eta(B)$.

(iii) For any countable family $\{A_n\}_{n\in\mathbb{N}} \subseteq 2^X$,

$$\eta\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) \le \sum_{n=1}^{\infty} \eta(A_n).$$

Then $A \subseteq X$ is called *η-measurable* if for all $Q \subseteq X$, $\eta(Q) = \eta(Q\cap A) + \eta(Q \cap A^c)$.

Theorem 3.6 (Carathéodory)**.** *Let $\mu\colon \mathfrak{h} \to [0,\infty]$ be a content on the semi-ring $\mathfrak{h} \subseteq 2^X$ and define for all $A \subseteq X$:*

$$\eta(A) := \inf\Big\{\sum_{n=1}^{\infty}\mu(A_n)\colon A_n \in \mathfrak{h}, A \subseteq \bigcup_{n=1}^{\infty} A_n\Big\} \qquad (*)$$

Then $\eta\colon 2^X \to [0,\infty]$ is an exterior measure and every $A \in \mathfrak{h}$ is η-measurable. Additionally, $\mathfrak{A}_\eta = \{A \subseteq X\colon A\ \eta\text{-measurable}\}$ is a σ-algebra and $\eta|_{\mathfrak{A}_\eta}$ is a measure. If μ is a premeasure then $\eta|_{\mathfrak{h}} = \mu$.

Example. The *exterior Lebesgue measure* is

$$\lambda^d_*(A) = \inf\Big\{\sum_{n=1}^{\infty}\lambda^d(A_n)\colon A_n \in J^d, A \subseteq \bigcup_{n=1}^{\infty} A_n\Big\}$$

$\mathfrak{A}_{\lambda^d_*}$ is the σ-algebra of *Lebesgue-measurable* sets. We have $\mathscr{B}(\mathbb{R}^d) \subsetneq \mathfrak{A}_{\lambda^d_*}$. $\lambda^d := \lambda^d_*|_{\mathfrak{A}_{\lambda^d_*}}$ is called *Lebesgue-measure* on $\mathbb{R}^d$ and $\lambda^d|_{\mathscr{B}(\mathbb{R}^d)}$ is called *Lebesgue-Borel-measure.*

Definition 3.7. A content $\mu\colon \mathfrak{h} \to [0,\infty]$ is called *σ-finite* if there exist countably many $A_1, A_2, \ldots \in \mathfrak{h}$ such that $\mu(A_n) < \infty$, $n \in \mathbb{N}$, and $\bigcup_{n\in\mathbb{N}} A_n = X$.

Theorem 3.8. *A σ-finite premeasure $\mu\colon \mathfrak{h} \to [0,\infty]$ can be uniquely extended to a measure on $\sigma(\mathfrak{h})$/on $\mathfrak{A}_\eta$ with η as in $(*)$.*

Example. The Lebesgue and Lebesgue-Stieltjes premeasures $\lambda^d\colon J^d \to [0,\infty]$ and $\mu_F\colon J^1 \to [0,\infty]$ are σ-finite. E.g.

$$\mathbb{R}^d = \bigcup_{n\in\mathbb{N}} (-n,n]^d.$$

3.1 Measurable Functions

Definition 3.9. Let $\mathfrak{A}$ be a σ-algebra on X. $f\colon X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ is called $\mathfrak{A}$-measurable if for all $a \in \mathbb{R}$, $f^{-1}((a, \infty]) \in \mathfrak{A}$.

Remark 3.10.

(a) If $f, g, f_1, f_2, \dots$ are measurable functions, $f+g$, fg, $\max_{1\leq i\leq n}\{f_1, \dots, f_n\}$, $\sup_{n\in\mathbb{N}} f_n$, $\inf_{n\in\mathbb{N}} f_n$, $\limsup_{n\to\infty} f_n$ and $\liminf_{n\to\infty} f_n$ are all measurable.

(b) $f\colon X \to [0, \infty]$ is measurable iff there exists a sequence of measurable step functions $\{u_n\}$, i.e. $u_n\colon X \to [0, \infty]$, $|u_n(X)| < \infty$ such that $u_n \nearrow f$ pointwise, i.e $u_1 \leq u_2 \leq \dots \leq f$ and $u_n(t) \to f(t)$ as $n \to \infty$. Indeed for a measurable function $f\colon X \to [0, \infty]$, define

$$u_n = \sum_{j=0}^{n2^n-1} \frac{j}{2^n}\chi_{\{j/2^n\leq f<(j+1)/2^n\}} + n\chi_{\{f\geq n\}}$$

Example. Any $f \in \mathscr{C}(\mathbb{R}^n, \mathbb{R})$ is measurable, if we take $\mathscr{B}(\mathbb{R}^n)$ as a σ-algebra on $\mathbb{R}^n$.

To define the Lebesgue integral we first define the integral of measurable step functions $u = \sum_{i=1}^n \alpha_i\chi_{A_i}$, $A_i = u^{-1}(\{\alpha_i\}) \in \mathfrak{A}$, $\alpha_i \geq 0$, with respect to the measure μ by

$$\int_X u \,\mathrm{d}\mu := \sum_{i=1}^n \alpha_i\mu(A_i) \in [0, \infty].$$

Now, taking a measurable function $f\colon X \to [0, \infty]$, pick measurable step functions $\{u_n\}$ with $u_N \nearrow f$ pointwise and define

$$\int_X f \,\mathrm{d}\mu = \lim_{n\to\infty} \int_X u_n \,\mathrm{d}\mu.$$

This is independent of the choice of $\{u_n\}$ which we will not prove here. It may happen that $\int_X f \,\mathrm{d}\mu = \infty$. We call f *μ-integrable* if $\int_X f \,\mathrm{d}\mu < \infty$. Finally, let $f\colon X \to \overline{\mathbb{R}}$ be measurable. Split $f = f_+ - f_-$ with $f_\pm \geq 0$. If at least one of the integrals $\int_X f_\pm \,\mathrm{d}\mu$ are finite, we define

$$\int_X f \,\mathrm{d}\mu = \int_X f_+ \,\mathrm{d}\mu - \int_X f_- \,\mathrm{d}\mu.$$

f is called μ-integrable if $\int_X f \,\mathrm{d}\mu \in \mathbb{R}$. Analogously, for $f\colon X \to \mathbb{C} \cup \{\infty\}$ define

$$\int_X f \,\mathrm{d}\mu = \int_X (\operatorname{Re} f)_+ \,\mathrm{d}\mu + i\int_X (\operatorname{Im} f)_+ \,\mathrm{d}\mu - \int_X (\operatorname{Re} f)_- \,\mathrm{d}\mu - i\int_X (\operatorname{Im} f)_- \,\mathrm{d}\mu$$

assuming all integrals are finite and defined. Note, that $f\colon X \to \mathbb{C} \cup \{\infty\}$ is integrable iff $|f|$ is integrable.

Remark. If $f\colon [0, 1] \to \mathbb{R}$ is Riemann-integrable then f is Lebesgue-integrable (with respect to λ^1) and

$$\int_0^1 f(t) \,\mathrm{d}t = \int_{[0,1]} f \,\mathrm{d}\lambda^1.$$

Convention. A statement $P(x)$ is said to hold μ-*almost everywhere* if $\mu(\{x\colon \neg P(x)\}) = 0$, e.g. "$f = g$ μ-almost everywhere" if $\mu(\{x\colon f(x) \neq g(x)\}) = 0$.

Remark.

(a) For all measurable $f\colon X \to [0,\infty]$, $\int_X f \,\mathrm{d}\mu = 0$ implies $f = 0$ μ-almost everywhere.

(b) For all integrable $f, g\colon X \to \mathbb{C} \cup \{0\}$, $f = g$ μ-almost everywhere implies

$$\int_X f \,\mathrm{d}\mu = \int_X g \,\mathrm{d}\mu.$$

3.2 p-integrable functions, $p \geq 1$

Definition 3.11. For a measure space $(X, \mathfrak{A}, \mu)$ we define

$$\mathscr{L}^p(X,\mu) = \left\{ f\colon X \to \mathbb{C} \cup \{\infty\} \colon f \text{ is measurable and } \int_X |f|^p \,\mathrm{d}\mu < \infty \right\}$$

and, for $f \in \mathscr{L}^p(X,\mu)$, we set

$$\|f\|_p = \left(\int_X |f|^p \,\mathrm{d}\mu \right)^{1/p}$$

Remark. $\|\cdot\|_p$ is only a semi-norm on $\mathscr{L}^p(X,\mu)$, since $\|f\|_p = 0$ only implies $f = 0$ μ-almost everywhere. Because of this we consider equivalence classes with respect to the equivalence relation

$$f \sim g \iff f = g \ \mu\text{-almost everywhere.}$$

Then $f = 0$ μ-almost everywhere is equivalent to $[f] = 0$.

Definition 3.12. For a measure space $(X, \mathfrak{A}, \mu)$ we define

$$L^p(X,\mu) = \mathscr{L}^p(X,\mu)/\!\sim\, = \{[f]\colon f \in \mathscr{L}^p(X,\mu)\}$$

and $\|[f]\|_p = \|f\|_p$. Then $\|\cdot\|_p$ is non-degenerate on $L^p(X,\mu)$.

Convention. One always writes f instead of $[f]$ for elements in $L^p(X,\mu)$. It should be clear from context when f is a function and when f is an equivalence class.

Our goal is to prove the following theorem:

Theorem 3.13 (Riesz-Fischer)**.** $\{L^p(X,\mu), \|\cdot\|_p\}$ *is a Banach space.*

For this we will prove that $L^p(X,\mu)$ is a vectorspace (1), $\|\cdot\|_p$ is a norm on $L^p(X,\mu)$ (2) and that any Cauchy sequence in $L^p(X,\mu)$ converges to an element in $L^p(X,\mu)$.

Proof of (1). Assume $f, g \in L^p(X,\mu)$. Note that for any $\alpha \in \mathbb{C}$, $\alpha\colon X \to \mathbb{C} \cup \{\infty\}, x \mapsto \alpha$ is measurable. Hence $\alpha f + g$ is measurable. Also,

$$\int_X |\alpha f + g|^p \,\mathrm{d}\mu \leq \int_X (|\alpha f| + |g|)^p \,\mathrm{d}\mu \leq \int_X (2\max\{|\alpha f|, |g|\})^p \,\mathrm{d}\mu \leq$$
$$\leq 2^p \int \max\{|\alpha|^p |f|^p, |g|^p \,\mathrm{d}\}\mu \leq 2^p \left(|\alpha|^p \int_X |f|^p \,\mathrm{d}\mu + \int_X |g|^p \,\mathrm{d}\mu \right) < \infty$$

Hence, $\alpha f + g \in L^p(X,\mu)$, i.e. $L^p(X,\mu)$ is a vectorspace. $\square$

Proposition 3.14 (Hölder's inequality)**.** *Let* $1 < p < \infty$*, and* $\frac{1}{p}+\frac{1}{q} = 1$*. If* $f \in L^p(X, \mu)$*,* $g \in L^q(X, \mu)$*, then* $fg \in L^1(X, \mu)$ *and*

$$\int_X |fg| \,\mathrm{d}\mu = \|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Proof. Recall that if $a, b \geq 0$, then

$$ab = \inf_{\varepsilon>0} \varepsilon^p \frac{a^p}{p} + \varepsilon^{-q} \frac{b^q}{q}$$

So for all $\varepsilon > 0$ and $t \in X$

$$|f(t)g(t)| = |f(t)||g(t)| \leq \varepsilon^p \frac{|f(t)|^p}{p} + \varepsilon^{-q} \frac{|g(t)|^q}{q}$$

Hence, for all $\varepsilon > 0$,

$$\int_X |fg| \,\mathrm{d}\mu \leq \varepsilon^p \frac{\|f\|_p^p}{p} + \varepsilon^{-q} \frac{\|g\|_q^q}{q}$$

and

$$\|fg\|_1 = \int_X |fg| \,\mathrm{d}\mu \leq \inf_{\varepsilon>0} \varepsilon^p \frac{\|f\|_p^p}{p} + \varepsilon^{-q} \frac{\|g\|_q^q}{q} = \|f\|_p \|g\|_q.$$

□

Corollary 3.15 (Minkowski's inequality)**.** *For* $p \geq 1$*, and* $f, g \in L^p(X, \mu)$*, we have*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

Proof. We have for $\frac{1}{p} + \frac{1}{q} = 1$ and $p > 1$

$$\int_X |f+g|^p \,\mathrm{d}\mu = \int_X |f+g||f+g|^{p-1} \,\mathrm{d}\mu \leq \int_X |f||f+g|^{p-1} \,\mathrm{d}\mu + \int_X |g||f+g|^{p-1} \,\mathrm{d}\mu \leq$$
$$\leq \|f\|_p \left\| |f+g|^{p-1} \right\|_q + \|g\|_p \left\| |f+g|^{p-1} \right\|_q = (\|f\|_p + \|g\|_p) \|f+g\|_p^{p-1}$$

since $\left\| |f+g|^{p-1} \right\|_q = \|f+g\|_p^{p-1}$. Dividing by $\|f+g\|_p^{p-1}$ yields Minkowski's inequality for $p > 1$. The cases $\|f+g\|_p = 0$ and $p = 1$ are trivial. □

For the proof of (3) recall proposition 2.5:

Lemma 3.16. *Let* X *be a normed space. The following are equivalent:*
(i) X *is a Banach space.*
(ii) Any absolutely convergent series is convergent.

Additionally we will need the following two important convergence results for Lebesgue integration (they "solve" the question: if $f_n(t) \to f(t)$ as $n \to \infty$ "for all t", is it true that $\int f_n \to \int f$ as $n \to \infty$?)

Theorem 3.17 (Beppo-Levi's theorem/Lebesgue's theorem on monotone convergence)**.** *Let* $(X, \mathfrak{A}, \mu)$ *be a measure space, and let* $f_1, f_2, \ldots : X \to [0, \infty]$ *be measurable, with* $f_1 \leq f_2 \leq \ldots$*. Let* $f(t) = \lim_{n\to\infty} f_n(t) \in [0, \infty]$*. Then* f *is measurable and*

$$\lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu = \int_X f \,\mathrm{d}\mu.$$

Theorem 3.18 (Lebesgue's theorem on dominated convergence)**.** *Let $f_1, f_2, \ldots : X \to \mathbb{C}$ be integrable and assume $f(t) = \lim_{n\to\infty} f_n(t)$ for μ-almost every t, and that f is measurable. Furthermore, assume there exists an integrable $g\colon X \to [0,\infty]$ such that $|f_n| \le g$ for all $n \in \mathbb{N}$. Then*

$$\lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu = \int_X f \,\mathrm{d}\mu.$$

Now we can prove the completeness of $L^p(X,\mu)$.

Proof of (3). Take $f_1, f_2, \dots \in L^p(X,\mu)$ such that $a = \sum_{n=1}^{\infty} \|f_n\|_p < \infty$. Let $\hat{g}(t) = \sum_{i=1}^{\infty} |f_i(t)|$, $t \in X$. Then $\hat{g}\colon X \to [0,\infty]$. Note that

$$\hat{g}(t) = \sup_{n\in\mathbb{N}} \sum_{i=1}^{n} |f_i(t)| = \lim_{n\to\infty} \hat{g}_n(t)$$

Where $\hat{g}_n(t) = \sum_{i=1}^{n} |f_i(t)|$. Hence, $\hat{g}_n$ and $\hat{g}$ are measurable. Also, $\hat{g}_n \in L^p(X,\mu)$, and

$$\|\hat{g}_n\|_p \le \sum_{i=1}^{n} \|f_i\|_p \le \sum_{i=1}^{\infty} \|f_i\|_p = a < \infty$$

for all $n \in \mathbb{N}$. By construction, $\hat{g}_n \nearrow \hat{g}$ as $n \to \infty$. Hence, $\hat{g}_n^p \nearrow \hat{g}^p$. Hence,

$$\int_X \hat{g}^p \,\mathrm{d}\mu = \lim_{n\to\infty} \int_X \hat{g}_n^p \,\mathrm{d}\mu = \lim_{n\to\infty} \|\hat{g}_n\|_p^p \le a^p < \infty.$$

Hence, $\hat{g} \in L^p(X,\mu)$. Since $\hat{g}\colon X \to [0,\infty]$, this implies $\hat{g}$ ist finite μ-almost everywhere, i.e., by possibly changing $\hat{g}$ on a set of measure 0, we get a finite-valued function $g\colon X \to [0,\infty)$ with $g(t) = \sum_{i=1}^{\infty} |f_i(t)|$ μ-almost everywhere. By Lemma 3.16 (for the Banach space $\mathbb{K}$), it follows that $f(t) := \sum_{i=1}^{\infty} f_i(t)$ is welldefined/finite for all $t \in X \setminus N$, $\mu(N) = 0$. Setting $f(t) = 0$ for all $t \in \mathbb{N}$ makes f measurable, $f\colon X \to \mathbb{K}$. It remains to show that $f \in L^p(X,\mu)$ and that $f = \sum_{i=1}^{\infty} f_i$ in $L^p(X,\mu)$, i.e. $\left\|\sum_{i=1}^{n-1} f_i - f\right\|_p \to 0$ as $n \to \infty$, i.e.

$$\int_X \Big|\sum_{i=n}^{\infty} f_i\Big|^p \,\mathrm{d}\mu \to 0, \quad n \to \infty.$$

By construction, $|f| \le \sum_{i=1}^{\infty} |f_i| = \hat{g}$ and

$$\int_X |f|^p \,\mathrm{d}\mu \le \int_X \hat{g}^p \,\mathrm{d}\mu \le a^p < \infty.$$

So, $f \in L^p(X,\mu)$. Finally, let

$$h_n = \Big|\sum_{i=n}^{\infty} f_i\Big|^p.$$

Then $h_n \to 0$ μ-almost everywhere as $n \to \infty$ and $0 \le h_n \le (\sum_{i=n}^{\infty} |f_i|)^p \le \hat{g}^p \in L^1(X,\mu)$. Now, by Lebesgue's theorem of dominated convergence,

$$\int_X h_n \,\mathrm{d}\mu \to \int_X 0 \,\mathrm{d}\mu = 0, \quad n \to \infty.$$

□

Remark. In the case $X = \mathbb{N}$, $\mathfrak{A} = 2^{\mathbb{N}}$ and μ the counting measure on $\mathbb{N}$, we have $L^p(X,\mu) = \ell_p(\mathbb{N})$. So, in fact, the proof of completeness of ℓ_p is contained in the above.

For $p = \infty$, the definition of $L^\infty(X,\mu)$ is slightly different (here, $(B(X,\mathbb{C}), d_\infty)$ is not the good concept).

Definition 3.19. Define

$$\mathscr{L}^\infty(X,\mu) = \{f\colon X \to \mathbb{C}\colon f \text{ is measurable and } \exists N \in \mathfrak{A}, \mu(N) = 0.\ f|_{X\smallsetminus N} \text{ is bounded}\}.$$

and $L^\infty(X,\mu) = \mathscr{L}^\infty(X,\mu)/\sim$ where again $f \sim g$ iff $f = g$ μ-almost everywhere. We set

$$\|[f]\|_\infty = \inf_{\substack{N\in\mathfrak{A}\\ \mu(N)=0}} \sup_{t\in X\smallsetminus N} |f(t)| = \inf_{\substack{N\in\mathfrak{A}\\ \mu(N)=0}} \|f|_{X\smallsetminus N}\|_\infty.$$

$\|[f]\|_\infty$ is called the *essential supremum* of f. It is "easy" to see that $L^\infty(X,\mu)$ is a vectorspace and $\|\cdot\|_\infty$ is a norm on $L^\infty(X,\mu)$, and that $(L^\infty(X,\mu), \|\cdot\|_\infty)$ is a Banach space.

Remark. Hölder's inequality holds for $p, q \in [1,\infty]$, $\frac{1}{p} + \frac{1}{q} = 1$ (with the convention $\frac{1}{\infty} = 0$).

4 Cornerstones of functional analysis

We return to the general abstract theory, to prove some of the most important results in functional analysis. Recall, for X a normed $\mathbb{K}$-vectorspace, $X' = B(X,\mathbb{K})$ is called the dual of X. There are two important questions about this space. Firstly, is $X' = \{0\}$? Secondly, what "is" X' for concrete examples of Banach spaces X?

Definition 4.1. Let E be an $\mathbb{R}$-vector space. A map $p\colon E \to \mathbb{R}$ is called a *sublinear functional* iff for $x, y \in E$

(i) $p(x+y) \le p(x) + p(y)$

(ii) $p(tx) = tp(x)$ for all $t \ge 0$.

Example. Any semi-norm and any norm is a sublinear functional

Theorem 4.2 (Hahn-Banach)**.** *Let E be an $\mathbb{R}$-vector space, $V_0 \subseteq E$ a linear subspace. Let $p\colon E \to \mathbb{R}$ be a sublinear functional, and $f_0\colon V_0 \to \mathbb{R}$ a linear form, such that $f_0(x) \le p(x)$ for all $x \in V_0$. Then there exists a linear form $f\colon E \to \mathbb{R}$ such that $f|_{V_0} = f_0$ and $f(x) \le p(x)$ for all $x \in E$.*

Proof. Idea: 1) Extend f_0 to "one dimension more" (preserving the bound) and 2) "keep going until done".

For step 1), let $x_1 \in E \smallsetminus V_0$ (this is nonempty, otherwise we are done) and define $V_1 = V_0 \oplus \operatorname{span} x_1 = \{x + \lambda x_1\colon x \in V_0, \lambda \in \mathbb{R}\} \subseteq E$ (linear subspace). For $x, y \in V_0$: $f_0(x) + f_0(y) = f_0(x+y) \le p(x+y) = p(x - x_1 + x_1 + y) \le p(x - x_1) + p(x_1 + y)$. Hence, $f_0(x) - p(x - x_1) \le p(x_1 + y) - f_0(y)$. Let $\alpha = \sup_{x\in V_0}(f_0(x) - p(x - x_1))$. Then $f_0(x) - p(x - x_1) \le \alpha$ for all $x \in E$, hence

(1) $f_0(x) - \alpha \leq p(x - x_1)$ for all $x \in V_0$

(2) $f_0(y) + \alpha \leq p(x_1 + y)$ for all $y \in V_0$.

Now, let $f_1 \colon V_1 \to \mathbb{R}$ be given by $f_1(x+\lambda x_1) = f_0(x)+\lambda\alpha$ for $x+\lambda x_1 \in V_1$ ($x \in V_0$, $\lambda \in \mathbb{R}$). Then f_1 is linear and $f_1|_{V_0} = f_0$. We still need to prove that $f_1(x + \lambda x_1) \leq p(x + \lambda x_1)$ for all $x \in V_0$, $\lambda \in \mathbb{R}$. Use (2) for $\lambda > 0$, $y \in V_0$

$$f_0\left(\frac{y}{\lambda}\right) + \alpha \leq p\left(\frac{y}{\lambda} + x_1\right)$$

So

$$f_1(y + \lambda x_1) = f_0(y) + \lambda\alpha = \lambda\left(f_0\left(\frac{y}{\lambda}\right) + \alpha\right) \leq \lambda p\left(\frac{y}{\lambda} + x_1\right) = p(y + \lambda x_1)$$

If $\lambda < 0$, then $-\lambda > 0$. Let $x \in V_0$. By (1),

$$f_0\left(\frac{x}{-\lambda}\right) - \alpha \leq p\left(\frac{x}{-\lambda} - x_1\right)$$

Hence

$$f_1(x + \lambda x_1) = f_0(x) + \lambda\alpha = -\lambda\left(f_0\left(\frac{x}{-\lambda}\right) - \alpha\right) \leq -\lambda p\left(\frac{x}{-\lambda} - x_1\right) = p(x + \lambda x_1)$$

Hence, $f_1 \colon V_1 \to \mathbb{R}$ is linear, $f_1|_{V_0} = f_0$ and $f_1(x) \leq p(x)$ for all $x \in V_1$.

For step 2), let $\mathcal{S}$ be the family of all pairs (V', f') with $V_0 \subseteq V' \subseteq E$, V' linear subspace, and $f' \colon V' \to \mathbb{R}$ with $f'|_{V_0} = f_0$ and $f'(x) \leq p(x)$ for all $x \in V'$. We define a partial ordering $\prec$ on $\mathcal{S}$ by $(V', f') \prec (V'', f'')$ iff $V' \subseteq V''$ and $f''|_{V'} = f'$. Let $\mathcal{T} \subseteq \mathcal{S}$ be totally ordered (i.e. for any $(V', f'), (V'', f'') \in \mathcal{T}$, either $(V', f') \prec (V'', f'')$ or $(V'', f'') \prec (V', f')$). Let $V^* = \bigcup_{V \in \mathcal{T}} V$. V^* is a linear subspace of E. Let $f^*(x) = f'(x)$ for $x \in V' \in \mathcal{T}$. This is well-defined since $\mathcal{T}$ is totally ordered. Now $(V', f') \prec (V^*, f^*)$ for all $(V', f') \in \mathcal{T}$. Hence, (V^*, f^*) is an upper bound for $\mathcal{T}$. In other words: Every totally ordered subfamily $\mathcal{T}$ of $\mathcal{S}$ has an upper bound. By Zorn's Lemma, $\mathcal{S}$ has a maximal element, i.e. there exists $(V, f) \in \mathcal{S}$ such that if $(V', f') \in \mathcal{S}$ satisfies $(V, f) \prec (V', f')$, then $(V, f) = (V', f')$. Note (by step 1), $V \equiv E$, and so $f \colon E \to \mathbb{R}$ is linear and $f_{V_0} = f_0$ and $f(x) \leq p(x)$ for all $x \in E$. ☐

Remark. Note that $-f(x) = f(-x) \leq p(-x)$, so $-p(-x) \leq f(x) \leq p(x)$ for all $x \in E$.

Theorem 4.3 (Hahn-Banach for semi-norms)**.** *Let E be a $\mathbb{K}$-vector space ($\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$) and $V_0 \subseteq E$ a linear subspace. Let $p \colon E \to \mathbb{R}$ be a semi-norm, and $f_0 \colon V_0 \to \mathbb{K}$ be a $\mathbb{K}$-linear form, with $|f_0(x)| \leq p(x)$ for all $x \in V_0$. Then there exists a $\mathbb{K}$-linear form $f \colon E \to \mathbb{K}$ such that $f|_{V_0} = f_0$ and $|f(x)| \leq p(x)$ for all $x \in E$.*

Proof. If $\mathbb{K} = \mathbb{R}$, then $|f(x)| \leq p(x)$ is equivalent to $-p(-x) \leq f(x) \leq p(x)$, since p is a seminorm, so the result follows from 4.2. If $\mathbb{K} = \mathbb{C}$: Consider the real linear form $u_0 = \operatorname{Re} f_0 \colon V_0 \to \mathbb{R}$ ($\operatorname{Re} f_0 \leq |f_0| \leq p$). By 4.2, there exists a real linear form $u \colon E \to \mathbb{R}$ with $u(x) \leq p(x)$ for all $x \in E$ and $u|_{V_0} = u_0 = \operatorname{Re} f_0$. Let $f \colon E \to \mathbb{C}$ be defined by $f(x) := u(x) - iu(ix) \in \mathbb{C}$ using that u is real linear, one gets that f is $\mathbb{C}$-linear and, using $z = \operatorname{Re} z - i \operatorname{Re}(iz)$, one has that $f|_{V_0} = f_0$. For $x \in E$, choose $\alpha \in \mathbb{C}$, $|\alpha| = 1$, such that $|f(x)| = \alpha f(x) = f(\alpha x) = u(\alpha x) \leq p(\alpha x) = |\alpha| p(x) = p(x)$. ☐

Theorem 4.4 (Hahn-Banach)**.** *Let X be a normed $\mathbb{K}$-linear vector space, let $V_0 \subseteq X$, and $f_0 \colon V_0 \to \mathbb{K}$, $f_0 \in V_0'$ (f_0 is a bounded linear form). Then f_0 has an extension $f \colon X \to \mathbb{K}$, $f|_{V_0} = f_0$, $f \in X'$ and $\|f\| = \|f_0\|$.*

Proof. Use Theorem 4.3 with $p(x) = \|f_0\| \cdot \|x\|$. □

Corollary 4.5. *Let X be a Banach space and $x \in X$, $x \neq 0$. Then there exists an $f \in X'$ such that $f(x) \neq 0$.*

Proof. Define $f_0(\alpha x) = \alpha\|x\|$ for $\alpha x \in \operatorname{span}\{x\} \equiv V_0$. Then there exists $f \colon X \to \mathbb{K}$, $f \in X'$, such that $f|_{V_0} = f_0$. In particular, $f(x) = f_0(x) = \|x\| \neq 0$. □

Remark. If $x \neq y$, $x, y \in X$, then $x - y \neq 0$, so there exists $f \in X'$ such that $f(x-y) \neq 0$, hence $f(x) \neq f(y)$. Hence, X' seperates points in X: If $f(x) = f(y)$ for all $f \in X'$, then $x = y$.

4.1 3 consequences of Baire's theorem

Recall Baire's theorem: If $M = \{A, d\}$ is a complete metric space and $\{V_n\}_{n\in\mathbb{N}}$ is a countable family of open, dense subsets of A, then $\bigcap_{n\in\mathbb{N}} V_n$ is also dense. On problem sheet 5 it was proven a corollary of Baire's theorem that a complete metric space is never the union of a countable number of nowhere dense, closed subsets.

Theorem 4.6 (Banach-Steinhaus/Principle of uniform boundedness)**.** *Let X be a Banach space, Y a normed space, I some index set, and for each $i \in I$ a bounded linear operator $T_i \colon X \to Y$. If $\sup_{i\in I} \|T_i x\| < \infty$ for all $x \in X$, then $\sup_{i\in I} \|T_i\| < \infty$.*

Proof. For $n \in \mathbb{N}$, let $E_n = \{x \in X \colon \sup_{i\in I} \|T_i x\| \leq n\}$. Then $X = \bigcup_{n\in\mathbb{N}} E_n$. Now, $E_n = \bigcap_{i\in I} \|T_i(\cdot)\|^{-1}([0, n])$, hence all E_n are closed, since $[0, n]$ is closed and $\|T_i(\cdot)\|$ is continuous because T_i is continuous. So, by Baire, there exists $n_0 \in \mathbb{N}$ such that E_{n_0} has an interior point $y \in E_{n_0}$, i.e. there is an $\varepsilon > 0$ such that $\|x - y\| \leq \varepsilon$ implies $x \in E_{n_0}$. Note, that E_{n_0} is symmetric, i.e. $z \in E_{n_0}$ implies $-z \in E_{n_0}$. Hence, $\|x - (-y)\| \leq \varepsilon$ implies $x \in E_{n_0}$. Also, E_{n_0} is convex, so $\|u\| \leq \varepsilon$ implies $u = \frac{1}{2}\big((u + y) + (u - y)\big) \in E_{n_0}$. Hence, $\|u\| \leq \varepsilon$ implies $u \in E_{n_0}$, that is $\|u\| \leq \varepsilon$ implies $\|T_i u\| \leq n_0$ for all $i \in I$. So, if $x \in X$, $\|x\| \leq 1$, then $\|\varepsilon x\| \leq \varepsilon$, so $\|T_i(\varepsilon x)\| \leq n_0$ for all $i \in I$. Hence, $\|x\| \leq 1$ implies $\|T_i x\| \leq n_0/\varepsilon$ for all $i \in I$, so $\|T_i\| \leq n_0/\varepsilon < \infty$ for all $i \in I$. □

Remark. There exist more general versions of this theorem, but the one given here is the most used.

Definition 4.7. A map between two metric spaces is called *open* iff the image of any open set is open.

Remark.
(a) Note the difference to "continuity".
(b) One cannot in general replace with "closed to closed".
(c) Clearly, a bijective map is open iff its inverse is continuous.

Lemma 4.8. *Let X, Y be normed spaces and $T\colon X \to Y$ linear. Then the following are equivalent:*

(i) T is open.

(ii) For all $r > 0$ there exists $\varepsilon > 0$ such that $B_\varepsilon(0) \subseteq T(B_r(0))$.

(iii) There exists $\varepsilon > 0$ such that $B_\varepsilon(0) \subseteq T(B_1(0))$.

Proof. To see $(i) \Rightarrow (ii)$ note that $T(B_r(0))$ is open in Y and $0 \in T(B_r(0))$. To prove $(ii) \Rightarrow (i)$, let $U \subseteq X$ be open, and $x \in U$. Then $Tx \in T(U)$. Since U is open, there exists $r > 0$ such that $B_r(x) \subseteq U$. Note that $B_r(x) = x + B_r(0) = x + rB_1(0)$. Hence, $x + B_r(0) \subseteq U$, so $Tx + T(B_r(0)) \subseteq T(U)$. Form (ii) we have $\varepsilon > 0$ such that $B_\varepsilon(0) \subseteq T(B_r(0))$, hence, $Tx + B_\varepsilon(0) \subseteq Tx + T(B_r(0)) \subseteq T(U)$. Now, $Tx + B_\varepsilon(0) = B_\varepsilon(Tx)$, so $Tx + B_\varepsilon(0)$ is open, contains Tx and is contained in $T(U)$. $(ii) \Leftrightarrow (iii)$ is clear. □

Remark. If $T\colon X \to Y$ is linear and open, then T is surjective.

Theorem 4.9 (Open mapping theorem)**.** *Let X and Y be Banach spaces, and assume $T \in B(X, Y)$ is surjective. Then T is open.*

Proof. We shall prove that (iii) in Lemma 4.8 holds. This is done in 2 steps: First we prove that there exists $\varepsilon_0 > 0$ such that $B_{\varepsilon_0} \subseteq \overline{T(B_1(0))}$. Since T is a surjection, $Y = \bigcup_{n\in\mathbb{N}} \overline{T(B_n(0))}$. Since Y is Banach, Baire's theorem implies that there exists $N \in \mathbb{N}$ such that $(\overline{T(B_N(0))})^\circ \neq \emptyset$, i.e. there exists $y_0 \in \overline{T(B_N(0))}$ and $\varepsilon > 0$ such that $\overline{B_\varepsilon(y_0)} \subseteq \overline{T(B_N(0))}$, in other words $\|z - y_0\| < \varepsilon$ implies $z \in \overline{T(B_N(0))}$ $(*)$. Now, $\overline{T(B_N(0))}$ is symmetric, hence $-y_0$ also satisfies $(*)$. Let $y \in Y$ with $\|y\| < \varepsilon$. Then $\|(y_0+y)-y_0\| < \varepsilon$, hence $y_0 + y \in \overline{T(B_N(0))}$. Similarly, $\|y\| < \varepsilon$ implies $-y_0 + y \in \overline{T(B_N(0))}$. Therefore, since $\overline{T(B_N(0))}$ is convex, we have $y = \frac{1}{2}((y_0 + y + (-y_0 + y)) \in \overline{T(B_N(0))}$, if $\|y\| < \varepsilon$. Hence, $B_\varepsilon(0) \subseteq \overline{T(B_N(0))}$. So, $B_{\varepsilon/N}(0) \subseteq \overline{T(B_1(0))}$. Then ε_0 is as claimed.

Now for the second step, let $\varepsilon_0 > 0$ be as above. We now prove that $B_{\varepsilon_0} \subseteq T(B_1(0))$. This will complete the proof. Let $y \in Y$ with $\|y\| < \varepsilon_0$. Take $\varepsilon > 0$ such that $\|y\| < \varepsilon < \varepsilon_0$ and write $\overline{y} = \frac{\varepsilon_0}{\varepsilon} y$. Then $\|\overline{y}\| < \varepsilon_0$, so $\overline{y} \in \overline{T(B_1(0))}$. Choose $\alpha \in (0, 1)$ such that $0 < \frac{\varepsilon}{\varepsilon_0}\frac{1}{1-\alpha} < 1$ and take $y_0 \in T(B_1(0))$ such that $\|\overline{y} - y_0\| < \alpha\varepsilon_0$. Since $y_0 \in T(B_1(0))$, there is a $x_0 \in B_1(0)$ such that $y_0 = Tx_0$. Now let $z_0 = \frac{\overline{y}-y_0}{\alpha}$. Then $\|z_0\| < \varepsilon_0$. So $z_0 \in B_{\varepsilon_0}(0) \subseteq \overline{T(B_1(0))}$. So there exists $y_1 \in T(B_1(0))$ such that $\|z_0 - y_1\| < \alpha\varepsilon_0$, that is $\|\overline{y} - (y_0 + \alpha y_1)\| \leq \alpha^2\varepsilon_0$. Repeat on $z_1 = \frac{\overline{y}-(y_0+\alpha y_1)}{\alpha^2}$, to get $y_2 = Tx_2 \in T(B_1(0))$ with $\|z_1 - y_2\| < \alpha\varepsilon_0$. Inductively, we get a sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq B_1(0)$ such that

$$\Big\|\overline{y} - \sum_{i=0}^{n} \alpha^i y_i\Big\| = \Big\|\overline{y} - T\Big(\sum_{i=0}^{n} \alpha^i x_i\Big)\Big\| < \alpha^{n+1}\varepsilon_0.$$

Since $\alpha \in (0, 1)$, and $\|x_i\| < 1$ for all $i \in \mathbb{N}$, the series $\sum_{i=0}^{\infty} \alpha^i x_i$ is absolutely convergent. Since X is Banach, the series $\sum_{i=0}^{\infty} \alpha^i x_i$ is convergent in X. Write $\overline{x} = \sum_{i=1}^{\infty} \alpha^i x_i \in X$. Since T is bounded,

$$T\Big(\sum_{i=0}^{n} \alpha^i x_i\Big) \xrightarrow{n\to\infty} T\overline{x}$$

in Y and by construction

$$T\Big(\sum_{i=0}^{n} \alpha^i x_i\Big) \xrightarrow{n\to\infty} \overline{y}.$$

Finally, let $x = \frac{\varepsilon}{\varepsilon_0}\overline{x}$. Then $Tx = y$. Also

$$\|x\| = \frac{\varepsilon}{\varepsilon_0}\|\overline{x}\| = \frac{\varepsilon}{\varepsilon_0}\Big\|\sum_{i=0}^{\infty}\alpha^i x_i\Big\| \leq \frac{\varepsilon}{\varepsilon_0}\sum_{i=0}^{\infty}\alpha^i \leq \frac{\varepsilon}{\varepsilon_0}\frac{1}{1-\alpha} < 1. \qquad \square$$

Corollary 4.10. *Let X and Y be Banach spaces, and assume $T \in B(X,Y)$ which is bijective. Then T is a homeomorphism.*

Corollary 4.11. *Let $\|\cdot\|$ and $\|\cdot\|'$ be two norms on the same vectorspace X, such that $\{X, \|\cdot\|\}$ and $\{X, \|\cdot\|'\}$ are both Banach. Assume there exists a constant $M > 0$ such that $\|x\| \leq M\|x\|'$ for all $x \in X$. Then $\|\cdot\|$ and $\|\cdot\|'$ are equivalent.*

Corollary 4.12. *Let X, Y be Banach spaces, and assume $T \in B(X,Y)$ is injective. Then $T^{-1}\colon \mathrm{R}(T) \to X$ is bounded iff $\mathrm{R}(T) \subseteq Y$ is closed.*

Definition 4.13. Let X, Y be normed spaces, $D \subseteq X$ a linear subspace, and $T\colon D \to Y$ a linear map (we write $D = \mathrm{dom}(T)$, $T\colon X \supseteq D \to Y$). We call T *closed* (a closed linear operator) iff for any sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq D$ such that $x_n \to x$ as $n \to \infty$ and $Tx_n \to y$ as $n \to \infty$ we have $x \in D$ and $Tx = y$.

Remark. Note the relation to continuity: If $\mathrm{dom}(T) = X$, look at:

(a) $x_n \to x$ as $n \to \infty$.

(b) $\{Tx_n\}$ is convergent.

(c) $Tx = y$.

Then T is continuous iff $(a) \Rightarrow (b) \wedge (c)$. T is closed iff $(a) \wedge (b) \Rightarrow (c)$.

Remark. A closed operator does not in general map closed sets to closed sets.

Definition 4.14. For linear $T\colon X \supseteq D \to Y$ we define the *graph* of T by

$$\mathrm{gr}(T) = \{(x, Tx) \colon x \in D\} \subseteq X \times Y.$$

Lemma 4.15. *Let X, Y, D, T be as in 4.14. Then*
(a) $\mathrm{gr}(T)$ is a linear subspace of $X \times Y$.
(b) T is a closed operator iff $\mathrm{gr}(T)$ is closed in $X \oplus_1 Y$ (here $\|(x,y)\|_1 = \|x\|_X + \|y\|_Y$).

Proof. This is left as an exercise.

Lemma 4.16. *Let X, Y be Banach spaces, $D \subseteq X$ a linear subspace, $T\colon X \supseteq D \to Y$ a closed operator. Then*
(a) $(D, \|\cdot\|')$ with $\|x\|' = \|x\|_X + \|Tx\|_Y$ is a Banach space. $\|\cdot\|'$ is called the graph *norm.*

(b) $T\colon (D, \|\cdot\|') \to (Y, \|\cdot\|_Y)$ *is bounded.*

Proof. Let $\{x_n\}_{n\in\mathbb{N}} \subseteq D$ be Cauchy with respect to $\|\cdot\|'$. Then $\{x_n\}$ is Cauchy with respect to $\|\cdot\|_X$, and $\{Tx_n\}_{n\in\mathbb{N}}$ is Cauchy (in Y) with respect to $\|\cdot\|_Y$. Hence, since X and Y are Banach, there exist $x \in X$, $y \in Y$ such that $x_n \to x$ as $n \to \infty$ and $Tx_n \to y$ as $n \to \infty$. Since T is closed, $x \in D$ and $y = Tx$. Then $\|x_n - x\|' = \|x_n - x\|_X + \|Tx_n - y\|_Y \to 0$ as $n \to \infty$. So, $x_n \to x$ as $n \to \infty$ with respect to $\|\cdot\|'$. (b) is trivial. ☐

Theorem 4.17. *Let X, Y be Banach spaces, $D \subseteq X$ a linear subspace, $T\colon X \supseteq D \to Y$ closed and surjective. Then T is open. If T is also bijective, then T^{-1} is continuous.*

Proof. By Lemma 4.16 and Theorem 4.9, $T\colon (D, \|\cdot\|') \to (Y, \|\cdot\|_Y)$ is open. Since $\|x\|_X \le \|x\|'$ for all $x \in D$, we have that any $\|\cdot\|_X$-open set is also $\|\cdot\|'$-open. So T is also open as a map $(D, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$. If T is also bijective, then T^{-1} is $(Y, \|\cdot\|_Y)$-$(X, \|\cdot\|_X)$-continuous. ☐

Theorem 4.18 (Closed graph theorem)**.** *Let X, Y be Banach spaces, and assume $T\colon X \to Y$ is linear and a closed operator. Then T is continuous.*

Proof. By Lemma 4.16(b), $T\colon X \to Y$ is continuous, when X is equipped with the graph norm, $\|x\|' = \|x\|_X + \|Tx\|_Y$. By corollary 4.11, $\|\cdot\|_X$ and $\|\cdot\|'$ are equivalent norms, since $\|x\|_X \le \|x\|'$ and $(X, \|\cdot\|)$ is Banach by assumption and $(X, \|\cdot\|')$ is Banach by 4.16(a). Therefore, T is also continuous with respect to $\|\cdot\|_X$. ☐

Remark. The theorem says a closed operator on all of a Banach space is automatically continuous. This, and the following consequence of Banach-Steinhaus illustrates why it is almost impossible to explicitly define a non-continuous linear operator on a Banach space.

Proposition 4.19. *Let X be a Banach space, Y a normed space, and let $T_n \in B(X, Y)$, $n \in \mathbb{N}$. Assume that $Tx := \lim_{n\to\infty} T_n x$ exists for all $x \in X$. Then T is linear and continuous.*

Proof. It is clear that T is linear. Since $\{T_n x\}_{n\in\mathbb{N}} \subseteq Y$ is convergent for all $x \in X$, $\{T_n x\}_{n\in\mathbb{N}} \subseteq Y$ is bounded, hence $\sup_{n\in\mathbb{N}} \|T_n x\|_Y < \infty$ for all $x \in X$. Hence, by Banach-Steinhaus, $\sup_{n\in\mathbb{N}} \|T_n\| = M < \infty$. It follows that $\|Tx\| = \lim_{n\to\infty} \|T_n x\| \le \lim_{n\to\infty} \|T_n\|\|x\| \le M\|x\|$, hence $T \in B(X, Y)$. ☐

Recall, that for a normed space X, $X' = B(X, \mathbb{K})$ is called the dual of X. Let $x \in X$, then $\|x\| = \sup\{|f(x)| \colon f \in X', \|f\| \le 1\} = \max\{|f(x)| \colon f \in X', \|f\| \le 1\}$.

Remark. Let $x \in X$ and define $\iota(x)\colon X' \to \mathbb{K}, f \mapsto f(x)$. Then ι is linear: $(\iota(x))(\lambda f + g) = (\lambda f + g)(x) = \lambda f(x) + g(x) = \lambda \iota(x)(f) + \iota(x)(g)$ and

$$\sup\{|(\iota(x))(f)| \colon f \in X', \|f\| \le 1\} = \sup\{|f(x)| \colon f \in X', \|f\| \le 1\} = \|x\| < \infty.$$

Hence, $\iota(x) \in B(X', \mathbb{K})$, i.e. $\iota(x) \in X''$, and $\|\iota(x)\|_{X''} = \|x\|_X$. Hence, $\iota\colon X \to X''$ is an isometrical embedding. ι is called the *canonical embedding.*

Definition 4.20. A subset $M \subseteq X$ (X normed) is called *weakly bounded* if for all $f \in X'$, $\sup_{x \in M} |f(x)| < \infty$. By the above, M is weakly bounded iff $\iota(M) \subseteq X''$ is pointwise bounded.

Proposition 4.21. *A weakly bounded set in a normed space is also bounded in the norm topology, i.e. there exists $R > 0$ such that $M \subseteq B_R(0)$.*

Proof. Use the principle of uniform boundedness. □

Definition 4.22. A normed space X is called *reflexive* if $\iota \colon X \to X''$ is surjective.

Remark. Any Hilbert space is reflexive.

Remark. Any reflexive space is complete.

Remark. If X is reflexive, and if $X \cong Y$, then Y is reflexive.

Remark. If X and Y are reflexive, $X \oplus_1 Y$ is reflexive.

Example. ℓ_p is reflexive for $1 < p < \infty$, since $(\ell_p)' = \ell_q$, $\frac{1}{p} + \frac{1}{q} = 1$, and hence $(\ell_p)'' = (\ell_q)' = \ell_p$. However, $(\ell_1)' = \ell_\infty$ and $(\ell_\infty)' \neq \ell_1$. Hence, ℓ_1 and ℓ_∞ are not reflexive. Furthermore, $(L^p)' = L^q$ if $1 < p < \infty$, which we will see later.

Definition 4.23.
(a) A sequence $\{x_n\}_{n \in \mathbb{N}} \subseteq X$ is said to *converge weakly* ($x_k \rightharpoonup x$ as $k \to \infty$ in X) if $f(x_k) \to f(x)$ as $k \to \infty$ for all $f \in X'$.
(b) A sequence $\{f_n\}_{n \in \mathbb{N}} \subseteq X'$ *converges weak** to $f \in X'$ (written $f_k \overset{*}{\rightharpoonup} f$) if $f_k(x) \to f(x)$ as $k \to \infty$ for all $x \in X$.
(c) Similarly, one defines the notion of Cauchy sequences (weak, weak*).
(d) A subset $M \subseteq X$ is called *weakly sequentially compact* if every sequence in M has a weakly convergent subsequence (with limit in M). Similarly for weak* sequentially compact.
(e) To avoid confusion with usual convergence, we call convergence with respect to the norm *strong convergence*.

Remark. $x_k \to x$ implies $x_k \rightharpoonup x$ since $|f(x_k) - f(x)| \leq \|f\| \|x_k - x\|$.

Remark. Since X is only canonically *embedded* in X'', weak convergence in X' is a priori stronger than weak* convergence.

Remark. One can, both for weak and weak* convergence, define this convergence by topologies ("complicated").

Remark. By the canonical embedding $\iota \colon X \to X''$ we have $x_k \rightharpoonup x$ as $k \to \infty$ in X iff $\iota(x_k) \overset{*}{\rightharpoonup} \iota(x)$ as $k \to \infty$ in X''.

Remark 4.24.

1. The weak limit of a sequence is unique (use Hahn-Banach). Also the weak* limit is unique.

2. Strong convergence implies weak convergence (and it also implies weak* convergence). The opposite is not true: take $X = \ell_p$, $X' = \ell_q$ with $\frac{1}{p} + \frac{1}{q} = 1$ and $e_n \in \ell_q = X'$. Then for all $x \in \ell_p$, $e_n(x) = \sum_{i\in\mathbb{N}} \delta_{in} x_i = x_n \to 0$ as $n \to \infty$. I.e. $\{e_n\}_{n\in\mathbb{N}} \subseteq X'$ and $e_n(x) \to 0$ as $n \to \infty$ for all $x \in X$. Hence $e_n \overset{*}{\rightharpoonup} 0$ as $n \to \infty$. But $\|e_n\|_{X'} = 1$, so $e_n \not\to 0$ as $n \to \infty$.
3. $x_k \rightharpoonup x$, $k \to \infty$, in X implies $\|x\| \le \liminf_{k\to\infty} \|x_k\|$.
4. $f_k \overset{*}{\rightharpoonup} f$, $k \to \infty$, in X' implies $\|f\| \le \liminf_{k\to\infty} \|f_k\|$.
5. The norm $\|\cdot\|\colon X \to \mathbb{R}$ or $\|\cdot\|\colon X' \to \mathbb{R}$ is weak/weak* lower semi-continuous.
6. Weak and weak* convergent sequences are bounded (in norm).

Theorem 4.25. *Let X be separable. Then the closed unit ball $\overline{B_1(0)} \subseteq X'$ is weak* sequentially compact, i.e. any bounded sequence in X' has a weak* convergent subsequence.*

Proof. Let $\{x_n\colon n \in \mathbb{N}\} \subseteq X$ be dense and let $\{f_k\}_{k\in\mathbb{N}} \subseteq X'$, with $\|f_k\| \le 1$, $k \in \mathbb{N}$. Then $\{f_k(x_n)\}_{k\in\mathbb{N}} \subseteq \mathbb{K}$ ($n \in \mathbb{N}$ fixed) is a bounded sequence in $\mathbb{K}$. By a diagonal argument (à la Cantor, but different) there exists a subsequence $\{f_{k_m}\}_{m\in\mathbb{N}}$ such that, for all $n \in \mathbb{N}$, $\{f_{k_m}(x_n)\}_{m\in\mathbb{N}}$ is convergent in $\mathbb{K}$, i.e. $\lim_{m\to\infty} f_{k_m}(x_n)$ exists for all $n \in \mathbb{N}$. Then, for all $y \in Y = \operatorname{span}\{x_n\colon n \in \mathbb{N}\} \subseteq X$ the limit $\lim_{m\to\infty} f_{k_m}(y)$ exists. Define $f(y) = \lim_{m\to\infty} f_{k_m}(y)$ for $y \in Y$. Then f is linear and $|f(y)| = \lim_{m\to\infty} |f_{k_m}(y)| \le \lim_{m\to\infty} \|f_{k_m}\|\|y\| \le \|y\|$ for all $y \in Y$. Then f has a unique extension to a bounded linear functional on X (again called f). So $f \in X'$, with $\|f\| \le 1$ and, for $x \in X$, $y \in Y$,

$$\big|(f - f_{k_m})(x)\big| \le \big|(f - f_{k_m})(x - y)\big| + \big|(f - f_{k_m})(y)\big| \le 2\|x - y\| + \big|(f - f_{k_m})(y)\big|.$$

The first term can be made arbitrarily small since $\overline{Y} = X$. The second term goes to 0 as $m \to \infty$ by definition of f. Hence $f_{k_m} \overset{*}{\rightharpoonup} f$ as $m \to \infty$. □

Definition 4.26. For $(n, \varphi, \varepsilon)$ with $n \in \mathbb{N}$, $\varphi = (\varphi_1, \dots, \varphi_n) \in (X')^n$ and $\varepsilon > 0$ define

$$U_{n,\varphi,\varepsilon} = \{x \in X\colon |\varphi_k(x)| < \varepsilon \text{ for } k = 1, \dots, n\}$$

and

$$\mathscr{T}_W = \{A \subseteq X\colon x \in A \implies x + U_{n,\varphi,\varepsilon} \subseteq A \text{ for some } U_{n,\varphi,\varepsilon}\}.$$

Then $\mathscr{T}_W$ is a topology on X and it is the weakest topology $\mathscr{T}$ on X such that all $f \in X'$ are continuous with respect to $\mathscr{T}$ as maps $f\colon X \to \mathbb{K}$. $\{X, \mathscr{T}_W\}$ is neither a normed nor a metric space, but (as in Lemma 2.2) the linear structure on X is $\mathscr{T}_W$ continuous. Finally, convergence in $\mathscr{T}_W$ is the same as weak convergence ($\{X, \mathscr{T}_W\}$ is a "locally convex topological vector space"). A similar construction works for the topology $\mathscr{T}_{W^*}$ on X' giving weak* convergence in X'.

Remark 4.27.
1. If X is reflexive then weak* and weak convergence in X' is the same.
2. If X is reflexive and $V \subseteq X$ is a closed subspace, then V is reflexive.
3. X is reflexive iff X' is reflexive.
4. If X' is separable, X is separable.

Theorem 4.28 (Banach-Alaoglu)**.** *If X is a reflexive Banach space, then every norm-bounded sequence has a weakly convergent subsequence, i.e. $\overline{B_1(0)}$ is weakly sequentially compact.*

Proof. Let $\{x_k\}_{k\in\mathbb{N}} \subseteq \overline{B_1(0)} \subseteq X$ and $Y = \overline{\operatorname{span}\{x_k : k \in \mathbb{N}\}} \subseteq X$. Then Y is reflexive and separable. Then $Y'' = \iota(Y)$ (where ι is the canonical embedding) is separable, so Y' is separable. Therefore, we can use 4.25 on Y' on the sequence $\{\iota(x_k)\}_{k\in\mathbb{N}} \subseteq Y''$, i.e. there exists $y \in Y''$ such that for a subsequence $\{\iota(x_{k_m})\}_{m\in\mathbb{N}}$, $\iota(x_{k_m})(f) \to y(f)$ for all $f \in Y'$. Let $x = \iota^{-1}(y) \in Y$. Then this means that $\iota(x_{k_m})(f) = f(x_{k_m}) \to y(f) = f(x)$ as $m \to \infty$ for all $f \in Y'$. Note that for $\varphi \in X'$, we have $\varphi|_Y \in Y'$. So it follows that $\varphi(x_{k_m}) \to \varphi(x)$ as $m \to \infty$ for all $\varphi \in X'$, i.e. $x_{k_m} \rightharpoonup x$ as $m \to \infty$ in X. □

Remark. In particular any Hilbert space is reflexive, so the closed unit ball in a Hilbert space is weakly sequentially compact.

5 Topics on operators

Definition 5.1. The *compact (linear) operators* from X to Y are defined by

$$K(X,Y) = \{T \in B(X,Y) : \overline{T(B_1(0))} \text{ is compact}\}.$$

Remark.

(i) If Y is Banach, then "$\overline{T(B_1(0))}$ compact" can be replaced by "$T(B_1(0))$ precompact".

(ii) That is, $T \in K(X,Y)$ iff T maps bounded sequences (in X) into sequences (in Y) which have a convergent subsequence.

(iii) For $k \in C(I^2)$, $I = [0,1]$,

$$(Kf)(x) = \int_0^1 k(x,y)f(y)\,\mathrm{d}y, \quad x \in I, f \in C(I)$$

defines a compact operator $K : C(I) \to C(I)$.

Proposition 5.2. *Let $T \in B(X,Y)$ and define, for $y' \in Y'$, $T'(y')(x) = y'(Tx)$. This defines a linear map $T' : Y' \to X'$ called the* adjoint *of T. We have $T' \in B(Y',X')$ with $\|T'\| = \|T\|$ and $\cdot' : B(X,Y) \to B(Y',X'), T \mapsto T'$ is an isometric embedding.*

Proof. We have $(T'y')(\lambda x_1+x_2) = y'(T(\lambda x_1+x_2)) = \lambda y'(Tx_1)+y'(Tx_2) = \lambda(T'y')(x_1)+(T'y')(x_2)$. Hence, $T'y' : X \to \mathbb{K}$ is linear. Also $|(T'y')(x)| = |y'(Tx)| \le \|y'\|\|Tx\| \le \|y'\|\|T\|\|x\|$, so $T'y' \in X'$. Hence, T' is well-defined. T' is linear, since $(T'(\lambda y_1'+y_2'))(x) = (\lambda y_1' + y_2')(Tx) + \lambda y_1'(Tx) + y_2'(Tx) = \lambda T'y_1'(x) + T'y_2'(x) = (\lambda T'y_1' + T'y_2')(x)$, i.e. $T'(\lambda y_1' + y_2') = \lambda T'y_1' + T'y_2'$.

From the above, one sees $\|T'y'\| \le \|T\|\|y'\|$, i.e. T' is bounded and $\|T'\| \le \|T\|$. On the other hand, for $\|y'\| \le 1$, $y' \in Y'$, $\|x\| \le 1$, $x \in X$, then

$$\|T'\| \ge \|T'y'\| \ge |(T'y')(x)| = |y'(Tx)|.$$

If $Tx \neq 0$, then by Hahn-Banach, there is a $\tilde{y}'$ such that $\|\tilde{y}'\| = 1$ and $\tilde{y}'(Tx) = \|Tx\|$. Hence, $\|T'\| \ge \|Tx\|$. Hence, $\|T'\| \ge \sup_{\|x\|\le 1} \|Tx\| = \|T\|$, so $\|T'\| = \|T\|$. □

Definition 5.3 (Hilbert space adjoint)**.** Let H be a Hilbert space, and let $\Phi\colon H \to H', y \mapsto \langle y, - \rangle$ be the map in Theorem 2.32 (Fréchet-Riesz), and let $T \in B(H)$. Then $T^* = \Phi^{-1}T'\Phi$ is called the *Hilbert space adjoint of* T. It satisfies

$$\langle T^*x, y\rangle = \langle x, Ty\rangle, \quad \forall x, y \in H.$$

T is called *selfadjoint* if $T^* = T$. Note, that T is assumed to be bounded. For unbounded operators, the definition of adjoint and therefore of selfadjointness is more complicated (for example in quantum mechanics).

Lemma 5.4 (algebraic properties)**.** *We have*

(1) $(\alpha T_1 + T_2)' = \alpha T_1' + T_2'$ *for* $T_1, T_2 \in B(X, Y)$ *and* $\alpha \in \mathbb{K}$.
(1)* $(\alpha T_1 + T_2)^* = \overline{\alpha} T_1^* + T_2^*$ *for* $T_1, T_2 \in B(H)$ *and* $\alpha \in \mathbb{K}$.
(2) $I' = I$ *for* $I \in B(X)$, $I\colon X \to X, x \mapsto x$.
(3) *For* $T_1 \in B(X, Y)$, $T_2 \in B(Y, Z)$, $(T_2T_1)' = T_1'T_2'$.
(4) *With* $\iota_X\colon X \to X''$ *and* $\iota_Y\colon Y \to Y''$ *the canonical embeddings and* $T \in B(X, Y)$, *we have* $T''\iota_X = \iota_Y T$.
(4)* *For* $T \in B(H)$, $T^{**} = T$.

Proposition 5.5. *Let* X, Y *be Banach spaces and* $T \in B(X, Y)$. *Then* $T^{-1} \in B(Y, X)$ *exists if and only if* $(T')^{-1} \in B(X', Y')$ *exists and, in this case,* $(T^{-1})' = (T')^{-1}$. *(or, if* $X = Y = H$ *a Hilbert space,* $(T^*)^{-1} = (T^{-1})^*$*).*

Definition 5.6. Let $T \in B(X)$ with a Banach space X over $\mathbb{C}$. We define the *resolvent set* of T by

$$\rho(T) = \{\lambda \in \mathbb{C}\colon\ \mathrm{N}(T - \lambda I) = 0 \text{ and } \mathrm{R}(T - \lambda I) = X\}$$

and the *spectrum* of T by

$$\sigma(T) = \mathbb{C} \smallsetminus \rho(T).$$

The spectrum can be split in three parts. The *point spectrum* is

$$\sigma_p(T) = \{\lambda \in \mathbb{C}\colon\ \mathrm{N}(T - \lambda I) \neq 0\}.$$

The *continuous spectrum* is

$$\sigma_c(T) = \{\lambda \in \mathbb{C}\colon\ \mathrm{N}(T - \lambda I) = 0 \text{ and } \mathrm{R}(T - \lambda I) \neq X, \text{ but } \overline{\mathrm{R}(T - \lambda I)} = X\}.$$

The *rest/residual spectrum* is

$$\sigma_r(T) = \{\lambda \in \mathbb{C}\colon \lambda \in \mathbb{C}\colon\ \mathrm{N}(T - \lambda I) = 0 \text{ and } \overline{\mathrm{R}(T - \lambda I)} \neq X\}.$$

Remark.

(1) Note that $\lambda \in \rho(T)$ if and only if $T - \lambda I\colon X \to X$ is bijective. This is equivalent to the existence of $R_\lambda(T) := (T - \lambda I)^{-1} \in B(X)$, called the *resolvent* of T (at λ).
(2) $\lambda \in \sigma_p(T)$ if and only if there exists $x \neq 0$ such that $Tx = \lambda x$. In this case, λ is called an *eigenvalue* and x is called an *eigenvector* ($x \in X$). However, in the cases where X is some space of functions — $\mathscr{C}(I)$, $L^p(\Omega)$, $\mathscr{C}^\alpha(I)$, $\mathscr{C}^{k,\alpha}(I)$, ... — such an X is normally called an eigenfunction. $\mathrm{N}(T - \lambda I)$ is called the *eigenspace* belonging to the eigenvalue λ. It is a T-invariant subspace, i.e. $T\,\mathrm{N}(T - \lambda I) \subseteq \mathrm{N}(T - \lambda I)$.

Remark. If f is an analytic function, i.e. f can be represented by a convergent power series, $f(x) = \sum_{n=0}^{\infty} a_n x^n$, we can define $f(T) = \sum_{n=0}^{\infty} a_n T^n$ (which is defined since $B(X)$ is Banach).

Proposition 5.7. *Let X be a Banach space, $T \in B(X)$ with $\|T\| < 1$. Then $(I - T)^{-1} \in B(X)$ and $(I - T)^{-1} = \sum_{n=0}^{\infty} T^n$ (the Neumann series) in $B(X)$.*

Proof. Let $S_k = \sum_{n=0}^{k} T^n$. Then, for $k < \ell$,

$$\|S_\ell - S_k\| = \Big\| \sum_{k<n\leq\ell} T^n \Big\| \leq \sum_{k<n\leq\ell} \|T^n\| \leq \sum_{k<n\leq\ell} \|T\|^n \leq \sum_{n=k+1}^{\infty} \|T\|^n \xrightarrow{k\to\infty} 0$$

Hence, $\{S_k\}$ is Cauchy in $B(X)$, so convergent. Let $S = \lim_{k\to\infty} S_k$ in $B(X)$ and for $k \to \infty$:

$$(I - T)S_k x = \sum_{n=0}^{k} (T^n - T^{n+1})x = x - T^{k+1}x \xrightarrow{k\to\infty} x$$

since $\|T^{k+1}x\| \leq \|T\|^{k+1}\|x\|$. On the other hand $(I - T)S_k x \to (I - T)Sx$ as $k \to \infty$. Hence, $S = (I - T)^{-1}$. □

Proposition 5.8. *Let $T \in B(X)$. Then $\rho(T) \subseteq \mathbb{C}$ is an open set, i.e. $\sigma(T) = \mathbb{C} \setminus \rho(T)$ is closed, and the* resolvent function *$\rho(T) \ni \lambda \mapsto R_\lambda(T) \in B(X)$ is a complex analytic map from $\rho(T)$ to $B(X)$ with $\|R_\lambda(T)\|^{-1} \leq d(\lambda, \sigma(T))$, i.e. for all $\lambda_0 \in \rho(T)$, there exists $r > 0$ such that*

$$R_\lambda(T) = \sum_{n=0}^{\infty} a_n (\lambda - \lambda_0)^n T^n$$

for all $\lambda \in B_r(\lambda_0)$.

Proof. Use that $(I - A)^{-1} = \sum_{n=0}^{\infty} A^n$ if $\|A\| < 1$ and $T - (\lambda - \mu)I = (T - \lambda I)(I - \mu R_\lambda(T)) =: (T - \lambda I)S(\mu)$. Then $S(\mu)$ is invertible if $|\mu|\|R_\lambda(T)\| < 1$. Hence, $R_{\lambda-\mu}(T) = S(\mu)^{-1}R_\lambda(T) = \sum_{k=0}^{\infty} \mu^k R_\lambda(T)^{k+1}$. □

Proposition 5.9. *Let X, Y be Banach spaces. Then the set of invertible operators in $B(X, Y)$ is an open set. If $X \neq 0$ and $Y \neq 0$, then for $S, T \in B(X)$, T invertible and $\|S - T\| < \|T^{-1}\|^{-1}$ implies S is invertible.*

Proof. Let $R = T - S$. Then $S = T(I - T^{-1}R) = (I - RT^{-1})T$ where $\|T^{-1}R\| < 1$ and $\|RT^{-1}\| < 1$. Now use 5.7. □

Definition 5.10. An operator $A \in B(X, Y)$ is called a *Fredholm* operator ("is Fredholm") iff

(i) $\dim \mathrm{N}(A) < \infty$.

(ii) $\mathrm{R}(A) \subseteq Y$ is closed.

(iii) $\operatorname{codim} \mathrm{R}(A) := \dim (Y/\mathrm{R}(A)) < \infty$.

The index of A is $\operatorname{ind}(A) = \dim \mathrm{N}(A) - \operatorname{codim} \mathrm{R}(A)$.

Theorem 5.11. *Let $T \in K(X)$. Then $A = I - T$ is a Fredholm operator with* $\operatorname{ind}(A) = 0$.

For compact operators, one has the following *spectral theorem for compact operators*:

Theorem 5.12 (Riesz-Schauder)**.** *For every operator $T \in K(X)$ one has*

(i) *$\sigma(T) \setminus \{0\}$ consists of countably (finite or infinitely) many eigenvalues, with 0 the only possible accumulation point. If $\sigma(T)$ consists of infinitely many elements, then it follows that $\overline{\sigma(T)} = \sigma_p(T) \cup \{0\}$.*

(ii) *For $\lambda \in \sigma(T) \setminus \{0\}$ one has $1 \leq n_\lambda = \max\{n \in \mathbb{N} \colon \mathrm{N}\left((T-\lambda I)^{n-1}\right) \neq \mathrm{N}\left((T-\lambda I)^n\right)\} < \infty$. n_λ is the* order *(or index) of λ and $\dim \mathrm{N}(T - \lambda I)$ is the* multiplicity *of λ.*

(iii) *(Riesz decomposition) For $\lambda \in \sigma(T) \setminus \{0\}$ one has $X = \mathrm{N}\left((T-\lambda I)^{n_\lambda}\right) \oplus \mathrm{R}\left((T-\lambda I)^{n_\lambda}\right)$. Both subspaces are closed, T invariant and $\mathrm{N}\left((T-\lambda I)^{n_\lambda}\right)$ is finite dimensional.*

(iv) $\sigma(T|_{\mathrm{R}((T-\lambda I)^{n_\lambda})}) = \sigma(T) \setminus \{\lambda\}$.

(v) *Let, for $\lambda \in \sigma(T) \setminus \{0\}$, E_λ be the projection on $\mathrm{N}\left((T-\lambda I)^{n_\lambda}\right)$ according to (iii). Then $E_\lambda E_\mu = \delta_{\mu,\lambda} E_\mu$ for $\lambda, \mu \in \sigma(T) \setminus \{0\}$.*

Corollary 5.13. *Let $T \in K(X)$ and $\lambda_0 \in \sigma(T) \setminus \{0\}$. Then the resolvent function $\lambda \mapsto R_\lambda(T)$ has an isolated pole of order n_{λ_0} at λ_0, i.e. the map $\lambda \mapsto (\lambda - \lambda_0)^{n_{\lambda_0}} R_\lambda(T)$ can be analytically continued at the point λ_0, and the value at λ_0 is not the zero operator.*

The fact that $\sigma(T) \setminus \{0\} \subseteq \sigma_p(T)$ can be formulated as follows:

Proposition 5.14 (Fredholm alternative)**.** *For compact T, either the equation $A_\lambda x = Tx - \lambda x = y$ has a unique solution for all $y \in X$ or the equation $Tx - \lambda x = 0$ has non-trivial solutions.*

Theorem 5.15 ("strong" Fredholm alternative)**.** *Let X be Banach, $T \in K(X)$, $\lambda \neq 0$. Then the equation $Tx - \lambda x = y$, $y \in X$, has a solution $x \in X$ iff $x'(y) = 0$ for all solutions $x \in X$ to the homogenous adjoint equation $T'x' - \lambda x' = 0$. The number of constraints on y (given by $x'(y) = 0$) is equal to the number of linearly independent solutions to the hoogenous equation $Tz - \lambda z = 0$ (i.e. to the dimension of* $\mathrm{N}(T - \lambda I)$*).*

Theorem 5.16 (Schauder)**.** *Let X, Y be Banach spaces and $T \in B(X,Y)$. Then $T \in K(X,Y)$ iff $T' \in K(Y', X')$.*

Remark. If $X = H$ a Hilbert space, $T \in K(X)$, $T = T^*$, then there exists an orthonormal system $\{e_n\}$ in H such that $Te_k = \lambda_k e_k$ for all k and $Tx = \sum \lambda_k \langle e_k, x\rangle e_k$.