BERGISCHE UNIVERSITÄT WUPPERTAL

Introduction to Functional Analysis

JOCHEN GLÜCK

Lecture notes
Summer term 2026

Version: August 23, 2026

Contents

Introduction

Notation

The following notation is used throughout the lecture notes:

Notation	Meaning
$\mathbb{N}$	$\{1,2,3,\ldots\}$
$\mathbb{N}_0$	$\{0,1,2,3,\ldots\}$
$\mathbb{K}$	$\mathbb{R}$ or $\mathbb{C}$

Some standard books on functional analysis

There are plenty of books available on functional analysis. We mention a few German and English ones in the following lists.

Books in German:

- *Funktionalanalysis* by Dirk Werner [12] is a classic among the German functional analysis book.

- The *Grundkurs Funktionalanalysis* by Winfried Kaballo [7] is very clearly and efficiently written. It starts each chapter with a list of introductory questions. This is where I got the inspiration to also do this in several of my lecture notes (for instance in the present ones).

  There is also an *Aufbaukurs Funktionalanalysis* by the same author [8]. As its title suggests, the contents of this book are far beyond an introductory course to functional analysis.

- *Einführung in die Funktionalanalysis* by Reinhold Meise and Dietmar Vogt. Dietmar Vogt is a Professor Emeritus at the University of Wuppertal.

Books in English

- *Functional analysis* by Walter Rudin [9] is a very popular English textbook on functional analysis.

- *Functional analysis, Sobolev spaces and partial differential equations* by Haim Brezis [1] presents an introduction to functional analysis with a focus on, well, partial differential equations.
- *Functional Analyis* by Kosaku Yosida [13] is a classical monograph on functional analysis and operator theory. Some of its contents are quite advanced.
- *Linear Operators. Part I* by Nelson Dunford and Jacob Schwartz [3] is the first part of a three volume monograph on linear operators and a real classic. (Unfortunately out of print, far too expensive, and with terrible typesetting.)

Version

This version of the lecture notes is from August 23, 2026.

Acknowledgements

I am grateful to Hannah Wähner for pointing out several errors in an earlier version of these lecture notes.

Chapter 1

Banach Spaces and Linear Operators

Throughout all chapters in these lecture notes, $\mathbb{K}$ will denote the real or the complex field, i.e. we assume throughout that

$$\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}.$$

1.0 A brief reminder on normed spaces

The contents of Section 1.0 have already been treated in the courses *Analysis 2* and *Analysis 3*, see [5, 6]. We include them here to make the lecture notes more self-contained, since most of functional analysis is based on those concepts.

Section 1.0 repeats some contents from the courses Analysis 2 and 3 which is important for the present course Introduction to Functional Analysis. Some parts of this material will be briefly recalled in the exercise session on Monday, 13 April.

Definition 1.0.1 (Normed space)**.**

(a) Let X be a vector space over $\mathbb{K}$. A **seminorm** on X is a map $\|\cdot\| : X \to [0, \infty)$ with the following two properties:

(I) **Absolute homogeneity:** One has $\|\alpha x\| = |\alpha| \, \|x\|$ for all $x \in X$ and all $\alpha \in \mathbb{K}$.

(II) **Triangle inequality:** One has $\|x + y\| \le \|x\| + \|y\|$ for all $x, y \in X$.

A seminorm on X is called a **norm** on X if it satisfies, in addition, the following property:

(III) **Positive definiteness:** If $x \in X$ and $\|x\| = 0$, then $x = 0$.

(b) A **normed space** is a pair $(X, \|\cdot\|)$, where X is a vector space over $\mathbb{K}$ and $\|\cdot\|$ is a norm on X.

In many cases it is convenient to simply say or write that "X is a normed space"; we will then implicitly assume that the norm is denoted by $\|\cdot\|$. If it is important or useful to stress the vector space X in the notation of the norm, we will e.g. write $\|\cdot\|_X$ instead of $\|\cdot\|$.

Note that a seminorm $\|\cdot\|$ always satisfies $\|0\| = 0$; this is a consequence of the absolute homogeneity since $\|0\| = \|0\cdot 0\| = |0|\,\|0\| = 0$. Whether $\|\cdot\|$ is positive definite (i.e. a norm) depends on whether the zero vector is the only vector that is sent to the number 0 by $\|\cdot\|$.

Remark 1.0.2 (Normed spaces are metric spaces)**.** Let X be a normed space and define $\mathrm{d}\colon X\times X\to[0,\infty)$ by

$$\mathrm{d}(x,y) := \left\|x-y\right\|$$

for all $x,y\in X$. Then it is easy to check that d is a metric. We say that this metric is **induced** by the norm on X.

Thus, every normed space automatically becomes a metric space. We endow every normed space with this metric, unless otherwise stated.

Definition 1.0.3 (Convergent sequences and Cauchy sequences)**.** Let (M,d) be a metric space.

(a) A sequence $(x_n)_{n\in\mathbb{N}}$ in M is said to **converge** to a point $x\in M$ if

$$\forall\varepsilon>0\,\exists n_0\in\mathbb{N}\,\forall n\in\mathbb{N}:\quad (n\ge n_0 \quad\Rightarrow\quad \mathrm{d}(x_n,x)<\varepsilon)$$

or, equivalently, if $\mathrm{d}(x_n,x)\to 0$.

In this case we write $x_n\to x$ as $n\to\infty$ or $x_n \overset{n\to\infty}{\to} x$ or $\lim_{n\to\infty} x_n = x$.[1]

(b) A sequence $(x_n)_{n\in\mathbb{N}}$ in M is said to be a **Cauchy sequence** if

$$\forall\varepsilon>0\,\exists n_0\in\mathbb{N}\,\forall n\in\mathbb{N}:\quad (n\ge n_0 \quad\Rightarrow\quad \mathrm{d}(x_n,x_{n_0})<\varepsilon)$$

or, equivalently, if

$$\forall\varepsilon>0\,\exists n_0\in\mathbb{N}\,\forall n,m\in\mathbb{N}:\quad (n,m\ge n_0 \quad\Rightarrow\quad \mathrm{d}(x_n,x_m)<\varepsilon).$$

The following continuity properties in normed spaces are straightforward to show.

Proposition 1.0.4 (Sequential continuity of the vector space operations)**.** *Let X be a normed space. Let (x_n) and (y_n) be sequences in X that converge to points $x,y\in X$, respectively, and let (α_n) be a sequence in $\mathbb{K}$ that converges to a point $\alpha\in\mathbb{K}$. Then*

$$x_n+y_n\to x+y \qquad \textit{and} \qquad \alpha_n x_n\to\alpha x.$$

Definition 1.0.5 (Banach spaces)**.**

(a) A metric space (M,d) is called **complete** if every Cauchy sequence in M converges to a point in M.

[1] Recall that the latter notation is justified since it is not difficult to show that limits of sequences are unique in metric spaces.

(b) A normed space X is called a **Banach space** if it is complete with respect to the metric induced by the norm.

Example 1.0.6 (Finite-dimensional spaces)**.** Let X be a normed space and assume that X is finite-dimensional. Then you know from your Analysis 2 course that X is a Banach space.[2]

Examples 1.0.7 (***L^p***-spaces and ***ℓ^p***-spaces)**.** Let $p \in [1,\infty)$ and let $(\Omega, \mathcal{A}, \mu)$ be a measure space.

(a) We recall from the Analysis 3 course that one defines

$$\|f\|_p := \Big(\int_\Omega |f|^p \,\mathrm{d}\mu\Big)^{1/p} \in [0,\infty]$$

for every measurable $f\colon \Omega \to \mathbb{K}$. Moreover, one defines

$$\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K}) := \big\{f\colon \Omega \to \mathbb{K} \mid f \text{ is measurable and } \|f\|_p < \infty\big\}$$

Then $\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ is a vector space with respect to pointwise addition and pointwise scalar multiplication and $\|\cdot\|_p$ is a seminorm on $\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$.

(b) The so-called **kernel**[3]

$$\ker\|\cdot\|_p := \big\{\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K}) \mid \|f\|_p = 0\big\}$$

is a vector subspace of $\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ and one defines the quotient space

$$L^p(\Omega,\mathcal{A},\mu;\mathbb{K}) := \mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})/\ker\|\cdot\|_p = \big\{\tilde f \mid f \in \mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})\big\},$$

where $\tilde f := f + \ker\|\cdot\|_p$ is the **equivalent class** of $f \in \mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ modulo $\ker\|\cdot\|_p$. Recall that two functions $f,g \in \mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ satisfy $\tilde f = \tilde g$ if and only if $f = g$ μ-almost everywhere.[4]

Moreover, the mapping $\|\cdot\|_p : L^p(\Omega,\mathcal{A},\mu;\mathbb{K}) \to [0,\infty)$, $\|\tilde f\|_p := \|f\|_p$ is well-defined and is a norm on $L^p(\Omega,\mathcal{A},\mu;\mathbb{K})$. With respect to this norm, $L^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ is a Banach space.

(c) In the special case of (b) where $\mathcal{A} = 2^\Omega$ and where $\mu = \#\colon 2^\Omega \to [0,\infty]$ is the counting measure, one uses the notation

$$\ell^p(\Omega;\mathbb{K}) := L^p(\Omega, 2^\Omega, \#;\mathbb{K}).$$

Observe that no subset of Ω has counting measure 0 except for the empty set. Hence, $\ker\|\cdot\|_p = \{0\}$ in this case and thus, the quotient space $L^p(\Omega,2^\Omega,\#;\mathbb{K})$ is the same as $\mathcal{L}^p(\Omega,2^\Omega,\#;\mathbb{K})$.

[2] More precisely, this follows from [5, Example 1.3.2 and Theorem 1.7.2] and from the fact that X is, as a vector space, isomorphic to $\mathbb{K}^d$ for $d := \dim X$.

[3] Beware however that this is not the kernel of a linear map since $\|\cdot\|_p$ is not linear.

[4] I.e. if there exists a set $N \in \mathcal{A}$ of measure $\mu(N) = 0$ such that $f(\omega) = g(\omega)$ for all $\omega \in \Omega \setminus N$.

If $\Omega = \mathbb{N}$, then functions $\Omega \to \mathbb{K}$ are the same as sequences in $\mathbb{K}$ with index set $\mathbb{N}$, so $\ell^p(\mathbb{N};\mathbb{K})$ consists of sequences in this case – precisely of those sequences $x = (x_k)_{k\in\mathbb{N}}$ in $\mathbb{K}$ that satisfy

$$\|x\|_p = \Big(\int_{\mathbb{N}} |x|^p \,\mathrm{d}\#\Big)^{1/p} = \Big(\sum_{k=1}^{\infty} |x_k|^p\Big)^{1/p} < \infty.$$

Examples 1.0.8 (Spaces of continuous functions)**.** Let (M,d) be a compact metric space. Then

$$\mathrm{C}(M;\mathbb{K}) := \big\{f\colon M \to \mathbb{K} \mid f \text{ is continuous}\big\}$$

is a vector space and the map $\|\cdot\|_\infty\colon \mathrm{C}(M;\mathbb{K}) \to [0,\infty)$ given by

$$\|f\|_\infty := \sup\{|f(\omega)| \mid \omega \in \Omega\}$$

for all $f \in \mathrm{C}(M;\mathbb{K})$ is a norm that turns $\mathrm{C}(M;\mathbb{K})$ into a Banach space.

In functional analysis, a linear map $T\colon X \to Y$ between two vector spaces over $\mathbb{K}$ is often called a **linear operator** or simply an **operator**.

Definition 1.0.9 (Operator norm)**.** Let X, Y be normed spaces over $\mathbb{K}$. For a linear operator $T\colon X \to Y$ we call

$$\|T\| := \|T\|_{Y\leftarrow X} := \sup\big\{\|Tx\|_Y \mid x \in X \text{ and } \|x\|_X \le 1\big\} \in [0,\infty]$$

the **operator norm of *T***.

Sometimes we will also use slightly different but analogous notation to make the norms on the domain and co-domain of T clear. For instance, let $c, d \in \mathbb{N}$ and $p, q \in [1,\infty]$. If $T\colon \mathbb{K}^c \to \mathbb{K}^d$ and if we endow the domain $\mathbb{K}^c$ with the norm $\|\cdot\|_p$ and the co-domain $\mathbb{K}^d$ with the norm $\|\cdot\|_q$, then we will often write

$$\|T\|_{q\leftarrow p}$$

for the operator norm of T.

Proposition 1.0.10 (Characterisation of the operator norm)**.** *Let X, Y be normed spaces over $\mathbb{K}$ and let $T\colon X \to Y$ be linear. Then one has*

$$\begin{aligned}\|T\|_{Y\leftarrow X} &= \sup\big\{\|Tx\|_Y \mid x \in X \textit{ and } \|x\|_X = 1\big\} \\ &= \inf\big\{L \in [0,\infty) \mid \forall x \in X\ \|Tx\|_Y \le L\|x\|_X\big\}\end{aligned}$$

where the supremum and the infimum are taken within the ordered set $[0,\infty]$, which means that $\inf\emptyset = \infty$ and $\sup\emptyset = 0$ in this case.[5]

[5]Beware that it also occurs very frequently in analysis that suprema and infima are taken within the ordered set $[-\infty,\infty]$, in which case one has $\sup\emptyset = -\infty$.

Proposition 1.0.11 (Characterisation of continuity and of the operator norm)**.** *Let X, Y be normed spaces over $\mathbb{K}$ and let $T\colon X \to Y$ be linear. The following are equivalent:*

(i) *T is continuous.*

(ii) *T is continuous at 0.*

(iii) *There exists a point $x_0 \in X$ such that T is continuous at x_0.*

(iv) *One has $\|T\|_{Y\leftarrow X} < \infty$.*

(v) *There exists a number $L \in [0,\infty)$ such that $\|Tx\|_Y \le L\,\|x\|_X$ for all $x \in X$.*

(vi) *T is Lipschitz continuous.*

Note that if the equivalent assertions in the preceding proposition are satisfied, then the smallest Lipschitz constant of T is precisely the operator norm $\|T\|_{Y\leftarrow X}$. Assertion (v) in the preceding proposition motivates part (a) of the following definition.

Definition 1.0.12 (The space of bounded linear operators)**.** Let X, Y be normed spaces over $\mathbb{K}$.

(a) A linear operator $T\colon X \to Y$ is called **bounded** if and only if it is continuous.

(b) We define

$$\mathcal{L}(X;Y) := \big\{T\colon X \to Y \mid T \text{ is linear and bounded}\big\}$$

and endow it with the operator norm.

The terminology **operator norm** is justified by the observation that it is indeed a norm on $\mathcal{L}(X;Y)$, a fact which is straightforward to check. Further elementary but very useful properties of the operator norm are listed in the following proposition.

Proposition 1.0.13 (Submultiplicativity of the operator norm)**.** *Let X, Y, Z be normed spaces over $\mathbb{K}$ and let $S\colon X \to Y$ and $T\colon Y \to Z$ be linear. In other words, we consider linear operators*

$$Z \xleftarrow{T} Y \xleftarrow{S} X.$$

(a) *For all $x \in X$ one has $\|Sx\| \le \|S\|\,\|x\|$ or, more precisely,*

$$\|Sx\|_Y \le \|S\|_{Y\leftarrow X}\,\|x\|_X$$

(b) *One has $\|TS\| \le \|T\|\,\|S\|$ or, more precisely,*

$$\|TS\|_{Z\leftarrow X} \le \|T\|_{Z\leftarrow Y}\,\|S\|_{Y\leftarrow X}\,.$$

Examples 1.0.14 (Operator norms of matrices)**.** Let $c, d \in \mathbb{N}$ and $A \in \mathbb{K}^{d\times c}$. As usual we interpret the matrix A as the linear map $\mathbb{K}^c \to \mathbb{K}^d$, $x \mapsto Ax$. We use the notation explained after Definition 1.0.9.

(a) One has

$$\|A\|_{1\leftarrow 1} = \max\Big\{ \|A_{\bullet,k}\|_1 \mid k \in \{1,\dots,c\}\Big\},$$

where $A_{\bullet,k} \in \mathbb{K}^d$ denotes the kth column of A.

(b) One has

$$\|A\|_{\infty\leftarrow\infty} = \max\Big\{ \|A_{j,\bullet}\|_1 \mid j \in \{1,\dots,d\}\Big\},$$

where $A_{j,\bullet} \in \mathbb{K}^{1\times c}$ denotes the jth row of A.

(c) Let $D \in \mathbb{K}^{d\times d}$ be a diagonal matrix. Let $p \in [1,\infty]$. Then

$$\|D\|_{p\leftarrow p} = \max\big\{ |D_{jj}| \mid j \in \{1,\dots,d\}\big\}.$$

Computing the norm $\|A\|_{p\leftarrow p}$ for a general matrix and a general p is usually a difficult task. There is no simple analytic formula and even numerically this is not easy if the dimension is high. On the other hand, there is indeed a good approach to compute the norm $\|A\|_{2\leftarrow 2}$. We will discuss this in Chapter 6.

1.1 Density and closedness of vector subspaces

Lecture 1 (Mo, 13 April) starts at the beginning of Section 1.1.

Introductory questions.

(a) There are infinitely many (even uncountably many) real numbers, but a computer has only finite memory. How can one actually work with real numbers on a computer?

(b) How can one represent functions $f\colon [0,1] \to \mathbb{R}$ on a computer?

(c) What does it mean that two functions $f, g\colon [0,1] \to \mathbb{R}$ are “close” to each other?

(d) Can every continuous function $f\colon [-1,1] \to \mathbb{R}$ be approximated by polynomials? And what does “approximated” even mean in this context?

We begin with a few elementary concepts and results in normed spaces that are related to density and closedness. At the end of this section we will prove two versions of the famous Stone–Weierstraß approximation theorem which gives sufficient criteria for a vector subspace to be dense in a space of continuous functions (Theorems 1.1.8 and 1.1.9).

Proposition 1.1.1 (Closed vector subspaces)**.** *Let X be a normed space and let $V \subseteq X$ be a vector subspace.*

(a) *The closure* $\overline{V}$ *is also a vector subspace of* X.

(b) *If* V *is a Banach space with respect to the norm induced by* X, *then* V *is closed in* X.

(c) *If* X *is a Banach space and* V *is closed in* X, *then* V *is also a Banach space with respect to the norm induced by* X.

Proof. (a) One has $0 \in V \subseteq \overline{V}$ and hence $0 \in \overline{V}$. Now let $x, y \in \overline{V}$ and $\alpha, \beta \in \mathbb{K}$. Then there exist sequences (x_n) and (y_n) in V that converge to x and y, respectively. It follows from Proposition 1.0.4 that

$$\underbrace{\alpha x_n + \beta y_n}_{\in V} \to \alpha x + \beta y,$$

so $\alpha x + \beta y \in \overline{V}$.

(b) *and* (c) These are general properties of subsets of metric spaces that are straightforward to check. □

Example 1.1.2 (Finite-dimensional subspaces)**.** Let X be a normed vector space. Then every finite-dimensional vector subspace $V \subseteq X$ is closed.

Proof. Let $V \subseteq X$ be finite-dimensional. According to Example 1.0.6, V is a Banach space with respect to any norm and hence, in particular, with respect to the norm induced by X. Thus, Proposition 1.1.1(b) shows that V is closed in X. □

Recall that a subset S of a metric space (M, d) is called **dense in M**, or more briefly **dense**, if $\overline{S} = M$.

Definition 1.1.3 (Separable space)**.** A metric space (M, d) is called **separable** if there exists a countable subset $S \subseteq M$ that is dense in M.

Recall form your Analysis 2 lecture that a subset $K \subseteq M$ of a metric space (M, d) is called **totally bounded** if for each $\varepsilon > 0$ there exists a finite set $F \subseteq M$ such that $K \subseteq \bigcup_{x \in F} B_{<\varepsilon}(x)$ [5, Definition 1.3.5]. Also recall that the subset $K \subseteq M$ is compact if and only if K is totally bounded and (K, d) is complete [5, Theorem 1.3.7 and Definition 1.3.8].

Alternatively, compactness can be characterized as follows:[6] a subset $K \subseteq M$ is compact if and only if the following holds: for every family $(U_j)_{j \in J}$ of open subsets of M that covers K, i.e. that satisfies $\bigcup_{j \in J} U_j \supseteq K$, there exist finitely many indices $j_1, \dots, j_n \in J$ such that $U_{j_1} \cup \dots U_{j_n} \supseteq K$.

Example 1.1.4 (Compact metric spaces are separable)**.** If a metric space is totally bounded, then it is separable. In particular, every compact metric space is separable.

[6] This alternative characterization, the so-called **Heine–Borel property**, is very useful, for instance in the proof of Theorem 1.1.8 below and in Exercise 1.2(b) on Homework Sheet 1.

Proof. Let (M, d) be a metric space and assume that M is totally bounded. Then for each $n \in \mathbb{N}$ there exists a finite set $S_n \subseteq M$ such that

$$\bigcup_{x \in S_n} B_{<1/n}(x) = M.$$

The set $S := \bigcup_{n\in\mathbb{N}} S_n$ is countable since it is a countable union of finite sets. Moreover, for each $y \in M$ and each $\varepsilon > 0$ there exists a point $x \in S$ such that $\mathrm{d}(x, y) < \varepsilon$. Thus, S is dense in M. □

For normed spaces we show in the next proposition that the vector space structure can be employed to characterize separability. This is sometimes useful to prove separability for concrete examples of normed spaces.

Proposition 1.1.5 (Characterisation of separable normed spaces)**.** *Let X be a normed spaces. The following are equivalent:*

(i) *The space X is separable, i.e. there exists a countable dense subset of X.*

(ii) *There exists a countable subset $C \subseteq X$ whose span*[7] $\operatorname{span} C$ *is dense in X.*

Proof. "(i) ⇒ (ii)" Let X be separable and let $S \subseteq X$ be a countable and dense subset of X. It follows from $S \subseteq \operatorname{span} S$ that

$$X = \overline{S} \subseteq \overline{\operatorname{span} S} \subseteq X,$$

so $X = \overline{\operatorname{span} S}$. Hence, (ii) holds with $C := S$.

"(ii) ⇒ (i)" We first consider the case where the scalar field $\mathbb{K}$ is equal to $\mathbb{R}$. Let $C \subseteq X$ be as in (ii) and let S be the set of all linear combinations of elements of C with rational coefficients, i.e.

$$S = \Big\{ \sum_{k=1}^{n} \alpha_k x_k \;\Big|\; n \in \mathbb{N}_0,\ x_1, \dots, x_n \in C,\ \alpha_1, \dots, \alpha_n \in \mathbb{Q} \Big\}.$$

Then S is countable and $\overline{S} = \overline{\operatorname{span} C} = X$, so S is a countable dense subset of X.

Now let the scalar field $\mathbb{K}$ be $\mathbb{C}$. Then the same argument works if one allows the coefficients in the construction of S to be from $\mathbb{Q}[\mathrm{i}] := \{\alpha + \mathrm{i}\beta \mid \alpha, \beta \in \mathbb{Q}\}$ instead of $\mathbb{Q}$ only. □

Examples 1.1.6 (Sequences spaces)**.**

(a) Let $p \in [1, \infty)$. Then $\ell^p(\mathbb{N}; \mathbb{K})$ is separable.[8]

[7] Recall that the **span** or **linear span** $\operatorname{span} S$ of a subset S of a vector space X is the smallest vector subspace of X that contains S. Also recall that $\operatorname{span} S$ consists of all linear combinations of elements of S.

[8] The definition of the space $\ell^p(\mathbb{N}; \mathbb{K})$ can be found in Example 1.0.7(c).

(b) On the other hand, let $\ell^\infty(\mathbb{N};\mathbb{K})$ denote the space of all bounded sequences $x = (x_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$. This is a Banach space when endowed with the norm $\|\cdot\|_\infty$ given by

$$\|x\|_\infty := \sup_{k\in\mathbb{N}} |x_k|.$$

for all $x \in \ell^\infty(\mathbb{N};\mathbb{K})$. This space is not separable.

(c) Consider the subset $c_{00}(\mathbb{N};\mathbb{K}) \subseteq \ell^\infty(\mathbb{N};\mathbb{K})$ given by

$$c_{00}(\mathbb{N};\mathbb{K}) := \Big\{x \in \ell^\infty(\mathbb{N};\mathbb{K}) \mid x_k = 0 \text{ for all but finitely many } k \in \mathbb{N}\Big\}.$$

Then $c_{00}(\mathbb{N};\mathbb{K})$ is a vector subspace of $\ell^\infty(\mathbb{N};\mathbb{K})$ that is not closed. The normed space $\big(c_{00}(\mathbb{N};\mathbb{K}), \|\cdot\|_\infty\big)$ is separable. It is not a Banach space.

Proof. (a) For every $n \in \mathbb{N}$ let $e^{(k)} \in \ell^p(\mathbb{N};\mathbb{K})$ denote the n-the canonical unit vector, i.e.

$$e^{(n)} = \big(0,\dots,0,1,0,\dots\big),$$

where the entry 1 occurs in the nth position and all other entries are 0. The set $C := \{e^{(n)} \mid \mathbb{N}\}$ is countable, so according to Proposition 1.1.5 it suffices that $\operatorname{span} C$ is dense in $\ell^p(\mathbb{N};\mathbb{K})$.[9] So let $x \in \ell^p(\mathbb{N};\mathbb{C})$. For each $n \in \mathbb{N}$ we set

$$x^{(n)} := \big(x_1, x_2, \dots, x_n, 0, 0, \dots\big) = \sum_{k=1}^n x_k e^{(k)} \in \operatorname{span} C.$$

Since $\sum_{k=1}^\infty |x_k|^p = \|x\|_p^p < \infty$, it follows that

$$\big\|x^{(n)} - x\big\|_p^p = \sum_{k=n+1}^\infty |x_k|^p \to 0$$

as $n \to \infty$. Hence, $x^{(n)} \to x$ as $n \to \infty$, i.e. $\overline{\operatorname{span} C} = \ell^p(\mathbb{N};\mathbb{K})$.

(b) The fact that $\ell^\infty(\mathbb{N};\mathbb{K})$ is a Banach space is a special case of Example 1.1.7 below; just take the measure space there to be $(\mathbb{N}, 2^{\mathbb{N}}, \#)$, where # denotes the counting measure.

Let us now show that $\ell^\infty(\mathbb{N};\mathbb{K})$ is not separable. To simplify the notation in the argument, we consider the sequences in $\ell^\infty(\mathbb{N};\mathbb{K})$ as bounded functions $\mathbb{N} \to \mathbb{K}$. For every $A \in 2^{\mathbb{N}}$ consider the indicator function $\mathbb{1}_A \in \ell^\infty(\mathbb{N};\mathbb{K})$. For two different sets $A_1, A_2 \in 2^{\mathbb{N}}$ the difference $\mathbb{1}_{A_1} - \mathbb{1}_{A_2}$ only takes the values $-1, 0, 1$, and at least one of the values $-1, 1$ is attained at least once. Hence, $\big\|\mathbb{1}_{A_1} - \mathbb{1}_{A_2}\big\|_\infty = 1$.

Now let $S \subseteq \ell^\infty(\mathbb{N};\mathbb{K})$ be dense. Then for each $A \in 2^{\mathbb{N}}$ there exists an $s_A \in S$ such that $\|s_A - \mathbb{1}_A\|_\infty < \frac{1}{2}$. But we just showed that the mutual distance of all the indicator functions $\mathbb{1}_A$ is 1, so $s_{A_1} \neq s_{A_2}$ for $A_1 \neq A_2$. This means that the map $2^{\mathbb{N}} \to S$, $A \mapsto s_A$ is injective. As $2^{\mathbb{N}}$ is uncountable, so is S.[10]

[9] We note in passing that actually $\operatorname{span} C = c_{00}(\mathbb{N};\mathbb{K})$.

[10] This argument is generalized a bit in Bonus Exercise 1.7 on Homework Sheet 1.

(c) It is easy to see that $c_{00}(\mathbb{N};\mathbb{K})$ is a vector subspace of $\ell^\infty(\mathbb{N};\mathbb{K})$. To see that it is not closed, consider the sequence $(x^{(n)})_{n\in\mathbb{N}}$ in $c_{00}(\mathbb{N};\mathbb{K})$ given by

$$x^{(n)} := \left(1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, 0, 0, \dots\right)$$

for each $n \in \mathbb{N}$, and let $x \in \ell^\infty(\mathbb{N};\mathbb{K})$ be given by $x_k = \frac{1}{k}$ for each $k \in \mathbb{N}$. Then we have $\left\|x^{(n)} - x\right\|_\infty = \frac{1}{n+1}$ for every $n \in \mathbb{N}$, so $x^{(n)} \to x$ with respect to $\|\cdot\|_\infty$ as $n \to \infty$. But $x \notin c_{00}(\mathbb{N};\mathbb{K})$, so $c_{00}(\mathbb{N};\mathbb{K})$ is not closed in $\ell^\infty(\mathbb{N};\mathbb{K})$. Proposition 1.1.1(b) thus implies that $c_{00}(\mathbb{N};\mathbb{K})$ is not a Banach space with respect to $\|\cdot\|_\infty$.

The space $c_{00}(\mathbb{N};\mathbb{K})$ is itself the span of the countable set of all canonical unit vectors, so it is separable by Proposition 1.1.5. □

Just as the spaces $\ell^p(\mathbb{N};\mathbb{K})$ are special case of L^p-spaces over general measure spaces, one can also consider the space $\ell^\infty(\mathbb{N};\mathbb{K})$ from Example 1.1.6(b) as a special case of L^∞-spaces over general measure spaces. We discuss now how those L^∞-spaces are defined.

Lecture 2 (Th, 16 April) starts with Example 1.1.7. Careful: In the lecture and in a former version of the lecture notes, a second formula for $\|f\|_{\text{ess},\infty}$ was mentioned here that contained a supremum. Unfortunately, the formula was complete nonsense.

Example 1.1.7 (The space $\boldsymbol{L^\infty}$)**.** Let $(\Omega, \mathcal{A}, \mu)$ be a measure space.

(a) For every measurable function $f\colon \Omega \to \mathbb{K}$ we define

$$\|f\|_{\text{ess},\infty} := \inf\left\{s \in [0,\infty) \mid \left|f\right|(\omega) < s \text{ for } \mu\text{-almost all } \omega \in \Omega\right\} \in [0,\infty],$$

where the infimum and the supremum are taken within the linearly ordered set $[0,\infty]$, which means that $\inf\emptyset = \infty$ and $\sup\emptyset = 0$.[11] We call the value $\|f\|_{\text{ess},\infty}$ the **essential sup norm** or the **essential ∞-norm** of f (although it is not really a norm since it can take the value ∞ and since it can be 0 for some non-zero functions).

It is not difficult to check that, for all measurable functions $f, g\colon \Omega \to \mathbb{K}$ and every $\alpha \in \mathbb{K}$ one has

$$\|f+g\|_{\text{ess},\infty} \le \|f\|_{\text{ess},\infty} + \|g\|_{\text{ess},\infty}, \qquad \|\alpha f\|_{\text{ess},\infty} = |\alpha|\,\|f\|_{\text{ess},\infty},$$

and $\|f\|_{\text{ess},\infty} = 0$ if and only if $f = 0$ almost everywhere. From those properties it follows immediately that

$$\mathcal{L}^\infty(\Omega, \mathcal{A}, \mu;\mathbb{K}) := \left\{f\colon \Omega \to \mathbb{K} \mid f \text{ is measurable and } \|f\|_{\text{ess},\infty} < \infty\right\}$$

is a vector subspace of the space of all measurable functions from Ω to $\mathbb{K}$.

(b) The kernel[12]

$$\ker\|\cdot\|_{\text{ess},\infty} := \left\{\mathcal{L}^\infty(\Omega, \mathcal{A}, \mu;\mathbb{K}) \mid \|f\|_{\text{ess},\infty} = 0\right\}$$

is a vector subspace of $\mathcal{L}^\infty(\Omega, \mathcal{A}, \mu;\mathbb{K})$ and one defines the quotient space

$$L^\infty(\Omega, \mathcal{A}, \mu;\mathbb{K}) := \mathcal{L}^\infty(\Omega, \mathcal{A}, \mu;\mathbb{K}) / \ker\|\cdot\|_{\text{ess},\infty} = \left\{\tilde{f} \mid f \in \mathcal{L}^\infty(\Omega, \mathcal{A}, \mu;\mathbb{K})\right\},$$

[11] Observe that $\sup\emptyset$ occurs if μ is the zero measure.

[12] As in Example 1.0.7 beware that this is not the kernel of a linear map since $\|\cdot\|_{\text{ess},\infty}$ is not linear!

where $\tilde f := f + \ker \|\cdot\|_{\mathrm{ess},\infty}$ is the **equivalence class** of a function $f \in \mathcal{L}^\infty(\Omega,\mathcal{A},\mu;\mathbb{K})$ modulo $\ker\|\cdot\|_{\mathrm{ess},\infty}$. One can check that two functions $f,g \in \mathcal{L}^\infty(\Omega,\mathcal{A},\mu;\mathbb{K})$ satisfy $\tilde f = \tilde g$ if and only if $f = g$ μ-almost everywhere.

Furthermore, the mapping $\|\cdot\|_{\mathrm{ess},\infty}\colon L^\infty(\Omega,\mathcal{A},\mu;\mathbb{K}) \to [0,\infty)$, that is defined by $\|\tilde f\|_{\mathrm{ess},\infty} := \|f\|_{\mathrm{ess},\infty}$ for every $\tilde f \in \mathcal{L}^\infty(\Omega,\mathcal{A},\mu;\mathbb{K})$, is well-defined and is a norm on $L^\infty(\Omega,\mathcal{A},\mu;\mathbb{K})$. This norm turns $L^\infty(\Omega,\mathcal{A},\mu;\mathbb{K})$ into a Banach space.[13]

These claims can be proved similarly as we did for $L^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ in [6, Section 3.2].

(c) In the special case of (b) where $\mathcal{A} = 2^\Omega$ and where $\mu = \#\colon 2^\Omega \to [0,\infty]$ is the counting measure, one uses the notation

$$\ell^\infty(\Omega;\mathbb{K}) := L^\infty(\Omega, 2^\Omega, \#;\mathbb{K}).$$

Observe that no subset of Ω has counting measure 0 except for the empty set. This implies that the norms $\|\cdot\|_{\mathrm{ess},\infty}$ and $\|\cdot\|_\infty$ on $\mathcal{L}^\infty(\Omega, 2\Omega, \#;\mathbb{K})$. Hence, $\ker\|\cdot\|_{\mathrm{ess},\infty} = \ker\|\cdot\|_\infty = \{0\}$ and thus, the quotient space $\ell^\infty(\Omega;\mathbb{K}) = L^\infty(\Omega,2^\Omega,\#;\mathbb{K})$ coincides with the space $\mathcal{L}^\infty(\Omega,2^\Omega,\#;\mathbb{K})$.

For $\Omega = \mathbb{N}$, this gives the space $\ell^\infty(\mathbb{N};\mathbb{K})$ that we introduced in Example (b).

In Example 1.1.6(a) we saw that, in some simple cases, one can show directly that a certain vector subspace is dense in a normed space. During the course we shall encounter more elaborate results that allow us to check density of certain vector subspaces. We start with the following two theorems; other density results will follow later.

Theorem 1.1.8 (Stone–Weierstraß approximation theorem: lattice version)**.** *Let* (M,d) *be a compact metric space and let* $A \subseteq \mathrm{C}(M;\mathbb{R})$ *satisfy the following properties:*

(1) *A is a vector subspace of* $\mathrm{C}(M;\mathbb{R})$.

(2) *One has* $\mathbb{1}_M \in A$.

(3) *A separates the points of M, i.e. for all $x,y \in M$ with $x \neq y$ there exists $f \in A$ such that $f(x) \neq f(y)$.*

(4) *One has $|f| \in A$ for all $f \in A$.*

Then A is dense in $\mathrm{C}(M;\mathbb{R})$.

[13] It is very common to actually write $\|\cdot\|_\infty$ instead of $\|\cdot\|_{\mathrm{ess},\infty}$ for the norm on the quotient space $L^\infty(\Omega,\mathcal{A},\mu;\mathbb{K})$, since this makes it easier to write down results simultaneously for all $p \in [1,\infty]$. We shall also use this notation later on. However, when using it one has to be very careful to avoid confusion with the same notation $\|\cdot\|_\infty$ for the pointwise supremum norm (rather than the essential supremum norm) of functions – so for the moment we stick to the more explicit notation $\|\cdot\|_{\mathrm{ess},\infty}$ until you had some time to digest all those concepts.

In the proof of the theorem we use the following notation that you already know from your Analysis 3 course: for a set M and two functions $f, g\colon M \to \mathbb{R}$ one defines the **pointwise supremum** $f \vee g\colon M \to \mathbb{R}$ and the **pointwise infimum** $f \wedge g\colon M \to \mathbb{R}$ by

$$(f \vee g)(x) := \max\{f(x), g(x)\} \qquad \text{and} \qquad (f \wedge g)(x) := \min\{f(x), g(x)\}$$

for all $x \in M$. One has the following relations to the pointwise modulus:

$$f \vee g = \frac{f + g + |f - g|}{2},$$
$$f \wedge g = \frac{f + g - |f - g|}{2}.$$

If M is endowed with a metric d and f, g are continuous, then so are $f \vee g$ and $f \wedge g$. The reason why we call Theorem 1.1.8 the **lattice version** of the Stone–Weierstraß theorem is that a vector subspace $A \subseteq \mathrm{C}(M;\mathbb{R})$ that satisfies $|f| \in A$ for all $f \in A$ is called a **sublattice** or, more precisely, a **vector sublattice** of $\mathrm{C}(M;\mathbb{R})$.

Proof of Theorem 1.1.8. Observe that it follows from assumptions (1) and (4) and from the formulas for $f \vee g$ and $f \wedge g$ before the proof that $f \vee g, f \wedge g \in A$ for all $f, g \in A$. Now let $h \in \mathrm{C}(M;\mathbb{R})$ and $\varepsilon > 0$. We proceed in several steps to show that there exists a function $f \in A$ that satisfies $\|f - h\|_\infty \le \varepsilon$

Step 1: We observe that for all $x, y \in M$ there exists $f_{x,y} \in A$ such that $f_{x,y}(x) = h(x)$ and $f_{x,y}(y) = h(y)$.

Indeed, if $x = y$ we can choose $f_{x,y} := h(x)\,\mathbb{1}_M$, which is in A due to assumptions (1) and (2). So let $x \neq y$ now. As A separates the points of M (assumption (3)), there exists $f \in A$ such that $f(x) \neq f(y)$. Set

$$f_{x,y} := \frac{h(y) - h(x)}{f(y) - f(x)}\big(f - f(x)\,\mathbb{1}_M\big) + h(x)\,\mathbb{1}_M.$$

This function is in A due to assumptions (1) and (2), and it satisfies $f_{x,y}(x) = h(x)$ and $f_{x,y}(y) = h(y)$, as claimed.

Step 2: We show that for all $x \in M$ there exists a function $f_x \in A$ such that $f_x(x) = h(x)$ and $f_x \ge h - \varepsilon\,\mathbb{1}_M$, where the inequality is meant pointwise.

So fix $x \in M$. For each $y \in M$ let $f_{x,y} \in A$ be a function that coincides with h at the points x and y; such a function exists according to Step 1. For each $y \in M$ the set

$$U_y := \{z \in M \mid f_{x,y}(z) > h(z) - \varepsilon\} \subseteq M$$

is open and contains y. This means that the sets U_y are an open cover of M. Due to the compactness of M we can thus find points $y_1, \dots, y_n \in M$ such that $U_{y_1} \cup \dots \cup U_{y_n} = M$ (see e.g. [5, Theorem 1.3.7 and Definition 1.3.8]). We set

$$f_x := f_{x,y_1} \vee \dots \vee f_{x,y_n}.$$

Then $f_x \in A$ by the observation from the beginning of the proof, and one has $f_x(x) = h(x)$ and $f_x(z) > h(z) - \varepsilon$ for all $z \in M$.

Step 3: We show that there exists a function $f \in A$ that satisfied $\|f - h\|_\infty \le \varepsilon$, as claimed.

For each $x \in M$ let $f_x \in A$ satisfy $f_x(x) = h(x)$ and $f_x \ge h - \varepsilon$; such a function exists according to Step 2. We set

$$V_x := \{z \in M \mid f_x(z) < h(z) + \varepsilon\}$$

for each $x \in M$. Each of the sets V_x is open and contains x, so the V_x are an open cover of M. Since M is compact we can find points $x_1, \dots, x_m \in M$ such that $V_{x_1} \cup \dots \cup V_{x_m} = M$. Set

$$f := f_{x_1} \wedge \dots \wedge f_{x_m}.$$

Then $f \in A$ and $h - \varepsilon \mathbb{1}_M \le f \le h + \varepsilon \mathbb{1}_M$, so $\|f - h\|_\infty \le \varepsilon$. □

Theorem 1.1.8 remains true if one replaces assumption (4) with the assumption that A be stable with respect to multiplication:

Theorem 1.1.9 (Stone–Weierstraß approximation theorem: algebra version)**.** *Let* (M, d) *be a compact metric space and let* $A \subseteq \mathrm{C}(M;\mathbb{R})$ *satisfy the following properties:*

(1) *A is a vector subspace of* $\mathrm{C}(M;\mathbb{R})$.

(2) *One has* $\mathbb{1}_M \in A$.

(3) *A separates the points of M, i.e. for all $x, y \in M$ with $x \ne y$ there exists $f \in A$ such that $f(x) \ne f(y)$.*

(4) *One has $fg \in A$ for all $f, g \in A$.*

Then A is dense in $\mathrm{C}(M;\mathbb{R})$.

Proof. We show that the closure $\overline{A}$ satisfies the assumption of Theorem 1.1.8. Then this theorem implies that $\overline{A}$ is dense in $\mathrm{C}(M;\mathbb{R})$, and as $\overline{A}$ is closed it follows that $\overline{A} = \mathrm{C}(M;\mathbb{R})$.

As A is a vector subspace of $\mathrm{C}(M;\mathbb{R})$, so is $\overline{A}$ according to Proposition 1.1.1(a). Similarly, one can conclude that $fg \in \overline{A}$ for all $f, g \in \overline{A}$. Moreover, one has $\mathbb{1}_M \in A \subseteq \overline{A}$. We show in Exercise 1.3 on Homework Sheet 1 that those properties together with the closedness of $\overline{A}$ imply that $|f| \in \overline{A}$ for all $f \in \overline{A}$.

Furthermore, as A separates the points of M, so does $\overline{A}$. Hence, the assumptions of Theorem 1.1.8 are satisfied, as claimed. □

Theorem 1.1.9 is frequently very useful. We will give some applications of this result in the exercises. For now, we use the theorem to show separability $\mathrm{C}(M;\mathbb{R})$ if (M, d) is a compact metric space:

Corollary 1.1.10 (Separability of spaces of continuous functions)**.** *Let* (M, d) *be a compact metric space. Then the Banach space* $\mathrm{C}(M,\mathbb{R})$ *is separable.*

Proof. As (M, d) is compact, it is separable according to Example 1.1.4, so there exists a countable and dense subset $S \subseteq M$. For each $s \in S$ define $f_s \in \mathrm{C}(M; \mathbb{R})$ by

$$f_s(z) := \mathrm{d}(z, s)$$

for all $z \in M$. The set $\{f_s \mid s \in S\}$ separates the points of M. Indeed, let $x, y \in M$ and $x \neq y$. Since S is dense in M, there exists a sequence (s_n) in S that converges to x. Hence, $f_{s_n}(x) = \mathrm{d}(x, s_n) \to 0$ and $f_{s_n}(y) = \mathrm{d}(y, s_n) \to \mathrm{d}(x, y) = \mathrm{d}(x, y) > 0$, so for some index n we have $f_{s_n}(x) \neq f_{s_n}(y)$.

Now set $C := \left\{\prod_{s \in F} f_s \mid F \subseteq S \text{ is finite}\right\}$. Observe that $\mathbb{1}_M \in C$ since the empty product of functions is $\mathbb{1}_M$ in $\mathrm{C}(M; \mathbb{R})$ (but this is not the crucial point here since otherwise we could simply unify C with $\{\mathbb{1}_M\}$). Moreover C is countable since S is countable. The span $A := \operatorname{span} C$ satisfies the assumptions (1) to (4) of Theorem 1.1.9 and is thus dense in $\mathrm{C}(M; \mathbb{R})$. So $\mathrm{C}(M; \mathbb{R})$ has a countable subset whose span is dense; according to Proposition 1.1.5 this implies that $\mathrm{C}(M; \mathbb{R})$ is separable. □

1.2 Linear operators and dual spaces

Lecture 3 (Mo, 20 April) starts with Section 1.2.

Introductory questions.

(a) When would you say that two normed spaces coincide?

(b) What is the difference between a column vector and a row vector?

It is a fundamental principle in most mathematical theory that one can use objects which satisfy a number of axioms, to construct new objects that satisfy the same axioms. The following proposition is a simple example for this principle.

Proposition 1.2.1 (Completeness of operator spaces)**.** *Let X, Y be normed spaces over $\mathbb{K}$. If Y is a Banach space, then so is $\mathcal{L}(X; Y)$.*

Note that it is not relevant for the proposition whether the normed space X is a Banach space or not.

Proof of Proposition 1.2.1. Let Y be a Banach space and let (T_n) be a Cauchy sequence in $\mathcal{L}(X; Y)$. Then (T_n) is bounded, i.e. $M := \sup_{n \in \mathbb{N}} \|T_n\| < \infty$. For every $x \in X$ it follows from

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

for all $n, m \in \mathbb{N}$ that $(T_n x)$ is a Cauchy sequence in Y. As Y is a Banach space, there exists a vector $Tx \in Y$ such that $T_n x \to Tx$. Clearly, the map $x \mapsto Tx$ is linear. Moreover, one has

$$\|Tx\| = \lim_n \|T_n x\| \leq M \|x\|$$

for every $x \in X$, so T is continuous and hence, $T \in \mathcal{L}(X; Y)$.

Finally we need to show that $T_n \to T$ with respect to the operator norm, so let $\varepsilon > 0$. There exists an $n_0 \in \mathbb{N}$ such that $\|T_n - T_m\| \le \varepsilon$ for all $n, m \ge n_0$. Thus, for all $x \in X$

$$\|(T_n - T)x\| = \lim_{m\to\infty} \|(T_n - T_m)x\| =\le \varepsilon \|x\|$$

for all $n \ge n_0$. This shows that $\|T_n - T\| \le \varepsilon$ for all $n \ge n_0$. □

Proposition 1.2.2 (Unique extension of bounded linear operators)**.** *Let X, Y be normed spaces over $\mathbb{K}$ and assume that Y is a Banach space. Let $V \subseteq Y$ be a dense vector subspace that we endow with the norm inherited from X.*

For every bounded linear operator $T\colon V \to Y$ there exists a unique bounded linear operator $\tilde{T}\colon X \to Y$ that extends T (i.e. that satisfies $\tilde{T}|_V = T$). Moreover, $\|\tilde{T}\| = \|T\|$.

Proof. Let $T\colon V \to Y$ be bounded.

Uniqueness: Let $\tilde{T}_1, \tilde{T}_2 \in \mathcal{L}(X;Y)$ be extensions of T. Let $x \in X$. Since V is dense in X be assumption, there exists a sequence (v_n) in V that converges to x. Since $\tilde{T}_1$ and $\tilde{T}_2$ are continuous from X to Y, it follows that

$$\tilde{T}_1 x = \lim_n \tilde{T}_1 v_n = \lim_n T v_n = \lim_n \tilde{T}_2 v_n = \tilde{T}_2 x.$$

Hence, $\tilde{T}_1 = \tilde{T}_2$.

Existence and norm estimate: We first observe that for every sequence (v_n) in V that converges to a point $x \in X$, the sequence $(T v_n)$ in Y converges to a point in Y. Indeed, it follows from

$$\|T v_n - T v_m\| \le \|T\| \, \|v_n - v_m\|$$

for all $n, m \in \mathbb{N}$ that $(T v_n)$ is a Cauchy sequence in Y. As Y is a Banach space by assumption, it follows that $(T v_n)$ converges in Y.

Next observe that, for a point $x \in X$ and a sequence (v_n) in V with $v_n \to x$, the limit $\lim_n T v_n \in Y$ only depends on x, but not on (v_n). Indeed, if (w_n) is another sequence in V that converges to x, then

$$\|T v_n - T w_n\| \le \|T\| \, \|v_n - w_n\| \to \|T\| \, \|x - x\| = 0,$$

so $\lim_n T v_n = \lim_n T w_n$.

Hence, we can set $\tilde{T} x := \lim_n T v_n$ for every $x \in X$, where (v_n) is any sequence in V that converges to x; note that such a sequence always exists since V is dense in X by assumption. We need to show that $\tilde{T} \in \mathcal{L}(X;Y)$ and that $\tilde{T}$ extends T.

If $v \in V$, then the constant sequence (v) converges to v, so $\tilde{T} = \lim_n T v = T v$; hence, $\tilde{T}$ does indeed extent T. Moreover, it is straightforward to check that $\tilde{T}$ is linear. To show the boundedness of $\tilde{T}$, let $x \in X$ and choose a sequence (v_n) in V that converges to x. Then $\|v_n\| \to \|x\|$ and hence,

$$\|\tilde{T} x\| = \lim_n \|T v_n\| = \lim_n \|T\| \, \|v_n\| = \|T\| \, \|x\|.$$

This shows that $\tilde{T}$ is indeed bounded and that $\|\tilde{T}\| \le \|T\|$. On the other hand, the inequality $\|\tilde{T}\| \ge \|T\|$ is clear since $\tilde{T}$ extends T. □

Note that the space $\mathbb{K}$ is a vector field over itself. Throughout, we endow it with the modulus $|\cdot|$, which is a norm on $\mathbb{K}$. As you now from course Analysis 1 this space is complete. In functional analytic terminology this means that $\mathbb{K}$ is a Banach space.

Definition 1.2.3 (Linear functionals and dual spaces)**.** Let X be a normed space over $\mathbb{K}$.

(a) A linear map $X \to \mathbb{K}$ is called a **linear functional**. If it is continuous we call it, in accordance with Definition 1.0.12(a), a **bounded linear functional**.

(b) The space

$$X' := \mathcal{L}(X;\mathbb{K})$$

of all bounded linear functionals on X called the **dual space of X**. We often denote its elements by x', y', and so on.

(c) For $x' \in X$ and $x \in X$ we often use the notation

$$\langle x', x \rangle := x'x = x'(x) \in \mathbb{K}.$$

The reasons why the notation $\langle x', x\rangle$ is often convenient will become apparent on several occasions when we actually use this notation. For the norm of a bounded linear functional $x' \in X'$ we often write $\|x'\|$ or $\|x'\|_{X'}$. Note that, since $\mathbb{K}$ is a Banach space, Proposition 1.2.1 implies that X' is a Banach space (no matter whether the normed space X itself is a Banach space).

Example 1.2.4 (Bounded linear functionals on $\boldsymbol{\ell^p}$)**.** Let $p, p' \in (1,\infty)$ be such that $\frac{1}{p} + \frac{1}{p'} = 1$. Let $g \in \ell^{p'}(\mathbb{N};\mathbb{K})$. For every $f \in \ell^p(\mathbb{N};\mathbb{K})$ one has $\sum_{n=1}^\infty |g_n f_n| \le \|f\|_p \|g\|_{p'} < \infty$, so we can define the map $\varphi_g \colon \ell^p(\mathbb{N};\mathbb{K}) \to \mathbb{K}$ by

$$\varphi_g(f) := \sum_{n=1}^\infty g_n f_n,$$

for every $f \in \ell^p(\mathbb{N};\mathbb{K})$. The map φ is a bounded linear functional on $\ell^p(\mathbb{N};\mathbb{K})$ – i.e. an element of the dual space $\big(\ell^p(\mathbb{N};\mathbb{K})\big)'$ – and its norm is $\|\varphi_g\| = \|g\|_{p'}$.

Proof. Consider the counting measure $\# \colon 2^{\mathbb{N}} \to [0,\infty]$. For every $f \in \ell^p(\mathbb{N};\mathbb{K})$ one has

$$\sum_{n=1}^\infty |g_n f_n| = \int_{\mathbb{N}} |f|\,|g| \,\mathrm{d}\# \le \Big(\int_{\mathbb{N}} |f|^p \,\mathrm{d}\#\Big)^{1/p} \Big(\int_{\mathbb{N}} |g|^{p'} \,\mathrm{d}\#\Big)^{1/p'} = \|f\|_p \|g\|_{p'},$$

where the inequality is Hölder's inequality [6, Theorem 3.2.2(e)] for the special case of the counting measure. So the series in the definition of $\varphi_g(f)$ is absolutely convergent. Clearly, $\varphi_g \colon \ell^p(\mathbb{N};\mathbb{K}) \to \mathbb{K}$ is linear. Moreover, one has

$$|\varphi_g(f)| \le \sum_{n=1}^\infty |g_n f_n| \le \|f\|_p \|g\|_{p'}$$

as we have just shown. Hence, φ_g is bounded and satisfies $\|\varphi_g\| \le \|g\|_{p'}$.

It remains to show the converse inequality $\|\varphi_g\| \ge \|g\|_{p'}$. If g is zero so is φ_g, so assume now that $g \ne 0$. Define a sequence $f = (f_n)_{n\in\mathbb{N}}$ as follows:

$$f_n := \begin{cases} \overline{g_n}\,|g_n|^{p'-2} & \text{if } g_n \ne 0, \\ 0 & \text{if } g_n = 0 \end{cases}$$

for all $n \in \mathbb{N}$. Then $f \ne 0$. One has $|f_n| = |g_n|^{p'-1}$ for every $n \in \mathbb{N}$, so

$$\|f\|_p^p = \sum_{n=1}^{\infty} |g_n|^{(p'-1)p} = \sum_{n=1}^{\infty} |g_n|^{p'} = \|g\|_{p'}^{p'} < \infty,$$

where we used that $\frac{1}{p} = 1 - \frac{1}{p'}$ and thus, $p' = (p'-1)p$. So $f \in \ell^p(\mathbb{N};\mathbb{K})$ and $\|f\|_p = \|g\|_{p'}^{p'/p} = \|g\|_{p'}^{p'-1}$. Moreover, we have $g_n f_n = |g_n|^{p'}$ for all $n \in \mathbb{N}$, so

$$|\varphi_g(f)| = \sum_{n=1}^{\infty} |g_n|^{p'} = \|g\|_{p'}^{p'} = \|g\|_{p'}\,\|g\|_{p'}^{p'-1} = \|g\|_{p'}\,\|f\|_p.$$

Since $f \ne 0$, it follows that $\|\varphi_g\| \ge \|g\|_{p'}$, as claimed. □

We will show in Theorem 1.2.6 below that every bounded linear functional on $\ell^p(\mathbb{N};\mathbb{K})$ is of the form above. To phrase this result in a particularly instructive way, we use the following notions.

Definition 1.2.5 ((Isometric) isomorphisms of normed spaces)**.** Let X, Y be normed spaces over $\mathbb{K}$.

(a) A linear operator $T\colon X \to Y$ is called **isometric** if $\|Tx\| = \|x\|$ for all $x \in X$.

(b) A linear operator T is called an **isomorphism of normed space** if T is bijective and both T and T^{-1} are continuous.

(c) A linear operator T is called an **isometric isomorphism of normed spaces** if T is bijective and isometric.

If X and Y are Banach spaces we sometimes say **isometric isomorphism of Banach spaces** instead of **isometric isomorphism of normed spaces**.

Note that it makes sense to called a bijective isometric linear map an **isomorphism** since its inverse map is automatically isometric, too.

Theorem 1.2.6 (The dual space of $\boldsymbol{\ell^p}$)**.** *Let $p, p' \in (1,\infty)$ satisfies $\frac{1}{p} + \frac{1}{p'} = 1$. For each $g \in \ell^{p'}(\mathbb{N};\mathbb{K})$ let $\varphi_g \in \big(\ell^p(\mathbb{N};\mathbb{K})\big)'$ be the functional from Example 1.2.4. Then the map*

$$\begin{aligned} T\colon \ell^{p'}(\mathbb{N};\mathbb{K}) &\to \big(\ell^p(\mathbb{N};\mathbb{K})\big)', \\ g &\mapsto \varphi_g \end{aligned}$$

is an isometric isomorphism of Banach spaces.

Lecture 4 (Th, 23 April) starts with the proof of Theorem 1.2.6.

Proof. One can readily check that T is linear, and we showed in Example 1.2.4 that $\|\varphi_g\| = \|g\|_{p'}$ for every $g \in \ell^{p'}(\mathbb{N};\mathbb{K})$, so T is isometric. It remains to show that T is surjectiv.

To this end, we use the following notation. For every $n \in \mathbb{N}$ consider the sequence $e^{(n)} = (0,\dots,0,1,0\dots)$ with index set $\mathbb{N}$, where the entry 1 occurs at position 1. Moreover, for every sequence $x = (x_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ we define he sequence

$$x^{(n)} := (x_1,\dots,x_n,0,0,\dots) = \sum_{k=1}^{n} x_k e^{(k)}.$$

Now let $\psi \in \left(\ell^p(\mathbb{N};\mathbb{K})\right)'$. Define sequence $g = (g_n)_{n\in\mathbb{N}}$ by setting

$$g_n := \langle \varphi, e^{(n)} \rangle \in \mathbb{K}$$

for each $n \in \mathbb{N}$. We first show that $g \in \ell^{p'}(\mathbb{N};\mathbb{K})$. Indeed, for every $n \in \mathbb{N}$ we have $g^{(n)} \in \ell^{p'}(\mathbb{N};\mathbb{K})$, and for each $f \in \ell^p(\mathbb{N};\mathbb{K})$,

$$\left\langle \varphi_{g^{(n)}}, f \right\rangle = \sum_{k=1}^{n} g_k f_k = \sum_{k=1}^{n} \left\langle \psi, e^{(k)} \right\rangle f_k = \langle \psi, f^{(n)} \rangle,$$

so $\left|\left\langle \varphi_{g^{(n)}}, f \right\rangle\right| \le \|\psi\| \, \|f^{(n)}\|_p \le \|\psi\| \, \|f\|_p$. This shows that $\left\|\varphi_{g^{(n)}}\right\| \le \|\psi\|$ for each $n \in \mathbb{N}$ and hence,

$$\|g\|_{p'} = \lim_{n\to\infty} \|g^{(n)}\|_{p'} = \lim_{n\to\infty} \left\|\varphi_{g^{(n)}}\right\| \le \|\psi\|,$$

where the first equality uses the monotone convergence theorem (for integrals with respect to the counting measure). So indeed, $g \in \ell^{p'}(\mathbb{N};\mathbb{K})$.

Finally, it is easy to check that $\langle \varphi_g, f \rangle = \langle \psi, f \rangle$ for all $f \in c_{00}(\mathbb{N};\mathbb{K})$. The proof of Example 1.1.6(a) shows that $c_{00}(\mathbb{N};\mathbb{K})$ is dense in $\ell^p(\mathbb{N};\mathbb{K})$, so the continuity of φ_g and ψ implies that $\varphi_g = \psi$. $\square$

A similar result as Theorem 1.2.6 also holds for L^p-spaces over general measure spaces, but this requires more preparation to prove. We will come back to (parts of) this result on various occasions throughout the course. Another interesting question that we will discuss occasionally is what happens in the cases $p = 1$, $p' = \infty$ and, conversely, $p' = \infty$, $p = 1$.

1.3 Nets, convergence, and series

Introductory questions.

(a) What precisely do we mean when we write $\lim_{x\to\infty} \frac{1}{x} = 0$? And how is this related to convergence of sequences?

(b) Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$. What does the notation $\sum_{n=1}^{\infty} a_n$ mean? Does the notation $\sum_{n\in\mathbb{N}} a_n$ mean the same thing?

(c) What properties of $\mathbb{N}$ are important when one studies convergence of sequences?

Definition 1.3.1 (Directed sets and nets)**.**

(a) Let I be a set. A relation $\leq$ on I is called a **direction** it it satisfies the following axioms:

(I) **Reflexivity:** For all $i \in I$ one has $i \leq i$.

(II) **Transitivity:** For all $i_1, i_2, i_3 \in I$ the following implication holds:

$$(i_1 \leq i_2 \quad i_2 \leq i_3) \qquad \Rightarrow \qquad i_1 \leq i_3$$

(III) **Directedness:** For all $i_1, i_2 \in I$ there exists $j \in I$ such that $i_1 \leq j$ and $i_2 \leq j$.

(b) A set I together with a direction $\leq$ on I is called a **directed set**.

(c) Let X be a set. A **net** in X is a family $(x_i)_{i\in I}$ where I is a non-empty directed set.

In other words: a net in X is a map x from a non-empty directed set I to X.

Examples 1.3.2 (Some directed sets and nets)**.** (a) The set $\mathbb{N}$ together with its usual order is directed. In particular, every sequence $(x_n)_{n\in\mathbb{N}}$ in a set X is a net.

(b) The set $\mathbb{R}$ together with its usual order is directed. In particular, every function f from $\mathbb{R}$ to a set X defines a net in X.

(c) Let S be a set and let $\mathcal{F}$ denote the set of all finite subsets of S. Then $\mathcal{F}$ is directed with respect to the set inclusion $\subseteq$. We shall see that this directed set is very instrumental to build useful nets.

Definition 1.3.3 (Convergent nets and Cauchy nets)**.** Let (M, d) be a metric space and let $(x_i)_{i\in I}$ be a net in M.

(a) Let $x \in M$. The net $(x_i)_{i\in I}$ is said to **converge to *x*** if

$$\forall \varepsilon > 0\, \exists i_0 \in I\, \forall i \in I: \qquad (i \geq i_0 \quad \Rightarrow \quad \mathrm{d}(x_i, x) < \varepsilon).$$

In this case x is called the **limit** of the net $(x_i)_{i\in I}$ and we write $x_i \xrightarrow{i} x$, or simply $x_i \to x$, or $x = \lim_i x_i$. If one wants to stress the index set one can also write $x_i \xrightarrow{i\in I} x$ or $x = \lim_{i\in I} x_i$.[14]

(b) The net $(x_i)_{i\in M}$ is said to **converge** or, more precisely, to **converge in *M*** if there exists a point $x \in M$ such that $(x_i)_{i\in I}$ converges to x.

[14] Note that the direction $\leq$ on the index set I is suppressed in each of those notations, but is extremely important for the definition of convergence. So if the direction is not clear from the context one has to mention explicitly which direction is used on I.

(c) The net $(x_i)_{i\in I}$ is said to be a **Cauchy net** if

$$\forall \varepsilon > 0\, \exists i_0 \in I\, \forall i \in I: \qquad (i \geq i_0 \quad \Rightarrow \quad \mathrm{d}(x_i, x_{i_0}) < \varepsilon)$$

or, equivalently, if

$$\forall \varepsilon > 0\, \exists i_0 \in I\, \forall i, j \in I: \qquad (i, j \geq i_0 \quad \Rightarrow \quad \mathrm{d}(x_i, x_j) < \varepsilon).$$

Observe that all those definitions are consistent with the usual definitions for sequences. The following proposition shows that it makes sense to talk about *the* limit x of a convergent net $(x_i)_{i\in I}$ and to use the notation $x = \lim_i x_i$. Its proof is also a first nice illustration of how the directed axiom of direct sets is used.

Proposition 1.3.4 (Limits of nets in metric spaces are unique)**.** *Let (M, d) be a metric space and let $(x_i)_{i\in I}$ be a net in M. Then $(x_i)_{i\in I}$ converges to at most one point in M.*

Proof. Let $x, y \in M$ and assume that $x_i \to x$ and $x_i \to y$. Let $\varepsilon > 0$. Then there exist indices $i_0, i_1 \in I$ such $\mathrm{d}(x_i, x) < \varepsilon/2$ for all $i \geq i_0$ and $\mathrm{d}(x_i, y) < \varepsilon/2$ for all $i \geq i_1$.

Since I is directed, there exists a point $j \in I$ that satisfies $j \geq i_0$ and $j \geq i_1$. Hence,

$$\mathrm{d}(x, y) \leq \mathrm{d}(x_j, x) + \mathrm{d}(x_j, y) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon;$$

here we used the triangle inequality and the symmetric of the metric d. Since $\varepsilon > 0$ was arbitrary, it follows that $\mathrm{d}(x, y) = 0$. So the positive definiteness of d implies $x = y$. □

Proposition 1.3.5 (Sequential completeness vs. net completeness)**.** *Let (M, d) be a metric space. The following are equivalent:*

(i) *The metric space (M, d) is complete, i.e. every Cauchy sequence in M converges.*

(ii) *Every Cauchy net in M converges.*

Proof. *"(ii) ⇒ (i)"* This implication is clear since every Cauchy sequence is also a Cauchy net.

"(i) ⇒ (ii)" Let $(x_j)_{j\in J}$ be a Cauchy net in M. Since J is directed we can then find a sequence of indices $j_1 \leq j_2 \leq j_3 \leq \dots$ in J such that $\mathrm{d}(x_j, x_{j_n}) \leq \frac{1}{n}$ for each $n \in \mathbb{N}$ and for all $j \geq j_n$. Then $(x_{j_n})_{n\in\mathbb{N}}$ is a Cauchy sequence in M and hence converges to a point $x \in M$ due to (i).

Note: In the lecture I gave a slightly different argument because I could not remember what precisely I had written here. Turns out that the argument here is a bit smoother.

Let us show that the net $(x_j)_{j\in J}$ converges to x. So fix $\varepsilon > 0$. Choose $m \in \mathbb{N}$ such that $\frac{2}{m} < \varepsilon$. For each $n \in \mathbb{N}$ with $n \geq m$ one has $\mathrm{d}(x_{j_m}, x_{j_n}) \leq \frac{1}{m}$, so $\mathrm{d}(x_{j_m}, x) \leq \frac{1}{m}$ since the metric is continuous. For every $j \in J$ with $j \geq j_m$ we thus obtain

$$\mathrm{d}(x_j, x) \leq \mathrm{d}(x_j, x_{j_m}) + \mathrm{d}(x_{j_m}, x) \leq \frac{1}{m} + \frac{1}{m} < \varepsilon.$$

This proves $x_j \to x$, as claimed. □

The non-trivial implication in the above proposition is, at first glance, a bit weird since the index sets of nets can be really huge compared to sequences (we shall see such examples later on). On the other hand, the implication is also very important since it shows that we do not need to care whether completeness of metric spaces is defined in terms if Cauchy sequences or in terms of Cauchy nets.

We will see in Chapter 4 that nets are extremely useful to study **topological spaces**, which are a very useful generalization of metric spaces. However, even in normed spaces, nets can be quite useful since they enable us to give a very natural definition of unconditional convergence of series.

Definition 1.3.6 (Convergence of series)**.** Let X be a normed space and $x \in X$. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X, let L be a set and let $(x_\ell)_{\ell\in L}$ be a family of vectors in X.

(a) The sequence $\left(\sum_{k=1}^n x_k\right)_{n\in\mathbb{N}}$ is called the **series** over $(x_n)_{n\in\mathbb{N}}$. To indicate that it converges to x, we write $\sum_{k=1}^\infty x_k = x$.

(b) The series $\left(\sum_{k=1}^n x_k\right)_{n\in\mathbb{N}}$ is called **absolutely convergent** if the series $\left(\sum_{k=1}^n \|x_k\|\right)_{n\in\mathbb{N}}$ converges in $\mathbb{R}$, i.e. if $\sum_{k=1}^\infty \|x_k\| < \infty$.

(c) Let $\mathcal{F}$ denote the set of all finite subsets of L, direct by the set inclusion $\subseteq$. The net $\left(\sum_{\ell\in F} x_\ell\right)_{F\in\mathcal{F}}$ is called the **unconditional series** over $(x_\ell)_{\ell\in L}$. To indicate that it converges to x, we write $\sum_{\ell\in L} x_\ell = x$.

(d) Let $\mathcal{F}$ be as in (d). We say that the unconditional series $\left(\sum_{\ell\in F} x_\ell\right)_{F\in\mathcal{F}}$ **converges absolutely** if the unconditional series $\left(\sum_{\ell\in F} \|x_\ell\|\right)_{F\in\mathcal{F}}$ converges in $\mathbb{R}$.

Lecture 5 (Mo, 27 April) is woll start with Definition 1.3.6(d).

For series with index set $\mathbb{N}$ – i.e. if the summands are a sequence $(x_n)_{n\in\mathbb{N}}$ – it is natural to wonder how convergence of the series over $(x_n)_{n\in\mathbb{N}}$ is related to convergence of the unconditional series over $(x_n)_{n\in\mathbb{N}}$:

Proposition 1.3.7 (Convergence of series)**.** *Let X be a normed space and $x \in X$, let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X, and let $x \in X$.*

(a) *If $\sum_{k\in\mathbb{N}} x_k = x$, then also $\sum_{n=1}^\infty x_k = x$.*

(b) *If $X = \mathbb{K}$ and $x_k \in [0,\infty)$ for all $k \in \mathbb{N}$, the converse implication in* (a) *also holds.*

Proof. (a) Assume that $\sum_{k\in\mathbb{N}} x_k = x$ and let $\varepsilon > 0$. Then there exists a finite set $F \subseteq \mathbb{N}$ such that $\left\|x - \sum_{k\in G} x_k\right\| < \varepsilon$ for all finite sets G with $F \subseteq G \subseteq \mathbb{N}$. By making F larger if necessary we may assume that $F \neq \emptyset$. Set $n_0 := \max F$. For every $n \in \mathbb{N}$ with $n \geq n_0$ we then have $F \subseteq \{1,\dots,n\}$ and thus, $\left\|x - \sum_{k=1}^n x_k\right\| = \left\|x - \sum_{k\in\{1,\dots,n\}} x_k\right\| < \varepsilon$. This proves that $\left(\sum_{k=1}^n x_k\right)_{n\in\mathbb{N}}$ converges to x, i.e. $\sum_{k=1}^\infty x_k = x$.

(b) Let $X = \mathbb{K}$, $x_k \in [0,\infty)$ for all $k \in \mathbb{N}$ and $\sum_{k=1}^\infty x_k = x$. Let $\varepsilon > 0$ and choose $n_0 \in \mathbb{N}$ such that $x - \sum_{k=1}^{n_0} x_k < \varepsilon$. For every finite set F that satisfies $\{1,\dots,n_0\} \subseteq F \subseteq \mathbb{N}$ the positivity of all summands implies

$$\sum_{k=1}^{n_0} x_k \leq \sum_{k\in F} x_k \leq \sum_{k=1}^{\max F} x_k \leq x,$$

so $|x - \sum_{k\in F} x_k| < \varepsilon$. Therefore, $\sum_{k\in\mathbb{N}} x_k = x$. □

Example 1.3.8 (An unconditional series in $\boldsymbol{\ell^p}$)**.** Let $p \in [1,\infty)$, abbreviate $\ell^p := \ell^p(\mathbb{N};\mathbb{K})$, and let $f \in \ell^p$. Let $\mathcal{F}$ denote the set of all finite subsets of $\mathbb{N}$, directed by set inclusion. The unconditional series

$$\Big(\sum_{k\in F} f_k e^{(k)}\Big)_{F\in\mathcal{F}}$$

converges to f (i.e. $f = \sum_{k\in\mathbb{N}} f_k e^{(k)}$). But if $p > 1$, there exists a vector $f \in \ell^p$ such that the series is not absolutely convergent.

Proof. Let $\varepsilon > 0$. Since $\sum_{k=1}^\infty |f_k|^p = \|f\|_p^p < \infty$, there exists a $k_0 \in \mathbb{N}$ such that $\sum_{k=k_0+1}^\infty |f_k|^p < \varepsilon^p$. Set $F_0 := \{1,\dots,k_0\} \in \mathcal{F}$. So for every $F \in \mathcal{F}$ with $F \supseteq F_0$ one has

$$\Big\|\sum_{k\in F} f_k e^{(k)} - f\Big\|_p \le \Big(\sum_{k=k_0+1}^{\infty} |f_k|_p\Big)^{1/p} \le \varepsilon,$$

which shows the convergence of the unconditional series to f.

Now let $p > 1$. The sequence f given by $f_n = \frac{1}{n}$ for every $n \in \mathbb{N}$. Then $f \in \ell^p$ since $\|f\|_p^p = \sum_{k=1}^\infty |f_k|^p = \sum_{k=1}^\infty \frac{1}{k^p} < \infty$, but $\sum_{k=1}^\infty \|f_k e^{(k)}\|_p = \sum_{k=1}^\infty |f_k| = \sum_{k=1}^\infty \frac{1}{k} = \infty$. So the series $\big(\sum_{k=1}^n \|f_k e^{(k)}\|_p\big)_{n\in\mathbb{N}}$ does not converge in $\mathbb{R}$ and hence, neither does the unconditional series $\big(\sum_{k\in F} \|f_k e^{(k)}\|_p\big)_{F\in\mathcal{F}}$ (Proposition 1.3.7(a)). □

We note in passing that if $p = 1$, then the unconditional series $\big(\sum_{k\in F} f_k e^{(k)}\big)_{F\in\mathcal{F}}$ in ℓ^1 is even absolutely convergent for every $f \in \ell^1$; more precisely, one has $\sum_{k\in\mathbb{N}} \|f_k e^{(k)}\|_1 = \|f\|_1$ as a consequence of Proposition 1.3.7(b).

One can characterize completeness of normed spaces by the property that every absolutely convergent series in X converges. More precisely, the following proposition holds. We already discussed the equivalence of (i) and (ii) in Analysis 3 [6, Proposition 3.2.17] and used it to prove the completeness of L^p-spaces [6, Theorem 3.2.16].

Proposition 1.3.9 (Completeness via series convergence)**.** *Let X be a normed space. The following assertions are equivalent:*

(i) *The space X is complete, i.e. a Banach space.*

(ii) *Every absolutely convergent series (with index set $\mathbb{N}$) in X converges.*

(iii) *Every absolutely convergent unconditional series (with arbitrary index set) in X converges.*

Proof. "(i) ⇒ (iii)" Assume that X is complete. Let L be a set, let $(x_\ell)_{\ell\in L}$ be family of vectors in X, and let $\mathcal{F}$ denote the set of all finite subsets of L, directed by set inclusion. Assume that the unconditional series $\big(\sum_{\ell\in F} x_\ell\big)_{F\in\mathcal{F}}$ is absolutely convergent, i.e. that the unconditional series $\big(\sum_{\ell\in F} \|x_\ell\|\big)_{F\in\mathcal{F}}$ converges in $\mathbb{R}$. Then one can readily check that $\big(\sum_{\ell\in F} x_\ell\big)_{F\in\mathcal{F}}$ is a Cauchy net in X, so it converges according to Proposition 1.3.5 since X is complete.

"(iii) ⇒ (ii)" Let (iii) hold. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X and assume that $\big(\sum_{k=1}^n x_k\big)_{n\in\mathbb{N}}$ is absolutely convergent, i.e. that the series $\big(\sum_{k=1}^n \|x_k\|\big)_{n\in\mathbb{N}}$ converges in $\mathbb{R}$. Let $\mathcal{F}$ denote the set of all finite subsets of $\mathbb{N}$, directed by set inclusion. It follows from Proposition 1.3.7(b) that the unconditional series $\big(\sum_{k\in F} \|x_k\|\big)_{F\in\mathcal{F}}$ converges in $\mathbb{R}$. Hence, the unconditional series $\big(\sum_{k\in F} x_k\big)_{F\in\mathcal{F}}$ in X converges according to (iii), i.e. there exists a point $x \in X$ such that $\sum_{k\in\mathbb{N}} x_k = x$. As shown in Proposition 1.3.7(a) this implies that $\sum_{k=1}^\infty x_k = x$, as claimed.

"(ii) ⇒ (i)" Assume (ii) and let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in X. We can find a strictly increasing sequence $n_1 < n_2 < n_3 < \dots$ in $\mathbb{N}$ such that $\|x_{n_{k+1}} - x_{n_k}\| \le \frac{1}{2^k}$ for all $k \in \mathbb{N}$. Define $y_1 := x_{n_1}$ and $y_k := x_{n_k} - x_{n_{k-1}}$ for all $k \ge 2$. Then $\sum_{k=1}^\infty \|y_k\| \le \|y_{n_1}\| + \sum_{k=2}^\infty \frac{1}{2^{k-1}} < \infty$, so it follows from (ii) that there exists a point $x \in X$ with $x_{n_k} = \sum_{\ell=1}^k y_\ell \to x$ as $k \to \infty$. So the subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ is convergent. But it is easy to check that every Cauchy sequence in a metric space that has a convergent subsequence is itself convergent. Hence, $(x_n)_{n\in\mathbb{N}}$ converges, so X is complete. □

1.4 Addendum: Spaces of signed measures

Addenda do not directly belong to the contents of the course, but contain additional information for interested readers.

Introductory questions.

(a) Consider a measurable space $(\Omega, \mathcal{A})$ and two finite measures $\mu, \nu \colon \mathcal{A} \to [0, \infty)$. Can we substract ν from μ? Why / why not?

(b) How would you mathematically describe a distribution of electrical charges? Does your suggestion include the possibility to describe a charge that is located at a single point?

(c) Can you write down three different bounded linear functionals on the Banach space $\mathrm{C}([0,1];\mathbb{R})$?

Consider a measurable space $(\Omega, \mathcal{A})$. The set of all, say finite, measures defined on $\mathcal{A}$ does not have a vector space structure: since measures are defined to take positive values only, one cannot multiply a measure by a negative number (or even a complex number). As is common in mathematics, if we cannot do something but would really like to do it, which just make it work by defining things in the "right" way. So let us now define measures that can take values in $\mathbb{R}$ (or even in $\mathbb{C}$).

Definition 1.4.1 (Scalar-valued measures)**.** Let $(\Omega, \mathcal{A})$ be a measurable space.[15]

(a) A map $\mu \colon \mathcal{A} \to \mathbb{K}$ is called a $\mathbb{K}$**-valued measure** if it is ***σ*-additive**, i.e. if it satisfies.

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

[15] In German: Messraum.

for every sequence $A_1, A_2, \dots \in \mathcal{A}$ of pairwise disjoint measurable sets. Here, the series is understood to be a limit of an unconditional series in the sense of Definition 1.3.6(d).[16]

We denote the set of all $\mathbb{K}$-valued measures on $(\Omega, \mathcal{A})$ by $\mathcal{M}(\Omega, \mathcal{A}; \mathbb{K})$.

An $\mathbb{R}$-valued measure is also called a **signed measure**.

(b) Let $\mu, \nu \in \mathcal{M}(\Omega, \mathcal{A}; \mathbb{K})$ and $\alpha \in \mathbb{K}$. Then we define $\mu + \nu \colon \mathcal{A} \to \mathbb{K}$ and $\alpha\mu \colon \mathcal{A} \to \mathbb{K}$ by

$$(\mu + \nu)(A) := \mu(A) + \nu(A) \qquad \text{and} \qquad (\alpha\mu)(A) := \alpha \cdot \mu(A)$$

for all $A \in \mathcal{A}$.

(c) Let $\mu, \nu \in \mathcal{M}(\Omega, \mathcal{A}; \mathbb{R})$. We write $\mu \le \nu$ if $\mu(A) \le \nu(A)$ for all $A \in \mathcal{A}$. The notation $\mu \ge \nu$ is defined analogously.

It is not difficult to check that, for a measurable space $(\Omega, \mathcal{A})$, $\mathbb{K}$-valued measures $\mu, \nu \in \mathcal{M}(\Omega, \mathcal{A}; \mathbb{K})$, and a scalar $\alpha \in \mathbb{K}$, the mappings $\mu + \nu$ and $\alpha\mu$ are $\mathbb{K}$-valued measures again. Moreover, this addition and scalar multiplication turns $\mathcal{M}(\Omega, \mathcal{A}; \mathbb{K})$ into a vector space over $\mathbb{K}$ whose zero element is the measure 0 that maps every set to $0 \in \mathbb{K}$.

At first glance one might wonder why the requirement $\mu(\emptyset) = 0$ does not occur in the definition of a $\mathbb{K}$-valued measure. However, since μ takes values in $\mathbb{K}$ (and thus cannot take the value ∞) this is actually a consequence of the σ-additivity:

Proposition 1.4.2 ($\mathbb{K}$-valued measures map $\emptyset$ to **0**). *Let $(\Omega, \mathcal{A})$ be a measurable space and $\mu \colon \mathcal{A} \to \mathbb{K}$ a $\mathbb{K}$-valued measure. Then $\mu(\emptyset) = 0$.*

Proof. In the σ-additivity property of μ (Definition 1.4.1(a)), choose all the sets A_n to be equal to $\emptyset$. It then follows from Proposition 1.3.7(a) that

$$\mathbb{K} \ni \mu(\emptyset) = \sum_{n \in \mathbb{N}} \mu(\emptyset) = \lim_{n \to \infty} \sum_{k=1}^{n} \mu(\emptyset) = \lim_{n \to \infty} n\mu(\emptyset).$$

This implies that $\mu(\emptyset) = 0$. □

The easiest way to construct $\mathbb{K}$-valued measures it to start with finite measures[17] and take linear combinations of them with coefficients from $\mathbb{K}$. One can even show that all $\mathbb{K}$-valued measures can be obtained in this way. For $\mathbb{R} = \mathbb{K}$ this is a consequence of Theorem 1.4.4 below. To keep this section concise and focused on the essential ideas, we refrain from also discussing the $\mathbb{C}$-valued case in detail.

[16] Note that this is very natural for the following reason. The union on the left hand side of the equation does obviously not depend on how the sets $A_1, A_2, \dots$ are enumerated. So even if we only required $\mu\big(\bigcup_{n\in\mathbb{N}} A_n\big) = \sum_{n=1}^{\infty} \mu(A_n)$ instead, this would automatically imply $\mu\big(\bigcup_{n\in\mathbb{N}} A_n\big) = \sum_{n=1}^{\infty} \mu(A_{\varphi(n)})$ for every bijection $\varphi \colon \mathbb{N} \to \mathbb{N}$. And according to Bonus Exercise 3.5(a) on Homework Sheet 3, this is equivalent to $\mu\big(\bigcup_{n\in\mathbb{N}} A_n\big) = \sum_{n\in\mathbb{N}} \mu(A_n)$.

[17] When we say **measure** instead of $\mathbb{K}$-**valued measure** we always mean a measure that takes values in $[0, \infty]$. Of course, a measure is finite if and only if it takes values in $[0, \infty)$ only.

Definition 1.4.3 (The total variation measure)**.** Let $(\Omega, \mathcal{A})$ be a measurable space and let $\mu\colon \mathcal{A} \to \mathbb{K}$ be an $\mathbb{R}$-valued measure.

(a) The map $|\mu| : \mathcal{A} \to [0, \infty]$,

$$|\mu|(A) := \sup\left\{ |\mu(B_1)| + |\mu(B_2)| \mid B_1, B_2 \in \mathcal{A},\ B_1 \cap B_2 = \emptyset,\ B_1 \cup B_2 = A \right\}$$

is called the **total variation measure of μ**.

(b) We call $\|\mu\| := \|\mu\|_{\mathrm{TV}} := |\mu|(\Omega)$ is called the **total variation norm of μ**.

As indicated before the definition, one can also define the total variation measure and the total variation norm for $\mathbb{C}$-valued measures (the formula of $|\mu|$ is a bit more involved then). To keep it brief we restrict ourselves to the case of real-valued measures, where many interesting ideas already occur.

Theorem 1.4.4 (Total variation measures and the space of measures)**.** *Let $(\Omega, \mathcal{A})$ be a measurable space.*

(a) *Let $\mu\colon \mathcal{A} \to \mathbb{R}$ be an $\mathbb{R}$-valued measure. For every $A \in \mathcal{A}$ one has*

$$|\mu|(A) := \sup\left\{ \mu(B_1) - \mu(B_2) \mid B_1, B_2 \in \mathcal{A},\ B_1 \cap B_2 = \emptyset,\ B_1 \cup B_2 = A \right\}$$

(b) *Let $\mu\colon \mathcal{A} \to \mathbb{R}$ be an $\mathbb{R}$-valued measure. Then the total variation measure $|\mu|$ is indeed a measure on $(\Omega, \mathcal{A})$ and it is finite, i.e. $\|\mu\| = |\mu|(\Omega) < \infty$.*

(c) *For all $\mathbb{R}$-valued measures $\mu, \nu\colon \mathcal{A} \to \mathbb{R}$ and all $\alpha \in \mathbb{R}$ one has*

$$\pm\mu \le |\mu|, \qquad |\mu + \nu| \le |\mu| + |\nu| \qquad \text{and} \qquad |\alpha\mu| = |\alpha|\,|\mu|.$$

Moreover, $|\mu| = 0$ if and only if $\mu = 0$.

(d) *Let $\mu\colon \mathcal{A} \to \mathbb{R}$ be a $\mathbb{R}$-valued measure. Then $\mu \ge 0$ if and only if $|\mu| = \mu$. In particular, $\|\mu\| = \mu(\Omega)$ if $\mu \ge 0$.*

(e) *The total variation norm is a norm on $\mathcal{M}(\Omega, \mathcal{A}; \mathbb{R})$ and turns $\mathcal{M}(\Omega, \mathcal{A}; \mathbb{R})$ into a Banach space.*

The proof of part (b) of the Theorem is a bit easier to write down if one uses the following lemma, which is nice to know anyway.

Lemma 1.4.5 (Characterization of measures)**.** *Let $(\Omega, \mathcal{A})$ be a measurable spaces. Let $\mu\colon \mathcal{A} \to [0, \infty]$ satisfy $\mu(\emptyset) = 0$. Then μ is a measure if and only if it satisfies the following properties:*

(1) *The map μ is σ-subadditive on disjoint sets, i.e. for every sequence $A_1, A_2, \dots \in \mathcal{A}$ of pairwise disjoint sets one has $\mu\big(\bigcup_{n=1}^\infty A_n\big) \le \sum_{n=1}^\infty \mu(A_n)$.*

(2) *The map μ is finitely superadditive on disjoint sets, i.e. one has $\mu(A \cup B) \geq \mu(A) + \mu(B)$ für all disjoint sets $A, B \in \mathcal{A}$.*

Proof. If μ is a measure, then it clearly satisfies (1) and (2). So assume now conversely that μ satisfies (1) and (2). We need to show that μ is σ-additive, so let $A_1, A_2 \dots \in \mathcal{A}$ be pairwise disjoint.

Due to (1) it only remains to show that $\mu\big(\bigcup_{n=1}^\infty A_n\big) \geq \sum_{n=1}^\infty \mu(A_n)$. To this end, let us first observe that μ is monotone: if $A, B \in \mathcal{A}$ satisfy $A \subseteq B$, then it follows from (2) and from the fact that μ takes values in $[0, \infty]$ that that $\mu(A) \leq \mu(A) + \mu(B \setminus A) \leq \mu(B)$.

By applying the mononicity and again (2), it follows for each $N \in \mathbb{N}$ that

$$\mu\big(\bigcup_{n=1}^\infty A_n\big) \geq \mu\big(\bigcup_{n=1}^N A_n\big) \geq \sum_{n=1}^N \mu(A_n).$$

Since the left hand side does not depend on N, it dominates $\sum_{n=1}^\infty \mu(A_n)$, as claimed. □

Proof of Theorem 1.4.4. (a) Fix a measurable set $A \in \mathcal{A}$ and set

$$\alpha := \sup\Big\{\mu(B_1) - \mu(B_2) \mid B_1, B_2 \in \mathcal{A},\ B_1 \cap B_2 = \emptyset,\ B_1 \cup B_2 = A\Big\}$$

The inequality $|A| \geq \alpha$ is clear. To prove the converse inequality $|A| \leq \alpha$, let $B_1, B_2 \in \mathcal{A}$ be disjoint and satisfy $B_1 \cup B_2$. We need to show that $\big|\mu(B_1)\big| + \big|\mu(B_2)\big| \leq \alpha$, and we do this by distinguishing four cases:

Case 1: $\mu(B_1) \geq 0$ and $\mu(B_2) \leq 0$.
In this case one has $\big|\mu(B_1)\big| + \big|\mu(B_2)\big| = \mu(B_1) - \mu(B_2) \leq \alpha$.

Case 2: $\mu(B_1) \leq 0$ and $\mu(B_2) \geq 0$.
In this case we have $\big|\mu(B_1)\big| + \big|\mu(B_2)\big| = \mu(B_2) - \mu(B_1) \leq \alpha$.

Case 3: $\mu(B_1) \geq 0$ and $\mu(B_2) \geq 0$.
In this case one has $\big|\mu(B_1)\big| + \big|\mu(B_2)\big| = \mu(A) - \mu(\emptyset) \leq \alpha$.

Case 4: $\mu(B_1) \leq 0$ and $\mu(B_2) \leq 0$.
In this case we have $\big|\mu(B_1)\big| + \big|\mu(B_2)\big| = \mu(\emptyset) - \mu(A) \leq \alpha$.

(b) Obviously, $\big|\mu\big|$ maps into $[0, \infty]$. Next we show that it is a measure.

It follows from $\mu(\emptyset) = 0$ (Proposition 1.4.2) that $\big|\mu\big|(\emptyset) = 0$. The fact that $\big|\mu\big|$ is σ-subadditive on disjoint sets is straightforward to prove from the definition of $\big|\mu\big|$. According to Lemma 1.4.5 we thus need to show that μ is finitely superadditive on disjoint sets.

To this end we use part (a) of the theorem. Let $A, \tilde{A} \in \mathcal{A}$ be disjoint. Also let $B_1, B_2 \in \mathcal{A}$ be disjoint and satisfy $B_1 \cup B_2 = A$ and let $\tilde{B}_1, \tilde{B}_2 \in \mathcal{A}$ be disjoint and satisfy $\tilde{B}_1 \cup \tilde{B}_2 = \tilde{A}$. Then[18]

$$\big(\mu(B_1) - \mu(B_2)\big) + \big(\mu(\tilde{B}_1) - \mu(\tilde{B}_2)\big) = \mu(B_1 \cup \tilde{B}_1) - \mu(B_2 \cup \tilde{B}_2) \leq \big|\mu\big|(A \cup \tilde{A}),$$

[18] Note that the first equality in this computation would, by the triangle inequality for the modulus in $\mathbb{R}$, become the inequality $\geq$ if we used directly the definition of $\big|\mu\big|$. But we need the converse inequality here; this is why part (a) of the proposition is useful here.

where the inequality holds since the two sets $B_1 \cup \tilde{B}_1$ and $B_2 \cup \tilde{B}_2$ are a measurable partition of $A \cup \tilde{A}$. By first taking the supremum over $B_1, \tilde{B}_1$ and then over $B_2, \tilde{B}_2$ we obtain $|\mu|(A) + |\mu|(\tilde{A}) \le |\mu|(A \cup \tilde{A})$, as claimed.

Finally, let us show that $|\mu|(\Omega) < \infty$. To this end, we first prove the following claim:

$(*)$ If $A \in \mathcal{A}$ satisfies $|\mu|(A) = \infty$, then there exist two disjoint measurable sets B_1, B_2 with $B_1 \cup B_2 = A$ that $|\mu(B_1)| \ge 1$ and $|\mu(B_2)| = \infty$.[19,20]

Indeed, let $|\mu|(A) = \infty$. Since $\mu(A) < \infty$ the definition of $|\mu|(A)$ implies that there exists disjoint measurable set B_1, B_2 such that $B_1 \cup B_2 = A$ and $|\mu(B_1)| + |\mu(B_2)| \ge 2\big(1 + |\mu(A)|\big)$. Then at least one of the summands on the left, say $|\mu(B_1)|$, satisfies $|\mu(B_1)| \ge 1 + |\mu(A)|$. So on the one hand, $|\mu(B_1)| \ge 1$, and on the other hand

$$|\mu(A)| = |\mu(B_1) + \mu(B_2)| \ge |\mu(B_1)| - |\mu(B_2)| \ge 1 + |\mu(A)| - |\mu(B_2)|,$$

so also $|\mu(B_2)| \ge 1$. Since we have already shown that $|\mu|$ is a measure, we also note that $\infty = |\mu|(A) = |\mu|(B_1) + |\mu|(B_2)$, so at least one of the values $|\mu|(B_1)$ and $|\mu|(B_2)$ is ∞. This proves the implication $(*)$.

Now assume for a contradiction that $|\mu|(\Omega) = \infty$. Then $(*)$ implies that we can find a measurable set $B_1 \subseteq \Omega$ such that $|\mu(B_1)| \ge 1$ and $|\mu|(\Omega \setminus B_1) = \infty$. Now we apply $(*)$ to the set $\Omega \setminus B_1$ to find a measurable set $B_2 \subseteq \Omega \setminus B_1$ such that $|\mu(B_2)| \ge 1$ and $|\mu|\big(\Omega \setminus (B_1 \cup B_2)\big) = \infty$. By iterating this we find a sequence $(B_n)_{n\in\mathbb{N}}$ of pairwise disjoint, measurable sets such that $|\mu(B_n)| \ge 1$ for every $n \in \mathbb{N}$. Hence, the net $\big(\sum_{n\in F} \mu(B_n)\big)_F$, where F runs through all finite subsets of $\mathbb{N}$, cannot be a Cauchy net and hence, it does not converge to a real number. But we know from the σ-additivity of μ that the net actually converges to $\mu\big(\bigcup_{n\in\mathbb{N}} \mu(B_n)\big)$; so we arrived at a contradiction.

(d) The inequality $\pm\mu \le |\mu|$, the subadditivity and the absolute homogeneity can easily be obtained from the definition of the total variation measure.

If $\mu = 0$, then it is also immediate from the definition of $|\mu|$ that $|\mu| = 0$. Conversely, if $|\mu| = 0$, then $\pm\mu \le 0$, which implies $\mu = 0$.

(d) If $\mu \ge 0$, it follows right from the definition of $|\mu|$ that $|\mu| = \mu$, and this in turn gives $\|\mu\| = |\mu|(\Omega) = \mu(\Omega)$.

So assume conversely that $|\mu| = \mu$. For every $B \in \mathcal{A}$ we then conclude that

$$\mu(B) + \mu(B^{\complement}) = \mu(\Omega) = |\mu|(\Omega) \ge |\mu(B)| + \left|\mu(B^{\complement})\right| \ge \mu(B^{\complement}),$$

so $\mu(B) \ge 0$. Hence, $\mu \ge 0$.

[19] Carefully note that the first modulus is taken of μ before substituting B_1, while the second modulus is taken simply of the number $\mu(B_2)$.

[20] Note that, as we are going to show $|\mu|(\Omega) < \infty$, there does actually not exists a set $A \in \mathcal{A}$ with $|\mu|(A) < \infty$. So this implication is purely hypothetical; we will use it in a contradiction argument to obtain $|\mu|(\Omega) < \infty$.

(e) Let us first show that the total variation norm $\|\cdot\|$ satisfies the norm axioms.

Let $\mu, \nu \in \mathcal{M}(\Omega, \mathcal{A}; \mathbb{R})$ and $\alpha \in \mathbb{R}$. If $\|\mu\| = 0$, then $|\mu|(\Omega) = 0$, so $|\mu| = 0$, and hence $\mu = 0$ according to (d). Moreover, it follows from (d) that

$$\|\mu + \nu\| = |\mu + \nu|(\Omega) \leq \big(|\mu| + |\nu|\big)(\Omega) = |\mu|(\Omega) + |\nu|(\Omega) = \|\mu\| + \|\nu\|$$

and

$$\|\alpha\mu\| = |\alpha\mu|(\Omega) = \big(|\alpha\mu|\big)(\Omega) = |\alpha|\,|\mu|(\Omega) = |\alpha|\,\|\mu\|.$$

So $\|\cdot\|$ is indeed a norm.

To show the completeness we use Proposition 1.3.9: let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{M}(\Omega, \mathcal{A}; \mathbb{R})$ such that $\sum_{n=1}^{\infty} \|\mu_n\| < \infty$. We need to show that the series $\big(\sum_{n=1}^{N} \mu_n\big)_{N\in\mathbb{N}}$ converges in $\mathcal{M}(\Omega, \mathcal{A}; \mathbb{K})$. To this end we proceed in two steps.

Step 1: We show the claim if $\mu_n \geq 0$ for each n.

In this case, define $\mu\colon \mathcal{A} \to [0, \infty]$ by setting $\mu(A) := \sum_{n=1}^{\infty} \mu_n(A)$ for all $A \in \mathcal{A}$.

We note that μ is a measure. Indeed, one clearly has $\mu(\emptyset) = 0$. And if $A_1, A_2, \dots \in \mathcal{A}$ are pairwise disjoint, then

$$\begin{aligned} \mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) &= \sum_{n=1}^{\infty} \mu_n\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mu_n(A_k) \\ &= \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \mu_n(A_k) = \sum_{k=1}^{\infty} \mu(A_k), \end{aligned}$$

where the equality between both lines follows, e.g., by applying Tonelli's theorem to the product measure $\# \otimes \#$, where $\#$ denotes the counting measure on $\mathbb{N}$.

Next we observe that the measure μ is finite. Indeed,

$$\mu(\Omega) = \sum_{n=1}^{\infty} \mu_n(\Omega) = \sum_{n=1} \|\mu_n\| < \infty$$

by assumption; the second equality uses the positivity of the μ_n.

Finally, we show that the sequence $\big(\sum_{n=1}^{N} \mu_n\big)_{N\in\mathbb{N}}$ converges to μ with respect to the total variation norm $\|\cdot\|$. Since all μ_n are positive, it follows for each $N \in \mathbb{N}$ that the measure $\mu - \sum_{n=1}^{N} \mu_n$ is positive. Thus,

$$\Big\|\mu - \sum_{n=1}^{N} \mu_n\Big\| = \big(\mu - \sum_{n=1}^{N} \mu_n\big)(\Omega) = \sum_{n=N+1}^{\infty} \mu_n(\Omega) = \sum_{n=N+1}^{\infty} \|\mu_n\| \to 0$$

as $N \to \infty$ since $\sum_{n=1}^{\infty} \|\mu_n\| < \infty$.

Step 2: We show the claim for general μ_n.

In this case we set $\nu_n := |\mu_n| - \mu_n \geq 0$ for each $n \in \mathbb{N}$. One has $\big\||\mu_n|\big\| = \|\mu_n\|$ and $\|\nu_n\| \leq 2\|\mu_n\|$ for each n. Hence, $\sum_{n=1}^{\infty} \big\||\mu_n|\big\| < \infty$ and $\sum_{n=1}^{\infty} \|\nu_n\| < \infty$. Thus, we can apply Step 1 to find measures $0 \le \lambda, \nu \in \mathcal{M}(\Omega, \mathcal{A}; \mathbb{R})$ such that $\sum_{n=1}^{N} |\mu_n| \to \lambda$ and $\sum_{n=1}^{N} \nu_n \to \nu$ as $N \to \infty$. Thus,

$$\sum_{n=1}^{N} \mu_n = \sum_{n=1}^{N} |\mu_n| - \sum_{n=1}^{n} \nu_\nu \overset{N\to\infty}{\to} \lambda - \nu. \qquad \square$$

Of course, the major purpose of a measure in analysis is to integrate functions against it. Since you already know from Analysis 3 how to integrate against positive measures, it is not easy to integration against $\mathbb{R}$-valued measures, too:

Definition 1.4.6 (Integration against signed measures)**.** Let $(\Omega, \mathcal{A})$ be a measurable space.

(a) For each signed measure $\mu \in \mathcal{M}(\Omega, \mathcal{A}; \mathbb{R})$ we define $0 \le \mu^+, \mu^- \in \mathcal{M}(\Omega, \mathcal{A}; \mathbb{R})$ as

$$\mu^+ := \frac{1}{2}(|\mu| + \mu) \qquad \text{and} \qquad \mu^- := \frac{1}{2}(|\mu| - \mu).$$

They are called the **positive part** and the **negative part** of μ.

(b) Let $f := \Omega \to \mathbb{R}$. We call f **integrable with respect to μ** if f is integrable with respect to μ^+ and μ^-.[21] In this case we define

$$\int_\Omega f \, \mathrm{d}\mu = \int_\Omega f \, \mathrm{d}\mu^+ - \int_\Omega f \, \mathrm{d}\mu^-.$$

It is not difficult to check that, in the situation of Definition 1.4.6(b), integrability of f with respect to μ^+ and μ^- is equivalent to integability of f with respect to $|\mu|$.

Spaces of signed measures can be used to represent the dual space of the space of continuous functions over a compact metric space. This is the famous Riesz representation theorem which we state now, although without proving it in this course.

Theorem 1.4.7 (The Riesz representation theorem for the dual space of $\mathbf{C}(M;\mathbb{R})$)**.** *Let (M, d) be a compact metric space and endow M with the Borel-σ-algebra $\mathcal{B}(M)$. The mapping $\Phi\colon \mathcal{M}(\Omega, \mathcal{B}(M); \mathbb{R}) \to \mathrm{C}(M;\mathbb{R})'$ given by*

$$\langle \Phi(\mu), f \rangle := \int_M f \, \mathrm{d}\mu \in \mathbb{R}$$

for all $\mu \in \mathcal{M}(\Omega, \mathcal{A}; \mathbb{R})$ and all $f \in \mathrm{C}(M;\mathbb{R})$, is an isometric isomorphism of normed spaces.

For a proof we refer, for instance, to [4, Korollar VIII.2.6].

[21] Note that this includes the condition that f be measurable.

Chapter 2

The Principles of Functional Analysis

2.1 Baire's theorem revisited

Introductory questions.

(a) Can every vector space over $\mathbb{K}$ be turned into a Banach space by choosing an appropriate norm?

(b) Is there a continuous function $f\colon [0,1] \to \mathbb{R}$ that is not differentiable at any point?

We briefly repeat the following two notions that you already know from the Analysis 3 course, see [6, Definition 1.5.6]

Definition 2.1.1 (Nowhere dense and meagre sets)**.** Let (M, d) be a metric space and $S \subseteq M$.

(a) The set S is said to be **nowhere dense** if its closure $\overline{S}$ has empty interior.

(b) The set S is set to be **meagre** if it can be written as a union of countably many subsets of M that are nowhere dense.

Recall that every countable union of meagre sets is again meagre, and that every subset of a meagre set is also meagre [6, Proposition 1.5.8]. Moreover, we proved that following equivalence in [6, Proposition 1.5.9]:

Proposition 2.1.2 (Characterisation of the Baire property)**.** *Let (M, d) be a metric space. The following assertions are equivalent.*

(i) *Every meagre subset of M has empty interior.*

(ii) *The complement of every meagre subset of M is dense in M.*

(iii) *If $C_1, C_2, \dots$ are closed subsets of M with empty interior, then $\bigcup_{n\in\mathbb{N}} C_n$ still has empty interior.*

(iv) *If $U_1, U_2, \dots$ are dense and open subsets of M, then $\bigcap_{n\in\mathbb{N}} U_n$ is still dense.*

Definition 2.1.3 (Baire spaces)**.** A metric space (M, d) is called a **Baire space** if it satisfies one (hence all) of the equivalent conditions in Proposition 2.1.2.

A very important class of Baire spaces are the complete metric spaces. We proved this in [6, Theorem 1.5.12] and state it again explicitly in the following theorem.

Theorem 2.1.4 (Complete Metric spaces are Baire)**.** *Every complete metric space* (M, d) *is a Baire space.*

Corollary 2.1.5 (Banach spaces with a countable basis)**.** *If X is an infinite-dimensional Banach space, then it does not have a countable basis.*

Proof. Let X be an infinite-dimensional Banach space and assume to the contrary that it has a countable basis $(x_n)_{n\in\mathbb{N}}$. Set $X_n := \operatorname{span}\{x_1, \dots, x_n\}$ for each $n \in \mathbb{N}$. Then $\bigcup_{n\in\mathbb{N}} X_n = X$.

But each of each vector subspaces X_n of X is finite-dimensional and thus closed according to Example 1.1.2. Since X is a Baire space by Theorem 2.1.4 it follows that it least one of the spaces X_n, say X_{n_0}, has non-empty interior.

But it is easy to check that a vector subspace with non-empty interior is always equal to the entire space, so $X_{n_0} = X$. Thus, X is finite-dimensional, a contradiction. $\square$

Example 2.1.6 (The space c_{00})**.** The space $c_{00}(\mathbb{N};\mathbb{K})$ is infinite dimensional and has a countable basis (consisting of the canonical unit vectors) Hence, it follows from Corollary 2.1.5 that there is no norm on $c_{00}(\mathbb{N};\mathbb{K})$ that turns this space into a Banach space.

2.2 The uniform boundedness principle

Lecture 6 (Th, 30 April) is planned to start with Section 2.2.

Introductory questions.

(a) Give an example of sequence $(f_n)_{n\in\mathbb{N}}$ in $C([0,1];\mathbb{R})$ such that $\sup_{n\in\mathbb{N}} \left|f_n(x)\right| < \infty$ for every $x \in [0,1]$ but $\sup_{n\in\mathbb{N}} \left\|f_n\right\|_\infty = \infty$.

(b) Let $(A_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{C}^{3\times 3}$ and assume that $\sup_{n\in\mathbb{N}} \|A_n x\|_1 < \infty$ for every $x \in \mathbb{C}^3$. Does it follows that $\sup_{n\in\mathbb{N}} \|A_n\|_{1\leftarrow 1} < \infty$?

Proposition 2.2.1 (Uniform boundedness of matrices)**.** *Let $\mathcal{T} \subseteq \mathbb{K}^{d\times d}$ and assume that $\sup_{T\in\mathcal{T}} \|Tx\|_{1\leftarrow 1} < \infty$ for each $x \in \mathbb{K}^d$. Then the set $\mathcal{T}$ is bounded in operator norm, i.e.* $\sup_{T\in\mathcal{T}} \|T\|_{1\leftarrow 1} < \infty$.

We note that the specific choice of the norms in the proposition is not important as all norms on $\mathbb{K}^d$ are equivalent [5, Theorem 1.7.2].[1] The reason why we endowed $\mathbb{K}^d$ with the 1-norm in the proposition is that it makes the computation in the following proof particularly simple to write down.

[1] We will come back to the notion **equivalence of norms** later in this functional analysis course, and discuss it in more detail then.

Proof of Proposition 2.2.1. For every $k \in \{1,\dots,d\}$ we let $e_k \in \mathbb{R}^d$ denote the kth canonical unit vector. Set

$$M := \max\bigl\{\sup_{T\in\mathcal{T}} \|Te_1\|_{1\leftarrow 1}, \dots, \sup_{T\in\mathcal{T}} \|Te_d\|_{1\leftarrow 1}\bigr\}.$$

Then $M < \infty$ by assumption. We are going to show that $\|T\|_{1\leftarrow 1} \le M$ for each $T \in \mathcal{T}$. To this end, fix $T \in \mathcal{T}$ and a vector $x \in \mathbb{K}^d$ of norm $\|x\|_1 \le 1$. Then

$$Tx = T\sum_{k=1}^d x_k e_k = \sum_{k=1}^d x_k Te_k, \qquad \text{hence} \qquad \|Tx\| \le \sum_{k=1}^d |x_k|\,\|Te_k\| \le M \underbrace{\sum_{k=1}^d |x_k|}_{=\|x\|_1 \le 1} \le M.$$

So indeed $\|T\|_{1\leftarrow 1} \le M$, as claimed. □

Observe that the proof of Proposition 2.2.1 relies very heavily on the finite-dimensional structure of $\mathbb{K}^d$. It is all the more surprising that the assertions actually remains true on general Banach spaces. This is the **uniform boundedness principle**, one of the fundamental principles of functional analysis:

Theorem 2.2.2 (Uniform boundedness principle)**.** *Let X, Y be normed spaces and assume that X is a Banach space. Let $\mathcal{T} \subseteq \mathcal{L}(X;Y)$ and assume that* $\sup_{T\in\mathcal{T}} \|Tx\| < \infty$ *for every $x \in X$. The $\mathcal{T}$ is bounded in operator norm, i.e.* $\sup_{T\in\mathcal{T}} \|T\| < \infty$.

Proof. For each $n \in \mathbb{N}$ we set

$$X_n := \bigl\{x \in X \mid \|Tx\| \le n \text{ for all } T \in \mathcal{T}\bigr\}.$$

Then X_n is a closed subset of X since all operators $T \in \mathcal{T}$ are bounded. (Note however that X_n will not be a vector subspace of X unless in very special cases.)

For every $x \in X$ the assumption $\sup_{T\in\mathcal{T}} \|Tx\|$ implies that there exists a number $n \in \mathbb{N}$ such that $\|Tx\| \le n$ for all $T \in \mathcal{T}$, i.e. $x \in X_n$. This shows that $\bigcup_{n\in\mathbb{N}} X_n = X$. Since X is complete, it is a Baire space by Theorem 2.1.4; so at least one of the sets X_n, say X_{n_0} has non-empty interior.

This means that there is a point $x_0 \in X_{n_0}$ and a radius $r > 0$ such that $B_{\le r}(x_0) \subseteq X_{n_0}$. Set $s := \sup_{T\in\mathcal{T}} \|Tx_0\|$ and note that $s < \infty$ by assumption. For every $x \in X$ of norm $\|x\| \le 1$ one has $x_0 + rx \in X_{n_0}$, and thus

$$\|Tx\| = \frac{1}{r}\|T(x_0 + rx - x_0)\| \le \frac{\|T(x_0+rx)\| + \|Tx_0\|}{r} \le \frac{n_0 + s}{r} =: M$$

The number M does not depend on x nor on T, and it follows that $\|T\| \le M$ for all $T \in \mathcal{T}$, which proves the claim. □

Definition 2.2.3 (Strong convergence)**.** Let X, Y be normed spaces and let $(T_j)_{j\in J}$ be a net of operators in $\mathcal{L}(X;Y)$. Also let $T \in \mathcal{L}(X;Y)$. We say that the net $(T_j)_{j\in J}$ **converges strongly** to T if the net $(T_j x)_{j\in J}$ converges to Tx for each $x \in X$.

Proposition 2.2.4 (Continuity of strong limits)**.** *Let X, Y be normed spaces and let $(T_j)_{j\in J}$ be a net of operators in $\mathcal{L}(X;Y)$ that is bounded, i.e. that satisfies $M := \sup_{j\in J} \|T_j\| < \infty$.*

Assume that, for each $x \in X$, the net $(T_j x)_{j\in J}$ converges to a vector in Y and define $T\colon X \to Y$ by $Tx := \lim_j T_j x$ for every $x \in X$. Then $T \in \mathcal{L}(X;Y)$ and $(T_j)_{j\in J}$ converges strongly to T.

Proof. It is straightforward to check that T is linear. Moreover, for every $x \in X$ one has

$$\|Tx\| = \lim_j \underbrace{\|T_j x\|}_{\le M\|x\|} \le M\|x\|,$$

where the equality follows from the continuity of the norm. Hence, T is bounded with $\|T\| \le M$. The strong convergence of $(T_j)_{j\in J}$ to T follows right from the definition of T. □

Corollary 2.2.5 (Banach–Steinhaus)**.** *Let X, Y be normed spaces and assume that X is a Banach space. Let $(T_n)_{n\in\mathbb{N}}$ be a sequence of bounded linear operator in $\mathcal{L}(X;Y)$ and assume that, for each $x \in X$, the sequence $(T_n x)_{n\in\mathbb{N}}$ converges to a vector in X. Define $T\colon X \to Y$ by $Tx := \lim_n T_n x$ for each $x \in X$. Then $T \in \mathcal{L}(X;Y)$ and $(T_n)_{n\in\mathbb{N}}$ converges strongly to T.*

Proof. For every $x \in X$ the sequence $(T_n x)_{n\in\mathbb{N}}$ in Y is convergent by assumption and is thus bounded. So the uniform boundedness principle (Theorem 2.2.2) implies that the sequence $(T_n)_{n\in\mathbb{N}}$ is bounded in operator norm. This, this claim follows from Proposition 2.2.4. □

Corollary 2.2.5 does not hold for nets $(T_j)_{j\in J}$ rather than sequences. Here is what goes wrong in the proof: for $x \in X$, the convergent net $(T_j x)_{j\in J}$ in Y might not be bounded, in general. Instead, the net $(T_j x)_{j\in J}$ is only eventually bounded in general – but the index from which on it is bounded might depend on x, so one cannot apply the uniform boundedness principle. We will discuss the existence of a counterexample in the exercises.

2.3 The open mapping theorem and the closed graph theorem

Introductory questions. As usual, endow $c_{00}(\mathbb{N};\mathbb{R})$ and $c_0(\mathbb{N};\mathbb{R})$ with the norm $\|\cdot\|_\infty$.

(a) Can you find an example of a bijective linear operator $T\colon c_{00}(\mathbb{N};\mathbb{R}) \to c_{00}(\mathbb{N};\mathbb{R})$ such that T is bounded but T^{-1} is not bounded?

(b) Can you find an example of a bijective linear operator $T\colon c_0(\mathbb{N};\mathbb{R}) \to c_0(\mathbb{N};\mathbb{R})$ such that T is bounded but T^{-1} is not bounded?

Definition 2.3.1 (Open maps)**.** Let (M, d) and (N, d) be metric spaces. A map $f\colon M \to N$ is called **open** if the image $f(U)$ is open in N for ever open set $U \subseteq M$.

Our first goal of this section is to show that every continuous linear operator between two Banach spaces is open (Theorem 2.3.3). To this end, we first show the following auxiliary result.

Lemma 2.3.2 (The little open mapping theorem)**.** *Let X, Y be normed spaces over $\mathbb{K}$, where X is a Banach space. Let $T \in \mathcal{L}(X;Y)$, let $r, s \in (0,\infty)$ and $\overline{TB_{<r}(0)} \supseteq B_{<s}(0)$. Then even $TB_{<r}(0) \supseteq B_{<s}(0)$.*

Proof. There is no loss of generality in assuming $s = 1$. Let $y \in Y$ satisfy $y \in B_{<1}(0)$. We have to find $x \in X$ such that $Tx = y$ and $x \in B_{<r}(0)$. For each $\delta \in (0,1)$ one has $B_{<1}(0) \subseteq \overline{TB_{<r}(0)} \subseteq TB_{<r}(0) + B_{<\delta}(0)$ and thus

$$B_{<\delta^k}(0) \subseteq TB_{<\delta^k r}(0) + B_{<\delta^{k+1}}(0) \quad \text{for all } k \in \mathbb{N}_0. \tag{2.3.1}$$

We may choose δ so small that $y_0 := \frac{y}{1-\delta} \in B_{<1}(0)$. We can thus recursively apply (2.3.1) to find sequences $(x_k)_{k\in\mathbb{N}_0}$ in X and $(y_k)_{k\in\mathbb{N}_0}$ in Y such that

$$y_k = Tx_k + y_{k+1}, \qquad x_k \in B_{<\delta^k r}(0), \qquad \text{and} \qquad y_{k+1} \in B_{<\delta^{k+1}}(0)$$

for all $k \in \mathbb{N}_0$. So $y_0 = \sum_{k=0}^n Tx_k + y_{n+1}$ and thus

$$y = (1-\delta)y_0 = T\sum_{k=0}^n (1-\delta)x_k + (1-\delta)y_{n+1}$$

for each $n \in \mathbb{N}_0$, which implies $T\sum_{k=0}^n (1-\delta)x_k \to y$ as $n \to \infty$. On the other hand, $\sum_{k=0}^\infty (1-\delta)\|x_k\| < (1-\delta)\sum_{k=0}^\infty r\delta^k = r$, so the completeness of the space X implies that $\sum_{k=0}^n (1-\delta)x_k$ converges to a point $x \in X$ as $n \to \infty$ (Proposition 1.3.9(ii)) Observe that $x \in B_{<r}(0)$. The continuity of T and the uniqueness of limits in Y gives $Tx = y$. □

Theorem 2.3.3 (Open mapping theorem)**.** *Let X, Y be Banach spaces over $\mathbb{K}$ and let $T \in \mathcal{L}(X;Y)$. Then either the range TX is meager in Y or T is surjective. In the latter case T is open.*

Lecture 7 (Mo, 04 May) is planned to start with Theorem 2.3.3.

Proof. First observe that TX cannot both be meager and equal to Y: since Y is complete and hence a Baire space, every meager subset of Y has empty interior.

Assume now that TX is not meager; we show that this implies that T is surjective and open. (Note that this implication gives all remaining assertions of the theorem, for if T is surjective then we have already seen in the previous paragraph that TX is not meager.) Let us first show that 0 is an interior point of $TB_{<1}(0)$. To this end, observe that

$$TX = T\bigcup_{n\in\mathbb{N}} B_{<n}(0) = \bigcup_{n\in\mathbb{N}} TB_{<n}(0).$$

As TX is not meager, it cannot be written as a countable union of nowhere dense sets, so there exists an $n \in \mathbb{N}$ such that $TB_{<n}(0)$ is not nowhere dense, i.e. $\overline{TB_{<n}(0)}$ has non-empty

interior. So we can find a point $y \in Y$ and a radius $r > 0$ such that $B_{<r}(y) \subseteq \overline{TB_{<n}(0)}$. Since, in particular, $y \in \overline{TB_{<n}(0)}$, there exists $x \in B_{<n}(0)$ such that $\|Tx - y\| < \frac{r}{2}$. Hence,

$$B_{<r/2}(0) = B_{<r/2}(y) - y \subseteq B_{<r}(y) - Tx \subseteq \overline{TB_{<n}(0)} - Tx = \overline{T(B_{<n}(0) - x)} \subseteq \overline{T(B_{<s}(0))}$$

for $s := n + \|x\|$. Since X is complete, we can apply the little open mapping theorem (Lemma 2.3.2) to conclude $B_{<r/2}(0) \subseteq TB_{<s}(0)$ and thus, $B_{<r/(2s)}(0) \subseteq TB_{<1}(0)$. So 0 is indeed an interior point of $TB_{<1}(0)$. According to Exercise 4.3(b) on Präsenzblatt 4 this implies that T is surjective and open. □

Corollary 2.3.4 (Continuous inverse theorem)**.** *Let X, Y be Banach spaces over $\mathbb{K}$ and let $T \in \mathcal{L}(X; Y)$ be bijective. Then $T^{-1}\colon Y \to X$ is continuous.*

Proof. Let $U \subseteq X$ be open. Then the pre-image $(T^{-1})^{-1}(U) = T(U)$ is open according to the open mapping theorem (Theorem 2.3.3). □

In the exercises we shall see how the continuous inverse theorem is related to the equivalence of norms on Banach spaces.

To prepare the next definition, consider normed spaces X, Y over $\mathbb{K}$. Unless otherwise stated we shall always assume that the the product $X \times Y$ is endowed with a norm such that convergence of sequences (and nets) is equivalent to convergence in both components. (Clearly such a norm exists; for instance, we may set $\|(x, y)\| := \|x\| + \|y\|$ for all $(x, y) \in X \times Y$.) It is not difficult to see that $X \times Y$ is a Banach space if and only if both X and Y are Banach spaces.

Definition 2.3.5 (Closed graph)**.** Let X, Y be normed space over $\mathbb{K}$. A linear map $T\colon X \to Y$ is said to have **closed graph** if its graph $\{(x, Tx) \mid x \in X\}$ is closed in $X \times Y$

Proposition 2.3.6 (Characterisation of operators with closed graph)**.** *Let X, Y be normed space over $\mathbb{K}$ and let $T\colon X \to Y$ be linear. The following assertions are equivalent:*

(i) *The map T has closed graph.*

(ii) *If a sequence (x_n) in X converges to a point $x \in X$ and (Ty_n) converges to a point $y \in Y$, then $Tx = y$.*

Proof. "(i) ⇒ (ii)" Assume that T has closed graph. Let (x_n) be a sequence in X that converges to a point $x \in X$ and assume that (Tx_n) converges to a point $y \in Y$. Then (x_n, Tx_n) is a sequence in the graph of T that converges to the point (x, y) in $X \times Y$. As the graph is closed, (x, y) is also located in the graph of T, so $y = Tx$.

"(ii) ⇒ (i)" Assume that (i) holds and let $\big((x_n, Tx_n)\big)$ be a sequence in the graph of T that converges to a point $(x, y) \in X \times Y$. Then $x_n \to x$ and $Tx_n \to y$, so (i) implies that $y = Tx$, i.e. the limit (x, y) is equal to (x, Tx) is is thus contained in the graph of T. □

Theorem 2.3.7 (Closed graph theorem)**.** *Let X, Y be Banach space over $\mathbb{K}$ and let $T\colon X \to Y$ be linear. If T has closed graph, then T is continuous.*

Proof. Endow $X \times Y$ with the norm given by $\|(x, y)\| := \|x\| + \|y\|$ for all $x \in X$ and all $y \in Y$. This turns $X \times Y$ into a Banach space. Note that the coordinate maps $P_X := X \times Y \to X$ and $P_Y : X \times Y \to Y$ given by

$$P_X(x, y) = x \qquad \text{and} \qquad P_Y(x, y) = y$$

for all $(x, y) \in X \times X$ are continuous.

Now assume that the graph $G := \{(x, Tx) \mid x \in X\} \subseteq X \times Y$ is closed. Then G is a Banach space with respect to the norm induced from $X \times Y$ (Proposition 1.1.1(c)). Observe that the restriction $P_X|_G : G \to X$ is bijective and hence, its inverse map that its inverse map $(P_X|_G)^{-1} : X \to G$ is continuous by the continuous inverse theorem (Corollary 2.3.4). Moreover, the inverse is given by $(P_X|_G)^{-1} x = (x, Tx)$ for all $x \in X$. Hence, $T : X \to Y$ is the composition of the continuous maps

$$X \overset{(P_X|_G)^{-1}}{\to} G \to X \times Y \overset{P_Y}{\to} Y,$$

where the arrow in the middle is the inclusion map from G into $X \times Y$ (and is thus isometric). Hence, T is continuous. □

2.4 Sublinear functionals, Zorn's lemma, and the Hahn–Banach extension theorem

Introductory questions.

(a) Let $X \neq \{0\}$ be a normed space. Does the dual space X' contain at least one nonzero element?

(b) Let X be a normed space. How are the norms on X and X' related?

The central concept in this section is the following generalization of linear functionals. We shall see towards the end of this section that this notion is very useful to prove extension results for bounded linear functionals.

Lecture 8 (Th, 07 May) is planned to start with Definition 2.4.1.

Definition 2.4.1 (Sublinear functionals)**.** Let X be a vector space over $\mathbb{K}$. A map $p : X \to \mathbb{R}$ is called **sublinear** if it satisfies the following two properties:

(I) **Subadditivity**: For all $x, y \in X$ one has $p(x + y) \leq p(x) + p(y)$.

(II) **Positive homogeneity**: For all $x \in X$ and all $\alpha \in [0, \infty)$ one has $p(\alpha x) = \alpha p(x)$.

Here are few simple examples of sublinear functional:

Examples 2.4.2 (Some sublinear functionals)**.**

(a) Let X be a vector space over $\mathbb{K}$. Every norm on X is a sublinear functional.

(b) Let X be a vector space over $\mathbb{K}$ and let $\varphi\colon X \to \mathbb{K}$ be linear. If $\mathbb{K} = \mathbb{R}$, then φ is sublinear. If $\mathbb{K} = \mathbb{C}$, then $\operatorname{Re} \circ \varphi$ is sublinear.

(c) Let (M, d) be a compact metric space. Then $\mathrm{C}(M;\mathbb{R}) \to \mathbb{R}$, $f \mapsto \|f^+\|_\infty$ is sublinear.

(d) Let $\alpha, \beta \in \mathbb{R}$ satisfy $\alpha \le \beta$ and define $p\colon \mathbb{R} \to \mathbb{R}$ by

$$p(x) = \begin{cases} \beta x & \text{if } x \ge 0, \\ \alpha x & \text{if } x < 0. \end{cases}$$

The following extension lemma for linear functionals that are dominated by sublinear functionals will turn out to be very useful.

Lemma 2.4.3. *Let V be vector subspace of an $\mathbb{R}$-vector space X, and $x_0 \in X \setminus V$. Let $p\colon X \to \mathbb{R}$ be sublinear and $\varphi\colon V \to \mathbb{R}$ linear such that $\varphi(v) \le p(v)$ for all $v \in V$. Then there exists a linear extension $\psi\colon W := V + \mathbb{R}x_0 \to \mathbb{R}$ of φ such that $\psi(w) \le p(w)$ for all $w \in W$.*

Proof. For all $u, v \in V$ one has $\varphi(u) + \varphi(v) = \varphi(u+v) \le p(u+v) \le p(u - x_0) + p(v + x_0)$, so

$$\varphi(u) - p(u - x_0) \le p(v + x_0) - \varphi(v).$$

Therefore, $\gamma := \inf_{v \in V} \big(p(v + x_0) - \varphi(v)\big)$ is an element of $\mathbb{R}$ and satisfies

$$\varphi(u) - p(u - x_0) \le \gamma \le p(u + x_0) - \varphi(u) \tag{2.4.1}$$

for all $u \in V$. Now, fix $v \in V$ and $\alpha \in \mathbb{R}$. If $\alpha > 0$ we use the right inequality in (2.4.1) for $u := v/\alpha$ and multiply it by α to get $\alpha\gamma \le p(v + \alpha x_0) - \varphi(v)$. If $\alpha < 0$ we use the left inequality in (2.4.1) for $u := -v/\alpha$ and multiply it by α to get $\alpha\gamma \le -\varphi(v) + p(v + \alpha x_0)$. So

$$\varphi(v) + \alpha\gamma \le p(v + \alpha x_0)$$

for all $v \in V$ and all $\alpha \in \mathbb{R}$ (for $\alpha = 0$ this inequality is clear by assumption). So we can simply set $\psi(v + \alpha x_0) := \varphi(v) + \alpha\gamma$ for all $v + \alpha x_0 \in W$. $\square$

A simple consequence of Lemma 2.4.3 is the following result.

Proposition 2.4.4 (Bounding a sublinear functional below by a linear functional)**.** *Let $p\colon \mathbb{R}^d \to \mathbb{R}$ be sublinear. Then there exists $y \in \mathbb{R}^d$ such that $y^{\mathrm{T}} x \le p(x)$ for all $x \in \mathbb{R}^d$.*

Proof. First consider the zero functional on the trivial subspace $\{0\}$ of $\mathbb{R}^d$. It is clearly dominated by p. Now we can use Lemma 2.4.3 d times to extend the functional first to $\operatorname{span}\{e^{(1)}\}$, then to $\operatorname{span}\{e^{(1)}, e^{(2)}\}$, and so on, until we end up with a linear functional $\psi\colon \mathbb{R}^d \to \mathbb{R}$ such that $\psi(x) \le p(x)$ for all $x \in \mathbb{R}^d$. It is a standard result in linear algebra that ψ can be represented by left multiplication with an $1 \times d$-matrix. $\square$

The main result of this section is the following famous theorem which shows that a similar result still holds in infinite-dimensions.

Theorem 2.4.5 (Hahn–Banach extension theorem)**.** *Let X be a vector space over $\mathbb{K}$ and let $p\colon X \to \mathbb{K}$ be sublinear. Let $V \subseteq X$ be a vector subspace and $\varphi\colon V \to \mathbb{K}$ a linear map that satisfies $\operatorname{Re}\varphi(v) \le p(v)$ for all $v \in V$. Then there exists a linear map $\psi\colon X \to \mathbb{K}$ that extends φ and that satisfies $\psi(x) \le \operatorname{Re} p(x)$ for all $x \in X$.*

If $\mathbb{K} = \mathbb{R}$ and $\dim X < \infty$, then Theorem 2.4.5 can be shown by essentially the same induction argument as Proposition 2.4.4. The key insight in the rest of this section is that Theorem 2.4.5 also holds on infinite-dimensional spaces. On such spaces one cannot simply copy the proof of Proposition 2.4.4. Recall from Corollary 2.1.5 that many spaces do not even have a countable basis, so a simple induction over a basis is not possible. Instead, we will use a very useful fundamental result called **Zorn's lemma** (Lemma 2.4.7). To formulate it, we first recall a few notions regarding partially ordered sets.

Recall that a **partially ordered set** (or briefly a **poset**) is a pair $(Z, \le)$, where Z is a set and $\le$ is **partial order** on Z, i.e. a relation on Z that satisfies the following axioms:

(I) **Reflexivity**: For all $x \in Z$ one has $x \le x$.

(II) **Anti-symmetry**: For all $x, y \in Z$ with $x \le y$ and $y \le x$ one has $x = y$.

(III) **Transitivity**: For all $x, y, z \in Z$ with $x \le y$ and $y \le z$ one has $x \le z$.

In the following definition we collect a number of notions in partially ordered sets that we will use in the sequel.

Definition 2.4.6 (Important notions in partially ordered sets)**.** Let $(Z, \le)$ be a partially ordered set and let $S \subseteq Z$.

(a) The set S is called a **chain**[2] if for all $a, b \in S$ one has $a \le b$ or $b \le a$.

(b) An element $u \in Z$ is called an **upper bound** of S if $u \ge s$ for all $s \in S$.

(c) An element $m \in Z$ is called a **maximal element** of Z if the following implication holds for all $z \in Z$:

$$z \ge m \qquad \Rightarrow \qquad z = m$$

(d) An element $\ell \in Z$ is called a **largest element** of Z if $\ell \ge z$ for all $z \in Z$.

Note that if Z has a largest element ℓ, then ℓ is also a maximal element of Z, and it is the only maximal element of Z. On the other hand, it can happen that Z does not have a largest element, but one or multiple maximal elements.

The following result is a consequence of the axioms of choice and is a fundamental result in logic and set theory. We do not prove it here, but we shall use it in the following.

Lemma 2.4.7 (Zorn's lemma)**.** *Let $(Z, \le)$ be a non-empty partially ordered set. If every non-empty chain $C \subseteq Z$ has an upper bound in Z, then Z has a maximal element.*

[2] In German: **Kette**

Zorn's lemma can, for instance, be used in linear algebra to show that every vector space has a basis.[3]

Proof of Theorem 2.4.5. We first treat the real case and then derive the complex case as a consequence.

The real case: Let $\mathbb{K} = \mathbb{R}$.

Let $\mathcal{G}$ denote the set of all graphs of linear extensions ψ of φ to vector subspaces $W \supseteq V$ of Z such that $\psi(w) \le p(w)$ for all $w \in W$. Then $\mathcal{G}$ is partially ordered with respect to set inclusion. If $\mathcal{C} \subseteq \mathcal{G}$ is a non-empty chain, then $\hat{G} := \bigcup_{G \in \mathcal{C}} G$ is again a graph of a linear functional $\hat{\psi}$ that is defined on an a vector subspace $\hat{W} \supseteq V$ of X; you prove this in Exercise 5.1 on Homework Sheet 5. For each $w \in \hat{W}$ one has $(w, \hat{\psi}(w)) \in \hat{G}$ and thus $(w, \psi(w)) \in G$ for one of the graphs $G \in \mathcal{C}$; this implies that $\psi(w) \le p(w)$. So the graph of ψ is an element of $\mathcal{G}$ and hence, $\mathcal{C}$ has an upper bound in $\mathcal{G}$. According to Zorn's lemma (Lemma 2.4.7) this implies that $\mathcal{G}$ has a maximal element $G_{\max}$.

Let $\psi_{\max}\colon W_{\max} \to \mathbb{R}$ the linear functional whose graph is $G_{\max}$. If $W_{\max} \ne X$, then Lemma 2.4.3 shows that we can extend $\psi_{\max}$ to a functional that is defined on a strictly larger vector subspace and that is still dominated by p; but this contradicts the maximality of $G_{\max}$. So $W_{\max} = X$ and hence, $\psi_{\max}$ is the desired extension of φ.

Lecture 9 (Mo, 11 May) is planned to start with the case $\mathbb{K} = \mathbb{C}$ in the proof of Theorem 2.4.5.

The complex case: Let $\mathbb{K} = \mathbb{C}$.

Define $\phi_{\mathbb{R}} := \operatorname{Re} \circ \phi\colon V \to \mathbb{R}$. If we consider X as an $\mathbb{R}$-vector space by restricting the scalar multiplication to real functional, then V is also a vector subspace of X over $\mathbb{R}$ and $\phi_{\mathbb{R}}$ is $\mathbb{R}$-linear; so the real case treated above gives as an $\mathbb{R}$-linear extension $\psi_{\mathbb{R}}\colon X \to \mathbb{R}$ that satisfies $\psi_{\mathbb{R}}(x) \le p(x)$ for all $x \in X$. Now we define $\psi\colon X \to \mathbb{C}$ by $\psi(x) := \psi_{\mathbb{R}}(x) - i\psi_{\mathbb{R}}(ix)$ for all $x \in X$. Note that $\operatorname{Re}\psi(x) = \psi_{\mathbb{R}}(x) \le p(x)$ for all $x \in \mathbb{R}$. Moreover ψ extends φ since

$$\psi(v) = \psi_{\mathbb{R}}(v) - i\psi_{\mathbb{R}}(iv) = \operatorname{Re}\varphi(v) - i\operatorname{Re}(i\varphi(v))$$

for all $v \in V$. To see that ψ is $\mathbb{C}$-linear, observe that ψ is $\mathbb{R}$-linear and that

$$\psi(ix) = \psi_{\mathbb{R}}(ix) - i\psi_{\mathbb{R}}(i^2 x) = \psi_{\mathbb{R}}(ix) + i\psi_{\mathbb{R}}(x) = i\big(\psi_{\mathbb{R}}(x) - i\psi_{\mathbb{R}}(ix)\big) = i\psi(x)$$

for all $x \in X$. □

Corollary 2.4.8 (Hahn–Banach extension theorem for continuous linear functionals)**.** *Let X be a normed space and $V \subseteq X$ a vector subspace which we endow with the norm inherited from X. Every bounded linear functional $v' \in V'$ can be extended to a bounded linear functional $x' \in X'$ of the same norm.*

[3]Probably you have seen a proof in your linear algebra course that every finite-dimensional vector spaces has a basis. Zorn's lemma is only needed to show that this is also true in the infinite-dimensional case.

Proof. For every $v \in V$ one has

$$\operatorname{Re} v'(v) \le \left|v'(v)\right| \le \left\|v'\right\| \|v\| .$$

Consider the sublinear functional $p\colon X \to \mathbb{R}$, $p(x) = \left\|v'\right\| \|x\|$. According to the Hahn–Banach extension theorem (Theorem 2.4.5) there exists a linear functional $\psi\colon X \to \mathbb{K}$ that extends v' and that satisfies $\operatorname{Re}\psi(x) \le p(x)$ for all $x \in X$.

Now fix a vector $x \in X$ and consider the number $\psi(x)$. We can find a scalar $\mu \in \mathbb{K}$ of modulus $\left|\mu\right| = 1$ such that $\left|\psi(x)\right| = \mu\psi(x)$. Hence,

$$\left|\psi(x)\right| = \psi(\mu x) \le p(\mu x) = \left\|v'\right\| \left\|\mu x\right\| = \left\|v'\right\| \|x\| .$$

So ψ is continuous and $\left\|\psi\right\| \le \left\|v'\right\|$. On the other hand, one clearly has $\left\|\psi\right\| \ge \left\|v'\right\|$ since ψ extends v'. So we obtain the claim for $x' := \psi$. $\square$

Recall that, for a functional $x' \in X'$ on a normed space X, the norm $\left\|x'\right\|$ is defined as

$$\left\|x'\right\| = \sup\left\{\left|\langle x', x\rangle\right| \mid x \in X,\ \|x\| \le 1\right\}.$$

We now derive from Corollary 2.4.8 that the same formula also holds with the roles of x and x' reversed:

Corollary 2.4.9 (The norm via continuous linear functionals)**.** *Let X be a normed space and let $x \in X$. Then*

$$\|x\| = \sup\left\{\left|\langle x', x\rangle\right| \mid x' \in X',\ \left\|x'\right\| \le 1\right\},$$

and the supremum is even a maximum. In particular, if $x \neq 0$, then there exists an $x' \in X'$ such that $\langle x', x\rangle \neq 0$.

Proof. The inequality $\ge$ follows from the estimate $\|x\| \left\|x'\right\| \ge \left|\langle x', x\rangle\right|$ for all $x' \in X'$.

To show the converse inequality $\le$ and that the supremum is even a maximum, consider the vector subspace $V := \operatorname{span}\{x\}$ of X and the continuous linear functional $v' \in V'$ given by $\langle v', \alpha x\rangle := \alpha \|x\|$ for all $\alpha \in \mathbb{K}$. Then $\left\|v'\right\| \le 1$ (the inequality is an equality unless $x = 0$) and $\langle v', x\rangle = \|x\|$. According to Corollary 2.4.8 there exists a functional $x' \in X'$ of norm $\left\|x'\right\| = \left\|v'\right\| \le 1$ that extends v'. In particular, $\left|\langle x', x\rangle\right| = \left|\langle v', x\rangle\right| = \|x\|$. $\square$

Corollary 2.4.9 shows that one can, to some extent, reverse the role of a normed space X with its dual space X'. This might not be too surprising if one thinks of the ℓ^p-spaces: for $p, p' \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{p'} = 1$ we proved in Theorem 1.2.6 that $(\ell^p)' \simeq \ell^{p'}$. But since the equality $\frac{1}{p} + \frac{1}{p'} = 1$ is symmetric with respect to p and p', one can of course swap their roles and thus also obtain $\left(\ell^{p'}\right)' \simeq \ell^p$. Later in the course will we investigate this theme in much greater generality, in particular in Section 5.4.

Chapter 3

Basic Geometry in Normed Spaces

3.1 Convex sets and the geometric Hahn–Banach theorems

Introductory questions.

(a) What is a **hyperplane** in $\mathbb{R}^d$? What is a hyperplane in a normed space X?

(b) Draw two non-empty compact convex subsets of $\mathbb{R}^2$. Can they be "separated by a straight line"? And what does that even mean?

Definition 3.1.1 (Convex sets)**.** Let X be a vector space over $\mathbb{K}$. A subset $C \subseteq X$ is called **convex** if, for all $x, y \in C$ and all $\lambda \in [0,1]$, one has $\lambda x + (1\lambda) y \in C$.

By an induction argument one can show that if C is a convex subset of a vector space over $\mathbb{K}$ and $x_1, \dots, x_n \in C$, then $\lambda_1 x_1 + \dots + \lambda_n x_n \in C$ for all $\lambda_1, \dots, \lambda_n \in [0,1]$ that satisfy $\lambda_1 + \dots + \lambda_n = 1$.

Theorem 3.1.2 (Hahn–Banach separation theorem for open sets)**.** *Let X be a normed space and let $\emptyset \neq U, C \subseteq X$ be convex and disjoint. If U is open, there exists a functional $x' \in X'$ and a number $\gamma \in \mathbb{R}$ such that*

$$\operatorname{Re}\langle x', u\rangle < \gamma \le \operatorname{Re}\langle x', c\rangle \qquad \textit{for all } u \in U,\ c \in C.$$

Proof. We proceed in two steps.

Step 1A: We show that claim if $C = \{c\}$ for some $c \in X$ and $\mathbb{K} = \mathbb{R}$.

After shifting the entire situation if necessary we may assume that $0 \in U$. Define

$$p\colon X \to \mathbb{R}, \quad x \mapsto \inf\{\lambda \in (0,\infty) \mid x \in \lambda U\}.$$

Note that the set under the infimum is always non-empty. This is clear for $x = 0$ and for $x \in X \setminus \{0\}$ one argues as follows: U is open and contains 0, one has $B_{\le r}(0) \subseteq U$ for some $r \in (0,\infty)$, so $x \in \frac{\|x\|}{r} U$ for all $x \in X$. In fact, this even shows that $p(x) \le \frac{\|x\|}{r}$ for all $x \in X$.

Let us prove that the function p is sublinear. One clearly has $p(\alpha x) = \alpha p(x)$ for all $x \in X$ and $\alpha \in [0,\infty)$. Now, let $x_1, x_2 \in X$ and let $\lambda_1, \lambda_2 \in (0,\infty)$ such that, for $k \in \{1,2\}$, one has $x_k \in \lambda_k U$, i.e. $x_k = \lambda_k u_k$ for some $u_k \in U$. So

$$x_1 + x_2 = (\lambda_1 + \lambda_2)\Big(\frac{\lambda_1}{\lambda_1 + \lambda_2} u_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} u_2\Big) \in (\lambda_1 + \lambda_2)U$$

since U is convex. Hence, $p(x_1 + x_2) \le \lambda_1 + \lambda_2$. By taking the infimum over all such λ_1 and λ_2 we obtain $p(x_1 + x_2) \le p(x_1) + p(x_2)$, as claimed.

Lecture 10 (Mo, 18 May) is starts with the argument for $p(c) \ge 1$ in the proof of Theorem 3.1.2.

We note that $p(c) \ge 1$: for every $\lambda \in (0,1]$ one cannot have $c \in \lambda U$ since otherwise $\frac{c}{\lambda} \in U$, which would imply that the convex combination c of $\frac{c}{\lambda}$ and 0 is also in U.

Now consider the one-dimensional vector subspace $V := \operatorname{span}\{c\}$ of X and the functional $v' \in V'$, given by $\langle v', \alpha c\rangle = \alpha$ for all $\alpha \in \mathbb{R}$. Observe that v' is dominated by p on V: for $\alpha \in [0,\infty)$ one has $\langle v', \alpha c\rangle = \alpha \le \alpha p(c) = p(\alpha c)$, and for $\alpha \in (-\infty, 0)$ one has $\langle v', \alpha c\rangle = \alpha < 0 \le p(\alpha c)$. So we can apply the Hahn–Banach extension theorem (Theorem 2.4.5), which gives us a linear functional $\psi\colon X \to \mathbb{R}$ that extends v' and satisfies $\psi(x) \le p(x) \le \|x\|/r$ for all $x \in X$. By applying the same inequality to $-x$ we obtain $|\psi(x)| \le \|x\|/r$ for all $x \in X$. This shows that ψ is bounded with norm $\|\psi\| \le \|x\|/r$, so let us set $x' := \psi \in X'$.

For every $u \in U$ one has $p(u) < 1$ since u is open, so

$$\langle x', u\rangle \le p(u) < 1 = \langle v', c\rangle = \langle x', c\rangle.$$

Step 1B: We show that claim if $C = \{c\}$ for some $c \in X$ and $\mathbb{K} = \mathbb{C}$.

We reduce this case to the real case similarly as in the proof of Theorem 2.4.5. So first consider X as normed space over $\mathbb{R}$. According to Step 1A there exists an $\mathbb{R}$-linear continuous functional $x'_{\mathbb{R}}\colon X \to \mathbb{R}$ and a number $\gamma \in \mathbb{R}$ such that $\langle x'_{\mathbb{R}}, u\rangle < \gamma \le \langle x'_{\mathbb{R}}, c\rangle$ for all $u \in U$. Define a $\mathbb{C}$-linear functional $x' \in X'$ by

$$\langle x', x\rangle := \langle x'_{\mathbb{R}}, x\rangle - i\,\langle x'_{\mathbb{R}}, ix\rangle$$

for all $x \in X$. Clearly, x' is continuous and $\mathbb{R}$-linear; one readily checks that it also satisfies $\langle x', ix\rangle = i\,\langle x', x\rangle$ for all $x \in X$, so we conclude that x' is $\mathbb{C}$-linear. Hence, x' is indeed an element of the dual space X' of the normed space X over $\mathbb{C}$. Since $\operatorname{Re}\langle x', x\rangle = \langle x'_{\mathbb{R}}, x\rangle$ for all $x \in X$, the functional x' satisfies $\operatorname{Re}\langle x', u\rangle < \gamma \le \operatorname{Re}\langle x', c\rangle$ for all $u \in U$.

Step 2: Now we consider the general case (and $\mathbb{K}$ can be $\mathbb{R}$ or $\mathbb{C}$ now).

Set $D := U - C$. It is easy to see that D is convex and open. Moreover, $0 \notin D$ since $U \cap C = \emptyset$. So by Step 1 we can find $x' \in X'$ such that $\operatorname{Re}\langle x', d\rangle < \operatorname{Re}\langle x', 0\rangle = 0$ for all $d \in D$. For all $u \in U$ and $c \in C$ this implies that $\operatorname{Re}\langle x', u - c\rangle < 0$, so

$$\operatorname{Re}\langle x', u\rangle < \operatorname{Re}\langle x', c\rangle.$$

As $x' \neq 0$, it is not difficult to check that $\operatorname{Re} x'(U)$ is open; so the claim follows by defining $\gamma := \inf_{c\in C} \operatorname{Re}\langle x', c\rangle$. □

To derive another version of the Hahn–Banach separation theorem in Corollary 3.1.5 below, we need the following notion.

Definition 3.1.3 (Distance to subsets)**.** Let (M, d) be a metric space and let $\emptyset \neq S \subseteq M$ and $x \in M$. Then the number

$$\operatorname{dist}(x, S) := \inf\left\{ \mathrm{d}(x, s) \mid s \in S \right\} \in [0, \infty)$$

is called the **distance of x to S**.

Let us collect two useful properties of the distance function in the next proposition.

Proposition 3.1.4 (Properties of the distance to subsets)**.** *Let (M, d) be a metric space and let $\emptyset \neq S \subseteq M$.*

(a) *For $x \in M$ one has $\operatorname{dist}(x, S) = 0$ if and only if $x \in \overline{S}$.*

(b) *The distance function $\operatorname{dist}(\cdot, S)\colon M \to [0, \infty)$ is Lipschitz continuous with Lipschitz constance 1.*

You prove Proposition 3.1.4 in Exercise 5.5 on Homework Sheet 5.

Corollary 3.1.5 (Hahn–Banach separation theorems for closed and compact sets)**.** *Let X be a normed space and let $\emptyset \neq Q, C \subseteq X$ be convex and disjoint. If Q is compact and C is closed, there exists a functional $x' \in X'$ and real numbers $\gamma_1 < \gamma_2$ such that*

$$\operatorname{Re}\langle x', q\rangle \leq \gamma_1 < \gamma_2 \leq \operatorname{Re}\langle x', c\rangle \qquad \text{for all } q \in Q,\ c \in C.$$

Proof. As C is closed, the distance function $\operatorname{dist}(\cdot, C)$ only vanishes on C according to Proposition 3.1.4(a). Hence, it does not vanish anywhere on Q as $Q \cap C = \emptyset$. Since Q is compact and $\operatorname{dist}(\cdot, C)$ is continuous (Proposition 3.1.4(b)), we conclude that

$$\varepsilon := \inf_{q \in Q} \operatorname{dist}(q, C) > 0.$$

Note that $\left\| q - c \right\| \geq \varepsilon$ for all $q \in Q$ and all $c \in C$. Hence, the set[1] $U := \{x \in X \mid \operatorname{dist}(x, Q)\} < \varepsilon$ does not intersect C. Moreover, U contains Q and it is open since $\operatorname{dist}(\cdot, Q)$ is continuous. Moreover, it is not difficult to check that the convexity of Q implies that U is also convex.

So we can now apply Theorem 3.1.2 to find a functional $x' \in X'$ and a number $\gamma_2 \in \mathbb{R}$ such that $\operatorname{Re}\langle x', q\rangle < \gamma_2 \leq \operatorname{Re}\langle x', c\rangle$ for all $c \in C$ and $q \in U$ (and thus, in particular, for all $q \in Q$). As $\operatorname{Re} \circ x'$ is continuous and Q is non-empty and compact, the set $\operatorname{Re} x'(Q) \subseteq \mathbb{R}$ is non-empty and compact and thus has a maximum γ_1. This proves the claim. $\square$

[1] Observe that we now consider the distance to Q instead of the distance to C; this choice turns out to be helpful at the end of the proof.

3.2 Linear projections and direct sums

Introductory questions.

(a) What picture do you think of when you hear the word "projection"?

(b) Consider a straight line L and a plane P in $\mathbb{R}^3$ such that $L \cap P = \{0\}$ and let $x \in \mathbb{R}^3$. What does it mean to "project x into P along L"?

Definition 3.2.1 (Linear projections and direct sums)**.** Let V be a vector space over some field and let X be a normed space over $\mathbb{K}$.

(a) A linear map $P\colon V \to V$ is called a **projection** if $P^2 = P$.

(b) Let $V_1, V_2 \subseteq V$ be vector subspaces. We say that V is the **direct sum of V_1 and V_2**, which we denote by $V = V_1 \oplus V_2$, if the linear map

$$V_1 \times V_2 \to V, \qquad (v_1, v_2) \mapsto v_1 + v_2$$

is bijective.

(c) Let $X_1, X_2 \subseteq X$ be vector subspaces. We say that X is the **topologically direct sum of X_1, X_2** if the linear map

$$X_1 \times X_2 \to X, \qquad (x_1, x_2) \mapsto x_1 + x_2$$

is an isomorphism of normed spaces.

To distinguish this notion more clearly from the one in (b), we sometimes call the concept in (c) **algebraically direct sum** rather than just **direct sum**.

Recall from linear algebra that, for two vector subspaces $V_1, V_2 \subseteq V$ of a vector subspace V, one has $V = V_1 \oplus V_2$ if and only if ones has $V = V_1 + V_2$ and $V_1 \cap V_2 = \{0\}$. (Beware, however, that this characterisation does not hold for three or more vector subspaces; we will discuss this in an exercise in more detail.)

Proposition 3.2.2 (Properties of projections)**.** *Let V be a vector space over $\mathbb{K}$. Let $P\colon V \to V$ be a linear projection.*

(a) *For each $x \in V$ one has $x \in PV$ if and only if $Px = x$.*

(b) *The map $\operatorname{id} - P$ is a projection, too. One has $\ker P = (\operatorname{id} - P)V$ and $\ker(\operatorname{id} - P) = PV$.*

(c) *One has $V = PV \oplus \ker P = PV \oplus (\operatorname{id} - P)V$. For every $v \in V$ the vectors $v_1 := Pv$ and $v_2 := (\operatorname{id} - P)v$ are the unique elements of PV and $(\operatorname{id} - P)V$ that satisfy $v = v_1 + v_2$.*

Proof. (a) Let $x \in V$. If $x = Px$, then obviously $x \in PV$. If, conversely, $x \in PV$, then $x = Pv$ for some $v \in V$, so $x = Pv = PPv = Px$.

Lecture 11 (Th, 21 May) starts with the proof of Proposition 3.2.2(b).

(b) The map $\mathrm{id} - P$ is a projection since

$$(\mathrm{id} - P)^2 = \mathrm{id}^2 - \mathrm{id}\,P - P\,\mathrm{id} + P^2 = \mathrm{id} - P.$$

To show the inclusion $\ker P \subseteq (\mathrm{id} - P)V$, let $x \in \ker P$ and observe that $x = (\mathrm{id} - P)x \in (\mathrm{id} - P)V$. To show the converse inclusion $\ker P \supseteq (\mathrm{id} - P)V$, let $x = (\mathrm{id} - P)v \in (\mathrm{id} - P)V$ for some $v \in V$. Then $Px = P(\mathrm{id} - P)v = (P - P^2)v = 0$, so $x \in \ker P$.

Finally, if one applies the equality $\ker P = (\mathrm{id} - P)V$ that we just proved to the projection $\mathrm{id} - P$ instead of P, we also obtain $\ker(\mathrm{id} - P) = PV$.

(c) We only need to prove that $V = PV \oplus \ker P$; the remaining assertions follow directly from this.

If $x \in PV \cap \ker P$, then it follows from (a) that $x = Px = 0$. On the other hand one has $v = Pv + (\mathrm{id} - P)v$, so $V = PV + (\mathrm{id} - PV)V$. □

Proposition 3.2.3 (Equality of projections)**.** *Let X be a vector space over $\mathbb{K}$ and let $P, Q\colon X \to X$ be linear projections. The following assertions are equivalent:*

(i) $P = Q$.

(ii) $PX = QX$ *and* $\ker P = \ker Q$.

(iii) $PX = QX$ *and* $PQ = QP$.

(iv) $\ker P = \ker Q$ *and* $PQ = QP$.

Proof. "(i) ⇒ (iv)" This implication is obvious.

"(iv) ⇒ (iii)" Let (iv) hold. We need to show that $PX = QX$, and for symmetry reasons it suffices to show that $PX \subseteq QX$. So let $x \in Px$. Then $x = Px$ and $(\mathrm{id} - Q)x \in \ker Q = \ker P$, which implies

$$(\mathrm{id} - Q)x = (\mathrm{id} - Q)Px = P(\mathrm{id} - Q)x = 0;$$

for the second equality was used that P and Q commute according to (iv). Thus, $x = Qx \in QX$.

"(iii) ⇒ (ii)" Let (iii) hold. We need to prove $\ker P = \ker Q$, and for symmetry reasons it suffices to show $\ker P \subseteq \ker Q$. So let $x \in \ker P$. As P is a projection, it acts as the identity on its range, so it follows from $PX = QX$ that $Qx = PQx = QPx = 0$.

"(ii) ⇒ (i)" Let (ii) hold and let $x \in X$. Then $Px \in PX = QX$ and $(\mathrm{id} - P)x \in \ker P = \ker Q$. Hence,

$$Qx = Q\big(Px + (\mathrm{id} - P)x\big) = QPx + Q(\mathrm{id} - P)x = Qx$$

as Q acts as the identity on QV. □

Corollary 3.2.4 (Projections and direct sums)**.** *Let X be a vector space over $\mathbb{K}$ and let $X_1, X_2 \subseteq X$ be vector subspaces such that $X = X_1 \oplus X_2$. For every $x \in X$ let $P_1 x \in X_1$ and $P_2 x \in X_2$ denote the unique vectors in those subspaces that satisfy $P_1 x + P_2 x = x$.*

Then $P_1, P_2 \colon X \to X$ are linear projections and $P_1 + P_2 = \mathrm{id}$. Moreover, P_1 is the unique linear projection $X \to X$ with range X_1 and kernel X_2, and P_2 is the unique linear projection $X \to X$ with range X_2 and kernel X_1.

Proof. The linearity of P_1 and P_2, the equation $P_1 + P_2 = \mathrm{id}$, and that equalities $P_1 X = X_1$ and $\ker P_1 = X_2$ are straightforward to check; we discuss the details in the exercises. The claimed uniqueness follows from Proposition 3.2.3(ii). □

The preceding corollary motivates the following definition:

Definition 3.2.5 (Projection onto a subspace along subspaces)**.** Let X be a vector subspace and let $X_1, X_2 \subseteq X$ be vector subspaces such that $X = X_1 \oplus X_2$. The linear projection $X \to X$ with range X_1 and kernel X_2 is called the **(linear) projection onto X_1 along X_2**.

Proposition 3.2.6 (Topological properties of projections and direct sum decompositions)**.** *Let X be a Banach space, let $X_1, X_2 \subseteq X$ be vector subspaces such that $X = X_1 \oplus X_2$, i.e. X is the algebraically direct sum of X_1 and X_2. Let $P_1, P_2 \colon X \to X$ be the projections onto X_1 along X_2 and vice versa. The following assertions are equivalent:*

(i) *At least one of the projections P_1, P_2 is continuous.*

(ii) *Both projections P_1, P_2 are continuous.*

(iii) *Both X_1 and X_2 are closed.*

(iv) *The space X is the topologically direct sum of X_1 and X_2. (Definition 3.2.1*(c)*).*

Proof. We will discuss the proof of this proposition on Homework Sheet 7. □

A particular prevalent class of examples are projections onto one-dimensional subsapces. To work with them efficiently, we first introduce the following general notation for operators of rank ≤ 1. Recall that the **rank** of a linear map is defined to be the dimension of its range.

Definition 3.2.7 (Notation for operators of rank at most one)**.** Let X, Y be normed spaces over $\mathbb{K}$, let $x' \in X'$ and $y \in Y$. Then we let $y \otimes x' \in \mathcal{L}(X, Y)$ denote the operator given by

$$(y \otimes x')x := y \, \langle x', x \rangle \qquad \text{for all } x \in X.$$

In the displayed formula in the above definition we wrote the scalar $\langle x', x \rangle$ on the right of the vector y because it makes the formula particularly intuitive, and also because it reflect very well what happens in finite dimensions when $y \in \mathbb{K}^d$ and $x' \in \mathbb{K}^{1 \times c}$. But of course, we can just as well write the scalar on the left, us usual, i.e. $(y \otimes x')x = \langle x', x \rangle \, y$.

Proposition 3.2.8 (Rank-**1** operators)**.** *Let X, Y, Z be normed spaces over $\mathbb{K}$.*

(a) *Let $x' \in X'$ and $y \in Y$. Then $\|y \otimes x'\| = \|y\| \, \|x'\|$.*

The operator $y \otimes x'$ is non-zero if and only if both y and x' are non-zero; in this case the operator has rank 1.

(b) *Conversely, let $T \in \mathcal{L}(X, Y)$ be an operator of rank 1. Then there exist non-zero vectors $x' \in X'$ and $y \in Y$ such that $T = y \otimes x'$.*

(c) *Let $x' \in X'$, $y \in Y$, $y' \in Y'$ and $z \in Z$. Then one has $(z \otimes y')\,(y \otimes x') = \langle y', y \rangle \, (z \otimes x')$.*

(d) *Now let $0 \neq x' \in X'$ and $0 \neq x \in X$. The rank-1 operator $x \otimes x' \in \mathcal{L}(X)$ is a projection if and only if $\langle x', x \rangle = 1$.*

Proof. (a) For every $x \in X$ one has

$$\|(y \otimes x')x\| = \left|\langle x', x \rangle\right| \, \|y\| \leq \|y\| \, \|x'\| \, \|x\|,$$

so $\|y \otimes x'\| = \|y\| \, \|x'\|$.

To show the converse inequality, note that there exists a sequence (x_n) in X such that $\|x_n\| \leq 1$ for all n and such that $\left|\langle x', x_n \rangle\right| \to \|x\|$. Hence,

$$\|y \otimes x'\| \geq \|(y \otimes x')x_n\| = \left|\langle x', x_n \rangle\right| \, \|y\| \to \|y\| \, \|x'\|.$$

From the norm equality that we just proved it is clear that $y \otimes x' = 0$ if and only if $y = 0$ or $x' = 0$.

Finally we note that, if y and x' are both non-zero, then the range of $y \otimes x'$ is precisely $\operatorname{span}\{y\}$, so $y \otimes x'$ has indeed rank 1.

(b) Let $T \in \mathcal{L}(X; Y)$ have rank 1. Then its range is one-dimensional, so there exists a vector $y \in Y \setminus \{0\}$ that spans TX. For every $x \in X$ there exists precisely one number $\alpha_x \in \mathbb{K}$ such that $Tx = \alpha_x y$. It is straightforward to show that α_x depends linearly on x. Moreover,

$$|\alpha_x| = \frac{\|Tx\|}{\|y\|} \leq \frac{\|T\|}{\|y\|} \, \|x\|$$

for all $x \in X$, which shows that the linear map $x' := X \to \mathbb{K}$, $x \mapsto \alpha_x$ is continuous and thus an element of X'. For every $x \in X$ one has $(y \otimes x')x = \langle x', x \rangle \, y = \alpha_x y = Tx$, so $T = y \otimes x'$. Since $T \neq 0$, it follows that also $x' \neq 0$.

(c) For every $x \in X$ one has

$$(z \otimes y')\,(y \otimes x')x = (z \otimes y')\,(y \langle x', x \rangle) = z \langle y', y \rangle \langle x', x \rangle = \langle y', y \rangle \, (z \otimes x')x.$$

(d) It follows from (c) that $(x \otimes x')^2 = \langle x', x \rangle \, (x \otimes x')$. As $x \otimes x' \neq 0$, the right hand side if equal to $x \otimes x'$ if and only if $\langle x', x \rangle = 1$. ☐

Sums of rank-1 maps turn out to be very useful to represent projections with finite rank and thus, in particular, projections on finite-dimensional spaces. Getting some practice with computing and drawing projections on $\mathbb{R}^d$ is very important in order to develop a proper intuition for projections. Thus, you can find plenty of exercises about this on Bonus Exercise Sheet A.

3.3 Quotient spaces

Introductory questions.

(a) Let U be a vector subspace of a vector space W over $\mathbb{K}$: Is there always a vector subspace $V \subseteq W$ such that $W = U \oplus V$?

(b) Let X be a closed vector subspace of a Banach space Z. Is there always a closed vector subspace Y such that $Z = X \oplus Y$?

Lecture 12 (Mo, 01 June) is planned to start with Section 4.1.

Definition 3.3.1 (Quotient spaces modulo closed vector subspaces)**.** Let X be a normed space over $\mathbb{K}$ and let $V \subseteq X$ be a vector subspace. The map $\|\cdot\|_/ : X/V \to [0,\infty)$ given by

$$\|x+V\|_/ := \operatorname{dist}(0, x+V) = \operatorname{dist}(x, V)$$

for all $x + V \in X/V$ is called the **quotient semi-norm** on X/V.

Proposition 3.3.2 (Properties of quotient spaces)**.** *Let X be a normed space over $\mathbb{K}$ and let $V \subseteq X$ be a vector subspace.*

(a) *The quotient semi-norm $\|\cdot\|_/ : X \to [0,\infty)$ is indeed a semi-norm, i.e. it is subadditive (i.e. it satisfies the triangle inequality) and it is absolutely homogeneous.*

(b) *Let $x \in X$. One has $\|x+V\|_/ = 0$ if and only if $x \in \overline{V}$.*

(c) *The quotient semi-norm $\|\cdot\|_/ : X \to [0,\infty)$ is a norm if and only if V is closed.*

Assume now in addition that V is closed, i.e. that $(X/V, \|\cdot\|_/)$ is a normed spaces.

(d) *The quotient map $q\colon X \to X/V$, $q(x) := x + V$ is linear and satisfies $q(B_{<1}(0)) = B_{<1}(0)$. In particular q is open and continuous with norm $\|q\| \le 1$.*[2]

(e) *If X is a Banach space, then X/V is a Banach space.*

Proof. (a) This is straightforward to check.

(b) One has $\operatorname{dist}(x,V) = 0$ if and only if there is a sequence (v_n) in V such that $\|v_n - x\| \to 0$ if and only if $x \in \overline{V}$.

(c) Let $\|\cdot\|_/$ be a norm. If $x \in \overline{V}$, then (d) implies $\|x+V\|_/ = 0$ and hence $x + V = 0$, so $x \in V$. Thus, V is closed.

Conversely, let V be closed and let $x \in X$ such that $\|x+V\|_/ = 0$. Then it follows from (d) that $x \in \overline{V} = V$, so $x + V = V = 0 + V$. Thus, $\|\cdot\|_/$ is indeed a norm.

Assume throughout the rest of the proof that V is closed and that $\|\cdot\|_/$ is thus a norm.

[2] And with $\|q\| = 1$ if $V \neq X$.

(d) Clearly q is linear. We show that $q(B_{<1}(0)) = B_{<1}(0)$. If $x \in X$ and $x \in B_{<1}(0)$, then $\|x\| < 1$. So $\|q(x)\|_{/} = \|x+V\|_{/} = \operatorname{dist}(x,V) \le \|x-0\| = \|x\| < 1$ and hence $q(x) \in B_{<1}(0)$. Conversely, let $x+V \in X/V$ and $x+V \in B_{<1}(0)$. Then $\operatorname{dist}(x,V) < 1$, so there exists a vector $v \in V$ such that $\|x-v\| < 1$. One has $x+V = q(x) = q(x-v)$, so $x+V \in q(B_{<1}(0))$.

From the inclusion $\subseteq$ in $q(B_{<1}(0)) = B_{<1}(0)$ it follows that q is continuous and $\|q\| \le 1$. Form the converse inclusion $\supseteq$ it follows that q is open (Exercise 4.3 on Präsenzblatt 4).

(e) Let X be complete. We use Proposition 1.3.9(i) and (ii).

Let (x_n+V) be a sequence in X/V such that $\sum_{n=1}^{\infty} \|x_n+V\|_{/} < \infty$. For each $n \in \mathbb{N}$ we can find a vector $v_n \in V$ such that $\|x_n - v_n\| \le \|x_n+V\|_{/} + \frac{1}{2^n}$. Then $\sum_{n=1}^{\infty} \|x_n - v_n\| < \infty$, so the completeness of X implies that the limit $x := \sum_{n=1}^{\infty}(x_n - v_n)$ exists in X. Hence,

$$\begin{aligned}\left\|(x+V) - \sum_{n=1}^{N}(x_n+V)\right\|_{/} &= \left\|\big(x - \sum_{n=1}^{N}(x_n - v_n)\big) + V\right\|_{/} \\ &\le \left\|x - \sum_{n=1}^{N}(x_n - v_n)\right\| \le \sum_{n=N+1}^{\infty} \|x_v - v_n\| \to 0\end{aligned}$$

as $N \to \infty$, so one has $x+V = \sum_{n=1}^{\infty}(x_n+V)$. □

There is a close connection between quotient spaces direct sum decompositions of vector spaces. If a vector space X satisfies $X = X_1 \oplus X_2$ for two vector subspaces $X_1, X_2 \subseteq X$, then X_1 is isomorphic to X/X_2 via the map $x_1 \mapsto x_1 + X_2$. In the following proposition is formulate this in terms of linear projections (which are, as you know from Proposition 3.2.2(c) and Corollary 3.2.4, an equivalent way to describe direct sum decompositions).

Proposition 3.3.3 (Direct summands vs. quotient spaces)**.**

(a) *Let X be a vector space and $P\colon X \to X$ a linear projection. Then the map $PX \to X/\ker P$, $x \mapsto x + \ker P$ is linear and bijective, i.e. an isomorphism of vector spaces.*

(b) *Let X be a normed space and $P\colon X \to X$ a bounded linear projection. Then the map $PX \to X/\ker P$, $x \mapsto x + \ker P$ is an isomorphism of normed spaces (where PX is endowed with the norm induced from X).*

Proof. (a) Let $r\colon PX \to X/\ker P$, $x \mapsto x + \ker P$, i.e. r is the restriction of the quotient map to the range PX of P. Clearly r is linear.

To show that r is injective, let $x \in PX$ and $x \in \ker r$. Then $x + \ker P = r(x) = 0 = 0 + \ker P$, so $x \in \ker P$, and hence $x = Px = 0$.

To show that r is surjective, let $x + \ker P \in X/\ker P$ for some $x \in V$: Since $Px - x \in \ker P$ one has

$$x + \ker P = Px + \ker P = r(Px),$$

so $x + \ker P$ is indeed located in the range of r.

(b) We again us the notation $r \colon PX \to X/\ker P$, $x \mapsto x + \ker P$. According to (a) r is linear and bijective. Moreover, r is a restriction of the quotient map $X \to X/\ker P$, so it is continuous according to Proposition 3.3.2(d). To show that the inverse map of r is also continuous, let (x_n) be a sequence in PX such that $r(x_n) = x_n + \ker P \to 0$. Then $\operatorname{dist}(x_n, \ker P) \to 0$, so there exists a sequence (y_n) in $\ker P$ such that $\|x_n - y_n\| \to 0$. Hence, $x_n = P(x_n - y_n) \to 0$, which proves that r^{-1} is indeed continuous. $\square$

Let X be a normed space and $V \subseteq X$ a closed vector subspace. Proposition 3.3.3(b) shows that if there exists a closed vector subspace $W \subseteq X$ such that X is the topologically direct sum of V and W, then considering the quotient space X/V is essentially the same as considering the space W. However, we shall see in the exercises that there are examples where there exists no closed vector subspace W such that X is the topologically direct sum of V and W. In such cases, the quotient space X/V can often serve as a very useful replacement for such a subspace W.

Chapter 4

A Taste of Topology

4.1 Topological spaces

Introductory questions.

(a) When is a set in a metric space called open? When is it called closed?

(b) Why are open sets important?

(c) Can you imagine something like a metric space without a metric? And what a weird question is that actually?

Definition 4.1.1 (Tological spaces, open and closed sets, neighbourhoods)**.**

(a) Let M be a set. A **topology** on M is a set $\tau \subseteq 2^M$ with the following properties:

   (I) $\emptyset \in \tau$ and $M \in \tau$.

   (II) τ is stable with respect to finite intersections, i.e. for all $U, V \in \tau$ one has $U \cap V \in \tau$.

   (III) τ is stable with respect to arbitrary unions, i.e. for every family $(U_j)_{j\in J}$ of sets $U_j \in \tau$ one has $\bigcup_{j\in J} U_j \in \tau$.

   A **topological space** is a pair (M,τ) where M is a set and τ is a topology on M.

(b) Let (M,τ) be a topological space. Then the elements of the topology τ are called the **open** subsets of M. A subset $C \subseteq \tau$ is called **closed** if its complement $C^\complement$ is open.

   Let $x \in M$ and $S \subseteq M$. The set S is called a **neighbourhood of x**[1] of there exists an open set U such that $x \in U \subseteq S$. The set S is called an **open neighbourhood of x** if S is itself open and $x \in S$

[1] In German: **Umgebung von x**.

For a topological space (M, τ) one often suppresses the topology τ in the notation and thus simply writes "Let M be a topological space."

Let M be a topological space. Since the union of arbitrarily many open sets is, by definition, open, it follows that the intersection of arbitrarily many closed sets is closed. Similarly, as the intersection of finitely many open sets is open, it follows that the union of finitely many closed sets is closed.

Examples 4.1.2 (Some topological spaces)**.**

(a) Let (M, d) space and let $\tau \subseteq 2^M$ be the set of all those subsets of M that are open in the sense of metric spaces.[2] Then τ is a topology on M; we call it the topology induced by the metric.

We endow every metric space, unless otherwise stated, with the topology induced by its metric.

(b) So in particular every normed space X automatically carries a topology that is induced by the metric that is, in turn, induced by the norm on X.

(c) Let M be a set. Then $\{\emptyset, M\}$ is a topology on M, the so-called **indiscrete topology**. The power set 2^M is also a topology on M, the so-called **discrete topology**.

It is easy to see that the discrete topology on M is induced by the discrete metric. On the other hand, the indiscrete topology on M is not induced by any metric of M has at least two points. This follows from the fact that every singleton $\{x\} \subseteq M$ is closed with respect to every metric, but not with respect to the indiscrete topology (unless $\#M \leq 1$).

(d) Let M be a set. We call a subset $U \subseteq F$ **co-finite** if its complement $U^{\complement} = M \setminus U$ is finite. The set

$$\tau := \{U \subseteq M \mid M \text{is co-finite}\} \cup \{\emptyset\}$$

can be easily check to be a topology on M.

If M is finite, τ is simply the discrete topology. But if M is infinite, we shall see that τ behaves quite weirdly.

Lecture 13 (Mo, 08 June; Pascal) starts with Definition 4.1.3.

Definition 4.1.3 (Closure, interior, and boundary)**.** Let M be a topological space and let $S \subseteq M$.

(a) The **interior of *S*** is defined to be the set

$$S^\circ := \bigcup_{\substack{U \subseteq S, \\ U \text{ open}}} U,$$

i.e. S° is the largest open subset of S.

[2] Recall that a subset U of a metric space M is called **open** if for every $x \in U$ there exists $\varepsilon > 0$ such that $B_{<\varepsilon}(x) \subseteq U$. Equivalently, for every $x \in U$ there exists $\varepsilon > 0$ such that $B_{\leq\varepsilon}(x) \subseteq U$.

(b) The **closure of *S*** is defined to be the set

$$\overline{S} := \bigcap_{\substack{C \supseteq S, \\ C \text{ closed}}} C,$$

i.e. $\overline{S}$ is the smallest closed superset of S.

(c) The **boundary of *S*** is the set $\partial S := \overline{S} \setminus S^\circ$.

Proposition 4.1.4 (Properties of interior and boundary)**.** *Let M be a topological space. Let $S, S_1, \dots, S_n \subseteq M$.*

(a) *If $S_1 \subseteq S_2$, then ${S_1}^\circ \subseteq {S_2}^\circ$ and $\overline{S_1} \subseteq \overline{S_2}$.*

(b) *One has $M \setminus S^\circ = \overline{M \setminus S}$ and $M \setminus \overline{S} = (M \setminus S)^\circ$.*

(c) *One has $\big(S_1 \cap \dots \cap S_n\big)^\circ = {S_1}^\circ \cap \dots \cap {S_n}^\circ$ and $\overline{\big(S_1 \cup \dots \cup S_n\big)} = \overline{S_1} \cup \dots \cup \overline{S_n}$.*

Proof. (a) *and* (b) This follows easily the definition of the interior and the closure.

(c) For each $j \in \{1, \dots, n\}$ one has $S_1 \cap \dots \cap S_n \subseteq S_k$ and thus, by (a), $(S_1 \cap \dots \cap S_n)^\circ \subseteq {S_k}^\circ$. This shows that

$$\big(S_1 \cap \dots \cap S_n\big)^\circ \subseteq {S_1}^\circ \cap \dots \cap {S_n}^\circ.$$

The converse inclusion holds since $S_1 \cap \dots S_n \supseteq {S_1}^\circ \cap \dots \cap {S_n}^\circ$ and the set on the right is open as intersections of finitely many open sets or open by the definition of a topology. ☐

Definition 4.1.5 (Hausdorff spaces)**.** A topological space M is called **Hausdorff** or ***T*₂** if for all distinct points $x, y \in M$ there exists disjoint open sets U, V such that $x \in U$ and $y \in V$.

The Hausdorff (or T_2) property is called a **separation axiom** because it says that any two distinct elements of the space can be **separated** by open sets. Separation axioms play an important role in topology and there are several more of them. Proposition 4.2.6 in the next subsection will demonstrate why the Hausdorff property is extremely important.

Examples 4.1.6 (Hausdorff and non-Hausdorff spaces)**.**

(a) Every metric space (M, d) is Hausdorff.

Indeed, let $x, y \in M$ such that $x \neq y$. For set $r := \frac{1}{2}\,\mathrm{d}(x, y)$ the balls $B_{<r}(x)$ and $B_{<r}(y)$ do not intersect and are open neighbourhoods of x and y, respectively.

(b) If a set M with at least two points carries the indiscrete topology, then it is not Hausdorff since there are no two disjoint non-empty open sets in this space.

(c) Let M be an infinite set and endowed with the topology of co-finite sets from Example 4.1.2(d). Two non-empty open sets in this space are never disjoint, so the space is not Hausdorff.

4.2 Convergence and continuity

Introductory questions.

(a) What are your first thoughts when you hear or read that a function is continuous?

(b) What is the weirdest behaviour of a sequence in $\mathbb{R}$ that you can imagine?

(c) What is the weirdest net in $\mathbb{R}$ that you can imagine?

Definition 4.2.1 (Eventual behaviour of nets)**.** Let $(x_j)_{j\in J}$ be a net in a set M and let $S \subseteq M$. We say that $(x_j)_{j\in J}$ **is eventually contained in *S*** or, more briefly, that $(x_j)_{j\in J}$ **is eventually in *S***, if there exists $j_0 \in J$ such that $x_j \in S$ for all $j \in J$ that satisfy $j \geq j_0$.

Definition 4.2.2 (Convergence of nets)**.** Let $(x_j)_{j\in J}$ be a net in a topological space M.

(a) Let $x \in M$. We say that $(x_j)_{j\in J}$ **converges to *x*** if for every open neighbourhood U of x the net $(x_j)_{j\in J}$ is eventually contained in U. In this case we also say that x is **a limit** of $(x_j)_{j\in J}$.

We denote this property by $x_j \overset{j\in J}{\to} x$ or $x_j \overset{j}{\to} x$, or simply by $x_j \to x$ if there is not risk of ambiguity.

(b) We say that $(x_j)_{j\in J}$ **converges** or, more precisely, **converges in *M*** if there exists a point $x \in M$ such that $(x_j)_{j\in J}$ converges to x.

Note that we consciously choose the article "a" in the phrase "a limit", because part (b) of the following example shows that limits in topological spaces need not be unique, in general (see however Proposition 4.2.6 below).

Examples 4.2.3 (Convergence of nets)**.** (a) Let $(x_j)_{j\in J}$ be a net in a metric space (M, d) and let $x \in M$. Then one can easily check that $x_j \to x$ in the sense of Definition 4.2.2 if and only if $x_j \to x$ in the sense of Definition 1.3.3.

So the convergence notion in topological spaces is consistent with the one that you already know in metric spaces.

(b) Endow $\mathbb{N}$ with the topology of co-finite sets (Example 4.1.2(d)) and consider the sequence $(n)_{n\in\mathbb{N}}$ (recall that every sequence is a net).

Let $k \in \mathbb{N}$ and let $U \subseteq \mathbb{N}$ be an open neighbourhood of k. Then U contains all but finitely many elements of $\mathbb{N}$, so the sequence $(n)_{n\in\mathbb{N}}$ is eventually contained in U.

This shows that $n \overset{n}{\to} k$. Note that $k \in \mathbb{N}$ was arbitrary, so the sequence $(n)_{n\in\mathbb{N}}$ converges to every point in $\mathbb{N}$. In particular, limits are not unique in this space.

Proposition 4.2.4 (Characterisation of the closure)**.** *Let M be a topological space, $S \subseteq M$, and $x \in M$. The following are equivalent:*

(i) *One has $x \in \overline{S}$.*

(ii) *For every open neighbourhood $U \subseteq M$ of x one has $U \cap S \neq \emptyset$.*

(iii) *There exists a net $(x_j)_{j\in J}$ in S such that $x_j \to x$.*

Proof. "(i) ⇒ (ii)" Assume that (ii) fails, i.e. there exists an open neighbourhoods $U \subseteq M$ of x such that $U \cap S = \emptyset$. Then $M \setminus U$ is closed and contains S, so $M \setminus U \supseteq \overline{S}$. Since $x \notin M \setminus U$, it follows that $x \notin \overline{S}$, so (i) does not hold.

"(ii) ⇒ (iii)" Let (ii) hold. Let $\mathcal{U}$ denote the set of all open neighbourhoods of x. Then $\mathcal{U}$ is a directed set with respect to the relation $\leq := \supseteq$ (since the intersection of two open neighbourhoods of x is again an open neighbourhood of x). For each $U \in \mathcal{U}$ there exists a point $x_U \in S \cap U$ according to (ii). So $(x_U)_{U\in\mathcal{U}}$ is a net in S that converges to the point x

"(iii) ⇒ (i)" Assume towards a contradiction that (iii) holds but (i) fails, i.e. there exists a net $(x_j)_{j\in J}$ in S that converges to x, but $x \in U := M \setminus \overline{S}$. Since U is open, there is an index $j \in J$ such that $x_j \in U$. Hence, $x_j \in U \cap S = \emptyset$, which is a contradiction. □

Proposition 4.2.5 (Closedness via convergent nets)**.** *Let M be a topological space and let $S \subseteq M$. The following assertions are equivalent:*

(a) *The set S is closed.*

(b) *For every net $(x_j)_{j\in J}$ in S and every $x \in M$ with $x_j \to x$ one has $x \in S$.*

Proof. "(a) ⇒ (b)" Let S be closed, let $(x_j)_{j\in J}$ be a net in S, let $x \in M$ and assume that $x_j \to x$. The it follows from Proposition 4.2.4 that $x \in \overline{S} = S$.

"(b) ⇒ (a)" Assume that (b) holds and let $x \in \overline{S}$. We need to show that $x \in S$. According to Proposition 4.2.4 there exists a net $(x_j)_{j\in J}$ in S that converges to x. Hence it follows from (b) that $x \in S$. □

Proposition 4.2.6 (Uniqueness of limits)**.** *A topological space M is Hausdorff if and only if each net in M converges to at most one point.*

Proof. "⇒" Let M be Hausdorff, let $(x_j)_{j\in J}$ be a net in M. Assume for a contradiction that $(x_j)_{j\in J}$ converges to two distinct points $x, y \in M$. Due to the Hausdorff property we can find disjoint open sets $U, V \subseteq M$ such that $x \in U$ and $y \in V$.

Hence, there exists indices $j_0, j_1 \in J$ such that $x_j \in U$ for all $j \geq j_0$ and $x_j \in V$ for all $j \geq j_1$. As J is directed, there exists a $j \in J$ such that $j \geq j_0$ and $j \geq j_1$. Thus, $x_j \in U \cap V = \emptyset$, which is a contradiction.

"⇐" Assume that M is not Hausdorff. Then there exist two distinct point $x, y \in M$ and for each open neighbourhood U of x and each open neighbourhood V of y we can find a point $z_{(U,V)} \in U \cap V$.

Let $\mathcal{U}$ and $\mathcal{V}$ denote the sets of all open neighbourhoods of x and y, respectively, and define a relation $\leq$ on the product $\mathcal{U} \times \mathcal{V}$ by

$$(U_1, V_1) \leq (U_2, V_2) \qquad :\Leftrightarrow \qquad U_1 \supseteq U_2 \quad \text{and} \quad V_1 \supseteq V_2.$$

This relation $\leq$ turns $\mathcal{U} \times \mathcal{V}$ into a directed set (since the intersection of two open neighbourhoods of x is again an open neighbourhood of x, and similarly for y). The net $\big(z_{(U,V)}\big)_{(U,V)\in\mathcal{U}\times\mathcal{V}}$ converges to both x and y. □

Lecture 14 (Th, 11 June; Pascal) starts with Definition 4.2.7.

Definition 4.2.7 (Continuity of functions)**.** Let M, N be topological spaces and consider a function $f\colon M \to N$.

(a) Let $x \in M$. The function f is called **continuous at x** if, for each open neighbourhood $V \subseteq N$ of $f(x)$ there exists an open neighbourhood $U \subseteq M$ of x such that $f(U) \subseteq V$.

(b) The function f is called **continuous** if it is continuous at every point $x \in M$.

For the special case of metric spaces you already know the following result from your Analysis 2 course.

Proposition 4.2.8 (Continuity via open preimages)**.** *Let M, N be topological spaces and let $f\colon M \to N$. The function f is continuous if and only if $f^{-1}(V)$ is open in M for every open set $V \subseteq N$.*

Proof. "$\Leftarrow$" Assume that the preimage under f is open for every open set in N. Let $x \in M$ and let $V \subseteq N$ be an open neighbourhood of $f(x)$. Then $U := f^{-1}(V)$ is open by assumption and contains x. Moreover, $f(U) \subseteq V$. So f is continuous at x.

"$\Rightarrow$" Let f be continuous and let $V \subseteq N$ be open. For each $x \in f^{-1}(V)$ one has $f(x) \in V$, so the continuity of f at x implies that there exists an open neighbourhood $U_x \subseteq M$ of x such that $f(U_x) \subseteq V$. Note that $x \in U_x \subseteq f^{-1}(V)$ for each $x \in f^{-1}(V)$, so $f^{-1}(V) = \bigcup_{x\in f^{-1}(V)} U_x$. The set on the right is open since it is a union of open sets. □

Proposition 4.2.9 (Continuity vs. net continuity)**.** *Let M, N be topological spaces and, let $x \in M$ and $f\colon M \to N$. The following are equivalent:*

(i) *The function f is continuous at x.*

(ii) *For each net (x_j) in M that converges to x, the net $(f(x_j))$ in N converges to $f(x)$.*

Proof. "(i) $\Rightarrow$ (ii)" Let f be continuous at x and let $(x_j)_{j\in J}$ be a net in M that converges to x. Let $V \subseteq M$ be an open neighbourhood of $f(x)$. By the continuity of f at x there exists an open neighbourhood U of x such that $f(U) \subseteq V$.

Since $x_j \to x$, there exists $j_0 \in J$ such that $x_j \in U$ for all $j \geq j_0$. Hence, $f(x_j) \in f(U) \subseteq V$ for all $j \geq j_0$, which shows that $f(x_j) \to f(x)$.

"(ii) $\Rightarrow$ (i)" Assume that f is not continuous at x and let $\mathcal{U}$ denote the set of all open neighbourhoods of x: There exists an open neighbourhood V of $f(x)$ such that $f(U) \not\subseteq V$ for each $U \in \mathcal{U}$. Hence, for each $\mathcal{U}$ we can find a point $x_U \in U$ such that $f(x_U) \not\in V$.

Endow $\mathcal{U}$ with the relation $\leq\, := \,\supseteq$. This turns $\mathcal{U}$ into a directed set and $(x_U)_{U\in\mathcal{U}}$ into a net that converges to x. But $\big(f(x_U)\big)_{U\in\mathcal{U}}$ does not converge to $f(x)$. □

Let M be a set and $\mathcal{E} \subseteq 2^M$. It is not difficult to check that the intersection of arbitrarily many topologies on M is again a topology. Hence, there exists a smallest topology on M that contains $\mathcal{E}$, that is given by the intersection of all topologies on M that contain $\mathcal{E}$. In contrast to the situation with σ-algebras that you know from your Analysis 3 course, the smallest topology that contains $\mathcal{E}$ is quite easy to describe:

Proposition 4.2.10 (Smallest topology)**.** *Let M be a set and $\mathcal{E} \subseteq 2^M$. Let $\tau \subseteq 2^M$ of the set of all (finite and infinite) unions of sets of the form*

$$E_1 \cap \dots \cap E_n,$$

where $n \in \mathbb{N}$ and $E_1, \dots, E_n \in \mathcal{E} \cup \{\emptyset, M\}$. Then τ is the smallest topology on M that contains $\mathcal{E}$ as a subset.

Proof. One only needs to check that τ is itself a topology, which is straightforward to do. □

Definition 4.2.11 (Initial topology)**.** Let M, Λ be sets with $\Lambda \neq \emptyset$. For each $\lambda \in \Lambda$ let N_λ be a topological space and $h_\lambda \colon M \to N_\lambda$ a map. Let $\tau \subseteq 2^M$ be the smallest topology on M that satisfies $h_\lambda^{-1}(V) \in \tau$ for all $\lambda \in \Lambda$ and all open sets $V \subseteq N_\lambda$. Then τ is called the **initial topology** of the family $(h_\lambda)_{\lambda \in \Lambda}$.

In other words, the initial topology of the family $(h_\lambda)_{\lambda \in \Lambda}$ is the smallest topology on M that makes all the maps $h_\lambda \colon M \to M_\lambda$ continuous.

Proposition 4.2.12 (Characterisation of the initial topology)**.** *Let M, Λ be sets with $\Lambda \neq \emptyset$. For each $\lambda \in \Lambda$ let N_λ be a topological space and $h_\lambda \colon M \to N_\lambda$ a map. Endow M with a topology τ. Then the following assertions are equivalent:*

(i) *The topology τ is the initial topology of the family $(h_\lambda)_{\lambda \in \Lambda}$.*

(ii) *The set τ consists of all unions of sets of the form*

$$h_{\lambda_1}^{-1}(V_1) \cap \dots \cap h_{\lambda_n}^{-1}(V_n),$$

where $n \in \mathbb{N}$, $\lambda_1, \dots, \lambda_n \in \Lambda$, and $V_1 \subseteq N_{\lambda_1}, \dots, V_n \subseteq N_{\lambda_n}$ are open.

(iii) *For every topological space L and every map $g \colon L \to M$ the following holds: the map g is continuous if and only if $h_\lambda \circ g \colon L \to N_\lambda$ is continuous for every $\lambda \in \Lambda$.*

(iv) *For every point $x \in M$ and every net $(x_j)_{j \in J}$ in M one has $x_j \to x$ if and only if $h_\lambda(x_j) \to h_\lambda(x)$ in N_λ for each $\lambda \in \Lambda$.*

Proof. "(i) ⇔ (ii)" This equivalence follows from Proposition 4.2.10.

"(i) ⇒ (iv)" Let (i) hold, i.e. let τ be the initial topology of $(h_\lambda)_{\lambda \in \Lambda}$. As we have already seen, this implies that (ii) also holds. Let $x \in M$ and let $(x_j)_{j \in J}$ be a net in M. We prove the equivalence stated in (iv):

"⇒" Assume that $x_j \to x$. Since (ii) holds, all the maps h_λ are continuous, so it follows that $h_\lambda(x_j) \to h_\lambda(x)$ for each $\lambda \in \Lambda$ (Proposition 4.2.9).

"⇐" Assume that $h_\lambda(x_j) \to h_\lambda(x)$ for all $\lambda \in \Lambda$. Let $U \subseteq M$ be an open neighbourhood of x. Then it follows from (ii) that there exists as $n \in \mathbb{N}$, indices $\lambda_1, \dots, \lambda_n$ and open sets $V_1 \subseteq N_{\lambda_1}, \dots, V_n \subseteq N_{\lambda_n}$ such that

$$x \in h_{\lambda_1}^{-1}(V_1) \cap \dots \cap h_{\lambda_n}^{-1}(V_n) \subseteq U.$$

Hence, $h_{\lambda_k}(x) \in V_k$ for all $k \in \{1, \dots, n\}$. Since $h_{\lambda_k}(x_j) to h_{\lambda_k}(x)$ for each such k, we can find indices $j_1, \dots, j_n \in J$ such that $h_{\lambda_k}(x_j) \in V_k$ for all $k \in \{1, \dots, n\}$ and all $j \geq j_k$.

As J is directed there exists an index $j_0 \in J$ such that $j_0 \geq j_1, \dots, j_n$. For each $j \geq j_0$ one thus get $h_{\lambda_k}(x_j) \in V_k$ for all $k \in \{1, \dots, n\}$ and thus

$$x_j \in h_{\lambda_1}^{-1}(V_1) \cap \dots \cap h_{\lambda_n}^{-1}(V_n) \subseteq U.$$

This proves that $x_j \to x$.

"(iv) ⇒ (iii)" Let (iv) hold, let L be a topological space, and let $g\colon L \to M$. We prove the equivalence stated in (iii):

"⇒" Let g be continuous and fix $\lambda \in \Lambda$. Let $w \in L$ and let $(w_j)_{j \in J}$ be a net in L that converges to w. As g is continuous, one has $g(w_j) \to g(w)$, so (iv) implies that $(h_\lambda \circ g)(w_j) = h_\lambda(g(w_j)) \to h_\lambda(g(w)) = (h_\lambda \circ g)(w)$. This proves that $h_\lambda \circ g$ is continuous.

"⇐" Assume that $h_\lambda \circ g$ is continuous for each $\lambda \in \Lambda$. Let $w \in L$ and let $(w_j)_{j \in J}$ be a net in L that converges to w. Then $h_\lambda(g(w_j)) = (h_\lambda \circ g)(w_j) \to (h_\lambda \circ g)(w) = h_\lambda(g(w))$ for each $\lambda \in \Lambda$. According to (iv) this implies that $g(w_j) \to g(w)$. Hence, g is continuous.

"(iii) ⇒ (i)" Let $\rho \subseteq 2^M$ denote the initial topology of $(h_\lambda)_{\lambda \in \Lambda}$, i.e. the smallest topology on M that makes all maps $h_\lambda\colon M \to N_\lambda$ continuous. We need to show that $\tau = \rho$.

Consider the identity map $\mathrm{id}\colon M \to M$. It is obviously continuous from (M, τ) to (M, τ), so it follows from (iii) that $h_\lambda = h_\lambda \circ \mathrm{id}$ is continuous from (M, τ) to N_τ for each $\lambda \in \Lambda$. This implies that the topology τ on M makes all the maps h_λ continuous, i.e. $\rho \subseteq \tau$.

On the other hand, the maps $h_\lambda \circ \mathrm{id} = h_\lambda$ is continuous from (M, ρ) to N_λ for each $\lambda \in \Lambda$ by the definition of ρ. So (iii) implies that $\mathrm{id}\colon M \to M$ is continuous from (M, ρ) to (M, τ). For each $U \in \tau$ one thus has $U = \mathrm{id}^{-1}(U) \in \rho$, so $\tau \subseteq \rho$. □

Lecture 15 (Mo, 15 June) is starts with Example 4.2.13.

Example 4.2.13 (Product topology)**.** Let Λ be a non-empty set and let N_λ be a topological space for each $\lambda \in \Lambda$. Set $M := \prod_{\lambda \in \Lambda} N_\lambda$ and let $p_\lambda\colon M \to N_\lambda$, $x \mapsto x_\lambda$ be the λth coordinate map for each $\lambda \in \Lambda$.

The initial topology $\tau \subseteq 2^M$ of $(p_\lambda)_{\lambda \in \Lambda}$ is called the **product topology** on M. According to Proposition 4.2.12 it has the following properties:

(a) The product topology τ consists of all unions of sets of the form

$$\{x \in M \mid x_{\lambda_1} \in V_1, \ldots, x_{\lambda_n} \in V_n\},$$

with $n \in \mathbb{N}$, $\lambda_1, \ldots, \lambda_n \in \Lambda$, and open sets $V_1 \subseteq N_{\lambda_1}, \ldots V_n \in\subseteq N_{\lambda_n}$.

(b) Let L be a topological space and $g\colon L \to M$. For each $\lambda \in \Lambda$ define the coordinate map $g_\lambda\colon L \to N_\lambda$, $g_\lambda(w) := (g(w))_\lambda$ for each $w \in L$. Then g is continuous if and only each g_λ is continuous.

(c) A net $(x_j)_{j\in J}$ converges to a point $x \in M$ if and only if $(x_j)_\lambda \xrightarrow{j} x_\lambda$ for every $\lambda \in \Lambda$.

4.3 Compactness

Introductory questions.

(a) When is a subset S of a metric space (M, d) called **compact**? Can you generalise this concept to topological spaces?

(b) Why is compactness relevant?

(c) How can you check whether a given set is compact?

Definition 4.3.1 (Compactness)**.** Let M be a topological space and let $K \subseteq M$. The set K is called **compact** if the following holds: For every set $J \neq \emptyset$ and every family $(U_j)_{j\in J}$ of open sets $U_j \subseteq M$ that satisfies $\bigcup_{j\in J} U_j \supseteq K$ there exists a finite subset $\emptyset \neq F \subseteq J$ such that $\bigcup_{j\in F} U_j \supseteq K$.

One sometimes summarizes the previous definition by saying that K is compact if and only if every open cover of K admits a finite sub-cover.

Theorem 4.3.2 (Compactness, the finite intersection property, and chains)**.** *Let M be a topological space and let $K \subseteq M$. The following assertions are equivalent:*

(i) *The set K is compact.*

(ii) *For every set $J \neq \emptyset$ and every family $(C_j)_{j\in J}$ of closed sets in M the following holds: if $K \cap \bigcap_{j\in F} C_j \neq \emptyset$ for each finite set $\emptyset \neq F \subseteq J$, then $K \cap \bigcap_{j\in J} C_j \neq \emptyset$.*

(iii) *For every set $J \neq \emptyset$ and every family $(C_j)_{j\in J}$ of closed sets in M, the following holds: If $K \cap C_j \neq \emptyset$ for each $j \in J$ and if $\{C_j \mid j \in J\}$ is a chain with respect to set inclusion, then $K \cap \bigcap_{j\in J} C_j \neq \emptyset$.*

Proof. "(i) ⇒ (ii)" Let (i) holds Let J be a non-empty set, let $(C_j)_{j\in J}$ be a family of closed sets in M, and assume that $K \cap \bigcap_{j\in F} C_j \neq \emptyset$ for each finite set $\emptyset \neq F \subseteq J$. Then $\bigcup_{j\in F}(M \setminus C_j) \not\supseteq K$ for each such F. Since $M \setminus C_j$ is open for each $j \in J$, the compactness of K implies that $\bigcup_{j\in J}(M \setminus C_j) \not\supseteq K$, so $K \cap \bigcap_{j\in J} C_j \neq \emptyset$.

"(ii) ⇒ (i)" Let (ii) hold. Let J be a non-empty set, let $(U_j)_{j\in J}$ be a family of open subsets of M that cover K. Then $K\cap\bigcap_{j\in J}(M\setminus U_j)=\emptyset$, so it follows from (ii) that there exists a finite set $\emptyset\neq F\subseteq J$ such that $K\cap\bigcap_{j\in J}(M\setminus U_j)=\emptyset$. For this F we obtain $\bigcup_{j\in J}U_j\supseteq K$.

"(ii) ⇒ (iii)" Let (ii) hold, let J be non-empty set and let $(C_j)_{j\in J}$ be a family of closed subsets of M such that $K\cap C_j\neq\emptyset$ for each $j\in J$ and such that $\{C_j \mid j\in J\}$ is a chain with respect to set inclusion. For every finite set $\emptyset\neq F\subseteq J$ there exists, due to the chain property, an index $j_0\in F$ such that $C_{j_0}\subseteq C_j$ for all $j\in F$, so $K\cap\bigcap_{j\in F}C_j=K\cap C_{j_0}\neq\emptyset$. Hence, (ii) implies $K\cap\bigcap_{j\in J}C_j\neq\emptyset$.

The proof of "(iii) ⇒ (ii)" is set theoretically a bit involved, and the implication is not used very frequently (we only included it for the sake of completeness). So we do not prove this implication in the lecture.

"(iii) ⇒ (ii)" Let (iii) hold. Let J be a non-empty set, let $(C_j)_{j\in J}$ be a family of closed sets in M, and assume that $K\cap\bigcap_{j\in F}C_j\neq\emptyset$ for each finite set $\emptyset\neq F\subseteq J$. We define

$$\mathcal{M}:=\Big\{I\subseteq J \mid K\cap\bigcap_{j\in I}C_j\neq\emptyset\Big\}\cup\{\emptyset\}\subseteq 2^J.$$

It suffices to show that $J\in\mathcal{M}$. By assumption, every finite subset of J is an element of $\mathcal{M}$. We now show that $\mathcal{M}$ is stable with respect to monotone unions, i.e. that one has $\cup\mathcal{L}\in\mathcal{M}$ for every non-empty chain $\mathcal{L}\subseteq\mathcal{M}$; Bonus Exercise 7.6 on Homework Sheet 7 in Analysis 3 then implies that $\mathcal{M}=2^J$ and thus $J\in\mathcal{M}$.

So let $\emptyset\neq\mathcal{L}\subseteq\mathcal{M}$ be a chain (with respect to set inclusion). Define $D_L:=\bigcap_{j\in L}C_j$ for each $L\in\mathcal{L}$. Then each D_L is a closed subset of M and satisfies $K\cap D_L\neq\emptyset$ since $L\in\mathcal{L}\subseteq\mathcal{M}$. Moreover, if $L_1,L_2\in\mathcal{L}$, then one has $L_1\subseteq L_2$ or vice versa. Thus implies $D_{L_1}\supseteq D_{L_2}$ or vice versa, so $\{D_L \mid L\in\mathcal{L}\}$ is a also chain with respect to set inclusion. Hence, (iii) implies

$$\emptyset\neq K\cap\bigcap_{L\in\mathcal{L}}D_L=K\cap\bigcap_{j\in\cup\mathcal{L}}C_j.$$

So indeed $\cup\mathcal{L}\in\mathcal{M}$, as claimed. □

Proposition 4.3.3 (Continuous images of compact sets)**.** *Let M and N be topological spaces, let $K\subseteq M$ be compact and let $f\colon M\to N$ be continuous. The $f(K)$ is also compact.*

Proof. This follows readily from the definition of compactness and from the fact that pre-images of open sets under a continuous function are open (Proposition 4.2.8). □

Definition 4.3.4 (Universal nets)**.** Let M be a set. A net $(x_j)_{j\in J}$ in M is called a **universal net** if, for each $S\subseteq M$, the net is eventually contained in S or eventually contained in $S^\complement$.

If a net $(x_j)_{j\in J}$ in a set M is eventually constant, i.e. if there exists $j_0\in J$ such that $x_j=x_{j_0}$ for all $j\geq j_0$, then one can readily see that $(x_j)_{j\in J}$ is universal. However, it is not obvious how to find universal nets that are not eventually constant, or whether they exist at all. Proposition 4.3.6 below show that there are actually plenty of them. However, their existence relies heavily on the axiom of choice, so one cannot "explicitly" wrote them down.

Definition 4.3.5 (Subnets)**.** Let S be a set and let $(x_j)_{j\in J}$ be a net in S.

(a) A net $(y_\ell)_{\ell\in L}$ in S is called a **subnet of $(x_j)_{j\in J}$** of there exists a map $\varphi\colon L \to J$ with the following properties:

  (I) One has $x_{\varphi(\ell)} = y_\ell$ for all $\ell \in L$.

  (II) The map φ is **majorizing** in J, i.e. for all $j_0 \in J$ there exists an $\ell_0 \in L$ such that $\varphi(\ell) \geq j_0$ for all $\ell \geq \ell_0$.

(b) A net $(y_\ell)_{\ell\in L}$ in S is called a **universal subnet of $(x_j)_{j\in J}$** if it is a universal net and, at the same time, a subnet of $(x_j)_{j\in J}$.[3]

We will discuss a number of examples of subnets in the exercises. It is easy to check that, if a net $(x_j)_{j\in J}$ in a topological space M converges to a point $x \in M$, then every subnet of $(x_j)_{j\in J}$ also converges to x.

The following result is very instrumental in working with nets. It can be derived from Zorn's lemma by employing so-called **ultrafilters**. The proof is by no means difficult, but it requires introducing a bit of extra terminology. So, to keep the lecture on track towards the more functional analytic topics in the next chapter, we outsource the proof to a bonus exercise.

Proposition 4.3.6 (Existence of universal subnets)**.** *Let S be a set. Every net in S has a universal subnet.*

Proposition 4.3.7 (Compactness via nets)**.** *Let M be a topological space and let $K \subseteq M$. The following are equivalent:*

(i) *K is compact.*

(ii) *Every net in K has a subnet that converges to a point in K.*

(iii) *Every universal net in K converges to a point in K.*

Proof. "(i) ⇒ (iii)" Let K be compact and let $(x_j)_{j\in J}$ be a universal net in K. For each $j \in J$ let C_j denote the closure of $\{x_k \mid k \in J, k \geq j\}$ in M. For every finite set $\emptyset \neq F \subseteq J$ there exists, be the directedness of J, an index $j_0 \in J$ such that $j_0 \geq j$ for all $j \in F$. Hence, $x_{j_0} \in K \cap \bigcap_{j\in F} C_j$, so that intersection is non-empty. According to Theorem 4.3.2 the compactness of K implies that there exists a point $x \in K \cap \bigcap_{j\in J} C_j$.

We show that $(x_j)_{j\in J}$ converges to x, so let $U \subseteq M$ be an open neighbourhood of x. For each $j \in J$ one has $x \in C_j = \overline{\{x_k \mid k \in J, k \geq j\}}$, so there exists a index $k \geq j$ such that $x_k \in U$ (Proposition 4.2.4). Hence, the net $(x_j)_{j\in J}$ is not eventually contained in $U^\complement$. As the net is universal, it follows that it is eventually contained U.

"(iii) ⇒ (ii)" Let (iii) hold and let $(x_j)_{j\in J}$ be a net in K. By Proposition 4.3.6 the net $(x_j)_{j\in J}$ has a universal subnet, and that subnet converges to a point in K according to (iii).

[3] So"universal" in "universal subnet" is indeed meant as an adjective that refers to the noun "net", not as an adverb that refers to the prefix "sub".

"(ii) ⇒ (i)" Assume that (ii) holds. To show that K is compact, we prove property (ii) in Theorem 4.3.2. So let J be a non-empty set and let $(C_j)_{j\in J}$ be a family of closed sets $C_j \subseteq M$ such that there exists a point $x_F \in K \cap \bigcap_{j\in F} C_j$ for each finite set $\emptyset \neq F \subseteq J$. We need to show that $K \cap \bigcap_{j\in J} C_j \neq \emptyset$.

The set $\{F \subseteq J \mid F \neq \emptyset, F \text{ is finite}\}$ is directed with respect to set inclusion, so $(x_F)_{F\in\mathcal{F}}$ is a net in K. It has, according to (ii), a subnet $(y_\ell)_{\ell\in L}$ that converges to a point $y \in K$. It suffices to prove that $y \in C_j$ for each $j \in J$, so fix $j \in J$.

According to Definition 4.3.5(a) there exists a majorizing map $\varphi\colon L \to \mathcal{F}$ such that $y_\ell = x_{\varphi(\ell)}$ for all $\ell \in L$. As φ is majorizing, there exists an index $\ell_0 \in L$ such that $\varphi(\ell) \subseteq \{j\}$ for all $\ell \geq \ell_0$. For each $\ell \geq \ell_0$ one thus has

$$y_\ell = x_{\varphi(\ell)} \in K \cap \bigcap_{i\in\varphi(\ell)} \subseteq C_j,$$

i.e. the net $(y_\ell)_{\ell\geq\ell_0}$ as contained in C_j. This net converges to y, and since C_j is closed we conclude that $y \in C_j$, as claimed. □

Lecture 16 (Th, 18 June) starts with Theorem 4.3.8.

Theorem 4.3.8 (Tychonoff: Compactness of product spaces)**.** *Let A be a non-empty set. For each $\alpha \in A$ let K_α be a compact topological space. Then $\prod_{\alpha\in A} K_\alpha$, endowed with the product topology, is also compact.*

Proof. Set $K := \prod_{\alpha\in A} K_\alpha$. For each $\alpha \in A$, let $p_\alpha\colon K \to K_\alpha$, $x \mapsto x(\alpha)$ denote the αth coordinate map.

Now let $(x_j)_{j\in J}$ be a universal net in K. According to Proposition 4.3.7 it suffices to show that this net converges to a point in K. For each $\alpha \in A$ the net $(x_j(\alpha))_{j\in J} = \big(p_\alpha(x_j)\big)_{j\in J}$ is a universal net in K_α (since it is the image of a universal net under a map, see Exercise ...). As K_α is compact by assumption, Proposition 4.3.7 implies that $(x_j(\alpha))_{j\in J}$ converges to a point $x(\alpha) \in K_\beta$. Set $x := (x(\alpha))_{\alpha\in A}$. Then $(x_j)_{j\in J}$ converges to x componentwise and hence with respect to the product topology. □

Chapter 5

Topologies in Functional Analysis

5.1 An (extremely) short glimpse at topological vector spaces

Introductory questions.

(a) What do the Sections 4.1–4.3 have to do with functional analysis?

(b) Do you know a norm on the vector space $\mathbb{R}^{[0,1]}$ such that the induced convergence notion is pointwise convergence?

Definition 5.1.1 (Topological vector spaces)**.** A **topological vector space** is a pair (X,τ), where X is a vector space over $\mathbb{K}$ and τ is a topology on X with the following properties:

(I) The addition $+\colon X \times X \to X$ is continuous when $X \times X$ is endowed with the product topology.

(II) The scalar multiplication $\cdot\colon \mathbb{K} \times X \to X$ is continuous when $\mathbb{K} \times X$ is endowed with the product topology (where $\mathbb{K}$ is endowed with the Euclidean topology, as usual).

As usual, we will often only write "Let X be a topological vector space" and thus suppress the topology τ in the notation.

Example 5.1.2 (Spaces with pointwise convergence)**.** Let S be a set and endow $\mathbb{K}^S$ with the pointwise addition and scalar multiplication, and with the product topology. This is a topological vector space.

In the preceding example, the coordinate map $\mathbb{K}^S \to \mathbb{K}$, $f \mapsto f(s)$ is continuous for each $s \in S$. Note how this is in contrast to Bonus Exercise 1.6(b) and (c) on Homework Sheet 1.

In Example 5.1.5 we shall see that not every "reasonable" notion of convergence stems from a topology. The following lemma is a preparation towards this example (and it also turns out to be useful in many other situations).

Lemma 5.1.3 (Hair splitting lemma for sequences)**.** *Let M be topological space, let (x_n) be a sequence in X and let $x \in X$. The sequence (x_n) converges to x if and only if every subsequence of (x_n) has a subsequence that converges to x.*

Proof. "⇒" This implication is clear.

"⇐" We show the contrapositive implication, so assume that (x_n) does not converges to x. Then there exists an open neighbourhood U of x such that (x_n) is not eventually contained in U. Thus, we can recursively construct indices $n_1 < n_2 < n_3 < \dots$ such that $x_{n_k} \not\in U$ for all $k \in \mathbb{N}$. In particular, $(x_{n_k})_{k\in\mathbb{N}}$ is a subsequence of (x_n) and is contained in $U^{\complement}$, so no subsequence of (x_{n_k}) converges to x. □

It is not difficult to show that an analogous version of the hair splitting lemma for nets is also true. To phrase the subsequent Example 5.1.5 in an efficient way, we use the following terminology.

Definition 5.1.4 (Almost everywhere convergence for equivalence classes of functions)**.** Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and $p \in [1,\infty]$. Let $\tilde{f} \in L^p := L^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ and let $(\tilde{f}_n)$ be a sequence in L^p. We say that $(\tilde{f}_n)$ **converges to $\tilde{f}$ almost everywhere** if the sequence (f_n) of representatives converges to the representative f almost everywhere (observe that this does not depend on the choice of the representatives).

The following example shows that the notion of almost everywhere convergence of sequences in L^p does not step from a topology, in general.

Example 5.1.5 (Almost everywhere convergence is not induced by a topology)**.** Let $p \in [1,\infty)$ and abbreviate $L^p := L^p([0,1), \mathcal{B}([0,1), \lambda_1; \mathbb{R})$. Consider the sequence $(\tilde{f}_n)$ in L^p given by the functions

$$\begin{aligned}
&f_1 := \mathbb{1}_{[0,1)}, \\
&f_2 := \mathbb{1}_{[0,\frac{1}{2})}, \quad f_3 := \mathbb{1}_{[\frac{1}{2},1)}, \\
&f_4 := \mathbb{1}_{[0,\frac{1}{4})}, \quad f_5 := \mathbb{1}_{[\frac{1}{4},\frac{1}{2})}, \quad f_6 := \mathbb{1}_{[\frac{1}{2},\frac{3}{4})}, \quad f_7 := \mathbb{1}_{[\frac{3}{4},1)}, \\
&\dots
\end{aligned}$$

So, more precisely speaking, for each $\ell \in \mathbb{N}_0$ and each $k \in \{0,\dots,2^\ell - 1\}$ we set $f_{2^\ell+k} := \mathbb{1}_{[\frac{k}{2^\ell},\frac{k+1}{2^\ell})}$

For each $x \in [0,1)$ the scalar sequence $(f_n(x))$ has a subsequence that converges to 0 and a subsequence that converges to 1, so $(f_n(x))$ does not converge in $\mathbb{R}$. So, the sequences of representatives (f_n) converges at no point in $[0,1)$; in particular, it does not converges almost everywhere, and thus the sequence $(\tilde{f}_n)$ in L^p does no converges almost everywhere according to Definition 5.1.4.

However, it easy to see that $(\tilde{f}_n)$ converges to 0 with respect to the norm $\|\cdot\|_p$. Now consider any subsequence $(\tilde{f}_{n_k})$. Of course, this subsequence also converges to 0 in norm, so according to the partial converse of the dominated convergence theorem given in Theorem B.2.3 in the Appendix,[1] it follows that $(\tilde{f}_{n_k})$ has a subsequence that converges to 0 alsmost everywhere.

So to sum up, every subsequence of $(\tilde{f}_n)$ has a subsequence that converges to 0 almost everyhwere. If almost everywhere convergence were induced by a topology on L^p, the hairsplitting Lemma 5.1.3 would imply that $(\tilde{f}_n)$ itself converges to 0 almost everywhere. But, as we have observed, it does not.

By the way, the same argument shows that pointwise almost everywhere convergence on $\mathcal{L}^p := \mathcal{L}^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ is not induced by a topology. This is in sharp contrast to pointwise convergence (i.e. without the qualifier "almost everywhere") on $\mathcal{L}^p$, which is induced by the initial topology of all the point evaluation maps $\mathcal{L}^p \to \mathbb{K}$, $f \mapsto f(\omega)$ for $\omega \in \Omega$.

5.2 Compactness and relative compactness in vector spaces

Introductory questions.

(a) Which subsets of $\mathbb{R}^d$ are compact? Which subsets of $\mathbb{R}^d$ have compact closure?

(b) Can you given an example of an infinite subset of $C([0,1];\mathbb{K})$ that is compact?

(c) Can you give an example of a subset of $C([0,1];\mathbb{K})$ that is bounded and closed but not compact?

Definition 5.2.1 (Relative compactness)**.** Let X be a topological Hausdorff space. A set $S \subseteq X$ is called **relatively compact** if its closure $\overline{S}$ is compact.[2]

The proof of the following proposition is a good exercise to get used to the concept of relative compactness. We thus discuss it in the exercises.

Proposition 5.2.2 (Simple properties of relatively compact sets)**.** *Let M, N be topological Hausdorff spaces.*

(a) *A set $S \subseteq M$ is relatively compact if and only if there exists a compact set $K \subseteq M$ such that $S \subseteq K$.*

(b) *If $S \subseteq M$ is relatively compact and $f\colon M \to N$ is continuous, then $f(S)$ is also relatively compact.*

Lecture 17 (Mo, 22 June) starts with Proposition 5.2.3.

Proposition 5.2.3 (Compactness versus relative compactness)**.** *Let M be a topological Hausdorff space. A set $S \subseteq M$ is compact if and only if it is closed and relatively compact.*

Proof. "$\Rightarrow$" If S is compact, then it is closed according to Exercise 10.1(a) on Präsenzblatt 10. Hence, $\overline{S} = S$, i.e. S has compact closure and is thus relatively compact.

[1] You have already used Theorem B.2.3 in the exercises.

[2] One should probably be careful when trying to define relative compactness also for non-Hausdorff spaces. See for instance *this MathOverflow question* by the lecturer and *this comment* on the question by Martin Väth.

"$\Leftarrow$" If S is closed and relatively compact, then $S = \overline{S}$ is compact. $\square$

Recall that we characterized compactness of subsets of topological spaces in terms of nets in Proposition 4.3.7. It is natural to ask whether a similar characterization holds for relative compactness. It turns out that this is not true in all topological spaces (see e.g. [11, Example 78] for a counterexample), but it is true in many important cases, namely for the following class of topological spaces:

Definition 5.2.4 (T3-spaces)**.** A topological space M is called a **T3-space** if it is Hausdorff (i.e. T2) and, in addition, the following separation axiom holds: For each closed set $C \subseteq M$ and each $x \in M \setminus C$ there exist disjoint open sets $U, V \subseteq M$ such that $x \in U$ and $C \subseteq V$.

Examples 5.2.5 (Important classes of T3-spaces)**.**

(a) Every metric space is T3.

(b) Every topological vector space that is Hausdorff is T3.

Proof. (a) Let (M, d) be a metric space; clearly, it is Hausdorff. Let $C \subseteq M$ be closed and $x \in C^{\complement}$. According to Proposition 3.1.4(a) one has $\varepsilon := \operatorname{dist}(x, C) > 0$, so $B_{\leq \varepsilon/2}(x)$ does not intersection C. Hence, $U := B_{<\varepsilon/2}(x)$ and $V := \big(B_{\leq \varepsilon/2}(x)\big)^{\complement}$ are disjoint open sets in M such that former contains x and the later contains C.

(b) Let X be a Hausdorff topological vector space. The continuity of the addition implies that, for every $z \in X$, the translation map $X \to X$, $y \mapsto y + z$ is a homeomorphims.[3] Hence, for every set $U \subseteq X$ and every vector $z \in X$, the set U is open if and only if $z + U$ is open. We will use this several times in the rest of the argument.

Let $C \subseteq X$ be closed and $x \in C^{\complement}$. Then $C^{\complement} - x$ is an open neighbourhood of 0. Due to the continuity of the addition $+\colon X \times X \to X$ at $(0,0)$, there exists an open neighbourhood $W \subseteq X \times X$ of $(0,0)$ such that $y_1 + y_2 \in C^{\complement} - x$ for all $(y_1, y_2) \in W$. According to the description of the product topology in Example 4.2.13(a) we can find open $U_1, U_2 \subseteq X$ such that $(0,0) \in U_1 \times U_2 \subseteq W$. So the set $U := U_1 \cap U_2 \subseteq X$ is an open neighbourhood of 0 and satisfies $U \times U \subseteq W$ and thus $U + U \subseteq C^{\complement} - x$, i.e. $U + U + x \subseteq C^{\complement}$.

Define $V := C - U = \{c - u \mid c \in C, u \in U\} = \bigcup_{c \in C}(c - U)$. Then V contains C and is open since $-U$ is open (here we used that scalar multiplication is continuous in topological vector spaces). On the other hand, $x + U$ is open and contains x. Finally, observe that $(x + U) \cap V = \emptyset$, for if y were an element of this intersection we could write it as $y = x + u_1 = c - u_2$ for elements $u_1, u_2 \in U$ and $c \in C$, which gives $U + U \ni x + u_1 + u_2 = c \in C$, a contradiction. $\square$

We note in passing a stronger result than stated in Examples 5.2.5 actually holds: each metric space and each Hausdorff topological vector space is even a so-called **T3$\frac{1}{2}$-space**, which is a very well-behaved class of topological spaces. However, we refrain from discussing this concept here.

[3] Recall from the exercises that a map between two topological spaces is called a **homeomorphism** if it is continuous and bijective and its inverse is also continuous.

Proposition 5.2.6 (Characterization of T3-spaces)**.** *A topological Hausdorff space is T3 if and only if the following formula holds for each closed set $C \subseteq M$:*

$$C = \bigcap_{C \subseteq V \text{ open}} \overline{V}$$

Proof. This follows easily from the definition of T3-spaces. □

The reason why we introduced T3-spaces is that, in those spaces, a similar characterisation as Proposition 4.3.7 also holds for relative compactness:

Proposition 5.2.7 (Relative compactness via subnets and universal nets)**.** *Let M be a topological space that is T3 and let $S \subseteq M$. The following assertions are equivalent:*

(i) *S is relatively compact.*

(ii) *Every net in S has a subnet that converges to a point in X.*

(iii) *Every universal net in S converges to a point in X.*

Proof. *"(i) ⇒ (iii)"* Let S be relatively compact and let (x_j) be a universal net in S. As $\overline{S}$ is compact and contains (x_j), it follows from Proposition 4.3.7 that (x_j) converges to a point in $\overline{S}$.

"(iii) ⇒ (ii)" This follows from the fact that every net has a universal subnet (Proposition 4.3.6).

"(ii) ⇒ (i)" Let (ii) holds. We show that the closure $\overline{S}$ satisfies condition (ii) in Theorem 4.3.2 and is thus compact. So let $j \in J$ be a non-empty index set and let $(C_j)_{j\in J}$ be a family of closed subsets of M such that $\overline{S} \cap \bigcap_{j\in F} C_j \neq \emptyset$ for each finite set $\emptyset \neq F \subseteq J$.

Let $L := \{(j,V) \mid j \in J,\ C_j \subseteq V \text{ open}\}$ and consider the family of closed sets $(\overline{V})_{(j,V)\in L}$ in M. For every finite set $\emptyset \neq F \subseteq L$ one has $\emptyset \neq \overline{S} \cap \bigcap_{(j,V)\in F} V$. According to the characterization of the closure $\overline{S}$ in Proposition 4.2.4(ii) we can thus find a point

$$x_F \in S \cap \bigcap_{(j,V)\in F} V \subseteq \overline{S} \cap \bigcap_{(j,V)\in F} \overline{V}$$

for each such F. Then (x_F) is a net in S, where F runs through the sets of all non-empty finite subsets of L. According to property (ii) the net (x_F) has a subnet $(z_i)_{i\in I}$ that converges to a point $x \in M$. For each $(j,V) \in L$ the net $(z_i)_{i\in I}$ is eventually contained in $\overline{S} \cap \overline{V}$, and thus $x \in \overline{S} \cap \overline{V}$. We conclude that

$$\emptyset \neq \bigcap_{(j,V)\in L} \overline{S} \cap \overline{V} = \overline{S} \cap \bigcap_{(j,V)\in L} \overline{V} = \overline{S} \cap \bigcap_{j\in J} \bigcap_{\substack{C_j \subseteq V, \\ V \text{ open}}} \overline{V} = \overline{S} \cap \bigcap_{j\in J} C_j,$$

where the last inequality holds since M is T3 (Proposition 5.2.6). So $\overline{S}$ is indeed compact. □

The reason why we introduced T3-spaces and proved Proposition 5.2.7 in the present section is that every Hausdorff topological vector space is T3 (Example (b)), so we now know that relative compactness in Hausdorff topological vector spaces can be characterized by the conditions in Proposition 5.2.7. In fact, we could have also shown this directly by a slightly different argument that does not rely on the notion T3-space – but then it would appear quite obscure why the same result also holds in metric spaces.

Next, let us discuss that (relative) compactness is not as ubiquitous in infinite-dimensional spaces as it is in finite-dimensional spaces. To show this we use the following auxiliary result.

Lemma 5.2.8 (Riesz-Lemma)**.** *Let X be a normed space and let $V \subsetneq X$ be a closed vector subspace. Then there are points in the unit sphere of X whose distance to V is almost* 1*; more precisely speaking, one has*

$$\sup_{x\in X,\, \|x\|=1} \operatorname{dist}(x,V) = 1.$$

Proof. *"≤"* For each $x \in X$ of norm $\|x\| = 1$ one has $\operatorname{dist}(x,V) \le \|x - 0\| = 1$.

"≥" As V is closed, the induced seminorm $\|\cdot\|_/$ on the quotient space X/V is a norm and since $V \neq X$, the quotient space X/V is non-zero. So there exists an $y \in X$ such that $\operatorname{dist}(y,V) = \left\|y + V\right\|_/ = 1$. Hence, we can find a sequence (v_n) in V such that $\left\|y - v_n\right\| \downarrow 1$. Set $x_n := \frac{y-v_n}{\|y-v_n\|} \in X$ for each index n. Then $\|x_n\| = 1$ for each n and

$$\operatorname{dist}(x_n,V) = \|x_n + V\|_/ = \frac{1}{\left\|y - v_n\right\|}\left\|y+V\right\|_/ = \frac{1}{\left\|y - v_n\right\|} \to 1,$$

which proves the supremum under consideration is indeed at least 1. □

Corollary 5.2.9 (Compactness of the unit ball)**.** *Let X be a normed spae. The following are equivalent:*

(i) *One has* $\dim X < \infty$.

(ii) *Every bounded subset of X is relatively compact.*

(iii) *Every closed and bounded subset of X is compact.*

(iv) *The closed unit ball $B_{\le 1}(0)$ is compact.*

Proof. *"(ii) ⇔ (iii)"* This follows from Proposition 5.2.3 and that fact that a subset of X is bounded if and only if its closure is bounded.

"(i) ⇒ (iii)" This is a result from Analysis 2.

"(iii) ⇒ (iv)" This implication is obvious.

"(iv) ⇒ (i)" We show the contrapositive implication. Let $\dim X = \infty$. Since every finite-dimensional vector subspace of X is closed (Example 1.1.2), we can use the Riesz lemma (Lemma 5.2.8) to construct a sequence $(x_n)_{n\in\mathbb{N}}$ in X that satisfies $\|x_n\| = 1$ and

$$\operatorname{dist}(x_n, \operatorname{span}\{x_k \mid k < n\}) \geq \frac{1}{2}$$

for all $n \in \mathbb{N}$. Hence, one has $\|x_n - x_m\| \geq \frac{1}{2}$ for all distinct $m, n \in \mathbb{N}$. Thus, no subsequence of (x_n) is Cauchy; in particular, no subsequence of (x_n) is convergent. As (x_n) is located in $B_{\leq 1}(0)$, it follows that $B_{\leq 1}(0)$ is not compact. □

The previous corollary shows that, in every infinite-dimensional normed space, there exists a bounded set that is not relatively compact. It thus becomes an interesting question to characterize those sets that are relatively compact (or compact) in important classes of spaces. One such result is the **Arzelà–Ascoli theorem** that we prove in Theorem 5.2.11 below. To formulate the theorem we need the following notion.

Lecture 18 (Th, 25 June) is planned to start with Definition 5.2.10.

Definition 5.2.10 (Equi-continuity)**.** Let M be a topological space and Z a normed vector space, and let $\mathcal{F}$ be a set of functions from M to Z.

(a) Let $x \in M$. The set $\mathcal{F}$ is called **equicontinuous**[4] at x if the following holds: for every $\varepsilon > 0$ there exists an (open) neighbourhood $U \subseteq M$ of x such that $\|f(y) - f(x)\| < \varepsilon$ for all $y \in U$ and all $f \in \mathcal{F}$.[5]

(b) The set $\mathcal{F}$ called **equicontinuous** if it is equicontinuous at every $x \in M$.

Theorem 5.2.11 (Arzelà–Ascoli: Relative compactness in **C(*M*; *Z*)**)**.** *Let M be a compact topological Hausdorff space, let Z be a normed space, and endow the space* $\mathrm{C}(M;Z)$ *of continuous functions from M to Z with the ∞-norm given by $\|f\|_\infty := \sup\{\|f(x)\|_Z \mid x \in M\}$ for all $f \in \mathrm{C}(M;Z)$. For every set $\mathcal{F} \subseteq \mathrm{C}(M;Z)$ the following are equivalent:*

(a) *The set $\mathcal{F}$ is relatively compact.*

(b) *The set $\mathcal{F}$ is equicontinuous and for every $x \in M$ the set $\mathcal{F}(x) := \{f(x) \mid f \in \mathcal{F}\}$ is relatively compact in Z.*

Proof. "(a) ⇒ (b)" Let $\mathcal{F}$ be relatively compact. Fix $x \in M$. To show that $\mathcal{F}$ is equicontinuous at x, it suffices to show that the closure $\overline{\mathcal{F}}$ is equicontinuous at x. So fix $\varepsilon > 0$. For every open neighbourhood $U \subseteq M$ of x, define the set

$$\mathcal{G}_U := \{f \in \mathrm{C}(M;Z) \mid \|f|_U - f(x)\,\mathbb{1}_U\|_\infty < \varepsilon\}.$$

This set is open since the map $\mathrm{C}(M;Z) \to \mathrm{C_b}(U;Z)$, $f \mapsto f|_U$ is continuous with respect to the ∞-norm on both spaces. Moreover, each $f \in \mathrm{C}(M;Z)$ is contained in $\mathcal{G}_U$

[4] In German: **gleichgradig stetig**

[5] The same definition also makes sense in the more general situation in which Z is a metric space instead of a normed space. In this case one has to replace $\|f(y) - f(x)\|$ with $\mathrm{d}(f(y), f(x))$, of course.

for some open neighbourhood of U of x since f is continuous at x. So the open sets $\mathcal{G}_U$ cover $C(M;Z)$ and thus, in particular, $\overline{\mathcal{F}}$ as U runs through the open neighbourhoods of x. Since $\overline{\mathcal{F}}$ is compact, we can find finitely many open neighbourhoods $U_1,\dots,U_n$ of x such that $\mathcal{G}_{U_1}\cup\dots\cup\mathcal{G}_{U_n}\supseteq\overline{\mathcal{F}}$.

The set $U_0 := U_1\cap\dots\cap U_n$ is also an open neighbourhood of x and for all $f\in\overline{\mathcal{F}}$ one has $\left\|f|_{U_0} - f(x)\,\mathbb{1}_{U_0}\right\|_\infty < \varepsilon$. This means precisely that $\left\|f(y)-f(x)\right\|_Z < \varepsilon$ for all $y\in U$ and all $f\in\overline{\mathcal{F}}$. So $\overline{\mathcal{F}}$ is indeed equicontinuous at x.

To see the second property claimed in (b), just observe that the point evaluation $C(M;Z)\to Z$, $f\mapsto f(x)$ is continuous, and thus sends relatively compact sets to relatively compact sets by Proposition 5.2.2(b).

"(b) $\Rightarrow$ (a)" Assume that (b) holds and let (f_j) be a universal net in $\mathcal{F}$. According to Proposition 5.2.7 and Example 5.2.5(a) (or (b)) it suffices to show that (f_j) converges uniformly to a function $g\in C(M;Z)$. For every $x\in M$, $(f_j(x))$ is a universal net in the relatively compact subset $\mathcal{F}(x)$ of Z. Hence, it converges to a point $g(x)\in Z$.

We now use the equicontinuity of $\mathcal{F}$ to show that g is continuous and that the pointwise convergence of (f_j) to g is even uniform, which proves the claim.

So let $\varepsilon>0$. For each $x\in M$ we can, by the equicontinuity of $\mathcal{F}$, find an open neighbourhood $U_x\subseteq M$ of x such that, for every $y\in U_x$, one has

$$\left\|f(y)-f(x)\right\|_Z\le\varepsilon\quad\text{for all } f\in\mathcal{F},\qquad\text{and hence also}\qquad \left\|g(y)-g(x)\right\|_Z\le\varepsilon.$$

This already shows that g is continuous. The open sets U_x cover M, so by the compactness of M we can find points $x_1,\dots,x_n\in M$ such that $U_{x_1}\cup\dots\cup U_{x_n}=M$. The pointwise convergence of f_j to g gives is an index j_0 such that $\left\|f_j(x_k)-g(x_k)\right\|_Z\le\varepsilon$ for all $j\ge j_0$ and all $k\in\{1,\dots,n\}$.

For every $y\in M$ there is $k\in\{1,\dots,n\}$ such that $y\in U_{x_k}$, and hence

$$\left\|f_j(y)-g(y)\right\|_Z\le\left\|f_j(y)-f_j(x_k)\right\|_Z+\left\|f_j(x_k)-g(x_k)\right\|_Z+\left\|g(x_k)-g(y)\right\|_Z\le 3\varepsilon$$

for every $j\ge j_0$. Thus, $\left\|f_j-g\right\|_\infty\le 3\varepsilon$ for all $j\ge j_0$. □

Example 5.2.12. Consider real numbers $a<b$ and endow $C([a,b];\mathbb{K})$ with the norm $\|\cdot\|_\infty$, as usual. Let $c\in(0,\infty)$. Then the set

$$\mathcal{F}:=\left\{f\in C^1([a,b];\mathbb{K})\mid \left|f(a)\right|\le c,\ \left\|f'\right\|_\infty\le c\right\}\subseteq C([a,b];\mathbb{K})$$

is relatively compact.

Proof. We check that $\mathcal{F}$ satisfies condition (b) in the Arzelà–Ascoli Theorem 5.2.11. Let $x\in[a,b]$.

For every $f\in\mathcal{F}$ we have

$$\left|f(x)\right|=\left|f(a)+\int_a^x f'(t)\,\mathrm{d}t\right|\le c+\int_a^b\left\|f'\right\|\,\mathrm{d}t=c(1+b-a).$$

So $\{f(x) \mid f \in \mathcal{F}\}$ is bounded and thus relatively compact in $\mathbb{K}$.

To show that $\mathcal{F}$ is equicontinuous at x, let $\varepsilon > 0$ and let $U := B_{<\varepsilon/c}(x)$, where the ball is taken within the interval $[a, b]$. For all $y, z \in U$ with $y \le z$ and all $f \in \mathcal{F}$ one has

$$\left|f(w) - f(v)\right| = \left|\int_v^w f'(t)\,\mathrm{d}t\right| \le \int_v^w c\,\mathrm{d}t = (w - v)c < \varepsilon.$$

So for all $f \in \mathcal{F}$ and all $y \in U$ one has $\left|f(y) - f(x)\right| < \varepsilon$ (take $(w, v) = (x, y)$ if $y \le x$ and $(v, w) = (y, x)$ if $x < y$). Hence, $\mathcal{F}$ is indeed equicontinuous at x. □

5.3 Dual operators and the weak* topology

Introductory questions.

(a) Consider a matrix $A \in \mathbb{R}^{d\times c}$ as a linear operator $\mathbb{R}^c \to \mathbb{R}^d$. How can one interpret the transposed matrix A^{T} from an operator theoretic point of view?

(b) Let (f_n) be a bounded sequence in $\ell^\infty(\mathbb{N};\mathbb{K})$. When is it true that $\langle f_n, g\rangle \to 0$ for every $g \in \ell^1(\mathbb{N};\mathbb{K})$?

Definition 5.3.1 (Dual operators)**.** Let X, Y be normed spaces over $\mathbb{K}$ and let $T\colon X \to Y$ be a bounded linear operator. The map $T'\colon Y' \to X'$, $y' \mapsto T'y' := y'T$ is called the **dual operator** of T; here, $y'T = y' \circ T\colon X \to \mathbb{K}$ denotes the composition of y' and T as usual.

The notation for dual operators interacts very well with the notation $\langle\cdot,\cdot\rangle$ for the application of bounded linear functions: In the situation of Definition 5.3.1 one has

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

for all $x \in X$ and $y' \in Y'$.

Proposition 5.3.2 (Properties of dual operators)**.** *Let W, X, Y be normed spaces over $\mathbb{K}$ and let $S, T\colon X \to Y$ be bounded linear operators.*

(a) *The dual operator $T'\colon Y' \to X'$ is also a bounded linear operator and $\|T'\| = \|T\|$.*

(b) *For all $\alpha, \beta \in \mathbb{K}$ one has $(\alpha S + \beta T)' = \alpha S' + \beta T'$.*

(c) *Let $R\colon W \to X$ also be a bounded linear operator. Then $(TR)' = R'T'$.*

(d) *Assume now that X and Y are Banach spaces. Then T is bijective if and only if its dual operator T' is bijective, and in this case one has $(T^{-1})' = (T')^{-1}$.*

Proof. (b) *and* (c) This can be checked by straightforward computations.

(a) Linearity of T' from the definition of the vector space operations on dual spaces. Moreover, for each $y' \in Y'$ one has $\|T'y'\| = \|y'T\| \le \|T\| \|y'\|$. This proves that T' is bounded and satisfies $\|T'\| \le \|T\|$.

To show the converse inequality, choose a sequence (x_n) in X that satisfies $\|x_n\| \le 1$ for all n and $\|Tx_n\| \to \|T\|$. According to Corollary 2.4.9 we can find, for each index n, an functional $y'_n \in Y'$ such that $\|y'_n\| \le 1$ and $|\langle y'_n, Tx_n\rangle| = \|Tx_n\|$. Hence,

Lecture 19 (Mo, 29 June) starts with the proof of the inequality $\|T'\| \ge \|T\|$ in Proposition 5.3.2(a).

$$\|T'\| \ge \|T'y'_n\| = \|y'_n T\| \ge |y'_n T x_n| = |\langle y'_n, Tx_n\rangle| = \|Tx_n\| \to \|T\|,$$

which proves $\|T'\| \ge \|T\|$.

(d) Let X, Y be Banach spaces. First assume that $T\colon X \to Y$ is bijective. Then $T^{-1}\colon Y \to X$ is also continuous by the continuous inverse theorem (Corollary 2.3.4). Observe that $T'\colon Y' \to X'$ and $(T^{-1})'\colon X' \to Y'$ For every $x' \in X'$ one has $T'(T^{-1})'x' = T'\big(x'T^{-1}\big) = x'T^{-1}T = x'$, so $T'(T^{-1})' = \mathrm{id}_{X'}$. By a similar computation one can readily check that $(T^{-1})'T' = \mathrm{id}_{Y'}$. So T' is indeed bijective with inverse $(T^{-1})'$.

The converse implication, i.e. that bijectivity of T' implies bijectivity of T, will be discussed on Homework Sheet 11. □

Example 5.3.3 (Dual operators in finite dimensions)**.** Consider a matrix $A \in \mathbb{K}^{d\times c}$ as a linear map $\mathbb{K}^c \to \mathbb{K}^d$. We identify linear functionals on $\mathbb{K}^d$ with row vectors in $\mathbb{K}^{1\times d}$, and likewise for $\mathbb{K}^c$. Let $T_d\colon \mathbb{K}^d \to \mathbb{K}^{1\times d}$, $T_d y := y^{\mathrm{T}}$ by the linear map that identifies $\mathbb{K}^d$ with its dual space $\mathbb{K}^{1\times d}$, and define T_c in the same way. For each $y \in \mathbb{K}^d$ one then has

$$A' T_d y = A' y^{\mathrm{T}} = y^{\mathrm{T}} A = (A^{\mathrm{T}} y)^{\mathrm{T}} = T_c A^{\mathrm{T}} y.$$

In other words, A' acts as A^{T} if we identify $\mathbb{K}^d$ and $\mathbb{K}^c$ with their dual spaces, respectively, by transposing column vectors into row vectors. More formally speaking, one can say that A' and A^{T} are intertwined by the canonical isomorphisms $T_c\colon \mathbb{K}^c \to \mathbb{K}^{1\times c}$ and $T_d\colon \mathbb{K}^d \to \mathbb{K}^{1\times d}$.

For every normed space X the dual space X' is also a normed space (even a Banach space), so it also has a dual space $(X')'$ which we simply denote as X'' and call the **bidual space** or the **second dual** of X.

Proposition 5.3.4 (Embedding into the bidual)**.** *Let X be a normed space. For each $x \in X$ let $\hat{x} \in X''$ denote the functional given by*

$$\langle \hat{x}, x'\rangle := \langle x', x\rangle.$$

*Then $\hat{x}$ is indeed a bounded linear functional on X' and $\|\hat{x}\| = \|x\|$. The map $J_X\colon X \to X''$, $x \mapsto \hat{x}$ is a linear isometry and is called the **canonical embedding** of X into X''.*

Proof. Let $x \in X$. Linearity of the map $\hat{x}\colon X' \to \mathbb{K}$ follows immediately from the definition of the vector space structure on $X' = \mathcal{L}(X;\mathbb{K})$. The norm equality $\|\hat{x}\| = \|x\|$ is precisely the formula for $\hat{x}$ in Corollary 2.4.9 (so $\|\hat{x}\| = \|x\|$ is a consequence of the Hahn–Banach extension theorem).

The linearity of the map J follows from the definition of the vector space structure on $X'' = \mathcal{L}(X';\mathbb{K})$ and from the fact that all elements of X' are linear. The fact that J is isometric is simply a rephrasing of the inequality $\|\hat{x}\| = \|x\|$ for all $x \in X$. $\square$

By the previous proposition we can consider every normed vector spaces a vector subspace of its bidual.

Example 5.3.5 (The embedding of $\boldsymbol{c_0}$ into its bidual space $\boldsymbol{\ell^\infty}$)**.** Abbreviate $c_0 := c_0(\mathbb{N};\mathbb{K})$, $\ell^1 := \ell^1(\mathbb{N};\mathbb{K})$, and $\ell^\infty := \ell^\infty(\mathbb{N};\mathbb{K})$. Consider the maps $\Phi\colon \ell^1 \to (c_0)'$ and $\Psi\colon \ell^\infty \to (\ell^1)'$ given by

$$\langle \Phi(g), f\rangle := \sum_{n=1}^{\infty} g_n f_n, \qquad \text{and} \qquad \langle \Psi(h), g\rangle := \sum_{n=1}^{\infty} h_n g_n$$

for all $f \in c_0$, $g \in \ell^1$, and $h \in \ell^\infty$ and let $I := c_0 \to \ell^\infty$ be the inclusion map.

(a) The maps Φ and Ψ are isometric isomorphisms of Banach spaces, and hence so is the map $\Phi'\colon (c_0)'' \to (\ell^1)'$ according to Proposition 5.3.2.

(b) The map $(\Phi')^{-1}\Psi\colon \ell^\infty \to (c_0)''$ is an isomorphism of Banach spaces, and the composition $(\Phi')^{-1}\Psi I\colon c_0 \to (c_0)''$ is the canonical embedding J_{c_0} from Proposition 5.3.4.

Proof. (a) The fact that Φ and Ψ are isometric isomorphisms of Banach spaces can be proved by similar arguments as we used for ℓ^p in the case $p \in (1,\infty)$ (Example 1.2.4 and Theorem 1.2.6).

(b) Since Φ' and Ψ are isometric isomorphisms, so is $(\Phi')^{-1}\Psi$. To prove that $(\Phi')^{-1}\Psi I = J_{c_0}$, first note that both operators map from c_0 to $(c_0)''$. Let $f \in c_0$ and $\varphi \in (c_0)'$. Since Φ is bijective there exists a sequence $g \in \ell^1$ such that $\Phi g = \varphi$. We have

$$\langle J_{c_0} f, \varphi\rangle = \langle \varphi, f\rangle = \langle \Phi g, f\rangle = \sum_{n=1}^{\infty} g_n f_n = \langle \Psi I f, g\rangle = \langle \Psi I f, \Phi^{-1}\varphi\rangle = \langle (\Phi^{-1})'\Psi I f, \varphi\rangle.$$

So indeed $J_{c_0} f = (\Phi^{-1})'\Psi I f$ for every $f \in c_0$ and hence, $J_{c_0} = (\Phi^{-1})'\Psi I = (\Phi')^{-1}\Psi I$, where the last equality follows from Proposition 5.3.2. $\square$

We phrased Example 5.3.5 very precisely, by explicitly writing down all isomorphisms. In practice one usually ignores those isomorphisms and just identifies $(c_0)'$ with ℓ^1 and $(\ell^1)'$ with ℓ^∞. Thus, one considers ℓ^∞ as the bidual space of c_0 and Example 5.3.5 says precisely that the inclusion map $c_0 \to \ell^\infty$ then becomes the canonical embedding of c_0 into its bidual.

Definition 5.3.6 (The weak* topology)**.** Let X be a normed space. The **weak***-topology on the dual space X' is the initial topology of the family $(\hat{x})_{x\in X}$ in X''.

Let us start our study of the weak* topology by listing some elementary properties in the following proposition.

Proposition 5.3.7. *Let X be a normed space and let w^* denote the weak*-topology on the dual space X'.*

(a) *Let $x' \in X'$. A net (x'_j) in X' converges to x' with respect to w^* if and only if $\left\langle x'_j, x \right\rangle \to \langle x', x\rangle$ holds for each $x \in X$.*

(b) *With respect to the weak* topology w^*, the dual space X' is a Hausdorff topological vector space (and is thus T3 according to Example 5.2.5*(b)*).*

(c) *The identity map $\mathrm{id}_{X'}$ is continuous from $(X', \|\cdot\|)$ to (X', w^*).*

Proof. (a) This is a general property of initial topologies, see Proposition 4.2.12(iv).

(b) It is straightforward to check that X' is a topological vector space with respect to the weak* topology. To show the Hausdorff property it suffices to show that weak* limits of nets are unique (Proposition 4.2.6). So let $(x'_j)_{j\in J}$ be a net in X' that weak* converges to two points $x', y' \in X$. For each $x \in X$ one has $\left\langle x'_j, x \right\rangle \to \langle x', x\rangle$ and $\left\langle x'_j, x \right\rangle \to \langle x', x\rangle$. Since $\mathbb{K}$ is Hausdorff, we conclude that $\langle x', x\rangle = \langle y', x\rangle$ for all $x \in X$ and hence, $x' = y'$.

(c) Let $(x'_j)_{j\in J}$ be a net in X' that converges with respect to the norm topology to a point $x' \in X'$. For every $x \in X$ one than has

$$\left|\left\langle x'_j, x \right\rangle - \langle x', x\rangle\right| \le \left\|x'_j - x'\right\| \|x\| \to 0,$$

so $x'_j \to x'$ with respect to the weak* topology. □

Theorem 5.3.8 (Banach–Alaoglu)**.** *Let X be a normed space. Then the closed unit ball $B_{\le 1}(0) \subseteq X'$ in the dual space is compact with respect to the weak* topology.*

Proof. Let (x'_j) be a universal net in $B_{\le 1}(0) \subseteq X'$. We need to show that it converges to a point in $B_{\le 1}(0)$.

For each $x \in X$ it follows from Exercise 9.5(a) on Homework Sheet 9 that $\left(\left\langle x'_j, x \right\rangle\right) = \left(\left\langle \hat{x}, x'_j \right\rangle\right)$ in a universal net in the compact set $B_{\le\|x\|}(0) \subseteq \mathbb{C}$, so it converges to a complex number $\varphi(x)$ of modulus $\left|\varphi(x)\right| \le \|x\|$. Clearly, the map $\varphi \colon X \to \mathbb{K}$ is linear and since it is bounded with norm $\left\|\varphi\right\| \le 1$. Hence, $x' := \varphi \in B_{\le 1}(0) \subseteq X'$. The very definition of $x' = \varphi$ implies that (x_j) is weak*-convergent to x'. □

Lecture 20 (Th, 02 July) starts with Examples 5.3.9.

Let us discuss the weak* topology on two concrete sequence spaces. In the following we identify $(c_0)'$ with ℓ^1 and $(\ell^1)'$ with ℓ^∞ – in contrast to Example 5.3.5 we no longer write down the isomorphisms between these spaces explicitly.

Examples 5.3.9 (The weak* topology on sequence spaces)**.** As in Example 5.3.5, abbreviate $c_0 := c_0(\mathbb{N};\mathbb{K})$, $\ell^1 := \ell^1(\mathbb{N};\mathbb{K})$, and $\ell^\infty := \ell^\infty(\mathbb{N};\mathbb{K})$.

(a) Let $(h_j)_{j\in J}$ be a bounded net in ℓ^∞ and let $h \in \ell^\infty$. Then $h_j \to h$ with respect to the weak*-topology on $\ell^\infty = (\ell^1)'$ if and only if $(h_j)_k \xrightarrow{j} h_k$ for each $k \in \mathbb{N}$.

(b) Let $(g_j)_{j\in J}$ be a bounded net in ℓ^1 and let $g \in \ell^1$. The $g_j \to g$ with respect to the weak* topology on $\ell^1 = (c_0)'$ if and only if $(g_j)_k \xrightarrow{j} g_k$ for each $k \in \mathbb{N}$.

Proof.

(a) "$\Rightarrow$" Assume htat $h_j \to h$ wit respect to the weak* topology and let $k \in \mathbb{N}$. For the k-th canonical unit vector $e_k \in \ell^1$ one then gets $(h_j)_k = \langle h_j, e_k\rangle \xrightarrow{j} \langle h, e_k\rangle = h_k$. (Note that we did not need the boundedness of the net (h_j) for this implication.

"$\Leftarrow$" Assume that $(h_j)_k \xrightarrow{j} h_k$ for each $k \in \mathbb{N}$ and let $g \in \ell^1$. Let $\varepsilon > 0$. We set $M = \sup_{j\in J} \|h_j\|_\infty < \infty$ and note that also $\|h\|_\infty \le M$. Since $g \in \ell^1$, there exists $k_0 \in \mathbb{N}$ such that $M\sum_{k=k_0+1}^\infty |g_k| \le \varepsilon$.
Moreover, there exists an index $j_0 \in J$ such that $|(h_j)_k - h_k||g_k| \le \varepsilon/k_0$ for all $j \ge j_0$. For each such j we thus obtain

$$\left|\langle h_j, g\rangle - \langle h, g\rangle\right| \le \sum_{k=1}^{k_0} \underbrace{|(h_j)_k - h_k||g_k|}_{\le \varepsilon/k_0} + \sum_{k=k_0+1}^{\infty} \underbrace{|(h_j)_k - h_k|}_{\le 2M}|g_k| \le 3\varepsilon.$$

(b) This can be shown by similar arguments as (a). $\square$

For what follows we need the following result from linear algebra. To focus further on the topological aspects of the theory, we do not prove the lemma in the lecture. Interested readers can find the proof in Appendix 5.5 in Lemma 5.5.1.

Lemma 5.3.10. *Let V be a vector space over $\mathbb{K}$, let $n \in \mathbb{N}$ and let $\varphi, \varphi_1, \dots, \varphi_n \colon V \to \mathbb{K}$ be linear. One has $\varphi \in \operatorname{span}\{\varphi_1, \dots, \varphi_n\}$ if and only if φ vanishes on $\ker\varphi_1 \cap \dots \cap \ker\varphi_n$.*

The following proposition gives two ways to identify which elements x'' in the bidual of a normed spaces X are in the range of the canonical embedding $J_X \colon X \to X''$.

Proposition 5.3.11 (Weak*-continuity of functionals)**.** *Let X be a normed space and let $J \colon X \to X''$ be its canonical embedding into its bidual. Let $x'' \in X''$. The following assertions are equivalent:*

(i) *There exists a point $x \in X$ such that $Jx = x''$.*

(ii) *The map $x'' \colon X' \to \mathbb{C}$ is continuous with respect to the weak* topoloy w^* on X'.*

(iii) *There exists a number $c \in [0, \infty)$ and finitely many points $x_1, \dots, x_n \in X$ such that*

$$\left|\langle x'', x'\rangle\right| \le c \max_{1\le k\le n} \left|\langle x', x_k\rangle\right| \qquad \textit{for all } x' \in X'.$$

Proof. *"(i) ⇒ (ii)"* Let (i) hold and let $x \in X$ such that $Jx = x''$. If (x'_j) is a net in X' that weak* converges to a point $x' \in X'$, then

$$\left\langle x'', x'_j \right\rangle = \left\langle x'_j, x \right\rangle \to \langle x', x \rangle = \langle x'', x' \rangle,$$

so $x'' \colon X' \to \mathbb{K}$ is indeed convergent with respect to the weak* topology.

"(ii) ⇒ (iii)" Let (ii) hold. Then $U := (x'')^{-1}(B^{\mathbb{K}}_{<1}(0)) \subseteq X'$ is a weak*-open neighbourhood of $0 \in X'$. It follows from the characterisation of initial topologies in Proposition 4.2.12(ii) there exists points $x_1, \dots, x_n \in X$ and open sets $V_1, \dots, V_n \subseteq \mathbb{K}$ for which one has $0 \in (Jx_1)^{-1}(V_1) \cap \dots \cap (Jx_n)^{-1}(V_n) \subseteq U$. In particular, $0 \in V_1 \cap \dots \cap V_n$, so there exists a number $\delta > 0$ such that $B^{\mathbb{K}}_{\le\delta}(0) \subseteq V_1 \cap \dots \cap V_n$.

Now consider an arbitrary point $x' \in X'$. If $\gamma > \max_{1\le k\le n} \left|\langle x', x_k \rangle\right|$, then $\left|\left\langle \frac{\delta}{\gamma} x', x_k \right\rangle\right| < \delta$ for each $k \in \{1, \dots, n\}$, then $\frac{\delta}{\gamma} x' \in (Jx_1)^{-1}(V_1) \cap \dots \cap (Jx_n)^{-1}(V_n) \subseteq U$ and thus one has $\left|\left\langle x'', \frac{\delta}{\gamma} x' \right\rangle\right| < 1$, so $\left|\langle x'', \delta x' \rangle\right| < \gamma$. This proves that

$$\left|\langle x'', x' \rangle\right| \le \frac{1}{\delta} \max_{1\le k\le n} \left|\langle x', x_k \rangle\right|,$$

as claimed.

"(iii) ⇒ (i)" Let (iii) hold. Then $\ker(Jx_1) \cap \dots \cap \ker(Jx_n) \subseteq \ker x''$, so Lemma 5.3.10 implies that x'' is a linear combination of $Jx_1, \dots, Jx_n$, and is thus in the range of J. □

We conclude this section with the following theorem and its subsequent corollary.

Theorem 5.3.12 (Goldstine: Weak* density in the bidual unit ball)**.** *Let X be a normed space and let $J_X \colon X \to X''$ denote the canonical embedding. Then $J_X B^X_{\le 1}(0)$ is weak* dense in $B^{X''}_{\le 1}(0)$.*

The proof of Theorem 5.3.12 also uses Lemma 5.3.10. It is not extremely difficult, but not very short either, so we outsource the proof for the interested reader into Addendum 5.5 where it can be found in Theorem 5.5.2.

Corollary 5.3.13. *Let X be a normed space and let $J_X \colon X \to X''$ be the canonical embedding. Then $J_X X$ is weak* dense in X''.*

Proof. Let $x'' \in X''$. If $x'' = 0$, there is nothing to prove, so let $x'' \neq 0$. By Goldstine's Theorem 5.3.12 there exists a net (x_j) in X such that $J_X x_j \to \frac{x''}{\|x''\|}$ with respect to the weak* topology. Hence $J_X\big(\|x'\|\, x_j\big) \to x''$ with respect to the weak* topology. □

From a theory building point of view, it is actually a bit odd to first discuss Goldstine's theorem and then derived Corollary 5.3.13 as a consequence: in fact, Corollary 5.3.13 can be proved directly by using a similar but simpler proof strategy as Goldstine's theorem. Anyway, to keep the rest of this section brief we chose to take the path above.

5.4 The weak topology and reflexive spaces

Let X be a normed space. In the previous section we defined the weak* topology in a dual space X' as the initial topology of the point evaluations at all $x \in X$. Similarly, we will now define a topology on X as the initial topology of all functionals $x' \in X'$.

Definition 5.4.1 (The weak topology)**.** Let X be a normed space. The **weak topology** w on X is the initial topology of the family $(x')_{x' \in X'}$ of all bounded linear functionals $x' : X \to \mathbb{K}$.

Proposition 5.4.2. *Let X be a normed space and let w denote the weak-topology on X.*

(a) *Let $x \in X$. A net (x_j) in X converges to x with respect to w if and only if $\langle x', x_j \rangle \to \langle x', x \rangle$ holds for each $x' \in X'$.*

(b) *With respect to the weak topology, the space X is a Hausdorff topological vector space (and is thus T3 according to Example 5.2.5*(b)*).*

(c) *The identity map* id_X *is continuous from $(X, \|\cdot\|)$ to (X, w).*

Proof. The proof is almost the same as for the weak* topology (Proposition 5.3.7), so we omit the details. □

Theorem 5.4.3 (Weak closedness of convex sets)**.** *Let X be a normed space and let $C \subseteq X$ be convex.*

(a) *The set C is closed in the norm topology if and only if it is closed in the weak topology.*

(b) *The closures of C in the norm topology and in the weak topology coincide, i.e. one has $\overline{C}^{\|\cdot\|} = \overline{C}^{w}$.*

Proof.

(a) “⇒” Let C be norm closed. For every $x \in C^{\complement}$ it follows from the Hahn–Banach separation theorem for closed and compact sets that there exists a functional $x'_x \in X'$ and a number $\gamma_x \in \mathbb{R}$ such that

$$\operatorname{Re}\langle x'_x, c \rangle \le \gamma_x < \operatorname{Re}\langle x'_x, x \rangle.$$

This implies that

$$C = \bigcap_{x \in C^{\complement}} (x'_x)^{-1}\big(\{\lambda \in \mathbb{C} \mid \operatorname{Re}\lambda \le \gamma_x\}\big).$$

Since each x'_x is continuous with respect to the weak topology on X, it follows that C is closed in the weak topology.

“⇐“ This implication is clear (and holds for general sets, not only for convex ones) since norm convergence implies weak convergence.

(b) It is not difficult to check that both closures $\overline{C}^{\|\cdot\|}$ and $\overline{C}^{w}$ are convex.

"$\subseteq$" Since every weakly closed set is also norm closed, the norm closure is an intersection of more sets than the weak closure.

"$\supseteq$" The norm closure $\overline{C}^{\|\cdot\|}$ is, of course, norm closed, and it is not difficult to check that it is also convex. Hence it is even weakly closed according to (a). So $\overline{C}^{\|\cdot\|} \supseteq C$ implies $\overline{C}^{\|\cdot\|} \supseteq \overline{C}^{w}$. □

Lecture 21 (Mo, 06 July) starts with Proposition 5.4.4.

Proposition 5.4.4 (Topologies on the dual space)**.** *Let X be a normed space. On the dual space X', consider the norm topology, the weak topology w and the weak* topology w^*. Then the maps*

$$(X', \|\cdot\|) \xrightarrow{\mathrm{id}} (X, w) \xrightarrow{\mathrm{id}} (X, w^*)$$

are continuous.

Proof. The continuity of the first map follows from Proposition 5.4.2(c), applied to the normed space space X'. To show that continuity of the second map, let (x'_j) be a net in X' that converges weakly to a point $x' \in X'$. Then one has $\left\langle x'', x'_j \right\rangle \to \langle x'', x' \rangle$ for each $x'' \in X''$. So in particular, for each $x \in X$,

$$\left\langle x'_j, x \right\rangle = \left\langle J_X x, x'_j \right\rangle \to \langle J_X x, x' \rangle = \langle x', x \rangle,$$

which shows that $x'_j \to x'$ with respect to the weak* topology. □

Definition 5.4.5 (Reflexive spaces)**.** A normed space X is called **reflexive** if the canonical embedding $J_X\colon X \to X''$ is surjective.

Recall that the dual space X' of a normed space X' is always a Banach space (by Proposition 1.2.1), and hence so is the bidual space X''. Since the canonical embedding $J_X\colon X \to X''$ is linear and isometric for every normed space X (Proposition 5.3.4), it follows that J_X is an isometric isomorphism of normed spaces if X is reflexive. Hence, every reflexive normed spaces X is isometrically isomorphic to the Banach space X'' and is thus itself a Banach space.

Theorem 5.4.6 (Characterisation of reflexive spaces)**.** *Let X be a Banach space. The following assertions are equivalent:*

(i) *X is reflexive.*

(ii) *The closed unit ball $B_{\le 1}(0)$ in X is compact with respect to the weak topology on X.*

(iii) *On the dual space X' the weak*-topology and the weak topology coincide.*

(iv) *The dual space X' is reflexive.*

Proof. Let $J_X\colon X \to X''$ denote the canonical embedding. We first make the following preliminary observation: A net (x_j) in X converges weakly to a point $x \in X$ if and only if $(J_X x_j)$ is weak* convergent to $J_X x$. In particular, J_X is continuous from (X, w) to $(X'', w*)$.

"(i) ⇒ (ii)" Let (i) hold, i.e. $J_X\colon X \to X''$ is bijective. According to the preliminary observation, J_X is then a homeomorphism from (X, w) to (X'', w^*). Since $B^{X''}_{\le 1}(0)$ is weak* compact by the Banach–Alaoglu Theorem 5.3.8, the set $J_X^{-1}\big(B^{X''}_{\le 1}(0)\big)$ is weakly compact in X; but the latter set is precisely $B^{X}_{\le 1}(0)$ since J_X is a linear isometry.

"(ii) ⇒ (i)" Let (ii) hold. Since J_X is continuous from (X, w) to (X'', w^*) according to the preliminary observation at the beginning of the proof, if follows that $J_X\big(B^{X}_{\le 1}(0)\big) \subseteq X''$ is weak* compact and thus, in particular, weak* closed. But at the same time it is a weak* dense subset of $B^{X''}_{\le 1}(0)$ according to Goldstine's Theorem 5.3.12. Hence, $J_X\big(B^{X}_{\le 1}(0)\big) = B^{X''}_{\le 1}(0)$, which implies that J_X is surjective.

"(i) ⇒ (iii)" The weak* topology on X' is by definition the initial topology of the family $(J_X x)_{x\in X}$ on X', and the weak topology on X' is by definition the initial topology of the family $(x'')_{x''\in X''}$. If X is reflexive and J_X is thus surjective, both families contain the same maps.

"(iii) ⇒ (iv)" Let (iii) hold The dual unit ball $B^{X'}_{\le 1}(0)$ is weak* compact by the Banach–Alaoglu Theorem 5.3.8 and is thus weakly compact according to (iii). We can now apply the implication (ii)⇒(i), which we aleady proved, to the space X' to conclude that X' is reflexive.

"(iv) ⇒ (i)" Let (iv) hold. In addition to $J_X\colon X \to X''$ we also consider its dual operator $J_X'\colon X''' \to X'$ as well as the canonical embedding $J_{X'}\colon X' \to X'''$. A brief computation shows that $J_X' J_{X'}\colon X' \to X'$ is just the identity map $\mathrm{id}_{X'}$. Since $J_{X'}$ is surjective according to (iv), it follows that J_X' is injective. Thus, J_X has (norm) dense range according to Exercise 12.2(a) on Homework Sheet 12.

But J_X is isometric and X is a Banach space, so the range of J_X is complete and thus closed in X''. Therefore, $J_X X = X''$. □

Examples 5.4.7.

(a) The space sequence spaces c_0, ℓ^1, and ℓ^∞ are not reflexive.

(b) Let $p \in (1, \infty)$. The sequence space ℓ^p is reflexive.

(c) More generally, let $p \in (1, \infty)$ and let $(\Omega, \mathcal{A}, \mu)$ be a measure space. The space $L^p := L^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ is reflexive.

Moreover, let $p' \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{p'} = 1$. Then the linear isometry

$$\Phi\colon L^{p'} \to (L^p)', \qquad \tilde g \mapsto \Big(\tilde f \mapsto \int_\Omega f g \,\mathrm{d}\mu\Big)$$

from Exercise 2.3 on Homework Sheet 2 is surjective, and is thus an isometric isomorphism of Banach spaces.

Proof. (a) As observed in Example 5.3.5, the bidual space $(c_0)''$ can be identified with ℓ^∞, and the canonical embedding J_{c_0} then simply become the inclusion map $c_0 \to \ell^\infty$. Hence, J_{c_0} is not surjective, so c_0 is not reflexive.

It thus follows from Theorem 5.4.6 that the dual space $(c_0)'$, which can be identified with ℓ^1, is not reflexive, either.

(b) Choose $p' \in (1,\infty)$ such that $\frac{1}{p} + \frac{1}{p'} = 1$. Then the dual space $(\ell^p)'$ can be identified with $\ell^{p'}$ according to Theorem 1.2.6, and the same theorem applied to the space $\ell^{p'}$ shows that the dual space $(\ell^{p'})'$ can be identified with ℓ^p. Hence, $(\ell^p)''$ is isometrically isomorphic to ℓ^p. The same argument as in the proof of Example 5.3.5 shows that, under those identifications, the canonical embedding J_{ℓ^p} becomes the identity map on ℓ^p and is thus surjective.

(c) We do not show the details of the arguments here, but we sketch two different options how one can prove the claimed properties:

Option 1: One first shows that Φ is surjective and then argues as in (b).

The main challenge in this method is to obtain the surjectivity of Φ, which uses the Radon–Nikodým theorem from Analysis 3 [6, Theorem 2.6.10] (which we did, however, not prove there). Afterwards one uses that $(L^p)'' \simeq (L^{p'})' \simeq L^p$ and checks that, via those identifications, the canonical embedding $L^p \to (L^p)''$ is just the identify map on L^p and is thus surjective; so L^p is reflexive.

Option 2: One uses geometric properties of the norm on L^p together with more abstract non-sense from Banach space theory.

More specifically, one shows that the norm $\|\cdot\|_p$ satisfies the so-called **Clarkson inequalities**, which in turn implies that L^p has, since $p \in (1,\infty)$, a properties called **uniform convexity**.[6] And one can show by general Banach space theory (i.e. without using the specific structure of L^p-spaces) that all uniformly convex Banach spaces are reflexive.

One can then obtain the surjectivity of Φ from the reflexivity of L^p and $L^{p'}$ as follows: Since $L^{p'}$ is complete and Φ is isometric, it follows that the range of Φ is closed. Assume towards a contradiction that Φ is not surjective. By the Hahn–Banach extension theorem – more precisely, by Exercise 12.1(a) on Homework Sheet 12 – find a functional $0 \neq \varphi \in (L^p)''$ that vanishes on $\Phi(L^{p'})$. Since L^p is reflexive there exists $f \in L^p$ such that $\varphi = J_{L^p} f$. Hence,

$$\langle \Phi(g), f \rangle = \langle \varphi, \Phi(g) \rangle = 0$$

for all $g \in L^{p'}$, which can easily be checked to be false by choosing an appropriate function $g \in L^{p'}$. □

[6] These concepts were introduced in the article [2] about 90 years ago.

One reason why reflexivity is a very useful property is the existence of proxima:

Definition 5.4.8 (Proxima)**.** Let X be a normed space, let $S \subseteq X$ and $x \in X$. A point $s_0 \in S$ is called a **proximum of *x* in *S*** if

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

Equivalently, $\|x - s\| \geq \|x - s_0\|$ for all $s \in S$.

Theorem 5.4.9 (Proxima in reflexive spaces)**.** *Let X be a reflexive Banach space and let $K \subseteq X$ be convex and closed. Then every $x \in X$ has a proximum in K.*

Proof. We may assume that $x = 0$, since otherwise we can shift the entire situation by $-x$. Choose a number $r > \operatorname{dist}(0, K)$ and set $K_r := K \cap B_{\leq r}(0))$. Then $\operatorname{dist}(0, K) = \operatorname{dist}(0, K_r)$. Since $K_r \subseteq K$ it hence suffices to show that 0 has a proximum in K_r. The set K_r is also convex and closed, and hence it is weakly closed according to Theorem 5.4.3. But at the same time K_R is contained in the set $B_{\leq r}(0) = rB_{\leq 1}(0)$ which is weakly compact due to the reflexivity of X (Theorem 5.4.6). Hence, K_R itself is weakly compact (Exercise 10.1(b) on Präsenzblatt 10).

Since the map $\|\cdot\|$ is norm continuous and convex, its restriction to the weakly compact convex set K_r has a global minimal point according to Exercise 12.3 on Homeworksheet 12. In other words, there exists a point $s_0 \in K$ such that

$$\|s_0 - 0\| \leq \|s_0\| \leq \|s\| = \|s - 0\|$$

for all $s \in K_r$, i.e. s_0 is a proximum of 0 in K_r. □

5.5 Addendum: Proof of Goldstine's theorem

In this addendum we provide a proof of Lemma 5.3.10 and a proof of Goldstine's Theorem 5.3.12. Let us state the lemma again and then prove it:

Lemma 5.5.1. *Let V be a vector space over $\mathbb{K}$, let $n \in \mathbb{N}$ and let $\varphi, \varphi_1, \dots, \varphi_n \colon V \to \mathbb{K}$ be linear. One has $\varphi \in \operatorname{span}\{\varphi_1, \dots, \varphi_n\}$ if and only if φ vanishes on $\ker \varphi_1 \cap \dots \cap \ker \varphi_n$.*

Proof. *"⇒"* This implication is clear.

"⇐" Assume that φ vanishes on $\ker \varphi_1 \cap \dots \cap \ker \varphi_n$. We proceed in two steps.

Step 1: We first prove the claim under the additional assumption that $\ker \varphi_1 \cap \dots \cap \ker \varphi_n = \{0\}$.

If $V = \{0\}$ then $\varphi = 0$, so there is nothing to prove, so assume that $V \neq \{0\}$. Consider the linear map $T \colon V \to \mathbb{K}^n$, $v \mapsto \big(\varphi_1(v), \dots, \varphi_n(n)\big)^{\mathrm{T}}$, One has $\ker T = \ker \varphi_1 \cap \dots \cap \ker \varphi_n = \{0\}$, so T is injective, and thus V is finite-dimensional with $m := \dim V \leq n$. We may, and shall, identify V with $\mathbb{K}^m$ and T with a matrix in $\mathbb{K}^{n \times n}$. Hence, every linear functional on $\mathbb{K}^m$ is a row vector in $\mathbb{K}^{1 \times m}$ and the functionals $\varphi_1, \dots, \varphi_n$ are then the rows of T.

The injectivity of T implies that the column rank of T is $\dim TV = \dim V = m$. Since the column rank of every matrix coincides with its row rank, it follows that the span $\operatorname{span}\{\varphi_1, \dots, \varphi_n\} \subseteq \mathbb{K}^{1\times m}$, has dimension m and thus coincides with $\mathbb{K}^{1\times m}$. In particular, the functional $\varphi \in \mathbb{K}^{1\times m}$ is a linear combination of $(\varphi_1, \dots, \varphi_n)$.

Step 2: We prove the claim in the general case.

So assume that φ vanishes on $W := \ker\varphi_1 \cap \dots \cap \ker\varphi_n$. Since all the functionals φ_k vanish on W, they induce well-defined linear functionals $(\varphi_k)_/ : V/W \to \mathbb{K}$, $v + W \mapsto \varphi_k(v)$ for $k \in \{1, \dots, n\}$. Since φ also vanishes on W (this is where we use the assumption!), we also obtain a well-defined linear functional $\varphi_/ : V/W \to \mathbb{K}$, $v + W \mapsto \varphi(v)$.

One has $\ker(\varphi_1)_/ \cap \dots \cap \ker(\varphi_n)_/ = 0$ in V/W, so it follows from Step 1 that there exist scalars $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ such that $\varphi_/ = \sum_{k=1}^n \alpha_k (\varphi_k)_/$. For each $v \in V$ we thus obtain

$$\varphi(v) = \varphi_/(v + W) = \sum_{k=1}^n \alpha_k (\varphi_k)_/(v + W) = \sum_{k=1}^n \alpha_k \varphi_k(v),$$

so $\varphi = \sum_{k=1}^n \alpha_k \varphi_k$, as claimed. □

The proof seems to come a bit more transparent if one state the result separately for the open and the closed unit balls.

Theorem 5.5.2 (Goldstine: Weak* density in the bidual unit ball)**.** *Let X be a normed space and let $J : X \to X''$ denote the canonical embedding*

(a) *$JB^X_{<1}(0)$ is weak* dense in $B^{X''}_{<1}(0)$.*

(b) *$JB^X_{\le 1}(0)$ is weak* dense in $B^{X''}_{\le 1}(0)$.*

Proof. (a) Let $x'' \in X''$, $\|x''\| < 1$. We proceed in two steps.

Step 1: Let $V \subseteq X'$ a finite-dimensional vector subspace. We show that there exists a vector $x_V \in X$ of norm $\|x_V\| < 1$ such that $\langle x'', v' \rangle = \langle v', x_V \rangle$ for all $v' \in V$.

If $V = \{0\}$ we can choose $x = 0$, so let $n := \dim V \in \mathbb{N}$. Choose a basis $(v'_1, \dots, v'_n)$ of V. The map

$$T : X \to \mathbb{K}^n, \qquad x \mapsto \begin{pmatrix} \langle v'_1, x \rangle \\ \vdots \\ \langle v'_n, x \rangle \end{pmatrix}$$

is surjective, since otherwise we could find a non-zero vector $\alpha \in \mathbb{K}^{1\times n}$ such that α^{T} vanishes on TX and hence $\alpha_1 v'_1 + \dots + \alpha_n v'_n = 0$, which contradicts the linear independence of $(v'_1, \dots, v'_n)$. Since T is surjective, $\big(\langle x'', v'_1 \rangle, \dots, \langle x'', v'_n \rangle\big)^{\mathrm{T}} \in \mathbb{K}^n$ is in its range, so there exists a point $x \in X$ such that $\langle x'', v' \rangle = \langle v', x \rangle$ for all $v' \in V$.

We are now going to modify x to obtain a vector that has the same property but also satisfies $\|x\| < 1$. To this end, consider the closed vector subspace

$$W := \bigcap_{v' \in V} \ker v' = \ker v_1' \cap \dots \cap \ker v_n'$$

of X. It suffices to show that $\operatorname{dist}(x, W) < 1$: then, there exists a point $w \in W$ such that $x_V := x - w$ has norm $\|x_V\| < 1$, and one has $\langle x'', v' \rangle = \langle v', x \rangle = \langle v', x_V \rangle$ for all $v' \in V$.

To prove that $\operatorname{dist}(x, W) < 1$, consider the quotient map $q\colon X \to X/W$ and recall that $\|q\| \le 1$. According to Corollary 2.4.9 there exists a functional $y' \in (X/W)'$ of norm $\|y'\| \le 1$ such that $|\langle y', x + W \rangle| = \|x + W\|_/$. The functional $x' := y' \circ q \in X'$ has norm ≤ 1, satisfies $|\langle x', x \rangle| = \|x + W\|_/$ and vanishes on W. It follows from Lemma 5.3.10 that x' is a linear combination $(v_1', \dots, v_m')$ and thus $x' \in V'$. Therefore,

$$\operatorname{dist}(x, W) = \|x + W\|_/ = |\langle x', x \rangle| = |\langle x'', x' \rangle| \le \|x''\| < 1,$$

as claimed.

Step 2: We construct a net in $JB_{<1}(0)X$ that weak*-converges to x''.

Let $\mathcal{V}$ denote the set of all finite-dimensional vector subspaces of X'. Then $\mathcal{V}$ is directed by set inclusion as $V_1 + V_2 \in \mathcal{V}$ for all $V_1, V_2 \in \mathcal{V}$. For each $V \in \mathcal{V}$ let $x_V \in B_{<1}(0)X$ be a vector whose existence we prove in Step 1. Then $(Jx_V)_{V \in \mathcal{V}}$ is a net in $JB_{<1}(0)X$. This net is weak*-convergent to x'' since, for each $x' \in X'$, Jx_V coincides with x'' on $\operatorname{span}\{x'\}$ for all $V \supseteq \operatorname{span}\{x'\}$.

(b) For the closed balls one has $B_{\le 1}^{X''}(0) \supseteq JB_{\le 1}^X(0)$ since J is isometric. Thus,

$$B_{\le 1}^{X''}(0) \supseteq \overline{JB_{\le 1}^X(0)}^{w*} \supseteq \overline{JB_{<1}^X(0)}^{w*} \supseteq \overline{B_{<1}^{X''}(0)}^{w*} \supseteq B_{\le 1}^{X''}(0),$$

where the first first inclusion holds since $B_{\le 1}^{X''}(0)$ is weak*-closed (as follows, for instance, from the Banach–Alaoglu Theorem 5.3.8), the second inclusion is clear, the third inclusion follows from (a), and the fourth inclusion follows from the weak*-continuity of scalar multiplication (Proposition 5.3.7(b)). □

Goldstine's Theorem 5.5.2 is actually a special case of a much more general result from the duality theory of topological vector spaces, the so-called **bipolar theorem**; see for instance [10, Section IV.1, pp. 123–126] for a discussion of this theorem.

5.6 Addendum: Sequential weak and weak* compactness

As we have seen in Chapter 4, sequences are not sufficient to study general topological spaces – one needs the more general concept of nets. On the other hand, in some cases it is very convenient to work with sequences instead, in particular if one wants to use results from measure theory as those typically work for sequences only.

Remarkably, there are two important results which show that it is sometimes sufficient to work with sequences instead of nets when one is dealing with the weak* and the weak topology. For the convenience of readers who might need those results – for instance in a course on partial differential equations – we include them here without proofs.

Theorem 5.6.1 (Metrizability of the weak* topology on the dual ball)**.** *Let X be a Banach space and assume that X is separable. Then there exists a metric on the dual unit ball $B' := B^{X'}_{\leq 1}(0)$ that induces the weak* topology on B. In particular, B' is* ***weak* sequentially compact****, i.e. every sequence in B' has a subsequence that converges to a point in B'.*

Proof. As X is separable, the weak* topology on B' is metrizable – see for instance [3, Theorem V.5.1] for the proof. Moreover, we know that B' is weak* compact by the Banach–Alaoglu theorem 5.3.8, and every compact metric space is sequentially compact. □

More surprisingly, it turns out that (relative) weak compactness in a Banach space X is also equivalent to (relative) weak sequential compactness – even if the space is not separable:

Theorem 5.6.2 (Eberlein–Smulian)**.** *Let X be a Banach space and $K \subseteq X$.*

(a) *The set K is compact with respect to the weak topology if and only if it is* ***weakly sequentially compact****, which means that every sequence in K has a subsequence that converges weakly to a point in K.*

(b) *The set K is relatively compact with respect to the weak topology if and only if it is* ***relatively weakly sequentially compact****, which means that every sequence in K has a subsequence that converges weakly to a point in X.*

Proof. See for instance [12, Theorem VIII.6.1]. □

Chapter 6

Hilbert Spaces

Hilbert spaces constitute a class of Banach spaces that is particularly well-behaved. The idea is to consider norms that are induced by so-called **inner products**[1] which has nice properties and helps a lot in analysing those spaces. There is a strong geometric aspect of Hilbert space theory since inner products are closely related to angles between vectors (you might already know this on the finite dimensional space $\mathbb{R}^d$).

6.1 Inner products and Hilbert spaces

Introductory questions.

(a) Do you remember how the standard inner product on $\mathbb{R}^d$ is defined?

(b) What does one even care about the standard inner product on $\mathbb{R}^d$?

Lecture 22 (Th, 09 July) starts with Definition 6.1.1.

Definition 6.1.1 (Sesquilinear forms)**.** Let X, Y be vector spaces over $\mathbb{K}$.

(a) A map $\varphi\colon X \to Y$ is called **anti-linear** if $\varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2)$ for all $x_1, x_2 \in X$ and $\varphi(\lambda x) = \overline{\lambda}\varphi(x)$ for all $x \in X$ and all $\lambda \in \mathbb{K}$.

Note that in the case $\mathbb{K} = \mathbb{R}$, the properties linear and anti-linear are equivalent.

Note: At this point in the lecture we only defined anti-linearity for maps $X \to \mathbb{K}$. However, we also need the notion for maps with values in a vector space, in particular in Theorem 6.2.6 below.

(b) A map $a\colon X \times X \to \mathbb{K}$ is called a **sesquilinear form** if $a(x, \cdot)\colon X \to \mathbb{K}$ is linear for each $x \in X$ and $a(\cdot, y)\colon X \to \mathbb{K}$ is **anti-linear** for each $y \in X$.[2]

If $\mathbb{K} = \mathbb{R}$, a sesquilinear form is also called a **bilinear form**.

---

[1] In German the notion **Skalarprodukt** is more common that **inneres Produkt**.

[2] So here we use the convention that sesquilinear forms are antilinear in the first argument and linear in the second argument. This convention is very common in physics, while it is more common in the mathematical literature to define it the other way round. The reason why we use the physical convention rather than the mathematical one is that…well…the physicists are right.

More seriously speaking, with the physical convention the standard inner product on $\mathbb{C}^d$ is $(x, y) \mapsto \overline{x}^{\mathrm{T}} y$ rather than $x^{\mathrm{T}}\overline{y}$, which leads to much easier, clearer, and less confusing notation when working with rank-one projections on $\mathbb{C}^d$.

(c) A sesquilinear form $a\colon X \times X \to \mathbb{K}$ is called **Hermitian** if $a(x,y) = \overline{a(y,x)}$ for all $x,y \in X$.

A very useful property of a sesquilinear form a is that, if the scalar field is complex or if its is real and a is Hermitian, then the values of a on the diagonal $\{(x,x) \mid x \in X\} \subseteq X \times X$ determine the values of a on the entire product. In other words, if we know $a(x,x)$ for all $x \in X$, then we automatically know $a(x,y)$ for all $x,y \in X$. This is the main point of the following proposition.

Proposition 6.1.2 (Polarisation identity)**.** *Let X be a vector space over $\mathbb{K}$ and let $a\colon X \times X \to \mathbb{K}$ be a sesquilinear form.*

(a) *If $\mathbb{K} = \mathbb{C}$, then $a(x,y) = \frac{1}{4}\sum_{k=0}^{3} i^k a\big(i^k x + y,\, i^k x + y\big)$ for all $x,y \in X$.*[3]

(b) *If $\mathbb{K} = \mathbb{R}$ and a is Hermitian, then $a(x,y) = \frac{1}{4}\big(a(x+y,x+y) - a(x-y,x-y)\big)$ for all $x,y \in X$.*

Proof. These formulas can be checked by straightforward computations. □

Definition 6.1.3 (Inner products and Hilbert spaces)**.** Let H be a vector space over $\mathbb{K}$.

(a) A **semi inner product** on H is a sesquilinear from $(\cdot \mid \cdot) : H \times H \to \mathbb{K}$ that satisfies the following two additional properties:

  (I) The sesquilinear form $(\cdot \mid \cdot)$ is Hermitian, i.e. one has $\big(x \mid y\big) = \overline{\big(y \mid x\big)}$ for all $x,y \in H$.

  (II) For all $x \in H$ one has $(x \mid x) \geq 0$.

  The semi inner product $(\cdot \mid \cdot)$ is called an **inner product** it it satisfies the following stronger version of (II):

  (III) For all $x \in H \setminus \{0\}$ one has $(x \mid x) > 0$.

(b) If $(\cdot \mid \cdot) : H \times H \to \mathbb{K}$ is a (semi) inner product on H; then we define the **induced (semi) norm** $\|\cdot\| : H \to [0,\infty)$ by $\|x\| := \sqrt{(x \mid x)}$ for all $x \in H$.

  (We will show in Proposition 6.1.4 below that this is indeed as (semi)norm.)

(c) A **pre-Hilbert space** is a pair $(H, (\cdot \mid \cdot))$, where H is a vector space over $\mathbb{K}$ and $(\cdot \mid \cdot) : H \times H \to \mathbb{K}$ is an inner product on H. Unless otherwise stated we always endow a Hilbert space with the norm induced by the inner product.

  A pre-Hilbert space is called a **Hilbert space** if it is complete with respect to the norm induced by its inner product.

[3] If one uses the mathematical convention that sesquilinear forms are anti-linear in their second argument, then the formula reads instead $a(x,y) = \frac{1}{4}\sum_{k=0}^{3} i^k a\big(x + i^k y,\, x + i^k y\big)$.

Let us make the following observation that is frequently useful for semi inner products $(\cdot \mid \cdot)$ on real or complex vector spaces H: for all $x, y \in H$ one has

$$\left(x \mid y\right) + \left(y \mid x\right) = 2\operatorname{Re}\left(x \mid y\right)$$

due to the symmetry of $(\cdot \mid \cdot)$. We will tacitly use this throughout

Proposition 6.1.4 (Cauchy–Schwarz inequality)**.** *Let H be a vector space over $\mathbb{K}$, let $(\cdot \mid \cdot) : H \times H \to \mathbb{K}$ a semi inner product, and $\|\cdot\| : H \to [0,\infty)$ the induced seminorm.*

(a) *For all $x, y \in H$ the **Cauchy–Schwarz inequality** $\left|\left(x \mid y\right)\right| \le \|x\| \left\|y\right\|$ hold.*

If $(\cdot \mid \cdot)$ is even an inner product, then the Cauchy–Schwarz inequality is an equality if and only if the pair (x, y) is linearly dependent.

(b) *The induced seminorm $\|\cdot\|$ is indeed a seminorm; it is a norm if and only if $(\cdot \mid \cdot)$ is an inner product.*

Proof. (a) Let $x, y \in H$. Choose a scalar $\mu \in \mathbb{K}$ of modulus $\left|\mu\right| = 1$ such that $\mu\left(x \mid y\right) = \left|\left(x \mid y\right)\right|$. For every $t \in (0,\infty)$ one has

$$\begin{aligned} 0 &\le \left\|x - t\mu y\right\|^2 = \left(x - t\mu y \mid x - t\mu y\right) = \|x\|^2 - t\left(x \mid \mu y\right) - t\left(\mu y \mid x\right) + t^2 \underbrace{\left\|\mu y\right\|^2}_{=\left\|y\right\|^2} \\ &= \|x\|^2 + t^2\left\|y\right\|^2 - 2t\operatorname{Re}\left(x \mid \mu y\right) = \|x\|^2 + t^2\left\|y\right\|^2 - 2t\left|\left(x \mid y\right)\right| \end{aligned}$$

and thus

$$2\left|\left(x \mid y\right)\right| \le \frac{1}{t}\|x\|^2 + t\left\|y\right\|^2.$$

If $\|x\| = 0$ or $\left\|y\right\| = 0$, this implies that $\left|\left(x \mid y\right)\right| = 0 = \|x\| \left\|y\right\|$. If $\|x\| \neq 0$ and $\left\|y\right\| \neq 0$, we can choose $t := \|x\| / \left\|y\right\|$ and thus obtain $2\left|\left(x \mid y\right)\right| \le 2\|x\| \left\|y\right\|$, as claimed.

Finally, let $(\cdot \mid \cdot)$ be an inner product and assume that the equality $\left|\left(x \mid y\right)\right| \le \|x\| \left\|y\right\|$ holds. If $x = 0$ or $y = 0$, then (x, y) is linearly dependent, so assume no that $x \neq 0$ and $y \neq 0$. Then $\|x\| > 0$ and $\left\|y\right\| > 0$ since $(\cdot \mid \cdot)$ is an inner product, so the above computations show, for the choice $t := \|x\| / \left\|y\right\|$, that actually $0 = \left\|x - t\mu y\right\|^2$. Hence, $x - t\mu y = 0$, which shows that (x, y) is linearly dependent.

(b) The positive (semi)definiteness is clear from the definition of a (semi) inner product. Moreover, one has

$$\|\lambda x\|^2 = (\lambda x \mid \lambda x) = \overline{\lambda}\lambda\,(x \mid x) = |\lambda|^2\,\|x\|^2$$

and thus $\|\lambda x\| = |\lambda|\,\|x\|$ for all $x \in H$ and all $\lambda \in \mathbb{K}$.[4]

[4] In fact, we had already tacitly used this in the proof of (a).

It remains to show the triangle inequality. Let $x, y \in H$. Then the Cauchy–Schwarz inequality gives

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + \|y\|^2 + 2\operatorname{Re}(x \mid y) \le \|x\|^2 + \|y\|^2 + 2\left|(x \mid y)\right| \\ &\le \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| = (\|x\| + \|y\|)^2,\end{aligned}$$

and thus $\|x+y\| \le \|x\| + \|y\|$. □

Examples 6.1.5 (Some (pre-)Hilbert spaces)**.**

(a) The space $\mathbb{K}^d$ with the standard inner product given by $(x \mid y) := \overline{x}^{\mathrm{T}} y$ is a Hilbert space. The norm induced by this inner product is the Euclidean norm.

(b) Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. The space $L^2 := L^2(\Omega, \mathcal{A}, \mu; \mathbb{K})$ is a Hilbert space with respect to the inner product given by $(\tilde{f} \mid \tilde{g}) := \int_\Omega \overline{f} g \,\mathrm{d}\mu$ for all $\tilde{f}, \tilde{g} \in L^2$. Note that this integrand $\overline{f} g$ is indeed integrable as a consequence of Hölder's inequality for $p = 2$. The norm induced by the inner product is precisely the usual 2-norm on L^2.

(c) As a special case of (b), the sequence space ℓ^2 is a Hilbert spaces with respect to the inner product given by $(f \mid g) := \sum_{k=1}^\infty \overline{f_k} g_k$ for all $f, g \in \ell^2$, where the sum converges absolutely (since it converges as a Lebesgue integral with respect to the counting measure).

6.2 Optimal approximation and orthogonality

For the proof of the following theorem we use the so-called **paralleogram identity**

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

that holds for all elements x, y of a pre-Hilbert space H; this is discussed in detail on Präsenzblatt 13.

Theorem 6.2.1 (Proxima in (pre-)Hilbert spaces)**.** *Let $C \subseteq H$ be a convex set in a pre-Hilbert space H and let $x \in H$.*

(a) *The point x has at most one proximum in C.*

(b) *If H is a Hilbert space and C is non-empty and closed, then x has precisely one proximum in C.*

Proof. Throughout the proof we may, and shall, assume that $x = 0$ (otherwise, shift the entire situation by $-x$). We make the following preliminary observation: if two points in C are almost as close to 0 as is possible within C, then those two points are close to each other. More precisely, for all $\delta > 0$ and $c, d \in C$ one has the implication

$$\|c\|^2, \|d\|^2 \le \operatorname{dist}(0, C)^2 + \delta^2 \qquad \Rightarrow \qquad \|c - d\| \le 2\delta.$$

Indeed, the premise of the implication yields

$$\operatorname{dist}(0,C)^2 + \left\|\frac{c-d}{2}\right\|^2 \le \left\|\frac{c+d}{2}\right\|^2 + \left\|\frac{c-d}{2}\right\|^2 \tag{6.2.1}$$

$$= \frac{1}{2}\|c\|^2 + \frac{1}{2}\|d\|^2 \le \operatorname{dist}(0,C)^2 + \delta^2, \tag{6.2.2}$$

where the first inequality follows from $\frac{c+d}{2} \in C$ since C is convex, and the equality follows from the parallelogram identity; hence, $\|c-d\| \le 2\delta$, as claimed.

Now we can prove the assertions of the theorem.

(a) Let $c, d \in C$ by proxima of $x = 0$. Then $\|c\| = \|d\| = \operatorname{dist}(0,C)$, so it follows from (6.2.1) that $\|c-d\| \le 2\delta$ for every $\delta > 0$ and hence, $c = d$.

(b) Let H be complete and assume that C is non-empty and closed. Choose a sequence (c_n) in C such that $\|c_n\|$ converges to $\operatorname{dist}(0,C)$ sufficiently fast to ensure that $\|c_n\|^2 \le \operatorname{dist}(0,C)^2 + \frac{1}{n^2}$ for each $n \in \mathbb{N}$. For all indices $1 \le m \le n$ it then follows from (6.2.1) that

$$\|c_m - c_n\| \le \frac{2}{m}.$$

So (c_n) is a Cauchy sequence in H and thus converges to a point $c \in H$ due to the completeness of H. As C is closed, we have $c \in C$. Clearly $\|c\| = \lim_n \|c_n\| = \operatorname{dist}(0,C)$, so c is a proximum of 0 in C. □

Lecture 23 (Mo, 13 July) starts with Definition 6.2.2.

Definition 6.2.2 (Orthogonality and orthogonal complements)**.** Let H be a pre-Hilbert space.

(a) Two elements $x, y \in H$ are called **orthogonal** or **perpendicular** if $\left(x \mid y\right) = 0$. We write $x \perp y$ to say that x and y are orthogonal.

(b) For a subset $S \subseteq H$ the set

$$S^\perp := \left\{x \in H \mid\ x \perp s \text{ for all } s \in S\right\}$$

is called the **orthogonal complement of S**.

Let us list a few simple properties of orthogonality and orthogonal complements in the following proposition.

Proposition 6.2.3. *Let H be a pre-Hilbert space.*

(a) *Orthogonality is a symmetric relation, i.e. for all $x, y \in H$ one has $x \perp y$ if and only if $y \perp x$.*

(b) *For each $S \subseteq H$ the orthogonal complement $S^\perp$ is a closed vector subspace of H.*

(c) *If $x, y \in H$ are orthogonal, then* ***Pythagoras' theorem*** $\left\|x+y\right\|^2 = \|x\|^2 + \left\|y\right\|^2$ *holds.*

Proof. Those properties are easy consequences of the definition of inner products and orthogonal complements. For the closedness of $S^\perp$ one needs the continuity of the inner product, which in turn follows from the Cauchy–Schwarz inequality. □

You will discuss a few further useful properties of orthogonal complements in the exercises. Orthogonality is closely related to proxima in vector subspaces, as the following proposition shows.

Proposition 6.2.4 (Proxima in vector subspaces)**.** *Let H be a pre-Hilbert space and $V \subseteq H$ a vector subspace. Let $x \in H$ and $v \in V$. The following are equivalent:*

(i) *The vector v is a (and thus the) proximum of x in V.*

(ii) *The difference $v - x$ is orthogonal to all elements of V.*

Proof. *"(ii) ⇒ (i)"* Let (ii) hold. For every $w \in V$ the vector $w - v$ is then orthogonal to $v - x$, which implies by Pythagoras' theorem

$$\|w - x\|^2 = \|(w - v) + (v - x)\|^2 = \|w - v\|^2 + \|v - x\|^2 \geq \|v - x\|^2,$$

so v is indeed a proximum of x in V.

"(i) ⇒ (ii)" Let (i) hold. For every $u \in V$ one has $v + u \in V$ and thus

$$\|v - x\|^2 \leq \|v + u - x\|^2 = \|v - x\|^2 + 2\,\mathrm{Re}\,(v - x \mid u) + \|u\|^2$$

and hence $0 \leq 2\,\mathrm{Re}\,(v - x \mid u) + \|u\|^2$. Note that the first summand scales linearly in u while the second summand scales quadratically in u. So by replycing u with tu for $t \in (0,\infty)$, dividing by t and finally letting $t \downarrow 0$, we obtain $0 \leq \mathrm{Re}\,(v - x \mid u)$ for all $u \in V$.

But then, this also holds if we replace u with $-u$ and hence, $0 = \mathrm{Re}\,(v - x \mid u) = 0$ for all $u \in V$. Finally, let $u \in V$ and choose $\mu \in \mathbb{K}$ of modulus $|\mu| = 1$ such that $\mu\,(v - x \mid u) = |(v - x \mid u)|$. Then

$$|(v - x \mid u)| = \mathrm{Re}\big(\mu\,(v - x \mid u)\big) = \mathrm{Re}\big(v - x \mid \mu u\big) = 0$$

since $\mu u \in V$. So $v - x$ is indeed perpendicular to each $u \in V$. □

Corollary 6.2.5. *Let H be a Hilbert space and let $V \subseteq H$ be a closed vector subspace. Then $H = V \oplus V^\perp$.*

Proof. We first show that $V \cap V^\perp = \{0\}$. Indeed, let $x \in V \cap V^\perp$. Then $x \perp x$, so $\|x\|^2 = (x \mid x) = 0$, which implies $x = 0$.

It remains to show that $H = V + V^\perp$, so let $x \in H$. Since V is closed and convex, Theorem 6.2.1(b) shows that x has a proximum v in V. According to Proposition (ii) $v - x$ is orthogonal to all elements of V, i.e. $v - x \in V^\perp$. Hence,

$$x = v + (x - v) \in V + V^\perp,$$

as claimed. □

If H is a pre-Hilbert space and $y \in H$, then the map $\big(y \mid \cdot\big)\colon H \to \mathbb{K}$ is linear and continuous, i.e. an element of the dual space H'. Indeed, the linearity holds since the inner product is sesquilinear by definition, and the continuous follows easily from the Cauchy–Schwarz inequality. The following theorem shows that, if H is a Hilbert space, all elements of H' are of this form.

Theorem 6.2.6 (Riesz–Fréchet representation theorem: the dual of a Hilbert space)**.** *Let H be a Hilbert space. The mapping*

$$D_H\colon H \mapsto H', \qquad y \mapsto \big(y \mid \cdot\big)$$

is an antilinear isometric bijection. It is called the ***Riesz–Fréchet isomorphism****.*[5]

Note that the RieFréchet isomorphism $D_H\colon H \to H'$ from Theorem 6.2.6 satisfies, by its very definition,

$$\big(y \mid x\big) = \langle D_H y, x\rangle \qquad \text{and hence} \qquad \big(D_H^{-1} y' \mid x\big) = \langle y', x\rangle \tag{6.2.3}$$

for all $x, y \in H$ and all $y' \in H'$.

Proof of Theorem 6.2.6. Antilinearity: The fact that D_H is antilinear follows from the antilinearity of the inner product in its first argument.

Isometry: To show that D_H is isometric, let $y \in H$. For $y = 0$ one has $D_H y = 0$, so let $y \neq 0$. For every $x \in H$ one has

$$\big|\langle D_H y, x\rangle\big| = \big|\big(y \mid x\big)\big| \le \|y\|\,\|x\|,$$

where the inequality is the Cauchy–Schwarz inequality. Hence, $\|D_H y\| \le \|y\|$. On the other hand,

$$\|D_H y\| \ge \left|\left\langle D_H y, \frac{y}{\|y\|}\right\rangle\right| = \left|\left(y \mid \frac{y}{\|y\|}\right)\right| = \|y\|,$$

so D_H is indeed isometric.

Surjectivity: Let $0 \neq x' \in H'$. Corollary 6.2.5 shows that $H = \ker x' \oplus (\ker x')^\perp$. Since x' is a non-zero linear functional, this implies that $(\ker x')^\perp$ is one-dimensional and is thus spanned by a vector $0 \neq x \in (\ker x')^\perp$. Observe that $0 \neq \|x\|^2 = (x \mid x)$, so $x \not\in \ker x'$. By replacing x with an appropriate multiple we may thus assume that $\langle x', x\rangle = 1$.

We claim that $\left(\frac{x}{\|x\|^2} \mid \cdot\right) = x'$. Indeed, both functional vanish on $\ker x'$, and they both map x to $\frac{1}{\|x\|^2}$, so the coincide on $\ker x' \oplus \operatorname{span}\{x\} = H$. So $D_H \frac{x}{\|x\|^2} = x'$. □

Observe that the completeness of H in Theorem 6.2.6 is only need for the proof of the surjectivity of D_H. If H is only a pre-Hilbert space, D_H is still antilinear and isometric.

[5]Observe that, if $\mathbb{K} = \mathbb{R}$ then D_H is an isometric of isomorphism of Banach spaces. If, however $\mathbb{K} = \mathbb{C}$ then D_H has all properties of an isometric isomorphism of Banach spaces except that it is antilinear rather than linear.

Examples 6.2.7 (The Riesz–Fréchet isomorphism in concrete spaces)**.**

(a) Endow $\mathbb{K}^d$ with its standard inner product given by $\left(x \mid y\right) = \overline{x}^{\mathrm{T}} y$ for all $x, y \in \mathbb{K}^d$ and recall that this inner product induces the Euclidean norm. As usual we identify the dual space $(\mathbb{K}^d)'$ with the space $\mathbb{K}^{1\times d}$ of row vector.s Then the Riesz–Fréchet isomorphism $D_{\mathbb{K}^d} \colon \mathbb{K}^d \to \mathbb{K}^{1\times d}$ is given by

$$D_{\mathbb{K}^d} x = (x \mid \cdot) = \overline{x}^{\mathrm{T}}$$

for all $x \in \mathbb{K}^d$.

(b) Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Since $\frac{1}{2} + \frac{1}{2}$ the space $L^2 := L^2(\Omega, \mathcal{A}, \mu; \mathbb{K})$ is, according to Example 5.4.7(c), isomorphic to its own dual space $\left(L^2\right)'$ via the map

$$L^2 \to (L^2)', \qquad \tilde{g} \mapsto \left(\tilde{f} \mapsto \int_\Omega g f \,\mathrm{d}\mu\right).$$

On the other hand, the Riest–Fréchet isomorphism D_{L^2} is given by

$$D_{L^2} \colon L^2 \to (L^2)', \qquad \tilde{g} \mapsto \left(\tilde{f} \mapsto \int_\Omega \overline{g} f \,\mathrm{d}\mu\right).$$

So both isomorphisms coincide in the real case, but they do not coincide in the complex case (which is no surprise since D_{L^2} is antilinear rather than linear).

Corollary 6.2.8. *Every Hilbert space is reflexive.*

Proof. We show that the Riesz–Fréchet isomorphism $D_H \colon H \to H'$ is a homeomorphism from (H, w) to (H', w^*). Since D_H maps $B^H_{\le 1}(0)$ bijectively to $B^{H'}_{\le 1}(0)$, it then follows from the Banach–Alaoglu theorem 5.3.8 that $B^H_{\le 1}(0)$ is weakly compact; hence, H is reflexive according to Theorem 5.4.6. So let (x_j) be a net in H and let $x \in H$. One has

$$\begin{aligned}
& x_j \to x \quad \text{weakly} \\
\Leftrightarrow\ & \left\langle y', x_j \right\rangle \to \left\langle y', x \right\rangle \quad \text{for all } y' \in H' \\
\Leftrightarrow\ & \left\langle D_H y, x_j \right\rangle \to \left\langle D_H y, x \right\rangle \quad \text{for all } y \in H \\
\Leftrightarrow\ & \left(y \mid x_j\right) \to \left(y \mid x\right) \quad \text{for all } y \in H \\
\Leftrightarrow\ & \left(x_j \mid y\right) \to \left(x \mid y\right) \quad \text{for all } y \in H \\
\Leftrightarrow\ & \left\langle D_H x_j, y \right\rangle \to \left\langle D_H x, y \right\rangle \quad \text{for all } y \in H \\
\Leftrightarrow\ & D_H x_j \to D_H x \quad \text{weak*},
\end{aligned}$$

so D_H is indeed a homeomorphism from (H, w) to (H', w^*). □

6.3 Operators on Hilbert spaces

Introductory questions.

(a) For a matrix $A \in \mathbb{C}^{d\times c}$, why does one sometimes consider A^{T} and sometimes $\overline{A}^{\mathrm{T}}$?

(b) Which matrices is $\mathbb{C}^{d\times d}$ are diagonalizable by a unitary transformation matrix?

Definition 6.3.1 (Adjoint operators)**.** Let H_1, H_2 be Hilbert spaces over $\mathbb{K}$. For each $k \in \{1,2\}$ let $D_{H_k}\colon H_k \to H_k'$ be the Riesz–Fréchet isomorphism For each bounded linear operator $T\colon H_1 \to H_2$ the operator

$$T^* := D_{H_1}^{-1} T' D_{H_2}\colon H_2 \to H_1$$

is called the **adjoint operator** of T.

Proposition 6.3.2 (Properties of adjoint operators)**.** *Let H_1, H_2, H_3 be Hilbert spaces over $\mathbb{K}$ and let $T \in \mathcal{L}(H_1; H_2)$ and $S \in \mathcal{L}(H_2; H_3)$.*

(a) *One has $T^* \in \mathcal{L}(H_2; H_1)$ with $\|T^*\| = \|T\|$.*

(b) *One has $\big(y \mid Tx\big) = \big(T^* y \mid x\big)$ for all $x \in H_1$ and all $y \in H_2$, and T^* is the only map $H_2 \to H_1$ with this property.*

Similarly, one has $\big(x \mid T^* y\big) = \big(Tx \mid y\big)$ for all $x \in H_1$ and all $y \in H_2$, and T^* is the only map from H_2 to H_1 with this property.

(c) *One has $(T^*)^* = T$ and $(ST)^* = T^* S^*$.*

(d) *The map $\mathcal{L}(H_1; H_2) \to \mathcal{L}(H_2; H_1)$, $S \mapsto S^*$ is antilinear and bijective (and isometric due to* (a)*).*

(e) *One has $\|T\|^2 = \|T^* T\| = \|T T^*\|$.*

Lecture 24 (Th, 16 July): is planned to start with the proof of Proposition 6.3.2.

Proof. (a) Since $D_{H_1}^{-1}$ and D_{H_2} are anti-linear and T' is linear, it follows that T^* is linear. Since the maps $D_{H_1}^{-1}$ and D_{H_2} are bijective and isometric one has

$$\left\| T^* \right\| = \left\| T' \right\| = \|T\|,$$

where the second equality holds according to Proposition 5.3.2(a).

(b) Let $x, y \in H$. We use the formula (6.2.3) for the Riesz–Fréchet isomorphism and thus obtain

$$\big(T^* y \mid x\big) = \Big(D_{H_1}^{-1} T' D_{H_2} y \mid x\Big) = \big\langle T' D_{H_2} y, x \big\rangle = \big\langle D_{H_2} y, Tx \big\rangle = \big(y \mid Tx\big),$$

as claimed. The second equality follows by taking complex conjugates and both sides.

The uniqueness are a simple consequence of Exercise 13.3(a) on Homework Sheet 13.

(c) *and* (d) The equality $(ST)^* = T^* S^*$ and the anti-linearity of $S \mapsto S^*$ follow from the analogous properties of dual operators in Proposition 5.3.2. The equality $(T^*)^*$ follows from (b): for all $x, y \in H$ one has

$$\big(y \mid Tx\big) = \big(T^* y \mid x\big) = \big(y \mid (T^*)^* x\big).$$

You will see on Homework Sheet 13 that this implies $T = (T^*)^*$. So the map $S \mapsto S^*$ is self-inverse and thus bijective.

(e) We first show the equality on the left. One has $\|T^*T\| \le \|T^*\| \, \|T\| = \|T\|^2$ due to (a). On the other hand, for each $x \in H$ it follows from the Cauchy–Schwarz inequality that

$$\|Tx\|^2 = \left|\left(T^*Tx \mid x\right)\right| \le \left\|T^*Tx\right\| \|x\| \le \left\|T^*T\right\| \|x\|^2$$

and thus $\|T\|^2 \le \|T^*T\|$. So indeed $\|T\|^2 = \|T^*T\|$.

To also obtain the equality $\|T\|^2 = \|TT^*\|$, apply the equality that we just shows to the operator T^* and use (a) and (c). □

Property (e) in Proposition (e), as harmless as it looks, is an essentially property of the operator norm on Hilbert spaces and is closely related to many important properties in operator theory on Hilbert spaces. It is, in fact, one the defining properties of so-called ***C*-algebras**, which you might learn about in the course *Functional Analysis 1*.

Example 6.3.3 (Adjoint operators in finite dimensions)**.** Endow $\mathbb{C}^c$ and $\mathbb{C}^d$ with the standard inner product and consider a matrix $A \in \mathbb{C}^{d\times c}$ as a linear operator $\mathbb{C}^c \to \mathbb{C}^d$. Then $A^* = \overline{A}^{\mathrm{T}}$.

Proof. Let $x \in \mathbb{C}^c$ and $y \in \mathbb{C}^d$. Then

$$\left(y \mid Ax\right) = \overline{y}^{\mathrm{T}} Ax = \overline{\overline{A}^{\mathrm{T}} y}^{\mathrm{T}} x = \left(\overline{A}^{\mathrm{T}} y \mid x\right).$$

It thus follows from Proposition 6.3.2(b) that $A^* = \overline{A}^{\mathrm{T}}$. □

Definition 6.3.4 (Self-adjoint, skew-adjoint and normal operators)**.** Let H be a Hilbert space and $A \in \mathcal{L}(H)$.

(a) The operator A is called **self-adjoint** or **Hermitian** if $A^* = A$. It is called **skew-adjoint** if $A^* = -A$.

(b) The operator A is called **normal** if A commutes with A^*, i.e. if $AA^* = A^*A$.

Observe that every self-adjoint and every skew-adjoint operator is normal.

Proposition 6.3.5 (The norm of a normal operator)**.** *Let H be a Hilbert space and let $A \in \mathcal{L}(H)$ be normal. Then $\|A\|^2 = \left\|A^2\right\|$.*

Proof. We first prove the claim under the stronger assumption that A be self-adjoint. In this case it follows from Proposition 6.3.2(e) that $\|A\|^2 = \|A^*A\| = \left\|A^2\right\|$.

Now let A be, more generally, normal. Observe that the operator A^*A is self-adjoint; by applying again Proposition 6.3.2(e) and the fact that we have already shown the claim for self-adjoint operator, we obtain

$$\left(\|A\|^2\right)^2 = \left(\left\|A^*A\right\|\right)^2 = \left\|(A^*A)^2\right\| = \left\|(A^2)^*A^2\right\| = \left\|A^2\right\|^2,$$

where the third equality follows from the normality of A.[6] $\square$

We close this section by briefly orthogonal projections and rank-1 operators in Hilbert spaces.

Definition 6.3.6 (Orthogonal projection)**.** Let H be a Hilbert space. An **orthogonal projection** on H is a projection $P \in \mathcal{L}(H)$ such that $x \perp y$ for all $x \in \ker P$ and all $y \in PH$.

The proof of the following nice characterization of self-adjoint projections is Exercise 14.5 in Homework Sheet 14.

Proposition 6.3.7 (Characterization of orthogonal projections)**.** *Let H be a Hilbert space and let $P \in \mathcal{L}(H)$ be a projection. The following assertions are equivalent:*

(i) *P is an orthogonal projection.*

(ii) *P is self-adjoint.*

(iii) *P is normal.*

(iv) $\|P\| \le 1$.

The following ideas are similar to what you have already seen on general normed spaces in Definition 3.2.7 and Proposition 3.2.8 – the main difference is that we do not work with elements of the dual space but only with elements from the Hilbert space itself now.[7]

Definition 6.3.8 (Hilbert space notation for operators of rank at most one)**.** Let H_1, H_2 be Hilbert spaces[8] and let $x_1 \in H_1$, $x_2 \in H_2$. We define the operator $x_2 \otimes x_1 \in \mathcal{L}(H_1; H_2)$ by

$$(x_2 \otimes x_1)x := x_2\,(x_1 \mid x) \qquad \text{for all } x \in H_1.$$

Example 6.3.9. Endow $\mathbb{C}^c$ and $\mathbb{C}^d$ with the standard inner product. As usual we identify linear operators in $\mathcal{L}(\mathbb{C}^c; \mathbb{C}^d)$ with matrices in $\mathbb{C}^{d\times c}$. Let $x_1 \in \mathbb{C}^d$ and $x_2 \in \mathbb{C}^d$. Then one has $x_2 \otimes x_1 = x_2 \overline{x_1}^{\mathrm{T}}$.

[6] Observe that one could also show the claim directly, without consider self-adjoint operators first: for normal A one has

$$\left(\|A\|^2\right)^2 = \left(\left\|A^*A\right\|\right)^2 = \left\|(A^*A)^*(A^*A)\right\| = \left\|(A^2)^*A^2\right\| = \left\|A^2\right\|^2,$$

where the third equality follows from the normality and each of the other three equalities uses Proposition 6.3.2(e). This is essentially the same argument, but one might look a bit more arbitrary and thus more difficult to remember than considering the simpler special case of self-adjoint A first.

[7] In view of the Riesz–Fréchet representation theorem (Theorem 6.2.6) this is, of course, possible. The main point here is that, in the Hilbert space seting, it is also quite natural and efficient to stay within the space rather than employing its dual space.

[8] This definition also makes sense on pre-Hilbert spaces, but we stick to Hilbert spaces throughout the rest of Section 6.3 since we discussed adjoint operators only in the Hilbert-space setting.

Proof. For each $x \in \mathbb{C}^c$ one has

$$(x_2 \otimes x_1)x = x_2\,(x_1 \mid x) = x_2(\overline{x_1}^{\mathrm{T}} x) = (x_2\overline{x_1}^{\mathrm{T}})x,$$

so indeed $x_2 \otimes x_1 = x_2\overline{x_1}^{\mathrm{T}}$. □

In the situation of Definition 6.3.8 we can also express the operator $x_2 \otimes x_1$ in the notation from Definition 3.2.7 if we use the functional $(x_1 \mid \cdot) \in H_1'$. One then has

$$x_2 \otimes x_1 := x_2 \otimes (x_1 \mid \cdot)\,.$$

It follows from this observation and from the Riesz–Fréchet representation theorem (Theorem 6.2.6) that the results from Proposition 3.2.8 hold analogously for the operator notation introduced in Definition 6.3.8. What makes the Hilbert space situation special is that one can now talk about adjoint operators, too:

Lecture 25 (Mo, 20 July) starts with Proposition 6.3.10.

Proposition 6.3.10 (Rank-**1** operators)**.** *Let H, H_1, H_2, H_3 be Hilbert spaces.*

(a) *Let $x_1 \in H_1$, $x_2, y_2 \in H_2$ and $y_3 \in H_3$. Then $(y_3 \otimes y_2)(x_2 \otimes x_1) = \big(y_2 \mid x_2\big)\,(y_3 \otimes x_1)$.*

(b) *Let $x_1 \in H_1$ and $x_2 \in H_2$. Then $(x_2 \otimes x_1)^* = x_1 \otimes x_2$.*

(c) *Let $u \in H$ be normalized, i.e. $\|u\| = 1$. Then $u \otimes u$ is a self-adjoint projection.*

Conversely, every orthogonal projection on H of rank-1 is of this form.

Proof. (a) This follows by a straightforward computation.

(b) For all $y_2 \in H_2$ and $y_1 \in H_1$ one has

$$\begin{aligned}\big((x_2 \otimes x_1)^* y_2 \mid y_1\big) &= \big(y_2 \mid (x_2 \otimes x_1)\acute{y}_1\big) = \big(y_2 \mid x_2\,(x_1 \mid y_1)\big)\\ &= \big(x_1 \mid y_1\big)\big(y_2 \mid x_2\big) = \big(x_1\,(x_2 \mid y_2) \mid y_1\big) = \big((x_1 \otimes x_2)y_2 \mid y_1\big)\,.\end{aligned}$$

This implies that $(x_2 \otimes x_1)^* = x_1 \otimes x_2$ (Exercise 13.3(a) on Homework Sheet 1).

(c) Let $u \in H$ be of norm 1. It follows from (a) that $u \otimes u$ is a projection and from (b) that $u \otimes u$ is self-adjoint. So $u \otimes u$ is an orthogonal projection according to Proposition 6.3.7.

Conversely, let $P \in \mathcal{L}(H)$ be an orthogonal rank-1 projection. It follows from Proposition 3.2.8(b) and from the Riesz–Fréchet representation theorem that there are vectors $v, w \in H \setminus \{0\}$ such that $P = v \otimes w$. Moreover, P is self-adjoint according to Proposition 6.3.7, so it follows from (a) that $P = v \otimes w = w \otimes v$. This implies that v is a multiple of w, i.e. there exists a scalar α such that $v = \alpha w$.[9] Since P is a projection, one can readily check that $1 = (w \mid v) = \alpha\,\|v\|^2$. So $\alpha \in (0,\infty)$ and by setting $u := v/\,\|v\|$ one gets

$$P = \alpha(v \otimes v) = \frac{1}{\|v\|^2}(v \otimes v) = u \otimes u,$$

as claimed. □

[9] Can you give the details of this argument?

In the following example we discuss how the unitary diagonalisation of normal matrices that you know from linear algebra, can be interpreted in three different ways, and how this is related to orthogonal rank-1 projections discussed in Proposition 6.3.10(c) above.

Note: The contents of Example 6.3.11 were spontaneously presented in the lecture and included in the lecture notes only afterwards. Hence, the example was not enumerated (and not even called an example) in the lecture.

Example 6.3.11 (Unitary diagnonalisation of matrices)**.** Endow $\mathbb{C}^d$ with the standard inner product and let $A \in \mathbb{C}^{d\times d}$ be normal, i.e. $A^*A = AA^*$ (see Definition 6.3.4(b) and Example 6.3.3). From your linear algebra course you know that this implies (and is in fact equivalent to) A being unitary diagonalisability.

The latter property can be thought of in three different ways:

(1) *Algebraically:* There exists a unitary matrix $U \in \mathbb{C}^{d\times d}$ (i.e. U is invertible and $U^{-1} = U^*$) and a diagonal matrix $D \in \mathbb{C}^{d\times d}$ such that $A = UDU^*$.

Recall that the diagonall entries D_{jj} of D are the eigenvalues of A and that the columns of U are corresponding eigenvectors.

(2) *Geometrically:* There exists a coordinate system of $\mathbb{C}^d$ with pairwise orthogonal axes, such that A only acts in the direction of each axis.

(3) *In terms of orthogonal rank-*1 *projections:* Observe that one has

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} = e_1 e_1^*,$$

and likewise for diaogonal matrices that have the entry 1 only once at other positions on the diagonal.

For every $k \in \{1,\dots,d\}$ the vector $u_k := Ue_k$ is the kth column of U and is thus an eigenvector for the eigenvalue D_{kk} of A. Note that $u_k^* u_j = (U^*U)_{kj} = \mathrm{id}_{kj}$ for all $k, j \in \{1,\dots,d\}$, so all the columns u_k are pairwise orthogonal and have Euclidean norm 1.[10]

Since $D = \sum_{k=1}^d D_{kk} e_k e_k^*$, the diagonalisation formula $A = UDU^*$ implies

$$A = U \sum_{k=1}^d D_{kk} e_k e_k^* U^* = \sum_{k=1}^d D_{kk} u_k u_k^* = \sum_{k=1}^d D_{kk} (u_k \otimes u_k);$$

for the last equality we used Example 6.3.9. So every normal matrix A can be written as a sum over orthogonal projections onto the eigenspaces of A, weighted with the corresponding eigenvalues.

[10] According to Definition 6.4.1 below this means that $(u_1,\dots,u_d)$ is an orthonormal system in $\mathbb{C}^d$; for dimension reasons, it is even an orthonormal basis of $\mathbb{C}^d$.

6.4 Orthonormal bases

Introductory questions.

(a) When you think of a basis of $\mathbb{R}^2$, how does it look like?

(b) When is a matrix $U \in \mathbb{R}^{d\times d}$ called orthogonal? Can you give several reasons why such matrices are important?

Definition 6.4.1 (Orthonormal systems and bases)**.** Let H be a pre-Hilbert space and consider a family of vectors $(u_j)_{j\in J}$ in H.

(a) The family $(u_j)_{j\in J}$ is called an **orthonormal system** if it is an orthogonal system and each vector u_j has norm 1.

(b) The family $(u_j)_{j\in J}$ is called an **orthonormal basis of *H*** if it is an orthonormal system and its span $\operatorname{span}\{u_j \mid j \in J\}$ is dense in H.

Note that a family $(u_j)_{j\in J}$ of vectors in a pre-Hilbert space H is an orthonormal system if and only if $\left(u_j \mid u_k\right) = \delta_{jk}$ for all $j,k \in J$, where δ_{jk} denotes the so-called **Kronecker symbol** that is defined as

$$\delta_{jk} := \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

Proposition 6.4.2 (Existence of an orthonormal basis)**.** *In every Hilbert space there exists an orthonormal basis.*

Proof. Let H be a Hilbert space. Let $\mathcal{E}$ be the set of all those subsets E of the unit sphere $\partial B_{\le 1}(0)$ in H whose elements are pairwise orthogonal. For every non-empty chain $\mathcal{C} \subseteq \mathcal{E}$ one can readily check that $\bigcup_{C\in\mathcal{C}} C \in \mathcal{E}$, so $\mathcal{C}$ has an upper bound in $\mathcal{E}$. Hence, $\mathcal{E}$ contains a maximal element E according to Zorn's lemma 2.4.7.

The maximality of E implies that $E^\perp = \{0\}$, so it follows from Exercise 14.1(f) on Homework Sheet 14 that $\operatorname{span} E$ is dense in H.(This is the only point in the argument where we use the completeness of H!). So $(e)_{e\in E}$ is an orthonormal basis of H. □

According to Example 6.1.5(c) the space $\ell^2(\mathbb{N};\mathbb{K})$ is a Hilbert space when endowed with the innner product given by $\left(f \mid g\right) := \sum_{k=1}^\infty \left(f_k \mid g_k\right)$ for all $f,g \in \ell^2(\mathbb{N};\mathbb{K})$. This can easily be generalized to more general index sets: for each set J the space $\ell^2(J;\mathbb{K})$ is a Hilbert space when endowed with the inner product given by

$$\left(f \mid g\right) = \sum_{j\in J} \overline{f_j} g_j$$

for all $f,g \in \ell^2(J;\mathbb{K})$. Part (d) of the following theorem shows that actually every Hilbert space is isometrically isomorphic to a space of the form $\ell^2(J;\mathbb{K})$.

Theorem 6.4.3 (Properties of orthonormal basis)**.** *Let H be a Hilbert space over $\mathbb{K}$ and let $(u_j)_{j\in J}$ be an orthonormal basis of H.*

(a) *Let $\mathcal{F}$ denote the set of all finite subsets of J, directed by set inclusion. Then the net $\left(\sum_{j\in F} u_j \otimes u_j\right)_{F\in\mathcal{F}}$ consists of orthogonal projections and converges strongly to* id_H.

(b) *For each $x \in H$ one has $x = \sum_{j\in J} u_j \left(u_j \mid x\right)$ and $\|x\|^2 = \sum_{j\in J} \left|\left(u_j \mid x\right)\right|^2$.*[11]

Lecture 26 (Th, 23 July) starts with Theorem 6.4.3(c).

(c) *For all $x, y \in H$ one has $\left(x \mid y\right) = \sum_{j\in J} \left(x \mid u_j\right)\left(u_j \mid y\right)$.*

(d) *One has the **Fischer–Riesz** representation theorem for H: the map*

$$\begin{aligned} U\colon H &\to \ell^2(J;\mathbb{K}), \\ x &\mapsto \left(\left(u_j \mid x\right)\right)_{j\in J} \end{aligned}$$

is an isometric isomorphism of Banach spaces.

(e) *The index set J is finite if and only if H is finite-dimensional, and in this case $\#J = \dim H$.*

The index set J is countable if and only if H is separable.

For the proof of the theorem we use the following two lemmas.

Lemma 6.4.4 (Convergence on total subsets)**.** *Let X, Y be normed spaces over $\mathbb{K}$, let $(T_j)_{j\in J}$ be a bounded net in $\mathcal{L}(X;Y)$ and $T \in \mathcal{L}(X;Y)$. Let $D \subseteq X$ be such that* $\operatorname{span} D$ *is dense in X. If $T_j x \to T x$ for all $x \in D$, then even $T_j x \to T x$ for all $x \in X$ (i.e. $T_j \to T$ strongly).*

Proof. It is clear that $T_j x \to T x$ for all $x \in \operatorname{span} D$. Now let $x \in X$ and let $(x_k)_{k\in\mathbb{N}}$ be a sequence in $\operatorname{span} D$ that converges to x. Set $M := \|T\| \vee \sup_{j\in J} \left\|T_j\right\|$ and note that $M < \infty$ by assumption.

Let $\varepsilon > 0$. There exists $k \in \mathbb{N}$ such that $M\|x_k - x\| \le \varepsilon$ and there exists $j_0 \in J$ such that $\left\|T_j x_k - T x_k\right\| \le \varepsilon$ for all $j \ge j_0$. For each $j \ge j_0$ one thus gets

$$\begin{aligned} \left\|T_j x - T x\right\| &\le \left\|T_j x - T_j x_k\right\| + \left\|T_j x_k - T x_k\right\| + \|T x_k - T x\| \\ &\le M\|x - x_k\| + \left\|T_j x_k - T x_k\right\| + M\|x_k - x\| \le 3\varepsilon, \end{aligned}$$

which proves that claim. ☐

Lemma 6.4.5 (Isometries between Hilbert spaces preserve the inner product)**.** *Let H_1, H_2 be Hilbert spaces over $\mathbb{K}$ and let $S \in \mathcal{L}(H_1;H_2)$. The operator S is isometric if and only if $\left(Sx \mid Sy\right) = \left(x \mid y\right)$ for all $x, y \in H_1$.*

Proof. “$\Leftarrow$“ This implication is clear.

“$\Rightarrow$” Assume S is isometric. Consider the map $a\colon H \times H \to \mathbb{K}$, $a(x,y) := \left(Sx \mid Sy\right)$. Then a and $(\cdot \mid \cdot)$ are Hermitian sesquilinear forms on H, and they coincide on the diagonal of $H \times H$ since

$$a(x,x) = (Sx \mid Sx) = \|Sx\|^2 = \|x\|^2 = (x \mid x)$$

for all $x \in H$. So it follows from the polarisation identity (Proposition 6.1.2(a) if $\mathbb{K} = \mathbb{C}$ and Proposition 6.1.2(a) if $\mathbb{K} = \mathbb{R}$) that $a = (\cdot \mid \cdot)$. ☐

[11] Recall that the sum symbols here denote limits of unconditional series.

Proof of Theorem 6.4.3. (a) For each $F \in \mathcal{F}$ we set $P_F := \sum_{j\in F} u_j \otimes u_j$. Let us first show that each such P_F is indeed an orthogonal projection. To this end it suffices to show that P_F is a self-adjoint projection (Proposition 6.3.7). By Proposition 6.3.10(c) P_F is a sum of self-adjoint rank-1 projections, so P_F is indeed self-adjoint. To see that P_F is also a projection, we compute

$$P_F^2 = \sum_{j\in F}\sum_{k\in F}(u_j\otimes u_j)(u_k\otimes u_k) = \sum_{j\in F}\sum_{k\in F}\underbrace{\left(u_j \mid u_k\right)}_{=\delta_{jk}}(u_j\otimes u_k) = \sum_{j\in F} u_j\otimes u_j = P_F,$$

where the second equality follows from Proposition 6.3.10(a).

It remains to show that the net $\left(P_F\right)_{F\in\mathcal{F}}$ converges strongly to id_H. Since every orthogonal projection has norm at most 1 (Proposition 6.3.7), the net $\left(P_F\right)_{F\in\mathcal{F}}$ is bounded. Moreover, for each $k\in J$ and each $F\in\mathcal{F}$ with $\{k\}\subseteq F$ one has

$$P_F u_k = \sum_{j\in F}(u_j\otimes u_j)u_k = \sum_{j\in F} u_j \underbrace{\left(u_j \mid u_k\right)}_{=\delta_{jk}} = u_k.$$

This implies that $P_F u_k \xrightarrow{F} u_k$. Since $(u_j)_{j\in J}$ is an orthonormal basis of H, the span of those vectors is, by definition, dense in H, so Lemma 6.4.4 implies that $P_F \to \mathrm{id}_H$ strongly.

(b) The equality $x = \sum_{j\in J} u_j \left(u_j \mid x\right)$ for all $x\in H$ is just a notation for the assertion of (a) (see Definition 1.3.6(d)).

To show the norm equality, let $x\in H$. For each $F\in\mathcal{F}$ one has

$$\begin{aligned}\|P_F x\|^2 &= \left\|\sum_{j\in F}(u_j\otimes u_j)x\right\|^2 = \left\|\sum_{j\in F} u_j\left(u_j\mid x\right)\right\|^2 = \left(\sum_{j\in F}\left(u_j\mid x\right)u_j \;\middle|\; \sum_{k\in F}(u_k\mid x)\,u_k\right)\\ &= \sum_{j\in F}\sum_{k\in F}\overline{\left(u_j\mid x\right)}\,(u_k\mid x)\underbrace{\left(u_j\mid u_k\right)}_{=\delta_{jk}} = \sum_{j\in F}\left|\left(u_j\mid x\right)\right|^2.\end{aligned}$$

Since $\|\cdot\|^2 : H\to[0,\infty)$ is continuous, it follows from (a) that

$$\|x\|^2 = \left\|\lim_F P_F x\right\|^2 = \lim_F \|P_F x\|^2 = \lim_F \sum_{j\in F}\left|\left(u_j\mid x\right)\right|^2 = \sum_{j\in J}\left|\left(u_j\mid x\right)\right|^2.$$

(d) It follows from (b) that U indeed maps H to $\ell^2(J;\mathbb{K})$ and is isometric. To prove that U is surjective, first $\alpha\in\ell^2(J;\mathbb{K})$ be an element that vanishes on all but finitely many indices, i.e. $\alpha\in c_{00}(J;\mathbb{K})$. Then $x := \sum_{j\in J}\alpha_j u_j$ is an element of H and $Ux=\alpha$ since

$$(Ux)_k = (u_k\mid x) = \sum_{j\in J}\alpha_j\left(u_k\mid u_j\right) = \alpha_k$$

for all $k\in J$. Hence, the range UH contains $c_{00}(J;\mathbb{K})$. Similarly as in the proof of Example 1.1.6(a) one can show that $c_{00}(J;\mathbb{K})$ is dense in $\ell^2(J;\mathbb{K})$. Hence, UH is dense in $\ell^2(J;\mathbb{K})$. On the other hand, since J is isometric on H is complete, the range JH is also complete and thus closed. Therefore, $UH = \ell^2(J;\mathbb{K})$.

(c) Since the map U from (d) is isometric, it preserves the inner product according to Lemma 6.4.5.

(e) The first part of the claim follows from (d) since $\ell^2(J;\mathbb{K})$ is finite-dimensional if and only if J is finite, and in this case $\dim \ell^2(J;\mathbb{K}) = \#J$.

The second part also follows from (d): if J is countable, then it is finite or in bijection to $\mathbb{N}$, so $\ell^2(J;\mathbb{K})$ is separable according to Example 1.1.6(a). On the other hand, if J is uncountable, then the family $(e_j)_{j\in J}$ if canonical unit vectors consists of uncountably many vectors that all have mutual distance $\sqrt{2}$, which implies that $\ell^2(J;\mathbb{K})$ is not separable (see the argument in the proof of Example 1.1.7(c) or Bonus Exercise 1.7(a) on Homework Sheet 1). □

Corollary 6.4.6 (Separable Hilbert spaces). *All separable infinite-dimensional Hilbert spaces over* $\mathbb{K}$ *are isometrically isomorphic.*

Proof. If H_1, H_2 are two such spaces, then they are both isometrically isomorphic to $\ell^2(\mathbb{N};\mathbb{K})$ according to Theorem 6.4.3(d). □

Here is one of the most classical example of an orthonormal basis of a Hilbert spaces. We present it in two different versions: on the complex unit circle and on the interval $[0,2\pi]$ (but the proof will show that both versions essentially say the same thing).

Examples 6.4.7 (Fourier series). For each $k \in \mathbb{Z}$, let $e_k\colon [0,2\pi] \to \mathbb{C}$, $s \mapsto e^{iks}$. Then the family $(\frac{1}{\sqrt{2\pi}}\tilde{e}_k)_{k\in\mathbb{Z}}$ is an orthonormal basis of $L^2 := L^2([0,2\pi], \mathcal{B}([0,2\pi]), \lambda_1; \mathbb{C})$.

For every $\tilde{f} \in L^2$ one thus has, according to Theorem (b),

$$\tilde{f} = \sum_{k\in\mathbb{Z}} \tilde{e}_k \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} e^{-iks} f(s)\,\mathrm{d}\mu(s),$$

where the sum denotes the L^2-limit of the corresponding unconditional series. This formula called the **Fourier series expansion** of $\tilde{f}$. The numbers $\frac{1}{\sqrt{2\pi}}\int_0^{2\pi} e^{-iks} f(s)\,\mathrm{d}\mu(s)$ are called the **Fourier coefficients** of $\tilde{f}$.

Proof. Let $j, k \in \mathbb{Z}$. If $j = k$, then

$$\left(\frac{1}{\sqrt{2\pi}}\tilde{e}_j \;\middle|\; \frac{1}{\sqrt{2\pi}}\tilde{e}_k\right) = \frac{1}{2\pi}\int_{[0,2\pi]} \underbrace{\overline{e^{iks}}e^{iks}}_{=1}\,\mathrm{d}\lambda_1(s) = 1.$$

On the other hand, if $j \neq k$, then

$$\begin{aligned}\left(\frac{1}{\sqrt{2\pi}}\tilde{e}_j \;\middle|\; \frac{1}{\sqrt{2\pi}}\tilde{e}_k\right) &= \frac{1}{2\pi}\int_{[0,2\pi]} \overline{e^{ijs}}e^{iks}\,\mathrm{d}\lambda_1(s) \\ &= \frac{1}{2\pi}\int_0^{2\pi} e^{i(k-j)s}\,\mathrm{d}s = \frac{1}{2\pi i(k-j)} \underbrace{e^{i(k-j)s}\Big|_{s=0}^{s=2\pi}}_{=1-1=0} = 0.\end{aligned}$$

So $(\frac{1}{\sqrt{2\pi}}\tilde{e}_k)_{k\in\mathbb{Z}}$ is indeed an orthonormal system. The density of its span is proved in Exercise 14.5(d) on Homework Sheet 14. □

Appendices

Appendix A

An Overview over Important Banach Spaces

A.1 Finite dimensional spaces

Every finite-dimensional normed space X has the following properties:

(a) **Completness:** X is a Banach space (Proposition 1.0.6).

(b) **Compactness properties:** A subset $S \subseteq \mathbb{K}^d$ is compact if and only if it is closed; this follows from [5, Theorems 1.3.10 and 1.7.2]. In particular, the closed unit ball $B_{\le 0}(1)$ is compact.

Conversely, if the closed unit ball in a normed space is compact, then the space is finite dimensional (add reference...).

(c) **Properties of the norm:** Any two norms on X are equivalent (this follows from [5, Theorem 1.7.2]).

(d) **Norm convergence:** If $X = \mathbb{K}^d$, then norm convergence in X (of sequences or nets) is equivalent to componentwise convergence [5, Theorem 1.3.13].[1]

(e) **Dual space:** If $X = \mathbb{K}^d$, then the dual space $(\mathbb{K}^d)'$ can be identified with the space $\mathbb{K}^{1\times d}$ of row vectors.

If $p, p' \in [1,\infty]$ such that $\frac{1}{p} + \frac{1}{p'} = 1$ and $\mathbb{K}^d$ is endowed with $\|\cdot\|_p$, then the operator norm of every functional $x' \in \mathbb{K}^{1\times d}$ is $\|x'\| = \|(x')^{\mathrm{T}}\|_{p'}$. This can be shown by the same arguments as in Example 1.2.4 and in Exercise 2.2 on Homework Sheet 2.

(f) **Weak and weak*-convergence, reflexivity:** One can show (although we have not done so in this course) that, on a finite-dimensional vector space over $\mathbb{K}$, all Hausdorff topologies that turn the space into a topological vector space, coincide.

[1] In this reference this is only shown for sequences, but the proof for nets is the same.

In particular, norm convergence and weak convergence on X coincide.

Each finite-dimensional Banach space X is reflexive (this follows, for instance, by using that each finite-dimensional normed space has has the same dimension as its dual space) and is thus the dual space of its dual space. So X carries a weak* topology w^*, and w^* also coincides with the norm topology.

(g) **Dense subspaces:** X does not have dense vector subspaces except for itself (since every vector subspace of X is finite-dimensional and thus closed according to Example 1.1.2).

(h) **Separability:** X is separable. This follows from Proposition 1.1.5 since X has a finite basis and the span of the basis is equal to X.

A.2 Sequence spaces

Let $p \in [1,\infty]$. We abbreviate $\ell^p := \ell^p(\mathbb{N};\mathbb{K})$. Note that $\ell^p = L^p(\mathbb{N}, 2^{\mathbb{N}}, \#; \mathbb{K})$ (Examples 1.0.7(c) and 1.1.7(c)). Properties of ℓ^p:

(a) **Completeness:** ℓ^p is a Banach space (Examples 1.0.7(c) and 1.1.7(c)).

(b) **Compactness properties:** For $p \in [1,\infty)$ a set $K \subseteq \ell^p$ is relatively compact if and only if K is bounded and

$$\sup_{x\in K} \sum_{j=n}^{\infty} |x_j|^p \to 0 \quad \text{as } n \to \infty.$$

See Exercisee 11.4 on Homework Sheet 11.

(c) **Properties of the norm:** The triangle inequality for the norm on ℓ^p is not obvious, but is a consequence of Hölder's inequality; see [6, Theorem 3.2.2(e) and (d)].

For $1 \le p \le q \le \infty$ one has the inclusions[2] $\ell^1 \subseteq \ell^p \subseteq \ell^q \subseteq \ell^\infty$ and the inclusion maps have norm 1. Hence,

$$\|x\|_\infty \le \|x\|_q \le \|x\|_p \le \|x\|_1$$

for every sequence $x = (x_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ (where some or all of those values can be ∞).

(d) **Norm convergence:** Norm convergence (of sequences or nets) implies componentwise convergence, but the converse is not true, in general.

Counterexample for the converse implication: The sequence $(e^{(n)})_{n\in\mathbb{N}}$ of canonical unit vectors.

[2] Here is an easy way to remember the direction of these inclusions: (i) remember that the inclusions point in the same direction on the entire ℓ^p-scale; (ii) observe that $\ell^1 \subseteq \ell^\infty$ (since every absolutely summable sequence converges to 0 and is thus bounded).

(e) **Dual space:** For $p \in [1,\infty)$ one has $(\ell^p)' \simeq \ell^{p'}$, where $p' \in (1,\infty]$ satisfies $\frac{1}{p} + \frac{1}{p'} = 1$.

For $p \in (1,\infty)$ (and hence also $p' \in (1,\infty)$) this is proved an Example 1.2.4 and Theorem 1.2.6. For $p = 1$ (and hence $p' = \infty$) precisely the same proof works.

However, the situation is different for $p = \infty$: The dual space of ℓ^∞ isometrically contains ℓ^1 (same proof as for other values of p), but the $(\ell^\infty)'$ is larger than ℓ^1 (Bonus Exercise 2.7 on Homework Sheet 2).

(f) **Weak and weak*-convergence, reflexivity:** Let $p' \in [1,\infty]$ with $\frac{1}{p} + \frac{1}{p'} = 1$.

(i) Let $p \in (1,\infty)$ (equivalently, $p' \in (1,\infty)$).

A bounded net (x_j) in ℓ^p converges weakly (equilvalently, weak*, as ℓ^p is reflexive) to a point $x \in \ell^p$ if and only if it converges componentwise to x.

(ii) End points of the ℓ^p-scale:

The space c_0, ℓ^1 and ℓ^∞ are not reflexiv (Example 5.4.7(a)).

A bounded net (x_j) in c_0 converges weakly to a point $x \in c_0$ if and only if it converges componentwise to x. This can be shown similarly as Example 5.3.9(a) – or it can be derived from that example by using the canonical embedding $c_0 \to (c_0)'' \simeq \ell^\infty$.

A bounded net (x_j) in ℓ^1 converges weak* to a point $x \in \ell^1$ if and only if it converges componentwise to x (Example 5.3.9(b)). Weak convergence of nets in ℓ^1 is not so easy to characterize.[3]

A bounded net (x_j) in ℓ^∞ converges weak* to a point $x \in \ell^\infty$ if and only if it converges pointwise to x (Example 5.3.9(a)).

The characterizations of weak or weak*-convergence that we just listed do not remain true for unbounded nets.

Warning: The above description of weak*-convergence in ℓ^1 holds if one identifies ℓ^1 with the dual space of c_0 in the usual way. However, there are also Banach spaces X that are not isomorphic to c_0 but whose dual space is isomorphic to ℓ^1. If we identify ℓ^1 with X' for such a Banach space X, then the weak* topology on ℓ^1 is different.

(g) **Dense subspaces:** The space c_{00} of all sequences that have only finitely many nonzero components, is dense in ℓ^p if $p \in [1,\infty)$ (see the proof of Example (a)), but not in ℓ^∞.

The closure of c_{00} in ℓ^∞ is the space c_0 of all sequences that converge to 0.

(h) **Separability:** The space ℓ^p is separable if $p \in [1,\infty)$ and not separable if $p = \infty$ (Example 1.1.6(a) and (b)).

[3] Fun fact: But weak convergence of nets sequences (!) in ℓ^1 can be characterized: it is equivalent to norm convergence! This is called the **Schur propeerty** of ℓ^1.

A.3 L^p-spaces

Let $p \in [1,\infty]$, let $(\Omega, \mathcal{A}, \mu)$ be a measure space and let $L^p := L^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$. The space is endowed with the norm $\|\cdot\|_p$ if $p \in [1,\infty)$ and with the essential supremum norm $\|\cdot\|_{\text{ess},\infty}$. See Examples 1.0.7 and 1.1.7 for a more detailed description of how those spaces and norms are defined.

(a) **Completeness:** The space L^p is complete with respect to $\|\cdot\|_p$; for $p \in [1,\infty)$ we proved this in Analysis 3 [6, Theorem 3.2.16]. For $p = \infty$ a similar argument can be used.

(b) **Compactness properties:** For $p \in [1,\infty)$ and the particular measure space $(\Omega, \mathcal{A}, \mu) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$, relative compactness in L^p is characterized by the so-called **Kolomogorov compactness criterion** (which we have not discussed in this course).

(c) **Properties of the norm:** The triangle inequality for the norm on L^p is a consequence of Hölder's inequality; see [6, Theorem 3.2.2(e) and (d)].

If the measure μ is finite – i.e. $\mu(\Omega) < \infty$ – and $1 \le p \le q \le \infty$, one has the inclusions[4] $L^1 \supseteq L^p \supseteq L^q \supseteq L^\infty$ and the inclusion maps are continuous; their norms can be explicitly estimate by using Hölder's inequality.

If μ is even a probability measure, it follows from Hölder's inequality that the inclusion maps have norm 1, so

$$\|f\|_\infty \ge \|f\|_q \ge \|f\|_p \ge \|f\|_1$$

for every measurable $f : \Omega \to \mathbb{K}$ in this case.

(d) **Norm convergence:** Norm convergence is related to (but in general not equivalent to) almost everywhere convergence by Lebesgue's dominated convergence theorem and its partial converse (Section B.2 in Appendix B).

(e) **Dual space:** Let $p' \in\in [1,\infty]$ such that $\frac{1}{p} + \frac{1}{p'} = 1$.

If $p \in (1,\infty)$ (equivalently, $p' \in (1,\infty)$, then $(L^p)' \simeq L^{p'}$ (Example 5.4.7(c)).

If the measure space $(\Omega, \mathcal{A}, \mu)$ is σ-finite, one can also show that $(L^1)' \simeq L^\infty$ (via essentially the same isomorphism an in the case $p \in (1,\infty)$). This is, however, no longer true for general measure spaces.

(f) **Weak and weak*-convergence, reflexivity:**

If $p \in (1,\infty)$, then L^p is reflexive (Example 5.4.7(c)).

One can show that the spaces L^1 and L^∞ are never reflexive unless they are finite-dimensional (for L^1 over a particular measure space this has been done in Exericse 13.2(d) on Homework Shet 13).

[4] To remember the direction of these inclusions: (i) remember that the inclusions point in the same direction on the entire L^p-scale; (ii) observe that $L^1 \supseteq L^\infty$, as every bounded measurable function is obviously integrable on a finite measure space.

(g) **Dense subspaces:** If μ is finite, Ω is endowed with a metric and $\mathcal{A}$ is the Borel σ-algebra $\mathcal{B}(\Omega)$ on Ω, then

$$\{\tilde{f} \mid f : \Omega \to \mathbb{K} \text{ is continuous and bounded}\}$$

is dense in L^p (Bonus Exercise 2.8(e) on Homework Sheet 2).

For σ-finite (rather than finite) measures one can sometimes show similar results. For instance, let $\Omega \subseteq \mathbb{R}^d$ be open and endow it with the Borel σ-algebra and the Lebesgue measure. Then

$$\{\tilde{f} \mid f : \Omega \to \mathbb{K} \text{ is continuous and } \exists K \subseteq \Omega \text{ compact such that } f|_{\Omega \setminus K} = 0\}$$

is dense in L^p.

(h) **Separability:** If $p \in [1,\infty)$, the measure μ is σ-finite, and the σ-Algebra $\mathcal{A}$ has a countable generator, then L^p is separable.

For finite measures μ this was proved in Bonus Exercise 2.8(c) on Homework Sheet 2. The σ-finite case can be reduced to the finite case: If μ is σ-finite, one can show that $L^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ is isometrically isomorphic to $L^p(\Omega,\mathcal{A},\nu;\mathbb{K})$ for the finite measure $\nu : \mathcal{A} \to [0,\infty]$ given by $\mathrm{d}\nu = h^p \,\mathrm{d}\mu$ for a function $h \in \mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ that is strictly positive μ-almost everywhere. This is called a **change of density** argument.

A.4 Spaces of continuous functions

In this section, let M be a compact topological Hausdorff space. We consider the space $\mathrm{C}(M;\mathbb{K})$ of continuous functions $M \to \mathbb{K}$, endowed with the norm $\|\cdot\|_\infty$.

(a) **Completeness:** The space $\mathrm{C}(M;\mathbb{K})$ is a Banach space. (For a discussion of the proof, see e.g. Exercise B.25 on the addition exercise sheet B; in the exercise only the special case $M = [0,1]$ is considered, but the proof is the same for general compact M.)

(b) **Compactness properties:** Relative compactness of subsets of $\mathrm{C}(M;\mathbb{K})$ is characterized by the Arzelà–Ascoli Theorem (Theorem 5.2.11).

(c) **Properties of the norm:** The norm on $\mathrm{C}(M;\mathbb{K})$ interacts well with the pointwise multiplication of functions; for instance, one has $\|fg\|_\infty \le \|f\|_\infty \|g\|_\infty$ and $\|f^n\|_\infty = \|f\|_\infty^n$ for all $f, g \in \mathrm{C}(M;\mathbb{R})$ and all $n \in \mathbb{N}_0$.

(d) **Norm convergence:** Coincides with the notions **uniform convergence** that you have first seen in your Analysis 1 course.

Dini's theorem says that pointwise convergence of an increasing (or decreasing) net in $\mathrm{C}(M;\mathbb{R})$ implies norm convergence if the limit is also continuous. The theorem was proved in Exercise 1.2 on Homework Sheet 1 for the special case where M is a metric space and where the net is a sequence; but the proof remains literally the same of M is a compact topological space and if one considers even nets instead of sequences.

(e) **Dual space:** If the compact space M is a metric space, the dual space of $\mathrm{C}(M;\mathbb{R})$ consists of all finite signed measures on the Borel σ-algebra $\mathcal{B}(M)$. This is discussed in detail in the Addendum in Section 1.4.

If the topology on M does not stem from a metric, the situation is a bit subtler: then $\mathrm{C}(M;\mathbb{R})'$ consists only of those signed Borel measures that satisfy a certain additional regularity property; we do not discuss this in detail here.

When the scalar field is $\mathbb{C}$ one has to consider $\mathbb{C}$-valued measures instead of signed measures only.

(f) **Weak and weak*-convergence, reflexivity:** A sequence (f_n) in $\mathrm{C}(M;\mathbb{K})$ converges weakly to a function $f \in \mathrm{C}(M;\mathbb{K})$ if and only if it is bounded and converges pointwise to f.

The implication "$\Rightarrow$" follows from the uniform boundedness principle and by testing against point evaluations. The converse implication "$\Leftarrow$" follows from the fact that every functional $\varphi \in \mathrm{C}(M;\mathbb{K})'$ can be represented as a $\mathbb{K}$-valued measure (see the previous point) together with a version of dominated convergence theorem for integrals with respect to $\mathbb{K}$-valued measures.

One can prove that $\mathrm{C}(M;\mathbb{K})$ is not reflexive unless M is finite (equivalently, $\mathrm{C}(M;\mathbb{K})$ is finite-dimensional).

For special cases such as e.g. $M = [0,1]$ this is not difficult to see: one can then easily find a sequence (f_n) in the unit ball of $\mathrm{C}(M;\mathbb{K})$ that do not have a weakly convergent subnet.

(g) **Dense subspaces:** The Stone–Weierstraß appromimation theorems give sufficient (but not necessary) criteria for a vector subspace of $\mathrm{C}(M;\mathbb{K})$ to be dense. You have seen three versions of the theorem:

The lattice version for real scalars (Theorem 1.1.8), the algebra version for real scalars (Theorem 1.1.9), and the algebra version for complex scalars (Exercise 14.5(a) in Homework Sheet 15).

(h) **Separability:** If the topology on M is induced by a metric, then $\mathrm{C}(M;\mathbb{K})$ is separable (for $\mathbb{K} = \mathbb{R}$ this is Corollary 1.1.10; the complex case can be derived from the real case).

One can show that the converse implication is also true, i.e. $\mathrm{C}(M;\mathbb{R})$ is separable if and only if the topology on M is induced by a metric.

Appendix B

Some Results from Measure Theory

B.1 The monotone convergence theorem and Fatou's lemma

In this section we briefly recall two results from Analysis 3, but we formulate them for almost everywhere pointwise convergence rather than pointwise convergence.

Theorem B.1.1 (Monotone convergence theorem)**.** *Let* $(\Omega, \mathcal{A}, \mu)$ *be a measure space. Let* $f, f_n : \Omega \to [0, \infty]$ *be measurable for all* $n \in \mathbb{N}$*. Assume that for every* $n \in \mathbb{N}$ *one has* $f_n \leq f_{n+1}$ *almost everywhere and that* $f_n \to f$ *pointwise almost everywhere as* $n \to \infty$*. Then* $\int_\Omega f \, \mathrm{d}\mu = \lim_{n\to\infty} \int_\Omega f_n \, \mathrm{d}\mu$*.*

Proof. By assumption we can find measurable sets $M, N_1, N_2, \dots \subseteq \Omega$ of measure 0 such that $f_n(\omega) \leq f_{n+1}(\omega)$ for all $n \in \mathbb{N}$ and all $\omega \in \Omega \setminus N_n$, and such that $f_n(\omega) \to f(\omega)$ for all $\omega \in \Omega \setminus M$. Set

$$N := M \cup \bigcup_{n\in\mathbb{N}} N_n.$$

Then N is also measurable and $\mu(N) = 0$, and for all $\omega \in \Omega \setminus N$ one has

$$f_n(\omega) \leq f_{n+1}(\omega) \quad \text{for all } n \in \mathbb{N} \qquad \text{and} \qquad f_n(\omega) \overset{n\to\infty}{\to} f(\omega).$$

So it follows from the pointwise version of the monotone convergence theorem [6, Theorem 2.2.3(b)] that

$$\int_\Omega f_n \, \mathrm{d}\mu = \int_\Omega f_n \, \mathbb{1}_{N^\complement} \, \mathrm{d}\mu \to \int_\Omega f \, \mathbb{1}_{N^\complement} \, \mathrm{d}\mu = \int_\Omega f \, \mathrm{d}\mu,$$

which proves the claim. □

Lemma B.1.2 (Fatou)**.** *Let* $(\Omega, \mathcal{A}, \mu)$ *be a measure space. Let* $f, f_n : \Omega \to [0, \infty]$ *be measurable for all* $n \in \mathbb{N}$*. Assume that* $f = \liminf_{n\to\infty} f_n$ *pointwise almost everywhere. Then* $\int_\Omega f \, \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_\Omega f_n \, \mathrm{d}\mu$*.*

Proof. This follows from the pointwise version of Fatou's lemma [6, Lemma 2.2.6] by a similar argument as the one that we used to derive Theorem B.1.1 follows from the pointwise version of the monotone convergence theorem. □

B.2 The dominated convergence theorem and its partial converse

Just as the monotone convergence theorem and Fatous' lemma, the dominated convergence theorem also holds in a version for almost everywhere convergence:

Theorem B.2.1 (Dominated convergence theorem)**.** *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Let $f, f_n : \Omega \to \mathbb{K}$ be measurable for all $n \in \mathbb{N}$ and let $h \colon \Omega \to [0, \infty]$ be integrable. Assume that for every $n \in \mathbb{N}$ one has $|f_n| \le h$ almost everywhere and that $f_n \to f$ pointwise almost everywhere as $n \to \infty$. Then each f_n is integrable, f is integrable, and $\int_\Omega f \, \mathrm{d}\mu = \lim_{n\to\infty} \int_\Omega f_n \, \mathrm{d}\mu$.*

Proof. Note that the integrability of h implies that $h(\omega) < \infty$ for almost all $\omega \in \Omega$. One can now derive theorem from the pointwise version of the dominated convergence theorem [6, Theorem 2.3.8] by a similar argument as in the proof of Theorem B.1.1 above. □

A nice consequence of the dominated convergence theorem is the result about norm convergence in L^p-spaces.

Corollary B.2.2. *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and let $p \in [1, \infty)$. Let $(\tilde{f}_n)$ be a sequence in $L^p := L^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ and $\tilde{f}, \tilde{h} \in L^p$. Assume that $f_n \to f$ pointwise almost everywhere and that for every $n \in \mathbb{N}$, $|f_n| \le h$ almost everywhere. Then $\tilde{f}_n \to \tilde{f}$ with respect to $\|\cdot\|_p$.*

Proof. This is proved in Exercise 6.2 on Präsenzblatt 6. □

The converse of Corollary B.2.2 is not true, in general. However, at least a partial converse does indeed hold – one only needs to choose an appropriate subsequence:

Theorem B.2.3 (Partial converse to the dominated convergence theorem)**.** *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and let $p \in [1, \infty)$. Let $(\tilde{f}_n)$ be a sequence in $L^p := L^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ that converges to a vector $\tilde{f} \in L^p$ with respect to $\|\cdot\|_p$.*

Then there exists $\tilde{h} \in L^p$ and a subsequence $(\tilde{f}_{n_k})$ of $(\tilde{f}_n)$ such that for every k, $|f_{n_k}| \le h$ almost everywhere and such that $f_{n_k} \to f$ pointwise almost everywhere.

For the proof we need the following auxiliary result:

Lemma B.2.4 (Functions that decrease to **0**)**.** *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and let $p \in [1, \infty)$. Let $(\tilde{g}_n)$ be a sequence in $L^p := L^p(\Omega, \mathcal{A}, \mu; \mathbb{R})$ such that, for every $n \in \mathbb{N}$, $g_{n+1} \le g_n$ almost everywhere. If $\tilde{g}_n \to 0$ with respect to $\|\cdot\|_p$, then $g_n \ge 0$ almost everywhere for every $n \in \mathbb{N}$ and $g_n \to 0$ almost everywhere as $n \to \infty$.*

Proof. After modifying all functions on a null set we may assume that the sequence (g_n) is pointwise decreasing. Then the sequence (g_n^-) of negative parts is pointwise increasing. For all $m, n \in \mathbb{N}$ with $m \le n$ we thus have so $g_m^- \le g_n^- \le |g_n|$ and thus

$$\|g_m^-\|_p \le \||g_n|\|_p = \|g_n\|_p \overset{n\to\infty}{\to} 0,$$

so $g_m^- = 0$ almost everywhere and hence, $g_m \ge 0$ almost everywhere.

After modifying, again, all functions on a null set we may thus assume now that, in addition, $g_n \geq 0$ pointwise for every $n \in \mathbb{N}$. So the sequence (g_n) in pointwise decreasing and bounded below by 0; hence, it converges pointwise to a measurable function $g\colon \Omega \to [0,\infty)$. It only remains to show that $g = 0$ almost everywhere. Fatou's lemma (Lemma B.1.2) implies

$$\|g\|_p^p = \int_\Omega g^p \,\mathrm{d}\mu \leq \liminf_{n\to\infty} \int_\Omega g_n^p \,\mathrm{d}\mu = \liminf_{n\to\infty} \|g_n\|_p^p = 0,$$

so $\|g\|_p = 0$ and hence $g = 0$ almost everywhere, as claimed. □

Proof of Theorem B.2.3. There is a subsequence $(\tilde{f}_{n_k})$ of $(\tilde{f}_n)$ such that $\|\tilde{f}_{n_k} - \tilde{f}\|_p \leq \frac{1}{2^k}$ for all $k \in \mathbb{N}$. Set $h_k := |f_{n_k} - f|$ for each $k \in \mathbb{N}$. Then

$$\|\tilde{h}_k\|_p = \|\tilde{f}_{n_k} - \tilde{f}\|_p \leq \frac{1}{2^k}$$

for each $k \in \mathbb{N}$. So we can define $\tilde{g}_k := \sum_{j=k}^\infty \tilde{h}_j \in L^p$ for each $k \in \mathbb{N}$, where the series is absolutely convergent and thus convergent in L^p. Note that $\|\tilde{g}_k\|_p \to 0$ as $k \to \infty$. For all $k \in \mathbb{N}$ we have $\tilde{g}_k = \tilde{g}_{k+1} + \tilde{h}_k$, and hence $g_k = g_{k+1} + h_k$ almost everywhere. This shows that $g_k \geq g_{k+1}$ almost everywhere. Therefore, Lemma B.2.4 implies that $g_k \to 0$ pointwise almost everywhere as $k \to \infty$ and that, for every k, $g_k \geq 0$ almost everywhere. For each k one has

$$|f_{n_k} - f| = h_k \leq h_k + g_{k+1} = g_k,$$

where the inequality holds almost everywhere. We thus conclude that $|f_{n_k} - f| \to 0$ almost everywhere as $k \to \infty$. □

Bibliography

[1] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Universitext. New York, NY: Springer, 2011.

[2] J. A. Clarkson. Uniformly convex spaces. *Trans. Am. Math. Soc.*, 40:396–414, 1936.

[3] Nelson Dunford and Jacob T. Schwartz. *Linear operators. Part I: General theory. With the assistance of William G. Bade and Robert G. Bartle.* Wiley Class. Libr. New York etc.: John Wiley &| Sons Ltd., repr. of the orig., publ. 1959 by John Wiley & Sons Ltd., Paperback ed. edition, 1988.

[4] Jürgen Elstrodt. *Maß- und Integrationstheorie.* Springer Spektrum, 8th enlarged and updated edition edition, 2018.

[5] Jochen Glück. Analysis 2. Vorlesungsmanuskript, Sommersemester 2025.

[6] Jochen Glück. Analysis 3. Vorlesungsmanuskript, Wintersemester 2025/26.

[7] Winfried Kaballo. *Grundkurs Funktionalanalysis.* Heidelberg: Spektrum Akademischer Verlag, 2011.

[8] Winfried Kaballo. *Aufbaukurs Funktionalanalysis und Operatortheorie. Distributionen, lokalkonvexe Methoden, Spektraltheorie.* Berlin: Springer Spektrum, 2014.

[9] Walter Rudin. *Functional analysis.* New York, NY: McGraw-Hill, 2nd ed. edition, 1991.

[10] Helmut H. Schaefer and M. P. Wolff. *Topological vector spaces.*, volume 3 of *Grad. Texts Math.* New York, NY: Springer, 2nd ed. edition, 1999.

[11] Lynn Arthur Steen and J. Arthur jun. Seebach. *Counterexamples in topology. 2nd ed.* Springer, 1978.

[12] Dirk Werner. *Funktionalanalysis.* Springer-Lehrb. Berlin: Springer Spektrum, 8th revised edition edition, 2018.

[13] Kosaku Yosida. *Functional analysis. 6th ed*, volume 123 of *Grundlehren Math. Wiss.* Springer, Cham, 1980.