Functional Analysis

Lecture Notes of Winter Semester 2017/18

These lecture notes are based on my course from winter semester 2017/18. I kept the results discussed in the lectures (except for minor corrections and improvements) and most of their numbering. Typically, the proofs and calculations in the notes are a bit shorter than those given in class. The drawings and many additional oral remarks from the lectures are omitted here. On the other hand, the notes contain a couple of proofs (mostly for peripheral statements) and very few results not presented during the course. With 'Analysis 1–4' I refer to the class notes of my lectures from 2015–17 which can be found on my webpage. Occasionally, I use concepts, notation and standard results of these courses without further notice.

I am happy to thank Bernhard Konrad, Jörg Bäuerle and Johannes Eilinghoff for their careful proof reading of my lecture notes from 2009 and 2011.

Karlsruhe, May 25, 2020 Roland Schnaubelt

Contents

Chapter 1. Banach spaces 2
1.1. Basic properties of Banach and metric spaces 2
1.2. More examples of Banach spaces 20
1.3. Compactness and separability 28

Chapter 2. Continuous linear operators 38
2.1. Basic properties and examples of linear operators 38
2.2. Standard constructions 47
2.3. The interpolation theorem of Riesz and Thorin 54

Chapter 3. Hilbert spaces 59
3.1. Basic properties and orthogonality 59
3.2. Orthonormal bases 64

Chapter 4. Two main theorems on bounded linear operators 69
4.1. The principle of uniform boundedness and strong convergence 69
4.2. Sobolev spaces 75
4.3. The open mapping theorem and invertibility 81

Chapter 5. Duality 86
5.1. The duals of sequence and Lebesgue spaces 86
5.2. The extension theorem of Hahn-Banach 91
5.3. Reflexivity and weak convergence 101
5.4. Adjoint operators 109

Chapter 6. The spectral theorem for compact self-adjoint operators 117
6.1. Compact operators 117
6.2. The spectral theorem 121

Bibliography 126

CHAPTER 1

Banach spaces

In these notes $X \neq \{0\}$ and $Y \neq \{0\}$ are always vector spaces over the field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

1.1. Basic properties of Banach and metric spaces

We start with the fundamental definitions of this course which connect the linear structure with convergence.

DEFINITION 1.1. *A* seminorm *on X is a map $p : X \to \mathbb{R}$ satisfying*

a) $p(\alpha x) = |\alpha|\, p(x)$ *(homogeneity),*

b) $p(x + y) \leq p(x) + p(y)$ *(triangle inquality)*

for all $x, y \in X$ and $\alpha \in \mathbb{F}$. If p fulfills in addition

c) $p(x) = 0 \implies x = 0$ *(definiteness)*

for all $x \in X$, then p is a norm. *One mostly writes $p(x) = \|x\|$ and $p = \|\cdot\|$. The pair $(X, \|\cdot\|)$ (or just X) is called a* normed vector space.

In view of Example 1.4(a), we interpret $\|x\|$ as the length of x and $\|x - y\|$ as the distance between x and y. Seminorms will only occur as auxiliary objects, see e.g. Proposition 1.8.

DEFINITION 1.2. *Let $\|\cdot\|$ be a seminorm on a vector space X. A sequence $(x_n)_{n\in\mathbb{N}} = (x_n)_n = (x_n)$ in X* converges *to a* limit *$x \in X$ if*

$$\forall\, \varepsilon > 0 \;\; \exists\, N_\varepsilon \in \mathbb{N} \;\; \forall\, n \geq N_\varepsilon : \quad \|x_n - x\| \leq \varepsilon.$$

We then write $x_n \to x$ as $n \to \infty$ or $x = \lim_{n\to\infty} x_n$. Moreover, (x_n) is a Cauchy sequence *in X if*

$$\forall\, \varepsilon > 0 \;\; \exists\, N_\varepsilon \in \mathbb{N} \;\; \forall\, n, m \geq N_\varepsilon : \quad \|x_n - x_m\| \leq \varepsilon.$$

A normed vector space $(X, \|\cdot\|)$ is a Banach space *if each Cauchy sequence in $(X, \|\cdot\|)$ converges in X. Then one also calls $(X, \|\cdot\|)$ or $\|\cdot\|$* complete.

In this section we discuss (and partly extend) various results from Analysis 2 whose proofs were mostly omitted in the lectures. We start with simple properties of norms and limits.

REMARK 1.3. Let $\|\cdot\|$ be a seminorm on a vector space X and (x_n) be a sequence in X. The following facts are shown as in Analysis 2, see e.g. Satz 1.2.

a) The vector 0 has the seminorm 0.

b) We have $\big| \|x\| - \|y\| \big| \le \|x-y\|$ and $\|x\| \ge 0$ for all $x, y \in X$.

c) If (x_n) converges, then it is a Cauchy sequence.

d) If (x_n) converges or is Cauchy, it is *bounded*; i.e., $\sup_{n\in\mathbb{N}} \|x_n\| < \infty$.

e) If $x_n \to x$ and $y_n \to y$ in X as $n \to \infty$ and $\alpha, \beta \in \mathbb{F}$, then the linear combinations $\alpha x_n + \beta y_n$ tend to $\alpha x + \beta y$ in X.

f) Limits are unique in the norm case: Let $\|\cdot\|$ be a norm. If $x_n \to x$ and $x_n \to y$ in X as $n \to \infty$ for some $x, y \in X$, then $x = y$. ◊

For our basic examples below and later use, we introduce some notation. Let X be a vector space and $S \neq \emptyset$ be a set. For maps $f, g : S \to X$ and numbers $\alpha \in \mathbb{F}$ one defines the functions

$$f + g : S \to X;\ (f+g)(s) = f(s) + g(s),$$
$$\alpha f : S \to X;\ (\alpha f)(s) = \alpha f(s).$$

It is easily seen that the set $\{f : S \to X\}$ becomes a vector space endowed with the above operations. Function spaces are always equipped with this sum and scalar multiplication. Let $X = \mathbb{F}$. Here one puts

$$fg : S \to \mathbb{F}; \quad (fg)(s) = f(s)g(s).$$

Let $\alpha, \beta \in \mathbb{R}$. We then write $f \ge \alpha$ ($f > \alpha$, respectively) if $f(s) \ge \alpha$ ($f(s) > \alpha$, respectively) for all $s \in S$. Similarly one defines $\alpha \le f \le \beta$, $f \le g$, and so on.

EXAMPLE 1.4. a) $X = \mathbb{F}^m$ is a Banach space for the norms

$$|x|_p = \begin{cases} \Big(\sum_{k=1}^m |x_k|^p \Big)^{\frac{1}{p}}, & 1 \le p < \infty, \\ \max\{|x_k| \,|\, k = 1, \dots, m\}, & p = \infty, \end{cases}$$

where $x = (x_1, \dots, x_m) \in \mathbb{F}^m$. Moreover, vectors v_n converge to x in $\mathbb{F}^m$ as $n \to \infty$ for each of these norms if and only if all components $(v_n)_k$ tend to x_k in $\mathbb{F}$ as $n \to \infty$. See Satz 1.4, 1.9 and 1.13 in Analysis 2. We always equip $X = \mathbb{F}$ with the absolute value $|\cdot|$ which coincides with each of the above norms.

b) Let $(X, \|\cdot\|)$ be a Banach space and K a compact metric space. Then the set

$$E = C(K, X) = \{f : K \to X \,|\, f \text{ is continuous}\}$$

endowed with the supremum norm

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

is a Banach space. We equip E with this norm, unless something else is specified.

Before proving the claim, we note that the above supremum is a maximum and thus finite, cf. Theorem 1.45 or Analysis 2, and that convergence in $\|\cdot\|_\infty$ is just uniform convergence from Analysis 1. In the special case $X = \mathbb{R}$, the norm $\|f - g\|_\infty$ is the maximal vertical distance between the graphs of $f, g \in E$. If $f \geq 0$ describes the temperature in K, for instance, then $\|f\|_\infty$ is the maximal temperature. Moreover, the closed ball $\overline{B}_E(f, \varepsilon)$ around f of radius $\varepsilon > 0$ consists of all functions g in E whose graph belongs to an 'ε–tube' around f.

PROOF. It is clear that E is a vector space. Let $f, g \in E$ and $\alpha \in \mathbb{F}$. Since $\|\cdot\|$ is a norm, we obtain

$$\|f\|_\infty = 0 \quad \Longrightarrow \quad \forall\, s \in K: \ f(s) = 0 \quad \Longrightarrow \quad f = 0,$$
$$\|\alpha f\|_\infty = \sup_{s\in K} \|\alpha f(s)\| = \sup_{s\in K} |\alpha|\, \|f(s)\| = |\alpha| \sup_{s\in K} \|f(s)\| = |\alpha|\, \|f\|_\infty,$$
$$\|f+g\|_\infty = \sup_{s\in K} \|f(s)+g(s)\| \leq \sup_{s\in K}(\|f(s)\|+\|g(s)\|) \leq \|f\|_\infty + \|g\|_\infty,$$

so that E is a normed vector space.

Take a Cauchy sequence (f_n) in E. For each $\varepsilon > 0$ there is an index $N_\varepsilon \in \mathbb{N}$ with

$$\|f_n(s) - f_m(s)\| \leq \|f_n - f_m\|_\infty \leq \varepsilon$$

for all $n, m \geq N_\varepsilon$ and $s \in K$. By this estimate, $(f_n(s))_n$ is a Cauchy sequence in X. Since X is complete, there exists the limit $f(s) := \lim_{n\to\infty} f_n(s)$ in X for each $s \in K$. Let $s \in K$ and $\varepsilon > 0$ be given. Take the index N_ε from above and $n \geq N_\varepsilon$. We then estimate

$$\|f(s) - f_n(s)\| = \lim_{m\to\infty} \|f_m(s) - f_n(s)\| \leq \limsup_{m\to\infty} \|f_m - f_n\|_\infty \leq \varepsilon.$$

Because N_ε does not depend on s, we can take the supremum over $s \in K$ and derive the inequality $\|f - f_n\|_\infty \leq \varepsilon$ for all $n \geq N_\varepsilon$.

To see that f belongs to E, fix an integer N with $\|f - f_N\|_\infty \leq \varepsilon$. The continuity of f_N yields a radius $\delta > 0$ such that $\|f_N(s) - f_N(t)\| \leq \varepsilon$ for all $s \in \overline{B}_K(t, \delta)$. For such s we deduce

$$\|f(s) - f(t)\| \leq \|f(s) - f_N(s)\| + \|f_N(s) - f_N(t)\| + \|f_N(t) - f(t)\| \leq 3\varepsilon;$$

i.e., $f \in E$. Summing up, (f_n) converges to f in E as required. □

c) Let $X = C([0, 1])$. We set

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

for $f \in X$. The number $\|f - g\|_1$ yields the area between the graphs of $f, g \in X$ (if $\mathbb{F} = \mathbb{R}$), and $\|f_n - f\|_1 \to 0$ is called convergence in the mean. For a mass density $f \geq 0$ of a substance, the integral $\|f\|_1$ is the total mass. We claim that $\|\cdot\|_1$ is an incomplete norm on X.

PROOF. Let $\alpha \in \mathbb{F}$ and $f, g \in X$. If $f \neq 0$, then there are numbers $0 \leq a < b \leq 1$ and $\delta > 0$ such that $|f(s)| \geq \delta$ for all $s \in [a, b]$ since f is continuous. It follows

$$\|f\|_1 \geq \int_a^b |f(s)|\, \mathrm{d}s \geq (b - a)\delta > 0.$$

Using standard properties of the Riemann integral, we also derive

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

$$\|f+g\|_1 = \int_0^1 |f(s)+g(s)|\,\mathrm{d}s \le \int_0^1 (|f(s)|+|g(s)|)\,\mathrm{d}s = \|f\|_1 + \|g\|_1.$$

As a result, $\|\cdot\|_1$ is a norm on X.[1]

To see its incompleteness, we consider the functions given by

$$f_n(s) = \begin{cases} 0, & 0 \le s \le \frac12 - \frac1n, \\ ns - \frac{n}{2} + 1, & \frac12 - \frac1n < s < \frac12, \\ 1, & \frac12 \le s \le 1, \end{cases}$$

for $n \in \mathbb{N}$ with $n \ge 2$. For $m \ge n \ge 2$ we compute

$$\|f_n - f_m\|_1 = \int_{\frac12-\frac1n}^{\frac12} |f_n(s) - f_m(s)|\,\mathrm{d}s \le \frac{1}{n} \longrightarrow 0$$

as $n \to \infty$; i.e., (f_n) is a Cauchy sequence for $\|\cdot\|_1$. There thus exists a limit f of (f_n) in $L^1([0,1])$ and a subsequence $(f_{n_j})_j$ tends to f pointwise a.e. by the Riesz–Fischer Theorem 5.5 in Analysis 3. On the other hand, f_n tends to the characteristic function $\mathbb{1}_{[0,1/2]}$ pointwise as $n \to \infty$, so that $f = \mathbb{1}_{[0,1/2]}$ a.e. and no representative of f is continuous. Since limits in $(X, \|\cdot\|_1)$ are unique, (f_n) does not converge in this space. □

d) Let $X = C(\mathbb{R})$ and $a < b$ in $\mathbb{R}$. One checks as in part b) that $p(f) = \sup_{s\in[a,b]} |f(s)|$ defines a seminorm on X. Moreover, if $p(f) = 0$, then $f = 0$ on $[a,b]$, but of course f does have to be the 0 function. ◊

The vector space $X = C([0,1])$ is infinite dimensional since the functions $p_n \in X$ given by $p_n(t) = t^n$ are linearly independent for $n \in \mathbb{N}$.

Before discussing further examples, we study fundamental 'topological' concepts in a general framework without vector space structure.

Definition 1.5. *A* distance *or* metric d *on a set $M \neq \emptyset$ is a map* $\mathrm{d} : M \times M \to \mathbb{R}$ *satisfying*

a) $\mathrm{d}(x,y) = 0 \iff x = y$ *(definiteness),*
b) $\mathrm{d}(x,y) = \mathrm{d}(y,x)$ *(symmetry),*
c) $\mathrm{d}(x,y) \le \mathrm{d}(x,z) + \mathrm{d}(z,y)$ *(triangle inequality)*

for all $x,y,z \in M$. The pair (M,d) (or just M) is called a metric space*. A sequence (x_n) in M* converges *to a* limit *$x \in M$ if*

$$\forall \varepsilon > 0 \;\; \exists N_\varepsilon \in \mathbb{N} \;\; \forall n \ge N_\varepsilon : \quad \mathrm{d}(x, x_n) \le \varepsilon,$$

in which case we write $x_n \to x$ as $n \to \infty$ or $x = \lim_{n\to\infty} x_n$. It is a Cauchy sequence *if*

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

[1] This part of the proof was omitted in the lectures.

The space (M, d) *(or* d*) is* complete *if each Cauchy sequence converges in* (M, d)*.*

A metric d automatically takes values in $[0, \infty)$. Indeed, for $x, y \in M$ the above properties yield

$$0 = \mathrm{d}(x, x) \leq \mathrm{d}(x, y) + \mathrm{d}(y, x) = 2\,\mathrm{d}(x, y).$$

In a metric space (M, d) we write $B(x, r) = B_M(x, r) = B_{\mathrm{d}}(x, r) = \{y \in X \mid \mathrm{d}(x, y) < r\}$ for the *open ball* with center $x \in M$ and radius $r > 0$, and $\overline{B}(x, r) = \{y \in X \mid \mathrm{d}(x, y) \leq r\}$ for the *closed ball.* We start with simple examples, cf. Beispiel 1.15 in Analysis 2.

Example 1.6. a) Let X be a normed vector space. Set $\mathrm{d}(x, y) = \|x - y\|$ for $x, y \in X$. Then Definition 1.5a) follows from Definition 1.1c), 1.5b) from 1.1a) with $\alpha = -1$, and 1.5c) from 1.1b). Convergence in $(X, \|\cdot\|)$ and in (X, d) are the same.

b) Let $N \subseteq M$ and d be a metric on M. Then $\mathrm{d}_N(x, y) = \mathrm{d}(x, y)$ for $x, y \in N$ defines the *subspace metric* d_N on N. One often writes d instead of d_N. For instance, let M be a normed vector space and d be given as in a). Here d_N is not a norm unless N is a linear subspace.

c) Let $M \neq \emptyset$ be any set. One defines the *discrete metric* on M by setting $\mathrm{d}(x, x) = 0$ and $\mathrm{d}(x, y) = 1$ for all $x, y \in M$ with $x \neq y$. It is easy to check that d is indeed a metric on M and that a sequence (x_n) converges to a point $x \in M$ in the discrete metric if and only if there is an index $m \in \mathbb{N}$ such that $x_n = x$ for all $n \geq m$.

The map $\delta : M^2 \to [0, \infty)$; $\delta(x, y) = 1$, satisfies properties b) and c) in Definition 1.5, but none of the implications in a); cf. Remark 1.3a).

d) Let (M_1, d_1) and (M_2, d_2) be metric spaces. On the product space $M = M_1 \times M_2$, we obtain a metric by setting $\mathrm{d}((x, y), (u, v)) = \mathrm{d}_1(x, u) + \mathrm{d}_2(y, v)$. A sequence (x_n, y_n) in M tends to $(x, y) \in M$ with respect to d if and only if $x_n \to x$ in M_1 and $y_n \to y$ in M_2 as $n \to \infty$.

e) Let M be the unit sphere in $\mathbb{R}^3$. The length of the smaller great circle through $x, y \in M$ defines a metric on M. ◊

We list basic properties of limits in metric spaces shown in Satz 1.16 of Analysis 2.

Remark 1.7. Let M be a metric space and (x_n) be a sequence in M. Then the following assertions hold.

a) If $x_n \to x$ in M as $n \to \infty$, then (x_n) is Cauchy.

b) If $x_n \to x$ and $x_n \to y$ in M for some $x, y \in X$, then $x = y$.

c) If (x_n) converges or is Cauchy, then it is *bounded*; i.e., there exist a point $z \in M$ and a radius $R > 0$ with $x_n \in \overline{B}(z, R)$ for all $n \in \mathbb{N}$. ◊

The next result describes how to construct a distance from a given sequence of seminorms. This procedure is often used in analysis.

PROPOSITION 1.8. *Let X be a vector space and p_j, $j \in \mathbb{N}$, be seminorms on X such that for each $x \in X \setminus \{0\}$ there is an index $k \in \mathbb{N}$ with $p_k(x) > 0$. Then*

$$\mathrm{d}(x,y) = \sum_{j=1}^{\infty} 2^{-j} \frac{p_j(x-y)}{1+p_j(x-y)}, \qquad x,y \in X,$$

defines a metric on X such that $\mathrm{d}(x_n,x) \to 0$ as $n \to \infty$ if and only if $p_j(x_n - x) \to 0$ as $n \to \infty$ for each $j \in \mathbb{N}$. For Cauchy sequences we have an analogous characterization.

PROOF. Note that the function $\varphi(t) = t/(1+t)$ increases strictly for $t \geq 0$, $\varphi(0) = 0$, and $\varphi(t) \in (0,1)$ for $t > 0$. In particular, the series in the statement converges in $[0,\infty)$. Let $x,y,z,x_n \in X$ for $n \in \mathbb{N}$.

1) We have $\mathrm{d}(x,y) = 0$ if and only if $p_j(x-y) = 0$ for all $j \in \mathbb{N}$ which is equivalent to $x = y$ by the assumption. Moreover, the identity $\mathrm{d}(x,y) = \mathrm{d}(y,x)$ follows from $p_j(x-y) = p_j(y-x)$ for each $j \in \mathbb{N}$. Using the monotonicity of φ, we further estimate

$$\begin{aligned}\mathrm{d}(x,z) &\leq \sum_{j=1}^{\infty} 2^{-j} \Big[\frac{p_j(x-y)}{1+p_j(x-y)+p_j(y-z)} + \frac{p_j(y-z)}{1+p_j(x-y)+p_j(y-z)}\Big] \\ &\leq \mathrm{d}(x,y) + \mathrm{d}(y,z).\end{aligned}$$

Thus, d is a metric on X.

2) Assume that $\mathrm{d}(x_n,x) \to 0$ as $n \to \infty$. Fix any $j \in \mathbb{N}$ and let $\varepsilon \in (0,1/2)$. Set $\eta = 2^{-j}\varepsilon$. There is an index $N_{\varepsilon,j} \in \mathbb{N}$ such that

$$\begin{aligned}2^{-j}\frac{p_j(x-x_n)}{1+p_j(x-x_n)} &\leq \mathrm{d}(x,x_n) \leq \eta = 2^{-j}\varepsilon, \\ p_j(x-x_n) &\leq \varepsilon(1+p_j(x-x_n)) \leq \varepsilon + \tfrac{1}{2}\, p_j(x-x_n), \\ p_j(x-x_n) &\leq 2\varepsilon\end{aligned}$$

for all $n \geq N_{\varepsilon,j}$; i.e., $p_j(x-x_n) \to 0$ as $n \to \infty$.

3) Conversely, assume that $p_j(x-x_n) \to 0$ as $n \to \infty$ for each $j \in \mathbb{N}$. Let $\varepsilon > 0$. Fix a number $J_\varepsilon \in \mathbb{N}$ with

$$\sum_{j=J_\varepsilon+1}^{\infty} 2^{-j} \leq \varepsilon.$$

We then find an index $N_\varepsilon \in \mathbb{N}$ such that $p_j(x-x_n) \leq \varepsilon$ for all $j \in \{1,\dots,J_\varepsilon\}$ and $n \geq N_\varepsilon$. It follows that

$$\mathrm{d}(x,x_n) \leq \sum_{j=1}^{J_\varepsilon} 2^{-j} p_j(x-x_n) + \sum_{j=J_\varepsilon+1}^{\infty} 2^{-j} \leq \varepsilon \sum_{j=1}^{J_\varepsilon} 2^{-j} + \varepsilon \leq 2\varepsilon$$

for all $n \geq N_\varepsilon$. The final assertion is similarly shown. □

We next see that the above approach fits to the 'uniform convergence on compact sets' known e.g. from complex analysis, which is used in several variants throughout analysis. In the following example we first recall important facts about the distance to sets which were shown in the exercises of Analysis 2.

EXAMPLE 1.9. a) Let X be Banach space and $A \subseteq X$ be non-empty. For $x \in X$ we define its distance to A by

$$\mathrm{d}(x, A) = \inf\{\|x - y\| \mid y \in A\}.$$

We also put $\mathrm{d}(x, \emptyset) = \infty$. The function $\mathrm{d}_A : X \to [0, \infty)$; $x \mapsto \mathrm{d}(x, A)$, has the following properties.

1) The map d_A is Lipschitz with constant 1. To check this claim, pick $x, y \in X$ and $z \in A$, where we may assume that $\mathrm{d}(x, A) \geq \mathrm{d}(y, A)$. We first note that $\mathrm{d}(x, A) \leq \|x-z\| \leq \|x-y\|+\|y-z\|$. Taking the infimum over $z \in A$, we deduce the inequality $\mathrm{d}(x, A) \leq \|x - y\| + \mathrm{d}(y, A)$ and hence $|\mathrm{d}(x, A) - \mathrm{d}(y, A)| \leq \|x - y\|$.

2) We have $\mathrm{d}(x, A) = 0$ if x belongs to A (choose $y = x$ in the definition). If A is closed, then d_A vanishes only on A. (Indeed, if $\mathrm{d}(x, A) = 0$, then there are vectors y_n in A with $\|x - y_n\| \to 0$ as $n \to \infty$, so that x is contained in A by the closedness.)

3) Let A be closed and $K \subseteq X$ be compact with $A \cap K = \emptyset$. Since d_A is continuous, Theorem 1.45 (see also Analysis 2) then yields a point $x_0 \in K$ with

$$\mathrm{dist}(K, A) := \inf_{x \in K} \mathrm{d}(x, A) = \mathrm{d}(x_0, A) > 0,$$

where we also use part 2).

Let $U \subseteq \mathbb{R}^m$ be open. For $j \in \mathbb{N}$ we set

$$K_j = \{s \in U \mid |s|_2 \leq j,\ \mathrm{d}(s, \partial U) \geq 1/j\}.$$

(If $U = \mathbb{R}^m$, then $K_j = U \cap \overline{B}(0, j)$.) Because of claim 1), these sets are closed and bounded, and hence compact by Bolzano-Weiserstraß. We also have the inclusions $K_j \subseteq K_{j+1} \subseteq U$, and the union of all K_j is U. Let $K \subseteq U$ be compact. Then K is contained in K_j for all integers $j \geq 1$ with $\mathrm{dist}(K, \partial U) \geq 1/j$ and $K \subseteq \overline{B}(0, j)$, which exist by 3).

Let $f \in E := C(U)$. We define $p_j(f) = \max_{s \in K_j} |f(s)|$ for $j \in \mathbb{N}$. As in Example 1.4d), every p_j is a seminorm. Since $\bigcup_j K_j = U$, only $f = 0$ satisfies $p_j(f) = 0$ for all $j \in \mathbb{N}$. Due to Proposition 1.8, the space E thus possesses the metric

$$\mathrm{d}(f, g) = \sum_{j=1}^{\infty} 2^{-j} \frac{\max_{s \in K_j} |f(s) - g(s)|}{1 + \max_{s \in K_j} |f(s) - g(s)|}.$$

Moreover, a sequence (f_n) tends to f in (E, d) if and only if $\max_{s \in K_j} |f_n(s) - f(s)| \to 0$ as $n \to \infty$ for each $j \in \mathbb{N}$. By the above observations, this property is equivalent to the convergence $\max_{s \in K} |f_n(s) - f(s)| \to 0$ as $n \to \infty$ for every compact subset $K \subseteq U$.

One can replace here the sets K_j by the closures L_k of any open subsets $O_k \subseteq U$ such that $L_k \subseteq U$ is compact and $\bigcup_{k\in\mathbb{N}} O_k = U$. (Note that every compact subset $K \subseteq U$ can be covered by finitely many of the sets O_k, see Theorem 1.37 or Analysis 2.)

Finally, (X, d) is complete. In fact, let (f_n) be Cauchy in (M, d). Proposition 1.8 yields that the restrictions $(f_n|_{K_j})_n$ are Cauchy in $C(K_j)$ for each $j \in \mathbb{N}$. These sequences have limits $f^{\langle j\rangle}$ in $C(K_j)$ by Example 1.4. Since $f_n|_{K_j}$ coincides with the restriction of $f_n|_{K_{j+1}}$ to K_j, the function $f^{\langle j\rangle}$ is the restriction of $f^{\langle j+1\rangle}$ to K_j. We can thus define a continuous map $f : U \to \mathbb{F}$ by setting $f(s) = f^{\langle j\rangle}(s)$ for any $j \in \mathbb{N}$ with $s \in K_j$. Then $(f_n|_{K_j})_n$ tends to $f|_{K_j}$ for each $j \in \mathbb{N}$, and hence (f_n) to f in (E, d).

b) Let $Y = C_b(\mathbb{R}) = \{f \in C(\mathbb{R}) \,|\, f \text{ is bounded}\}$ and d_Y be the subspace metric of d in part a). Then (Y, d_Y) is not complete. Indeed, take the functions $f_n \in Y$ given by $f_n(s) = |s|$ if $|s| \le n$ and $f_n(s) = n$ otherwise. Then f_n converges in (X, d) to the function f given by $f(s) = |s|$. Therefore (f_n) is a Cauchy sequence in (Y, d_Y). But it cannot have a limit g in Y, since then g would be equal to $f \notin Y$. $\Diamond$

The next definitions belong to the most basic ones in analysis; below we characterize them in terms of sequences.

Definition 1.10. *Let M be a metric space. A subset $O \subseteq M$ is called* open *if for each $x \in O$ there is a radius $r_x > 0$ with $B(x, r_x) \subseteq O$. Moreover, $\emptyset$ is open by definition. A subset $N \subseteq M$ is a* neighborhood *of a point $x_0 \in M$ if there is a radius $r_0 > 0$ with $B(x_0, r_0) \subseteq N$. A subset $A \subseteq M$ is called* closed *if its complement $M \setminus A$ is open.*

We first illustrate these concepts by simple examples.

Example 1.11. Let (M, d) be a metric space, $x \in M$, and $r > 0$.

a) The ball $B(x, r)$ is open. In fact, for each point $y \in B(x, r)$ we define the radius $\rho = r - \mathrm{d}(x, y) > 0$. Let $z \in B(y, \rho)$. It follows

$$\mathrm{d}(z, x) \le \mathrm{d}(z, y) + \mathrm{d}(y, x) < \rho + \mathrm{d}(x, y) = r;$$

i.e., $z \in B(x, r)$. This means that $B(y, \rho)$ belongs to $B(x, r)$ as claimed.

b) The ball $\overline{B}(x, r)$ is closed.[2] Indeed, for any given $y \in M \setminus \overline{B}(x, r)$ we set $R = \mathrm{d}(x, y) > r$. Let $z \in B(y, R - r)$. We then estimate

$$R = \mathrm{d}(x, y) \le \mathrm{d}(x, z) + \mathrm{d}(z, y) < \mathrm{d}(x, z) + R - r,$$

and thus $\mathrm{d}(x, z) > r$. By this fact, the ball $B(y, R - r)$ is contained in $M \setminus \overline{B}(x, r)$, so that $M \setminus \overline{B}(x, r)$ is open and $\overline{B}(x, r)$ is closed.

c) The sets $\emptyset$ and M are both closed and open. As in part b) with $r = 0$, we see that the set $\{x\}$ is closed for every $x \in M$.

[2] This proof was omitted in the lectures.

d) We equip $\mathbb{Z}$ with the subspace metric $\mathrm{d}(k,l) = |k-l|$ of $\mathbb{R}$. Then the open ball $B(k,1)$ is equal to the singleton $\{k\}$ for each $k \in \mathbb{Z}$, which is also closed by part c). $\Diamond$

PROPOSITION 1.12. *Let (M,d) be a metric space and $A, O \subseteq M$. Then we have the following characterizations.*

a) The subset A is closed in M if and only if for each sequence (x_n) in A with $x_n \to x$ in M as $n \to \infty$, the limit x belongs to A.

b) The subset O is open in M if there does not exist a sequence (x_n) in $M \setminus O$ converging to a point $x \in O$.

PROOF.[3] By definition, $O \subseteq M$ is open if and only if for each $x \in O$ there is a radius $r > 0$ such that each $y \in M$ with $\mathrm{d}(x,y) < r$ already belongs to O. This statement is equivalent to the fact that no $x \in O$ can be the limit of a sequence in $M \setminus O$; i.e., assertion b) holds. Part a) follows from b) by taking complements in M. $\square$

The above characterization of closedness is often employed in these lectures, for instance in the following useful fact.

COROLLARY 1.13. *Let (M,d) be a complete metric space and $A \subseteq M$. The set A is closed if and only if it is complete for the subspace metric d_A. In particular, if X is a Banach space and $Y \subseteq X$ a linear subspace, then Y is closed in X if and only if it is a Banach space for the restriction $\|\cdot\|_Y$ of the norm $\|\cdot\|_X$ to Y.*

PROOF. First, let A be closed. Take a Cauchy sequence (x_n) in A with respect to d_A. Since (M,d) is complete, this sequence has a limit x in M. The point x belongs to A by Proposition 1.12; i.e., (x_n) tends to x in (A,d_A).

Second, let (A,d_A) be complete. Take a sequence (x_n) in A with a limit x in (M,d). This sequence is Cauchy in (M,d) and hence in (A,d_A). By assumption, it then possesses a limit y in A which has to be equal to x, so that x is contained in A. Proposition 1.12 thus yields the closedness of A. The addendum is a direct consequence. $\square$

We can now discuss several typical examples.

EXAMPLE 1.14. Let $X = C([0,1])$ be endowed with $\|\cdot\|_\infty$.

a) In $\mathbb{R}$ the set $S = (0,1]$ is neither closed nor open. Indeed, the points $1 + \frac{1}{n}$ do not belong to S, but tend to 1 as $n \to \infty$, and the numbers $\frac{1}{n} \in S$ converge to 0 as $n \to \infty$. On the other hand, S is a neighborhood of $\frac{1}{2}$ since it contains the open interval $(\frac{1}{4}, \frac{3}{4})$.

b) The subset $\mathbb{Z}$ is closed in $\mathbb{R}$ because it is the union of the closed sets $\{k\}$ for $k \in \mathbb{Z}$, see Example 1.11 and Proposition 1.15.

c) Let $Y = \{f \in X \mid f(0) = 0\}$. Clearly, Y is a linear subspace. Let (f_n) in Y tend to f in X. It follows that $0 = f_n(0) \to f(0)$ as $n \to \infty$,

[3] This proof was omitted in the lectures.

hence $f(0) = 0$ and $f \in Y$. Proposition 1.12 and Corollary 1.13 thus imply that Y is closed and that Y is a Banach space for $\|\cdot\|_\infty$.

d) i) The set $O = \{f \in X \,|\, f > 0\}$ is open in X. In fact, let $f \in O$. We then have $\min_{s\in[0,1]} f(s) = f(s_0) =: \delta > 0$ for some $s_0 \in [0,1]$. Take $g \in X$ with $\|f - g\|_\infty < \delta$. We now estimate

$$g(s) = f(s) + g(s) - f(s) \ge \delta - \|f - g\|_\infty > 0$$

for all $s \in [0,1]$, so that $g \in Y$. This means that the ball $B(f,\delta)$ is contained in O, and hence O is open.

ii) The set $A = \{f \in X \,|\, f \ge 0\}$ is closed in X. Indeed, take a sequence (f_n) in A with a limit f in X. For each $s \in [0,1]$ the numbers $f_n(s) \ge 0$ converge to $f(s)$ as $n \to \infty$, so that $f \ge 0$ and $f \in A$. Proposition 1.12 now yields the claim.

iii) The set $C = \{f \in X \,|\, f(0) > 0 \text{ and } f(1) \ge 0\}$ is neither open nor closed in X. Again this fact follows from Proposition 1.12: The functions $f_n = \frac{1}{n}\mathbb{1}$ belong to C and converge in X to $0 \notin C$ which shows that C is not closed. Moreover, the functions g_n given by $g_n(s) = 1 - (1+\frac{1}{n})s$ do not belong to C, but they have the limit $g(s) = 1 - s$ in X which is an element of C, so that C is not open.

e) i) Let $E = C_0(\mathbb{R}) := \{f \in C(\mathbb{R}) \,|\, f(s) \to 0 \text{ as } s \to \pm\infty\}$. Clearly, E is a linear subspace of $C_b(\mathbb{R}) := \{f \in C(\mathbb{R}) \,|\, f \text{ is bounded}\}$. In Exercise 2.1 it is checked that $C_b(\mathbb{R})$ is a Banach space for the supremum norm, compare also Example 1.4. Let (f_n) be a sequence in E having a limit f in $C_b(\mathbb{R})$. Take some $\varepsilon > 0$. Fix an index N with $\|f - f_N\|_\infty \le \varepsilon$. Since $f_N \in E$, there is a number $s_\varepsilon \ge 0$ such that $|f_N(s)| \le \varepsilon$ for all $s \in \mathbb{R}$ with $|s| \ge s_\varepsilon$. For such s, we then estimate

$$|f(s)| \le |f(s) - f_N(s)| + |f_N(s)| \le 2\varepsilon.$$

This means that f belongs to E, and hence E is a Banach space for $\|\cdot\|_\infty$ by Proposition 1.12 and Corollary 1.13.

ii) We can show that $A = \{f \in E \,|\, f \ge 0\}$ is closed in E as in part d). However, the set $V = \{f \in E \,|\, f > 0\}$ and also A are not open in E. Indeed, we look at the function $f \in V \subseteq A$ given by $f(s) = s^{-2}$ for $|s| \ge 1$ and $f(s) = 1$ for $s \in (-1,1)$. We take maps $\varphi_n \in E$ with $0 \le \varphi_n \le \frac{1}{n}$ and $\varphi_n(n) = \frac{1}{n}$ for $2 \le n \in \mathbb{N}$. Then $f_n := f - \varphi_n$ belongs to E but not to $A \supseteq V$ because $f_n(n) < 0$. Since also $\|f - f_n\|_\infty \le \frac{1}{n}$ for every n, Proposition 1.12 yields the claim.

f) Let $\ell > 0$. The set $L = \{f \in X \,|\, \forall\, t,s \in [0,1] : |f(t) - f(s)| \le \ell\,|t-s|\}$ of functions with Lipschitz constant less or equal ℓ is closed in X, again because of Proposition 1.12: Even if functions $f_n \in L$ converge only pointwise to some f as $n \to \infty$, we conclude

$$|f(t) - f(s)| = \lim_{n\to\infty} |f_n(t) - f_n(s)| \le \ell\,|t-s|$$

for all $t,s \in [0,1]$, so that the limit f belongs to L.

On the other hand, the set $D = \{f \in C^1([0,1]) \mid \|f'\|_\infty \le 1\}$ is not closed in X. To check this fact, we take the maps $f_n \in D$ given by $f_n(s) = ((s-\frac{1}{2})^2 + \frac{1}{n})^{1/2}$ for $n \in \mathbb{N}$. They tend in X to the function $s \mapsto f(s) = |s - \frac{1}{2}|$, which is not differentiable at $\frac{1}{2}$. $\Diamond$

We recall the permanence properties of openness and closedness.

Proposition 1.15. *Let M be a metric space. Then the following assertions hold.*

a) The union of an arbitrary collection of open sets in M is open. The intersection of finitely many open sets in M is open.

b) The intersection of an arbitrary collection of closed sets in M is closed. The union of finitely many closed sets in M is closed.

Proof.[4] Let be $\mathcal{C}$ be a collection of open sets O in M. Take $x \in V := \bigcup_{O\in\mathcal{C}} O$. Then there is a set $O' \in \mathcal{C}$ containing x. Since O' is open, we have a radius $r > 0$ with $B(x,r) \subseteq O' \subseteq V$. Therefore V is open. Let $O_1, \cdots, O_n \subseteq M$ be open and $x \in D := O_1 \cap \cdots \cap O_n$. Again, there are radii $r_j > 0$ such that $B(x, r_j) \subseteq O_j$ for each $j \in \{1, \ldots, n\}$. Setting $\rho := \min\{r_1, \ldots, r_n\} > 0$, we arrive at $B(x,\rho) \subseteq D$, so that D is open. Assertion b) follows from a) by taking complements. $\square$

The finiteness assumptions in the above result are needed, as seen by easy examples: The sets $(0, 1 + \frac{1}{n})$ are open in $\mathbb{R}$ for each $n \in \mathbb{N}$, but their intersection $\bigcap_{n\in\mathbb{N}}(0, 1+\frac{1}{n}) = (0,1]$ is not open in $\mathbb{R}$. The sets $[0, 1-\frac{1}{n}]$ are closed in $\mathbb{R}$ for each $n \in \mathbb{N}$, but their union $\bigcup_{n\in\mathbb{N}}[0, 1-\frac{1}{n}] = [0,1)$ is not closed in $\mathbb{R}$.

We now construct the 'nearest' open or closed set for a given $N \subseteq M$ in a canonical way.

Definition 1.16. *Let M be a metric space and $N \subseteq M$. We define*

a) the interior $N^\circ = \operatorname{int} N$ *of N in M by $N^\circ = \bigcup\{O \subseteq M \mid O$ open in $M, O \subseteq N\}$,*

b) the closure $\overline{N} = \operatorname{cls} N$ *of N in M by $\overline{N} = \bigcap\{A \subseteq M \mid A$ closed in $M,\ A \supseteq N\}$,*

c) the boundary ∂N *of N in M by $\partial N = \overline{N} \setminus N^\circ = \overline{N} \cap (M \setminus N^\circ)$.*

The set N is called dense *in M if $\overline{N} = M$. An element x of N° is an* interior point *of N, and $x \in \overline{N}$ an* adherent point. *Moreover, $x \in M$ is said to be an* accumulation point *of N if there is a sequence in $N \setminus \{x\}$ converging to x. If $x \in N$ is not an accumulation point of N, then it is* isolated *in N.*

The above concepts can be characterized in various ways, in particular using sequences or balls.

Proposition 1.17. *Let M be a metric space M and $N \subseteq M$. Then the following assertions are true.*

[4] This proof was omitted in the lectures.

a) i) N° *in the largest open subset of* N*.*

ii) N *is open if and only if* $N = N^\circ$*.*

iii) $N^\circ = \{x \in M \mid \exists\, r > 0 \text{ with } B(x,r) \subseteq N\} =: N_1$
$= \{x \in M \mid \nexists\, (x_n) \text{ in } M \setminus N \text{ with } x_n \to x,\ n \to \infty\} =: N_2$*.*

b) i) $\overline{N}$ *in the smallest closed subset of* M *containing* N*.*

ii) N *is closed if and only if* $N = \overline{N}$*.*

iii) $\overline{N} = \partial N \,\dot\cup\, N^\circ$*.*

iv) $\overline{N} = \{x \in M \mid \exists\, (x_n) \text{ in } N \text{ with } x_n \to x,\ n \to \infty\} =: N_3$*.*

c) i) ∂N *is closed.*

ii) $\partial N = \{x \in M \mid \exists\, x_n \in N,\ y_n \notin N \text{ s.t. } x_n \to x,\ y_n \to x,\ n \to \infty\}$*.*

d) N *is dense in* M *if and only if for each* $x \in M$ *there are* $x_n \in N$ *with* $x_n \to x$ *as* $n \to \infty$*.*

Proof.[5] a) The inclusion $N^\circ \subseteq N$ follows from the definition of N°, and N° is open by Proposition 1.15. If $N^\circ \subseteq O \subseteq N$ for an open set O, then $O \subseteq N^\circ$ due to Definition 1.16, so that $O = N^\circ$. We have shown claim i) which implies ii). In iii), we deduce $N_1 \subseteq N^\circ$ from the definition of N°, and $N^\circ \subseteq N_2$ is a consequence of Proposition 1.12 and the openness of N°. If $x \notin N_1$, then there is a sequence (x_n) in $M \setminus N$ converging to x; i.e., $x \notin N_2$. Hence, assertion iii) holds.

b) Assertions i) and ii) can be shown as in part a), and Definition 1.16 yields iii). Part iii) in a) implies that $N_3 = M \setminus (M \setminus N)^\circ$, and thus N_3 is closed. Clearly, N is a subset of N_3. From Proposition 1.12 and the closedness of $\overline{N}$, we infer the inclusion $N_3 \subseteq \overline{N}$. Assertion iv) is now a consequence of i).

c) and d) follow from parts a) and b) and the definition of ∂N. □

We note that N_1 and N_3 are the most important descriptions of the interior and the closure, respectively. We add a typical consequence.

Corollary 1.18. *Let* X *be a normed vector space.*

a) If $Y \subseteq X$ *is a linear subspace, then also* $\overline{Y}$ *is a linear subspace.*

b) If $Y \subseteq X$ *is convex, then also* $\overline{Y}$ *is convex.*

Proof. Let Y be convex. Take $x, y \in \overline{Y}$ and $t \in [0,1]$. Proposition 1.17 yields vectors x_n and y_n in Y with $x_n \to x$ and $y_n \to y$ in X as $n \to \infty$. By assumption, the convex combinations $tx_n + (1-t)y_n$ belong to Y, and they converge to $tx + (1-t)y$. Assertion b) then follows from Proposition 1.17. Part a) is shown similarly. □

Let M be a metric space and $f : M \to X$. The *support* of f is

$$\operatorname{supp} f = \operatorname{supp}_M f = \operatorname{cls}_M \{s \in M \mid f(s) \neq 0\}.$$

We now illustrate the above concepts by a series of standard examples.

[5]This proof was omitted in the lectures.

Example 1.19. a) Using the decimal representation of the reals, see Example 3.18 of Analysis 1, one sees that $\mathbb{Q}$ is dense in $\mathbb{R}$ and has empty interior, so that $\partial\mathbb{Q} = \overline{\mathbb{Q}} \setminus \mathbb{Q}^\circ = \mathbb{R}$.

By Example 1.14, the subset $\mathbb{Z}$ is closed in $\mathbb{R}$. Moreover, every intersection $B(k,1) \cap \mathbb{Z}$ is equal to $\{k\}$, so that each $k \in \mathbb{Z}$ is isolated in $\mathbb{Z}$ and $\mathbb{Z}$ has no interior points. If we equip $\mathbb{Z}$ with the subspace metric, then the closed ball $\overline{B}_{\mathbb{Z}}(k,1) = \{k-1, k, k+1\}$ differs from the closure $\overline{B_{\mathbb{Z}}(k,1)} = \{k\}$ of $B_{\mathbb{Z}}(k,1)$ for each $k \in \mathbb{Z}$.

b) In $\mathbb{R}$, the largest of open subset of $(0,1]$ is $(0,1) = (0,1]^\circ$, and $[0,1] = \overline{(0,1]}$ is the smallest closed set containing $(0,1]$. Hence, $\{0,1\}$ is the boundary of $(0,1]$.

c) The set P of polynomials is dense in $X = C([0,1])$ by Weierstraß' approximation Theorem 5.14 from Analysis 3. Since $P \subseteq C^k([0,1])$, also the subspace $C^k([0,1])$ is dense in $C([0,1])$ for every $k \in \mathbb{N}$.

d) Let X be a normed vector space, $x \in X$, and $r > 0$. We have $\overline{B(x,r)} = \overline{B}(x,r)$ and $\partial B(x,r) = \partial \overline{B}(x,r) = \{y \in X \mid \|x-y\| = r\} =: S$, compare part a).

Proof.[6] Proposition 1.17 and Example 1.11 show the relations $B(x,r) \,\dot\cup\, \partial B(x,r) = \overline{B(x,r)} \subseteq \overline{B}(x,r)$. The boundaries $\partial B(x,r)$ and $\partial \overline{B(x,r)}$ are thus contained in $S = \overline{B}(x,r) \setminus B(x,r)$. Take $y \in S$. Then the vectors $y_n = y - \frac{1}{n}(y-x)$ belong to $B(x,r)$ and $z_n = y + \frac{1}{n}(y-x)$ to $X \setminus \overline{B}(x,r)$ for all $n \in \mathbb{N}$, and they converge to y as $n \to \infty$. Consequently, y is an element of $\partial B(x,r)$ and $\partial \overline{B}(x,r)$ due to Proposition 1.17. The assertions easily follow. □

e) Let $X = C([0,1])$ and $N = \{f \in X \mid f > 0 \text{ on } [0, \frac{1}{2}),\ f \geq 0 \text{ on } [\frac{1}{2}, 1]\}$. We then have $N^\circ = \{f \in X \mid f > 0\} =: O$, $\overline{N} = \{f \in X \mid f \geq 0\} =: A$, and $\partial N = \{f \in X \mid f \geq 0,\ \exists\, s \in [0,1] \text{ with } f(s) = 0\} =: R$.

Proof. By Example 1.14, the set O is open and A is closed in X. Proposition 1.17 thus yields the inclusions $O \subseteq N^\circ \subseteq N \subseteq \overline{N} \subseteq A$. Take $f \in N \setminus O$. There is a point $s_0 \in [\frac{1}{2}, 1]$ with $f(s_0) = 0$. Choose functions φ_n in X satisfying $0 \leq \varphi_n \leq 1/n$ and $\varphi_n(s_0) > 0$ for $n \in \mathbb{N}$. The maps $f_n := f - \varphi_n$ then belong to $X \setminus N$ and converge to f in X as $n \to \infty$. Because of Proposition 1.17, the function f is thus not contained in N° and hence $O = N^\circ$. Next, let $g \in A$. The maps $g_n = g + \frac{1}{n}\mathbb{1}$ are elements of N and tend to g in X as $n \to \infty$, so that $g \in \overline{N}$ and $A = \overline{N}$. Definition 1.16 now implies that $\partial N = R$. □

f) Let $X = C([0,1])$ and put $C(0,1) := C((0,1))$. The closure of

$$C_c(0,1) := \{f \in C(0,1) \mid \exists\, 0 < a_f \leq b_f < 1 : \operatorname{supp} f \subseteq [a_f, b_f]\}$$

in X is given by

$$C_0(0,1) = C_0((0,1)) := \{f \in C(0,1) \mid \exists \lim_{t\to 0} f(t) = 0 = \lim_{t \to 1} f(t)\}.$$

[6] This proof was omitted in the lectures.

We consider $C_c(0,1)$ and $C_0(0,1)$ as subspaces of X by extending functions by 0 to $[0,1]$. In particular, $C_0(0,1)$ is a Banach space for the supremum norm by Corollary 1.13.

PROOF. As in Example 1.14 we see that $C_0(0,1)$ is closed in X. Since $C_c(0,1) \subseteq C_0(0,1)$, Proposition 1.17 yields the inclusion $\overline{C_c(0,1)} \subseteq C_0(0,1)$. Let $f \in C_0(0,1)$. Choose functions $\varphi_n \in C_c(0,1)$ such that $0 \le \varphi_n \le 1$ and $\varphi_n(t) = 1$ for $\frac{1}{n} \le t \le 1 - \frac{1}{n}$ and $n \ge 2$. The products $f_n := \varphi_n f$ then belong to $C_c(0,1)$. We estimate

$$\|f - f_n\|_\infty = \sup_{0 \le t \le \frac{1}{n}, 1-\frac{1}{n} \le t \le 1} |1 - \varphi_n(t)|\,|f(t)| \le \sup_{0 \le t \le \frac{1}{n}, 1-\frac{1}{n} \le t \le 1} |f(t)|.$$

The right hand side tends to 0 as $n \to \infty$ because of $f \in C_0(0,1)$. As a result, $C_0(0,1)$ is a subset of $\overline{C_c(0,1)}$ as needed. □

g) The sets L and D from Example 1.14 or $C^1([0,1])$ have no interior points in $X = C([0,1])$. Indeed, take a function f from one of these sets. We use the maps $\varphi_n \in X$ given by $\varphi_n(s) = \sqrt{s}$ for $s \in [0, \frac{1}{n})$ and $\varphi_n(s) = \sqrt{1/n}$ for $s \in [\frac{1}{n}, 1]$, which are not Lipschitz for each $n \in \mathbb{N}$ and tend to 0 in X as $n \to \infty$. The differences $f_n = f - \varphi_n$ then converge to f in X and do not belong to L, D, resp. $C^1([0,1])$. ◊

Though subsets $N \subseteq M$ can be equipped with the restriction of the metric d in M, properties of $S \subseteq N$ may change when passing form d to d_N as indicated in the next facts, see Bemerkung 1.25 in Analysis 2.

REMARK 1.20. Let (M, d) be a metric space and $N \subseteq M$ be endowed with the subspace metric d_N. A set $C \subseteq N$ is called *relatively open* (resp., *relatively closed*) if it is open (resp., closed) in (N, d_N), and analogously for the other concepts introduced above.

a) The open balls in (N, d_N) are given by

$$B_N(x,r) = \{y \in N \,|\, r > \mathrm{d}_N(x,y) = \mathrm{d}(x,y)\} = B(x,r) \cap N.$$

b) A subset $S \subseteq N$ is relatively open (resp., closed) if and only if there is an open (resp., closed) subset S' of M with $S = S' \cap N$.

c) The set N is open (resp., closed) in M if and only if openness (resp., closedness) in N and M coincide.

d) The open unit ball in $N = \mathbb{R}^2_+$ for $|\cdot|_2$ is given by $B_N(0,1) = \{(x,y) \in \mathbb{R}^2 \,|\, x, y \ge 0,\ x^2 + y^2 < 1\}$. ◊

After having discussed convergence and topological concepts, we now study the class of functions which preserve these structures.

DEFINITION 1.21. *Let (M, d) and (M', d') be metric spaces.*

a) Let $D \subseteq M$, $x_0 \in M$ be an accumulation point of D, $y_0 \in M'$, and $f : D \to M'$. We say that $f(x)$ converges *to y_0 as $x \to x_0$ if for* **every** *sequence (x_n) in $D \setminus \{x_0\}$ with limit x_0 in M we have $f(x_n) \to y_0$ in M' as $n \to \infty$. We then write $y_0 = \lim_{x \to x_0} f(x)$ or $f(x) \to y_0$ as $x \to x_0$.*

b) Let $f : M \to M'$ and $x_0 \in M$. The map f is called continuous at x_0 *if for* **every** *sequence (x_n) in M with $x_n \to x_0$ in M, we have $f(x_n) \to f(x_0)$ in M' as $n \to \infty$. If f is continuous at every $x_0 \in M$, it is said to be* continuous (on M). *We write $C(M, M') = \{f : M \to M' \,|\, f$ is continuous$\}$ and put $C(M) := C(M, \mathbb{F})$. If $f \in C(M, M')$ is bijective with $f^{-1} \in C(M', M)$, then f is an* homeomorphism.

In other words, continuity means that $f(x_0) = \lim_{x \to x_0} f(x)$ if x_0 is an accumulation point of M. (If x_0 is isolated in M, then f is automatically continuous at x_0 since only eventually constant sequences can converge to x_0 in M in this case.) We have formulated part a) in the above definition a bit more general to admit functions which are not everywhere defined or possess a limit different from $f(x_0)$. To define continuity for functions $f : D \to M'$ on a subset $D \subseteq M$, one has to replace (M, d) by the metric space (D, d_D), cf. Proposition 1.23 below.

We first discuss a couple of simple examples. In the next chapter we investigate continuous linear maps in great detail.

Example 1.22. a) Every distance $d : M \times M \to \mathbb{R}$ is continuous, where $M \times M$ is endowed with the metric from Example 1.6. In fact, let $(x_n, y_n) \to (x, y)$ in $M \times M$ as $n \to \infty$. Using the inequalities

$$\begin{aligned} \mathrm{d}(x_n, y_n) - \mathrm{d}(x, y) &\le \mathrm{d}(x_n, x) + \mathrm{d}(x, y_n) - \mathrm{d}(x, y) \\ &\le \mathrm{d}(x_n, x) + \mathrm{d}(x, y) + \mathrm{d}(y, y_n) - \mathrm{d}(x, y) \\ &= \mathrm{d}(x, x_n) + \mathrm{d}(y, y_n), \\ \mathrm{d}(x, y) - \mathrm{d}(x_n, y_n) &\le \mathrm{d}(x, x_n) + \mathrm{d}(x_n, y_n) + \mathrm{d}(y_n, y) - \mathrm{d}(x_n, y_n) \\ &= \mathrm{d}(x, x_n) + \mathrm{d}(y, y_n), \end{aligned}$$

we deduce that $|\mathrm{d}(x_n, y_n) - \mathrm{d}(x, y)|$ tends to 0 as $n \to \infty$.[7]

b) Let X be a normed vector space. Then the maps $\mathbb{F} \times X \to X$; $(\alpha, x) \mapsto \alpha x$, and $X \times X \to X$; $(x, y) \mapsto x + y$, are continuous (cf. the proof of Corollary 1.18).

c) Let $X = C([0, 1])$ and $\varphi : \mathbb{F} \to \mathbb{F}$ be Lipschitz on $\overline{B}(0, r) \subseteq \mathbb{F}$ with constant L_r for every $r \ge 0$. Set $(F(u))(s) = \varphi(u(s))$ for all $s \in [0, 1]$ and $u \in X$. Since $F(u)$ belongs to X, we have defined a map $F : X \to X$. We claim that F is Lipschitz on all balls of X and thus continuous.

Proof. Let $u, v \in \overline{B}_X(0, r)$ for some $r > 0$. For each $s \in [0, 1]$ we then have $|u(s)|, |v(s)| \le r$ and hence

$$\begin{aligned} \|F(u) - F(v)\|_\infty &= \sup_{s \in [0,1]} |\varphi(u(s)) - \varphi(v(s))| \le \sup_{s \in [0,1]} L_r \, |u(s) - v(s)| \\ &= L_r \, \|u - v\|_\infty. \end{aligned}$$

[7] This proof was omitted in the lectures.

Let $u_n \to u$ in X. Then $R := \sup_{n\in\mathbb{N}} \|u_n\|_\infty$ is finite and $\|u\|_\infty \le R$. It follows

$$\|F(u) - F(u_n)\|_\infty \le L_R \|u - u_n\|_\infty \longrightarrow 0, \qquad n \to \infty. \qquad \square$$

We recall the permanence properties of continuity from Satz 1.29 in Analysis 2.

PROPOSITION 1.23. *Let (M, d), (M', d') and (M'', d'') be metric spaces and $x_0 \in M$. Then the following assertions are true.*

a) Let $f : M \to M'$ continuous at x_0 and $h : M' \to M''$ at $f(x_0)$. Then the composition $h \circ f : M \to M''$ is continuous at x_0.

b) Let $D \subseteq M$, $x_0 \in D$ and $f : M \to M'$ be continuous at x_0. Then the restriction $f_{|D} : D \to M'$ is continuous at x_0. The converse implication is true if D is open in M.

c) Let Y be a normed vector space, $f, g : M \to Y$ be continuous at x_0, and $\alpha, \beta \in \mathbb{F}$. Then the linear combination $\alpha f + \beta g : M \to Y$ is continuous at x_0. If $Y = \mathbb{F}$, then also the product $fg : M \to \mathbb{F}$ is continuous at x_0.

The openness of D is needed in the second part of b), as seen by the example $M = \mathbb{R}$, $D = [0, 1]$, $x_0 = 0$ and $f = \mathbb{1}_{[0,1]}$. One can nicely characterize continuity in terms of open or closed sets.

PROPOSITION 1.24. *Let (M, d) and (M', d') be metric spaces, $x_0 \in M$, and $f : M \to M'$.*

a) The following assertions are equivalent.

(i) *The map f is continuous at x_0.*
(ii) $\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in M$ *with* $\mathrm{d}(x, x_0) < \delta :\ \mathrm{d}'(f(x), f(x_0)) < \varepsilon$.
(iii) *If V is a neighborhoood of $f(x_0)$ in M', then $f^{-1}(V)$ is a neighborhood of x_0 in M.*

b) The following assertions are equivalent.

(i) *The map f is continuous on M.*
(ii) *If $O \subseteq M'$ is open, then $f^{-1}(O)$ is open in M.*
(iii) *If $A \subseteq M'$ is closed, then $f^{-1}(A)$ is closed in M.*

PROOF.[8] a) Let (iii) be wrong. Then there exists a neighborhood V of $f(x_0)$ in M' and elements x_n of $M \setminus f^{-1}(V)$ converging to x_0 as $n \to \infty$. Hence, $f(x_n)$ is not contained in V for all $n \in \mathbb{N}$ so that $(f(x_n))$ cannot tend to $f(x_0)$. Therefore assertion (i) is false.

Let (iii) be true. Set $V = B(f(x_0), \varepsilon)$ for any given $\varepsilon > 0$. Assumption (iii) yields a radius $\delta > 0$ with $B(x_0, \delta) \subseteq f^{-1}(V)$. For every point x in $B(x_0, \delta)$, the image $f(x)$ thus is an element of $V = B(f(x_0), \varepsilon)$, which is the content of (ii).

Let (ii) be true. Let $x_n \to x$ as $n \to \infty$. Take $\varepsilon > 0$. Due to (ii), we have $\mathrm{d}'(f(x_n), f(x_0)) < \varepsilon$ for all sufficiently large n; i.e., $f(x_n) \to f(x_0)$ as $n \to \infty$, and (i) has been shown.

[8]This proof was omitted in the lectures.

b) Let (i) be valid. Take a closed subset $A \subseteq M'$. Choose a sequence (x_n) in $f^{-1}(A)$ with a limit x in M. By continuity, the images $f(x_n)$ in A tend to $f(x)$. Since A is closed, $f(x)$ is contained in A and hence x in $f^{-1}(A)$. Proposition 1.12 now implies that $f^{-1}(A)$ is closed, and so (iii) is true. Assertion (ii) follows from (iii) by taking complements.

Let (ii) hold. Choose any $x \in M$. Take a neighborhood V of $f(x)$ in M'. We thus have a radius $r > 0$ with $B(x, r) \subseteq V$. By (ii), the preimage $f^{-1}(B(f(x), r))$ is open in M. As it contains x and belongs to $f^{-1}(V)$, the latter is a neighborhood of x in M. The continuity of f at x is now a consequence of part a), and (i) has been proved. □

We recall variants of continuity from Analysis 2.

DEFINITION 1.25. *In the framework of Proposition 1.24, a function $f : M \to M'$ is called* uniformly continuous *if it fulfills statement (ii) in b) for all $x_0 \in M$ with a radius $\delta = \delta_\varepsilon > 0$ not depending on x_0.*

The map f is a called Lipschitz (continuous) *if there is a constant $L \geq 0$ such that $\mathrm{d}'(f(x), f(y)) \leq L\, \mathrm{d}(x, y)$ for all $x, y \in M$.*

Bemerkung 1.34 in Analysis 2 says that Lipschitz continuity implies uniform continuity which in turns yields continuity. Moreover, the converses of these implications are wrong in general.

The above proposition can be used to show openness or closedness.

EXAMPLE 1.26. Let $X = C([0, 1])$, $s \in [0, 1]$, and $\varphi_s : X \to \mathbb{F}$; $\varphi_s(g) = g(s)$. The map φ_s is continuous since $g_n(s) \to g(s)$ in $\mathbb{F}$ if $g_n \to g$ in X as $n \to \infty$. By Proposition 1.24, the preimage $\varphi_s^{-1}(B) = \{g \in X \,|\, g(s) \in B\}$ is open (resp. closed) in X if $B \subseteq \mathbb{F}$ is open (resp. closed). Using also Proposition 1.15, we see that

$$U = \{f \in X \,|\, \exists\, s \in [0, 1] : f(s) \in B_s\} = \bigcup_{s \in [0,1]} \varphi_s^{-1}(B_s)$$

is open if all sets $B_s \subseteq \mathbb{F}$ are open and that

$$A = \{f \in X \,|\, \forall\, s \in [0, 1] : f(s) \in B_s\} = \bigcap_{s \in [0,1]} \varphi_s^{-1}(B_s)$$

is closed if all set $B_s \subseteq \mathbb{F}$ are closed. (Compare Example 1.14.) ◇

Quite often one has several norms on a vector space, so that one needs concepts to compare them.

DEFINITION 1.27. *Let $\|\cdot\|$ and $|||\cdot|||$ be norms on a vector space X. If there is a constant $C > 0$ such that $\|x\| \leq C\, |||x|||$ for all $x \in X$, then $|||\cdot|||$ is called* finer *or* stronger *than $\|\cdot\|$ (and $\|\cdot\|$ is* coarser *or* weaker *than $|||\cdot|||$). In this case one says that the norms are* comparable. *If there are constants $c, C > 0$ such that*

$$c\, |||x||| \leq \|x\| \leq C\, |||x|||$$

for all $x \in X$, then the norms are called equivalent.

The above notions are again characterized by means of sequences and open or closed sets.

Proposition 1.28. *Let $\|\cdot\|$ and $|||\cdot|||$ be norms on a vector space X.*

a) The following assertions are equivalent.

(i) *The norm $\|\cdot\|$ is coarser than $|||\cdot|||$.*
(ii) *There is a constant $C > 0$ such that $\overline{B}_{|||\cdot|||}(x, r/C) \subseteq \overline{B}_{\|\cdot\|}(x, r)$ for all $x \in X$ and $r > 0$.*
(iii) *If (x_n) in X converges for $|||\cdot|||$, then it converges for $\|\cdot\|$.*
(iv) *If a set $A \subseteq X$ is closed for $\|\cdot\|$, then it is closed for $|||\cdot|||$.*
(v) *If a set $O \subseteq X$ is open for $\|\cdot\|$, then it is open for $|||\cdot|||$.*

In this case the limits in (iii) are equal.

b) The norms are equivalent if and only if (iii), (iv) or (v) becomes an equivalence, or if there are constants $C, c > 0$ such that $\overline{B}_{|||\cdot|||}(x, r/C) \subseteq \overline{B}_{\|\cdot\|}(x, r) \subseteq \overline{B}_{|||\cdot|||}(x, r/c)$ for all $x \in X$ and $r > 0$.

c) Let the norms be equivalent. Then $(X, \|\cdot\|)$ is complete if and only if $(X, |||\cdot|||)$ is complete.

Proof. a) Let (i) be true. Let $x \in X$, $r > 0$ and $y \in \overline{B}_{|||\cdot|||}(x, r/C)$, where $C > 0$ is taken from Definition 1.27. We then obtain the bound

$$\|x - y\| \le C\, |||x - y||| \le r,$$

so that y belongs to $\overline{B}_{\|\cdot\|}(x, r)$. Therefore claim (ii) is valid.

The implication '(ii) $\Rightarrow$ (iii)' and the equality of the limits are a consequence of the definition of convergence.

Let (iii) hold. Choose a set $A \subseteq X$ which is closed for $\|\cdot\|$. Take a sequence (x_n) in A with a limit $x \in X$ for $|||\cdot|||$. Because of (iii), the vectors x_n also converge for $\|\cdot\|$ so that x belongs to A by the closedness. As result, A is also closed for $|||\cdot|||$. The implication '(iv) $\Rightarrow$ (v)' follows by taking complements.

Let statement (v) be true. The ball $B_{\|\cdot\|}(0, 1)$ is then open for $|||\cdot|||$, too. We can thus find a radius $r > 0$ with $B_{|||\cdot|||}(0, r) \subseteq B_{\|\cdot\|}(0, 1)$. Take $x \in X \setminus \{0\}$. (The case $x = 0$ is clear.) The vector $r(2\,|||x|||)^{-1}\, x$ belongs to $B_{|||\cdot|||}(0, r)$ and hence to $B_{\|\cdot\|}(0, 1)$; i.e., $1 > r(2\,|||x|||)^{-1}\, \|x\|$ which yields (i) with $C := 2/r$.

The assertions b) und c) are shown similarly. □

We illustrate the above notions with a few typical examples. They show in particular that on function spaces one may have non-comparable (natural) norms and thus different notions of limits, which is a first fundamental difference to the finite dimensional situation.

Example 1.29. a) On a finite dimensional vector space X all norms are equivalent by Satz 1.51 in Analysis 2. On $X = \mathbb{F}^m$ we have the more precise result

$$|x|_q \le |x|_p \le m^{\frac{1}{p} - \frac{1}{q}}\, |x|_q$$

for all $x \in \mathbb{F}^m$ and $1 \le p \le q \le \infty$. (See Satz 1.8 in Analysis 2.)

b) Let $X = C([0,1])$ and $0 < w \in X$. We set $\|f\|_w := \|wf\|_\infty$ on X. As in Example 1.4 one checks that this defines a (*weighted*) norm on X. Observe that $\delta := \inf_{s\in[0,1]} w(s) > 0$. We estimate

$$\delta\,|f(s)| \le w(s)\,|f(s)| \le \|w\|_\infty\,|f(s)|$$

for all $f \in X$ and $s \in [0,1]$. Taking the supremum over $s \in [0,1]$, we deduce the equivalence of $\|\cdot\|_\infty$ and $\|\cdot\|_w$.

c) Let $X = C([0,1])$. Since $\|f\|_1 \le \|f\|_\infty$ for $f \in X$, the supremum norm is finer than the 1-norm on X. On the other hand, Example 1.4 and Proposition 1.28 show that these norms are not equivalent on X because only $\|\cdot\|_\infty$ is complete. This fact can directly be checked using the functions $f_n \in X$ given by $f_n(s) = 1 - ns$ for $0 \le s \le \frac{1}{n}$ and $f_n(s) = 0$ for $\frac{1}{n} \le s \le 1$, since $\|f_n\|_\infty = 1$ and $\|f_n\|_1 = \frac{1}{2n}$ for all $n \in \mathbb{N}$.

d) The 1-norm and the supremum norm on $C_c([0,\infty)) := \{f \in C([0,\infty)) \,|\, \operatorname{supp} f \text{ is compact}\}$ are not comparable. Indeed, by means of functions as f_n in part c), one sees that $\|\cdot\|_1$ cannot be finer than $\|\cdot\|_\infty$. Conversely, take maps $g_n \in C_c([0,\infty))$ with $0 \le g_n \le 1$ and $g_n = 1$ on $[0,n]$ for $n \in \mathbb{N}$. Then $\|g_n\|_\infty = 1$ and $\|g_n\|_1 \ge n$ for all n, so that $\|\cdot\|_\infty$ cannot be finer than $\|\cdot\|_1$. Here one can replace the space $C_c([0,\infty))$ by $\{f \in C_b([0,\infty)) \,|\, \|f\|_1 < \infty\}$. ♢

1.2. More examples of Banach spaces

A) Sequence spaces. Let $s = \{x = (x_j) \,|\, x_j \in \mathbb{F} \text{ for all } j \in \mathbb{N}\}$ be the space of all sequences in $\mathbb{F}$. A distance on s is given by

$$\mathrm{d}(x,y) = \sum_{j=1}^{\infty} 2^{-j} \frac{|x_j - y_j|}{1 + |x_j - y_j|}\,, \qquad x = (x_j) \in s,\; y = (y_j) \in s.$$

For $v_n = (v_{nj})_j$ and $x = (x_j)$ in s, we have $\mathrm{d}(v_n, x) \to 0$ as $n \to \infty$ if and only $v_{nj} \to x_j$ for each $j \in \mathbb{N}$, as $n \to \infty$. These facts follow from Proposition 1.8 with $p_j(x) = |x_j|$. The completeness of (s, d) is proved as in Example 1.9. For $x \in s$, one defines the supremum norm

$$\|x\|_\infty = \sup_{j\in\mathbb{N}} |x_j| \;\in [0,\infty]$$

and introduces by

$$\begin{aligned}
\ell^\infty &= \{x \in s \,|\, \|x\|_\infty < \infty\},\\
c &= \{x \in s \,|\, \exists \lim_{j\to\infty} x_j\} \;\subseteq \ell^\infty,\\
c_0 &= \{x \in s \,|\, \lim_{j\to\infty} x_j = 0\} \;\subseteq c,\\
c_{00} &= \{x \in s \,|\, \exists\, m_x \in \mathbb{N} \text{ such that } x_j = 0 \text{ for all } j > m_x\} \;\subseteq c_0
\end{aligned}$$

the spaces of *bounded, converging, null* and *finite sequences*, respectively. We note that c_{00} is the linear hull of the unit sequences $e_n = (\delta_{kn})_k$ for $n \in \mathbb{N}$, where $\delta_{nn} = 1$ and $\delta_{kn} = 0$ for $k \ne n$.

For $1 \le p < \infty$ and $x \in s$, one further sets

$$\|x\|_p = \Big(\sum_{j=1}^{\infty} |x_j|^p\Big)^{\frac{1}{p}} \in [0,\infty] \qquad \text{and} \qquad \ell^p = \{x \in s \mid \|x\|_p < \infty\},$$

where $\infty^p := \infty$. Observe that

$$|x_k| \le \|x\|_p \qquad \text{for all } \; k \in \mathbb{N},\; 1 \le p \le \infty,\; x = (x_j) \in \ell^p. \tag{1.1}$$

We also put

$$p' = \begin{cases} \infty, & p = 1, \\ \frac{p}{p-1}, & 1 < p < \infty, \\ 1, & p = \infty. \end{cases}$$

For $p, q \in [1, \infty]$, we have the properties

$$\frac{1}{p} + \frac{1}{p'} = 1; \quad p'' = p; \quad p' = 2 \iff p = 2; \quad p \le q \iff q' \le p'. \tag{1.2}$$

We stress that a sequence belongs to ℓ^p if a well-defined quantity is finite, whereas one has to establish convergence to obtain $x \in c$ or $x \in c_0$. In this lecture we use sequence spaces mostly to illustrate results by relatively simple examples. They further serve as state spaces for systems which can be described by countably many numbers.

Proposition 1.30. *Let $p \in [1,\infty]$, $x, y \in \ell^p$ and $z \in \ell^{p'}$. Then the following assertions are true.*

a) $xz \in \ell^1$ and $\|xz\|_1 = \sum_{j=1}^{\infty} |x_j z_j| \le \|x\|_p \|z\|_{p'}$ (Hölder's inequality).

b) $x+y \in \ell^p$ and $\|x+y\|_p \le \|x\|_p + \|y\|_p$ (Minkowski's inequality).

c) ℓ^p is a Banach space; c and c_0 are closed subspaces of ℓ^∞.

Proof. Satz 1.4 of Analysis 2 yields

$$\sum_{j=1}^{J} |x_j z_j| \le \Big(\sum_{j=1}^{J} |x_j|^p\Big)^{\frac{1}{p}} \Big(\sum_{j=1}^{J} |z_j|^{p'}\Big)^{\frac{1}{p'}} \le \|x\|_p \, \|z\|_{p'}$$

for $J \in \mathbb{N}$ if $p \in (1, \infty)$, and similarly for $p \in \{1, \infty\}$. Assertion a) then follows by taking the supremum over $J \in \mathbb{N}$. Part b) is proven analogously.

c) 1) One shows that ℓ^∞ is a Banach space as in Exercise 2.1 or Example 1.4, and the closedness of c_0 as in Example 1.14e). Let $v_n = (v_{nj})_j$ be sequences in c which converge to v in ℓ^∞ as $n \to \infty$. Set $\xi_n = \lim_{j\to\infty} v_{nj}$ in $\mathbb{F}$ for each $n \in \mathbb{N}$. Since $|\xi_n - \xi_m| \le \|v_n - v_m\|_\infty$ for $n, m \in \mathbb{N}$, we have a limit ξ of $(\xi_n)_n$ in $\mathbb{F}$. The sequences $v_n - \xi_n \mathbb{1} \in c_0$ then tend in ℓ^∞ to $v - \xi\mathbb{1}$ as $n \to \infty$, which thus belongs to c_0. Therefore, v has the limit ξ, and so c is closed.

2) Let $p \in [1, \infty)$, $x \in \ell^p$, and $\alpha \in \mathbb{F}$. Clearly, also the sequence αx is an element of ℓ^p and it fulfills $\|\alpha x\|_p = |\alpha|\, \|x\|_p$. If $\|x\|_p = 0$, then $x_j = 0$ for all j by (1.1). Hence, ℓ^p is a normed vector space in view of assertion b).

Let $v_n = (v_{nj})_j$ for $n \in \mathbb{N}$. Assume that $(v_n)_n$ is a Cauchy sequence in ℓ^p. By (1.1) its components yield a Cauchy sequence $(v_{nj})_n$ in $\mathbb{F}$ for each $j \in \mathbb{N}$. Their limits in $\mathbb{F}$ are denoted by x_j for $j \in \mathbb{N}$, and we set $x = (x_j) \in s$. Let $\varepsilon > 0$. Take the index $N_\varepsilon \in \mathbb{N}$ from Definition 1.2 of a Cauchy sequence. For all $J \in \mathbb{N}$ and $n \geq \mathbb{N}_\varepsilon$, it follows

$$\sum_{j=1}^{J} |v_{nj} - x_j|^p = \lim_{m\to\infty} \sum_{j=1}^{J} |v_{nj} - v_{mj}|^p \leq \|v_n - v_m\|_p^p \leq \varepsilon^p.$$

Letting $J \to \infty$, we deduce

$$\sum_{j=1}^{\infty} |v_{nj} - x_j|^p \leq \varepsilon^p$$

for all $n \geq N_\varepsilon$. As a result, $v_n - x$ belongs to ℓ^p and converges to 0 in ℓ^p as $n \to \infty$; i.e., $x = x - v_n + v_n \in \ell^p$ and $v_n \to x$ in ℓ^p. □

Unless something else is said, we endow ℓ^p with the p–norm, and c_0 with $\|\cdot\|_\infty$. The sequence spaces are ordered with increasing p.

Proposition 1.31. *For exponents* $1 < p < q < \infty$ *and sequences* $x \in s$*, we have*

$$c_{00} \subsetneqq \ell^1 \subsetneqq \ell^p \subsetneqq \ell^q \subsetneqq c_0 \subsetneqq \ell^\infty \quad \text{and} \quad \|x\|_\infty \leq \|x\|_q \leq \|x\|_p \leq \|x\|_1,$$
$$\overline{c_{00}}^{\ell^p} = \ell^p \quad \text{for } 1 \leq p < \infty \qquad \text{and} \qquad \overline{c_{00}}^{\|\cdot\|_\infty} = c_0.$$

Proof. It is clear that $c_{00} \subsetneqq \ell^1$ and $\ell^q \subseteq c_0 \subsetneqq \ell^\infty$ for all $q < \infty$. Set $y_k = k^{-\frac{1}{p}}$. Then $(y_k) \notin \ell^p$, but $(y_k) \in \ell^q \cap c_0$ for all $1 \leq p < q < \infty$. The sequence $(z_k) = (1/\ln(k+1))$ belongs to c_0, but not to ℓ^q since

$$\sum_{k=1}^{\infty} \frac{1}{(\ln(k+1))^q} \geq \int_2^\infty \frac{\mathrm{d}s}{(\ln s)^q} = \int_{\ln 2}^\infty t^{-q}\mathrm{e}^t \,\mathrm{d}t \geq c_q \int_{\ln 2}^\infty \mathrm{e}^{t/2}\,\mathrm{d}t = \infty$$

for a constant $c_q > 0$. (Use the transformation $t = \ln s$.)

By Satz 1.8 of Analysis 2 the asserted inequalities are true for the truncated sequences $(x_1, \dots, x_J)$. They then follow by taking the supremum over $J \in \mathbb{N}$. Hence, the first part is shown.

For the final claim, take $x \in \ell^p$ if $p \in [1, \infty)$ and $x \in c_0$ if $p = \infty$. We use the finite sequences $v_n = (x_1, \dots, x_n, 0, \dots)$. The remainder $x - v_n = (0, \dots, 0, x_{n+1}, \dots)$ tends to 0 in $\|\cdot\|_p$ as $n \to \infty$, showing the asserted density. □

B) Spaces of holomorphic functions. Let $U \in \mathbb{C}$ be open. A function $f : U \to \mathbb{C}$ is called holomorphic if the limit

$$\lim_{w\to z} \frac{f(w) - f(z)}{w - z} =: f'(z)$$

exists in $\mathbb{C}$ for each $z \in U$. We denote the space of such functions by $H(U)$. We recall that a function is holomorphic on U if and only if at each $z \in U$ it is given by a power series on a ball $B(z, r_z) \subseteq U$. (See

Satz 1.3 and Theorem 2.12 in Analysis 4.) We use holomorphic maps only in very few peripherical results.

EXAMPLE 1.32. Let $U \subseteq \mathbb{C}$ be open. We equip $H(U)$ with the metric of uniform convergence on compact sets from Example 1.9, identifying U with a subset of $\mathbb{R}^2$ as usual. Let (f_n) be a Cauchy sequence in $H(U)$. By Example 1.9 this sequence has a limit f in $C(U, \mathbb{C})$. Weierstraß' approximation Theorem 2.20 of Analysis 4 now shows that f is holomorphic, and so $H(U)$ is complete.

We equip the set $H^\infty(U) = \{f \in H(U) \,|\, f \text{ bounded}\}$ with the supremum norm. It is a linear subspace of $C_b(U, \mathbb{C})$, which is a Banach space for $\|\cdot\|_\infty$ by Exercise 2.1. Let (g_n) be a sequence in $H^\infty(U)$ with limit g in $C_b(U, \mathbb{C})$. We see as above that g is holomorphic. Hence, $H^\infty(U)$ is closed in $C_b(U, \mathbb{C})$, and thus a Banach space by Corollary 1.13. We note that $H^\infty(\mathbb{C})$ consists only of constant functions by Liouville's Theorem 2.16 of Analysis 4. $\Diamond$

C) L^p spaces. We look for a Banach space of functions with respect to the norm $\|f\|_p = (\int |f|^p \,\mathrm{d}\mu)^{1/p}$ for $p \geq 1$. In the following we recall the relevant definitions and a few facts from Analysis 3.

1) A *σ–algebra* $\mathcal{A}$ on a set $S \neq \emptyset$ is a collection of subsets A of S satisfying the following properties.

a) $\emptyset \in \mathcal{A}$.
b) If $A \in \mathcal{A}$, then its complement $S \setminus A$ is contained in $\mathcal{A}$.
c) If $A_k \in \mathcal{A}$ for all $k \in \mathbb{N}$, then their union $\bigcup_{k\in\mathbb{N}} A_k$ belongs to $\mathcal{A}$.

Observe that the power set $\mathcal{P}(S) = \{A \,|\, A \subseteq S\}$ is a σ–algebra over S.

Let M be a metric space and $\mathcal{O}(M) = \{O \subseteq M \,|\, O \text{ is open}\}$. The smallest σ–algebra on M that contains $\mathcal{O}(M)$ is given by

$$\mathcal{B}(M) := \{A \subseteq M \,|\, A \in \mathcal{A} \text{ for each } \sigma\text{–algebra } \mathcal{A} \supseteq \mathcal{O}(M)\}.$$

It is called the *Borel σ–algebra* on M, and one says that $\mathcal{O}(M)$ *generates* $\mathcal{B}(M)$. We write $\mathcal{B}_m$ instead of $\mathcal{B}(\mathbb{R}^m)$ and endow $\mathbb{C}^m$ with $\mathcal{B}_{2m}$. We stress that $\mathcal{B}_m$ is a strict subset of $\mathcal{P}(\mathbb{R}^m)$. For Borel sets $B \in \mathcal{B}_m$ one has $\mathcal{B}(B) = \{A \in \mathcal{B}_m \,|\, A \subseteq B\} = \{A' \cap B \,|\, A' \in \mathcal{B}_m\}$.

On $M = [0, \infty]$ one considers the σ–algebra generated $\mathcal{B}([0, \infty)) \cup \{\infty\}$. It is the Borel σ–algebra for a metric on $[0, \infty]$ discussed in the exercises of Analysis 2. On $[-\infty, \infty]$ one proceeds similarly.

2) Let $\mathcal{A}$ be a σ–algebra on S. A *(positive) measure* μ on $\mathcal{A}$ is a map $\mu : \mathcal{A} \to [0, \infty]$ such that $\mu(\emptyset) = 0$ and

$$\mu\Big(\bigcup_{k\in\mathbb{N}} A_k\Big) = \sum_{k=1}^{\infty} \mu(A_k) \quad \text{for all pairwise disjoint } \ A_k \in \mathcal{A},\ k \in \mathbb{N}.$$

The tripel $(S, \mathcal{A}, \mu)$ is called a *measure space*. It is *finite* if $\mu(S) < \infty$, and *σ–finite* if there are sets S_k in $\mathcal{A}$ such that $\mu(S_k) < \infty$ for all $k \in \mathbb{N}$ and $\bigcup_{k\in\mathbb{N}} S_k = S$. We present a few examples.

i) Given $s \in S$, we define $\delta_s(A) = 1$ if $s \in A$ and $\delta_s(A) = 1$ if $s \notin A$. Then the *point measure* δ_s is a finite measure on $\mathcal{P}(S)$.

ii) Let $p_k \in [0, \infty)$ for all $k \in \mathbb{N}$ be given. For $A \subseteq \mathbb{N}$, we set

$$\mu(A) = \sum_{k \in A} p_k = \sum_{k=1}^{\infty} p_k \delta_k,$$

which is a σ–finite measure on $\mathcal{P}(\mathbb{N})$. It is finite if (p_k) belongs to ℓ^1. If $p_k = 1$ for all $k \in \mathbb{N}$, then we obtain the *counting measure* $\zeta(A) = \#A$.

iii) On $\mathcal{B}_m$ there is exactly one measure $\lambda = \lambda_m$ such that $\lambda(J)$ is the usual volumen for each interval J in $\mathbb{R}^m$. It is σ–finite and called *Lebesgue measure*. Let $B \in \mathcal{B}_M$. The restriction of λ to $\mathcal{B}(B)$ is a σ–finite measure, also called Lebesgue measure $\lambda_B = \lambda$. Unless otherwise specified, we endow Borel sets B in $\mathbb{R}^m$ with $\mathcal{B}(B)$ and λ.

iv) The collection $\mathcal{L}_m = \{B \cup N \,|\, B \in \mathcal{B}_m,\ N \subseteq N' \text{ for some } N' \in \mathcal{B}_m \text{ with } \lambda(N') = 0\}$ is a σ–algebra on $\mathbb{R}^m$, and $\lambda(B \cup N) := \lambda(B)$ defines a measure on $\mathcal{L}_m$; the *completion* of the Lebesgue measure.

3) Let $\mathcal{A}$ and $\mathcal{B}$ be σ–algebras on sets S and T, respectively. A map $f : S \to T$ is called *measurable* if $f^{-1}(B) \in \mathcal{A}$ for every $B \in \mathcal{B}$. Besides the usual permance properties, the pointwise limit of measurable functions is measurable (if $T \in \{\mathbb{F}, [-\infty, \infty]\}$, say). An $\mathbb{F}^m$–valued function is measurable if and only if all its components are measurable.

Let $A \subseteq S$. The characteristic function $\mathbb{1}_A : S \to \mathbb{F}$ is measurable if and only if A belongs to $\mathcal{A}$. Linear combinations of measurable characteristic functions are called *simple functions*. If a function with values in $\mathbb{F}$ or $[-\infty, \infty]$ is measurable, then also its absolute value, its real and imaginary part, and its positive $f_+ = \max\{0, f\}$ and negative part $f_- = -\max\{0, -f\}$ (if the range is not $\mathbb{C}$) are measurable. For $\mathcal{A} = \mathcal{B}_m$ and $\mathcal{B} = \mathcal{B}_k$ measurability is a rather weak condition.

4) Let $(S, \mathcal{A}, \mu)$ be a measure space. A non-negative simple function f can be written as $f = \sum_{j=1}^{k} y_k \mathbb{1}_{A_j}$ for some $y_j \geq 0$ and disjoint $A_j \in \mathcal{A}$. Its integral is given by

$$\int_S f \, \mathrm{d}\mu = \sum_{j=1}^{k} y_k \mu(A_j).$$

Next, let $f : S \to [0, \infty]$ be measurable. One can approximate f monotonically by simple functions $f_n : S \to [0, \infty)$. This fact allows us to define its *integral* by

$$\int_S f \, \mathrm{d}\mu = \sup_{n \in \mathbb{N}} \int_S f_n \, \mathrm{d}\mu \ \in [0, \infty].$$

The function f is called *integrable* if its integral is finite. We stress that for non-negative measurable functions one obtains a convenient integration theory (one only has to avoid negative factors) without requiring integrability. For instance, the map $f \to \int f \, \mathrm{d}\mu$ is monotone, additive and positive homogeneous. The integral is zero if and only if $\mu(\{f \neq 0\}) = 0$.

A measurable function $f : S \to \mathbb{F}$ is *integrable* if the non-negative measurable maps $(\operatorname{Re} f)_\pm$ and $(\operatorname{Im} f)_\pm$ are integrable, which is equivalent to the integrability of $|f|$. So besides the measurability of f one only has to check that the number $\int_S |f| \,\mathrm{d}\mu$ is finite. In this case one introduces the *integral* of f by setting

$$\int_S f \,\mathrm{d}\mu = \int_S f(s) \,\mathrm{d}\mu(s) := \int_S (\operatorname{Re} f)_+ \,\mathrm{d}\mu - \int_S (\operatorname{Re} f)_- \,\mathrm{d}\mu + \mathrm{i} \int_S (\operatorname{Im} f)_+ \,\mathrm{d}\mu - \mathrm{i} \int_S (\operatorname{Im} f)_- \,\mathrm{d}\mu.$$

One writes $\mathrm{d}x$ instead of $\mathrm{d}\lambda$ or $\mathrm{d}\lambda(x)$. (In this case, the above integral coincides with the Riemann–integral if f is a piecewise continuous function on $[a, b]$.) The integral is linear, monotone (for $\mathbb{F} = \mathbb{R}$), and satisfies the inequality $|\int f \,\mathrm{d}\mu| \le \int |f| \,\mathrm{d}\mu$. It vanishes if $\mu(\{f \neq 0\}) = 0$.

Let $f : S \to \mathbb{F}^m$ be measurable. The map $|f|_2$ is integrable if and only if all components $f_k : S \to \mathbb{F}$ are measurable. One then defines its integral by

$$\int_S f \,\mathrm{d}\mu = \Big(\int_S f_k \,\mathrm{d}\mu \Big)_k.$$

It has similar properties as in the scalar-valued case.

5) Let $p \in [1, \infty)$. For measurable $f : S \to \mathbb{F}$, we define the quantity $\|f\|_p^p = \int_S |f|^p \,\mathrm{d}\mu \in [0, \infty]$ and the set

$$\mathcal{L}^p(\mu) = \mathcal{L}^p(S) = \mathcal{L}^p(S, \mathcal{A}, \mu) = \{f : S \to \mathbb{F} \,|\, \text{measurable}, \ \|f\|_p < \infty\}.$$

Again, to show that f is contained in $\mathcal{L}^p(\mu)$ one only has to check its measurability and that an integral for a non-negative function is finite. We note that $\|\cdot\|_p$ is a seminorm on the vector space $\mathcal{L}^p(\mu)$. We list again a few basic examples.

i) For every function $f : S \to \mathbb{F}$ and $s \in S$, we have $\int_S f \,\mathrm{d}\delta_s = f(s)$.

ii) We have $\mathcal{L}^p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \zeta) = \ell^p$ and $\int_{\mathbb{N}} f \, d\zeta = \sum_{j=1}^{\infty} f(j)$.

iii) Let $B \in \mathcal{B}_m$. We usually write $\mathcal{L}^p(B)$ instead of $\mathcal{L}(B, \mathcal{B}(B), \lambda)$, where $p \in [1, \infty)$. A measurable function $f : B \to \mathbb{F}$ is an element of $\mathcal{L}^p(B)$ if and only if its 0–extension $\widetilde{f}$ belongs to $\mathcal{L}^p(\mathbb{R}^m)$, and we have

$$\int_B f \,\mathrm{d}x := \int_B f \,\mathrm{d}\lambda_B = \int_{\mathbb{R}^m} \widetilde{f} \,\mathrm{d}x.$$

6) It would be nice if $\mathcal{L}^p(\mu)$ was a Banach space, but unfortunately one has $\|\mathbb{1}_N\|_p = 0$ if $\mu(N) = 0$. We are led to the following concept.

Let $(S, \mathcal{A}, \mu)$ be a measure space. A set $N \in \mathcal{A}$ is called a *null set* if $\mu(N) = 0$. A property which holds for all $s \in S \setminus N$ and a null set N is said to hold *almost everywhere (a.e.)* are for *almost all (a.a.)* s.

A countable union of null sets is a null set, and $M \in \mathcal{A}$ is a null set if it is contained in a null set. Uncountable unions of null sets can even have measure ∞; for instance, we have $\mathbb{R} = \bigcup_{x\in\mathbb{R}}\{x\}$. In $(\mathbb{R}^m, \mathcal{B}_m, \lambda)$ hyperplanes, countable subsets, or graphs of measurable functions are null sets. On the other hand, $\emptyset$ is the only null set in $(\mathbb{N}, \mathcal{P}(N), \zeta)$, whereas in $(S, \mathcal{P}(S), \delta_s)$ with $s \in S$ the null sets are those $A \subseteq S$ that do not contain s.

We further introduce the set of *null functions*

$$\mathcal{N} = \{f : S \to \mathbb{F} \,|\, f \text{ is measurable}, \;\; f = 0 \text{ a.e.}\}$$

which is a linear subspace of $\mathcal{L}^p(\mu)$ for every $1 \le p < \infty$. We can thus define the vector space

$$L^p(\mu) = L^p(S) = L^p(S, \mathcal{A}, \mu) := \mathcal{L}^p(\mu)/\mathcal{N} = \{\hat{f} = f + \mathcal{N} \,|\, f \in \mathcal{L}^p(\mu)\}.$$

To avoid that $L^p(\mu) = \{0\}$, we always assume $\mu(S) > 0$. We set

$$\|f + \mathcal{N}\|_p := \|f\|_p, \qquad \text{resp.} \qquad \int_S \hat{f}\, d\mu := \int_S f\, d\mu,$$

for each $\hat{f} \in L^p(\mu)$, resp. $\hat{f} \in L^1(\mu)$, and any representative f of $\hat{f}$. These definitions do not depend on the choice of the representative. The Riesz–Fischer Theorem 5.5 in Analysis 3 says that $(L^p(\mu), \|\cdot\|_p)$ is a Banach space. Unless something else is said, we endow $L^p(\mu)$ with the p–norm.

In $L^p(\mu)$, we write $\hat{f} \ge 0$ if $f \ge 0$ a.e. for some representative f of $\hat{f}$. It then follows that $g \ge 0$ a.e. for all representatives g of $\hat{f}$. Analogous statements hold for the relations '=' or '>.'

One usually identifies $\hat{f}$ with any representative f and $L^p(\mu)$ with $\mathcal{L}^p(\mu)$, provided that one only deals with properties not depending on the representative. For instance, by $f \mapsto f(0)$ one can **not** define a map on $L^p(\mathbb{R}, \mathcal{B}_1, \lambda)$.

As noted in Example 1.4, the 1–norm of a mass density $u \ge 0$ represents the total mass of the corresponding object. Consider a velocity field $v : S \to \mathbb{R}^3$ of a fluid with density $\rho > 0$. Then the kinetic energy of the fluid is given by $\frac{1}{2}\int \rho\, |v|_2^2 \,\mathrm{d}x$. Other p-norms occur in the context of energies for nonlinear material laws. We next illustrate the above spaces as in Example 1.14.

Example 1.33. a) Let $p \in [1, \infty)$, $(S, \mathcal{A}, \mu)$ be a measure space, and $g \in L^p(\mu)$. Then the set $E = \{f \in L^p(\mu) \,|\, f \ge g \text{ a.e.}\}$ is closed in $L^p(\mu)$.

Proof. 1) Let[9] $(\hat{f}_n)$ converge to $\hat{f}$ in $L^p(\mu)$. Choose representatives f_n of $\hat{f}_n$, f of $\hat{f}$, and g of $\hat{g}$. Then there are null sets N_n such that $f_n(s) \ge g(s)$ for all $s \in S \setminus N_n$ and $n \in \mathbb{N}$. Moreover, the Riesz–Fischer Theorem 5.5 of Analysis 3 provides a subsequence and a null set N such that $f_{n_j}(x) \to f(s)$ for all $s \in S \setminus N$ as $j \to \infty$. Therefore $f(s) \ge g(s)$ for all s not contained in the nullset $\bigcup_n N_n \cup N$. This means that $\hat{f} \in E$, and so E is closed. □

[9]As an exception we do not identify $\hat{f}$ and f in this proof.

b) In part a) we take the measure space $(B, \mathcal{B}(B), \lambda)$ for some $B \in \mathcal{B}_m$. Then E possesses no interior points and thus $E = \partial E$ in $L^p(B)$.

PROOF. Let $f \in E$. There is an integer $k \in \mathbb{N}$ such that $B_0 = \{s \in B \,|\, 0 \le f(s) - g(s) \le k\}$ is not a null set (since othewise $f - g = \infty$ a.e.). Let $(C_j^n)_{j\in\mathbb{N}}$ be a countable covering of $\mathbb{R}^m$ with closed cubes of volume $\frac{1}{n}$ which have disjoint interiors. Because $0 < \lambda(B_0) = \sum_{j=1}^{\infty} \lambda(B_0 \cap C_j^n)$, for each $n \in \mathbb{N}$ there is an index j_n such that the measure of the set $A_n = B_0 \cap C_{j_n}^n$ is contained in $(0, \frac{1}{n}]$. Put $f_n = f - (k+1)\mathbb{1}_{A_n}$. The maps f_n then belong to $L^p(B) \setminus E$ and $\|f - f_n\|_p \le (k+1)\lambda(A_n)^{1/p} \le (k+1)n^{-1/p}$ for all $n \in \mathbb{N}$. Hence, f is not an interior point of E. $\Box$.

For a measure space $(S, \mathcal{A}, \mu)$, we further introduce the space

$$\begin{aligned} \mathcal{L}^\infty(\mu) &= \mathcal{L}^\infty(S) = \mathcal{L}^\infty(S, \mathcal{A}, \mu) \\ &:= \{f : S \to \mathbb{F} \,|\, f \text{ is measurable and bounded a.e.}\} \end{aligned}$$

Let $f : S \to \mathbb{F}$ be measurable. Its *essential supremum norm* is given by

$$\|f\|_\infty = \operatorname*{ess\,sup}_{s\in S} |f(s)| := \inf\{c \ge 0 \,|\, |f(s)| \le c \text{ for a.e. } s \in S\} \in [0, \infty].$$

Of course, f belongs to $\mathcal{L}^\infty(\mu)$ if and only if its essential supremum norm is finite. For the Lebesgue measure on $B \in \mathcal{B}_m$ this definition coincides with the usual supremum norm if f is continuous and if $\lambda(B \cap B(x, r)) > 0$ für alle $x \in B$ und $r > 0$.

Indeed, suppose there was a point $x \in B$ with $\delta := |f(x)| - \operatorname{ess\,sup}_S |f| > 0$. Since f is continuous, there exists a radius $r > 0$ such that $|f| \ge \frac{\delta}{2} + \operatorname{ess\,sup}_S |f|$ on the set $B \cap B(x, r)$ which has positive measure by assumption. This is impossible.

For $B = [0, 1]$ the map $f = a\mathbb{1}_{\{0\}}$ has the essential supremum norm 0 for all $a \in [0, \infty]$. If $B = \{0\} \cup [1/2, 1]$, the same is true, but here f is continuous on B. Moreover, the function given by $f(s) = \frac{1}{s}$ for $s > 0$ and $f(0) = 0$ is not essentially bounded on $B = [0, \infty)$. We also set

$$L^\infty(\mu) = \mathcal{L}^\infty(\mu)/\mathcal{N} \qquad \text{and} \qquad \|f + \mathcal{N}\|_\infty = \|f\|_\infty.$$

The importance of this space becomes clear later, see e.g. Theorem 5.4.

PROPOSITION 1.34. *Let $(S, \mathcal{A}, \mu)$ be a measure space. Then $L^\infty(\mu)$ endowed with $\|\cdot\|_\infty$ is a Banach space.*

PROOF. Let $f_k \in \mathcal{L}^\infty(\mu)$ and $\alpha_k \in \mathbb{F}$ for $k \in \{1, 2\}$. Then there are numbers $c_k \ge 0$ and null sets N_k such that $|f_k(s)| \le c_k$ for all $s \in S \setminus N_k$. We thus obtain $|\alpha_1 f_1(s) + \alpha_2 f_2(s)| \le |\alpha_1|\, c_1 + |\alpha_2|\, c_2$ for every $s \notin N_1 \cup N_2 =: N$, where N is a null set. The linear combination $\alpha_1 f_1 + \alpha_2 f_2$ then belongs to $\mathcal{L}^\infty(\mu)$ and $\mathcal{L}^\infty(\mu)$ is a vector space. Taking the infimum over such c_k, we also see that $\|f_1 + f_2\|_\infty \le \|f_1\|_\infty + \|f_2\|_\infty$. It is clear that $\|\alpha_1 f_1\|_\infty = |\alpha_1|\, \|f_1\|_\infty$. Moreover, $\mathcal{N}$ is a linear subspace of $\mathcal{L}^\infty(\mu)$ and we have $\|f_1\|_\infty = \|f_2\|_\infty$ if $f_1 = f_2$ a.e.. Therefore $L^\infty(\mu)$ is a vector space, $\|\hat{f}\|_\infty$ is well defined, and it is a norm on $L^\infty(\mu)$.

Let $(\hat{f}_n)$ be Cauchy in $L^\infty(\mu)$. Fix representatives f_n of $\hat{f}_n$. For every $j \in \mathbb{N}$ there is an index $k(j) \in \mathbb{N}$ such that $\|f_n - f_m\|_\infty \le \frac{1}{j}$ for all $n, m \ge k(j)$. Hence, the set

$$N_{n,m,j} = \{s \in S \,|\, |f_n(s) - f_m(s)| > \tfrac{2}{j}\}$$

has measure 0 for these integers, and so $N := \bigcup_{n,m\ge k(j), j\in\mathbb{N}} N_{n,m,j}$ is also a the null set. For $s \in S \setminus N$, we then obtain

$$|f_n(s) - f_m(s)| \le 2/j \qquad \text{for all} \ \ n, m \ge k(j) \ \ \text{and} \ \ j \in \mathbb{N}.$$

There thus exists $f(s) = \lim_{n\to\infty} f_n(s)$ in $\mathbb{F}$ for all $s \in S \setminus N$. We set $f(s) = 0$ for all $s \in N$. Then the map $f : S \to F$ is measurable. Let $\varepsilon > 0$ and take $j \ge 1/\varepsilon$. It follows that

$$\|\hat{f}_n - \hat{f}\|_\infty \le \sup_{s\in S\setminus N} |f_n(s) - f(s)| = \sup_{s\in S\setminus N} \lim_{m\to\infty} |f_n(s) - f_m(s)| \le \frac{2}{j} \le 2\varepsilon$$

if $n \ge k(j)$. As a result, the equivalence class $\hat{f} = \hat{f}_n + \hat{f} - \hat{f}_n$ belongs to $L^\infty(\mu)$ and $\hat{f}_n \to \hat{f}$ in $L^\infty(\mu)$ as $n \to \infty$. □

We finally recall *Hölder's inequality* and an immediate consequence, see Satz 5.1 and Korollar 5.2 of Analysis 3.

Proposition 1.35. *Let $(S, \mathcal{A}, \mu)$ be a measure space, $p \in [1, \infty]$, $f \in L^p(\mu)$, and $g \in L^{p'}(\mu)$. Then the following assertions hold.*

a) We have $fg \in L^1(\mu)$ and $\|fg\|_1 \le \|f\|_p \|g\|_{p'}$.

b) If $\mu(S) < \infty$ and $1 \le p < q \le \infty$, then $L^q(\mu) \subseteq L^p(\mu)$ and

$$\|f\|_p = \Big(\int_S \mathbb{1}\, |f|^p \, d\mu\Big)^{\frac{1}{p}} \le \|\mathbb{1}\|_{r'}^{\frac{1}{p}} \, \| \, |f|^p\|_r^{\frac{1}{p}} = \mu(S)^{\frac{1}{p}-\frac{1}{q}} \|f\|_q,$$

where $r = q/p > 1$ and $r' = q/(q-p)$ by (1.2).

The inclusion in Proposition 1.35b) is strict in general, as shown by the function $(0, 1) \ni t \mapsto f(t) = t^{-1/q}$. Similarly, one checks that there is no inclusion between $L^q(\mathbb{R}^m)$ and $L^p(\mathbb{R}^m)$ if $p \ne q$.

1.3. Compactness and separability

Compactness is one of the most important concepts in analysis, as one could already see in the Analysis lectures. We first define this notion and some variants in a metric space.

Definition 1.36. *Let M be a metric space and $K \subseteq M$.*

a) K is compact *if every* open covering *$\mathcal{C}$ of K (i.e., a collection $\mathcal{C}$ of open sets $O \subseteq M$ with $K \subseteq \bigcup\{O \,|\, O \in \mathcal{C}\}$) contains a finite subcovering $\{O_1, \dots, O_m\} \subseteq \mathcal{C}$ with $K \subseteq O_1 \cup \dots \cup O_m$.*

b) K is sequentially compact *if every sequence (x_n) in K possesses a subsequence converging to some $x \in K$.*

c) K is relatively compact *if $\overline{K}$ is compact.*

d) K is totally bounded *if for each $\varepsilon > 0$ there are a number $m \in \mathbb{N}$ and points $x_1, \dots, x_m \in M$ such that $K \subseteq B(x_1, \varepsilon) \cup \dots \cup B(x_m, \varepsilon)$.*

We stress that sequential compactness gives some convergence for free, and compactness provides finiteness (which often transfers into uniformity). Clearly, compactness is just relatively compactness plus closedness, and it implies total boundedness using the covering $\mathcal{C} = \{B(x,\varepsilon) \,|\, x \in K\}$.

In the definition of total boundedness, one can equivalently require that the centers x_j belong to K. To show this, take $\varepsilon > 0$ and the points $x_1, \dots, x_m \in M$ from Definition 1.36 with ε replaced by $\varepsilon/2$. If $B(x_j, \varepsilon/2) \cap K$ is empty, we drop this element x_j. Otherwise we replace it by some y_j in $B(x_j, \varepsilon/2) \cap K$. Then $B(x_j, \varepsilon/2)$ is contained in $B(y_j, \varepsilon)$ so that also the balls $B(y_j, \varepsilon)$ cover K.

If K is totally bounded, then it is bounded as a subset of the ball $B(x_1, r)$ with radius $r = \varepsilon + \max_{j\in\{1,\dots,m\}} \mathrm{d}(x_1, x_j)$, where x_j are the points from part d) of the above definition.

As in Satz 2.18 of Analysis 1 one sees that a sequence in a metric space M has a subsequence converging to some x in M if and only if x is an accumulation point of the sequence.

We first show that in metric spaces the seemingly unrelated concepts of sequential compactness and compactness are equivalent. This astonishing fact goes back to *Heine*, *Borel*, and others.

THEOREM 1.37. *Let (M, d) be metric space and $K \subseteq M$. Then K is compact if and only if it sequentially compact.*

PROOF. 1) Let K be compact. Suppose K was not sequentially compact. There thus exists a sequence (x_n) in K without an accumulation point in K. In other words, for each $y \in K$ there is a radius $r_y > 0$ such that the open ball $B(y, r_y)$ contains only finitely many of the members x_n. Since $K \subseteq \bigcup_{y\in K} B(y, r_y)$ and K is compact, there are centers $y_1, \dots, y_m \in K$ such that $K \subseteq B(y_1, r_{y_1}) \cup \dots \cup B(y_m, r_{y_m})$. This inclusion is impossible because (x_n) has infinitely many members, and hence K must be sequentially compact.

2) Let K be sequentially compact and let $\mathcal{C}$ be an open covering of K. We suppose that no finite subset of $\mathcal{C}$ covers K. Due to Lemma 1.38 below, for each $n \in N$ there are finitely many open balls of radius $1/n$ covering K. For every $n \in \mathbb{N}$, we can find a ball $B_n = B(x_n, 1/n)$ such that $B_n \cap K$ is not covered by finitely many sets from $\mathcal{C}$ (since we would otherwise obtain a finite subcovering). Because K is sequentially compact, there is an accumulation point $\widehat{x} \in K$ of (x_n). There further exists an open set $\widehat{O} \in \mathcal{C}$ with $\widehat{x} \in \widehat{O}$, and hence $B(\widehat{x}, \varepsilon) \subseteq \widehat{O}$ for some $\varepsilon > 0$. We then obtain an index $N \in \mathbb{N}$ such that $\mathrm{d}(x_N, \widehat{x}) < \varepsilon/2$ and $N \geq 2/\varepsilon$. Each $x \in B_N$ thus satisfies the inequality

$$\mathrm{d}(x, \widehat{x}) \leq \mathrm{d}(x, x_N) + \mathrm{d}(x_N, \widehat{x}) < \tfrac{1}{N} + \tfrac{\varepsilon}{2} \leq \varepsilon;$$

i.e., $B_N \subseteq B(\widehat{x}, \varepsilon) \subseteq \widehat{O}$. By this contradiction, K is compact. □

We have isolated the following lemma from the above proof.

Lemma 1.38. *Let (M, d) be a metric space, $K \subseteq M$, and each sequence in K have an accumulation point. Then K is totally bounded.*

Proof. Suppose that K was not totally bounded. There would thus exist a number $r > 0$ such that K cannot be covered by finitely many balls of radius r. As a result, there exists a center $x_1 \in K$ such that $K \not\subseteq B(x_1, r)$. Take any point $x_2 \in K \setminus B(x_1, r)$. We then have $\mathrm{d}(x_2, x_1) \geq r$. Since K cannot be contained in $B(x_1, r) \cup B(x_2, r)$, we find an element x_3 of $K \setminus (B(x_1, r) \cup B(x_2, r))$ implying that $\mathrm{d}(x_3, x_1) \geq r$ and $\mathrm{d}(x_2, x_1) \geq r$. Inductively, we obtain a sequence (x_n) in K with $\mathrm{d}(x_n, x_m) \geq r > 0$ for all $n > m$. By assumption, this sequence has an accumulation point x. This means that there are infinitely members x_{n_j} in $B(x, r/2)$ and thus $\mathrm{d}(x_{n_j}, x_{n_i}) < r$ for all $i \neq j$. This is impossible, so that K is totally bounded. □

We now show that total boundedness yields relative compactness if the metric is complete. This fact is often used to check compactness, see e.g. Proposition 1.46.

Corollary 1.39. *Let (M, d) be a metric space and $N \subseteq M$. Then the following assertions are equivalent.*[10]

a) *N is relatively compact.*
b) *Each sequence in N has an accumulation point (belonging to $\overline{N}$).*
c) *N is totally bounded and $\overline{N}$ is complete.*

Proof. The implication "a)⇒b)" follows from Theorem 1.37 applied to $\overline{N}$.

Let statement b) be true. Lemma 1.38 yields that N is totally bounded. Take a Cauchy sequence (x_n) in $\overline{N}$. There are points $\widetilde{x}_n \in N$ such that $\mathrm{d}(x_n, \widetilde{x}_n) \leq 1/n$ for every $n \in \mathbb{N}$. By b), there exists a subsequence $(\widetilde{x}_{n_j})_j$ with a limit x in $\overline{N}$. Let $\varepsilon > 0$. We can then find an index $j = j(\varepsilon)$ such that $n_j \geq 1/\varepsilon$, $\mathrm{d}(\widetilde{x}_{n_j}, x) \leq \varepsilon$ and $\mathrm{d}(x_n, x_{n_j}) \leq \varepsilon$ for all $n \geq n_j$. For such n, we estimate

$$\mathrm{d}(x_n, x) \leq \mathrm{d}(x_n, x_{n_j}) + \mathrm{d}(x_{n_j}, \widetilde{x}_{n_j}) + \mathrm{d}(\widetilde{x}_{n_j}, x) \leq 3\varepsilon,$$

so that assertion c) has been shown.

Let N be totally bounded and $\overline{N}$ be complete. Take a sequence (x_n) in $\overline{N}$. We choose points $\widetilde{x}_n \in N$ as in the previous paragraph. By assumption, N is covered by finitely many balls B_j^1 in M of radius 1. We can then find an index j_1 and a subsequence $(\widetilde{x}_{\nu_1(k)})_k$ in $B_{j_1}^1$. There further exist finitely many balls B_j^2 in M of radius $1/2$ which cover N. Again, there is an index j_2 and a subsequence $(\widetilde{x}_{\nu_2(k)})_k$ of $(\widetilde{x}_{\nu_1(k)})_k$ which belongs to $B_{j_2}^2$. By induction, for every $m \in \mathbb{N}$ we obtain a subsequence $(\widetilde{x}_{\nu_m(k)})_k$ of $(\widetilde{x}_{\nu_{m-1}(k)})_k$ that is contained in a ball $B_{j_m}^m$ of radius $1/m$. To define a diagonal sequence, we set $n_m = \nu_m(m)$ for

[10] In the lectures we showed a slightly weaker result.

$m \in \mathbb{N}$. Note that the points $\widetilde{x}_{n_m}$ and $\widetilde{x}_{n_p}$ are elements of $B^m_{j_m}$ for $p \geq m$ since n_p and n_m are contained in ν_m. Take $\varepsilon > 0$. Choose a number $m \in \mathbb{N}$ with $m, n_m \geq 1/\varepsilon$. It follows that

$$\mathrm{d}(x_{n_m}, x_{n_p}) \leq \mathrm{d}(x_{n_m}, \widetilde{x}_{n_m}) + \mathrm{d}(\widetilde{x}_{n_m}, \widetilde{x}_{n_p}) + \mathrm{d}(\widetilde{x}_{n_p}, x_{n_p}) \leq \tfrac{1}{n_m} + \tfrac{2}{m} + \tfrac{1}{n_p} \leq 4\varepsilon$$

for all $p \geq m$. Because $\overline{N}$ is complete, the Cauchy sequence $(x_{n_m})_m$ has a limit x in $\overline{N}$. Statement c) is thus true by Theorem 1.37. □

We state two important necessary conditions for compactness.

Corollary 1.40. *Let K be a compact subset of a metric space M. Then K is closed and bounded.*

Proof. The boundedness follows from Corollary 1.39 and the remarks after Definition 1.36. Let points $x_n \in K$ tend to some x in M as $n \to \infty$. Since K is sequentially compact, there is a subsequence $(x_{n_j})_j$ with a limit y in K. Hence, $x = y$ belongs to K and K is closed. □

We now discuss whether there is a converse of the above result.

Example 1.41. a) A subset K of a finite dimensional normed vector space X is compact if and only if K is closed and bounded, because of Theorem 1.43 in Analysis 2 and Theorem 1.37. In particular, closed balls are compact in finite dimensions.

b) Let $X = \ell^p$ and $1 \leq p \leq \infty$. We then have $\|e_n\|_p = 1$ and $\|e_n - e_m\|_p = 2^{\frac{1}{p}}$ if $n \neq m$, so that (e_n) has no converging subsequence. As a result, the closed (and bounded) unit ball in ℓ^p is not compact.

c) Let $X = C([0,1])$. For $n \in \mathbb{N}$ we define the functions $f_n \in X$ by

$$f_n(t) = \begin{cases} 2^{n+1}t - 2, & 2^{-n} \leq t \leq \frac{3}{2} \cdot 2^{-n}, \\ 4 - 2^{n+1}t, & \frac{3}{2} \cdot 2^{-n} < t \leq 2^{-n+1}, \\ 0, & \text{otherwise}. \end{cases}$$

Then $\|f_n\|_\infty = 1$ and $\|f_n - f_m\|_\infty = 1$ for $n \neq m$, so that again the closed unit ball is not compact. ◊

The next theorem shows that these simple examples are typical. This fact is arguably the most important difference between finite and infinite dimensional normed vector spaces, but see Theorem 1.50.

Theorem 1.42. *Let X be a normed vector space. The closed unit ball $B = \overline{B}(0,1)$ in X is compact if and only if $\dim X < \infty$.*

Proof. If $\dim X < \infty$, then B is compact by Example 1.41a). Let $\dim X = \infty$. Take any $x_1 \in X$ with $\|x_1\| = 1$. Set $U_1 = \operatorname{lin}\{x_1\}$. Since U_1 is finite dimensional, U_1 is closed in X due to Lemma 1.43 below and $X \neq U_1$ because of $\dim X = \infty$. Lemma 1.44 thus yields a vector x_2 in X with $\|x_2\| = 1$ and $\|x_2 - x_1\| \geq \frac{1}{2}$. Then $U_2 = \operatorname{lin}\{x_1, x_2\} \neq X$ is also closed, and from Lemma 1.44 we obtain an element x_3 of X with

$\|x_3\| = 1$, $\|x_3 - x_2\| \geq \frac{1}{2}$, and $\|x_3 - x_1\| \geq \frac{1}{2}$. Inductively, we can now construct a sequence (x_n) in B such that $\|x_n - x_m\| \geq \frac{1}{2}$ for all $n \neq m$; i.e., B is not sequentially compact. □

We still have to prove two lemmas, the seond one is due to *F. Riesz.*

LEMMA 1.43. *Let Y be a finite dimensional subspace of a normed vector space X. Then Y is closed in X.*

PROOF. Choose a basis $B = \{b_1, \dots, b_m\}$ of Y. Let $\tilde{y} = (y_1, \dots, y_m) \in \mathbb{F}^m$ be the (uniquely determined) vector of coefficients with respect to B for a given $y \in Y$. Set $|||\tilde{y}||| := \|y\|$. This gives a norm on $\mathbb{F}^m$, which is complete by Satz 1.51 of Analysis 2. Let $v_n \in Y$ converge to some x in X. Then $(\tilde{v}_n)$ is a Cauchy sequence in $\mathbb{F}^m$, and so there exists a limit $\tilde{z} = (z_1, \dots, z_m) \in \mathbb{F}^m$ of $(\tilde{v}_n)$ for $|||\cdot|||$. Set $z = z_1 b_1 + \dots + z_m b_m \in Y$. It follows that $\|v_n - z\| = |||\tilde{v}_n - \tilde{z}||| \to 0$ as $n \to \infty$, and thus $x = z$ belongs to Y; i.e., Y is closed. □

LEMMA 1.44. *Let X be a normed vector space, $Y \neq X$ be a closed linear subspace, and $\delta \in (0,1)$. Then there exists a vector $\bar{x} \in X$ with $\|\bar{x}\| = 1$ and $\|\bar{x} - y\| \geq 1 - \delta$ for all $y \in Y$.*

PROOF. Take any $x \in X \setminus Y$. Since $X \setminus Y$ is open, we have $d := \inf_{y \in Y} \|x - y\| > 0$. Hence, $d < d/(1-\delta)$ and there is a vector $\bar{y} \in Y$ with $\|x - \bar{y}\| \leq d/(1-\delta)$. Set $\bar{x} = \frac{1}{\|x - \bar{y}\|}(x - \bar{y})$. Let $y \in Y$. We then obtain $\|\bar{x}\| = 1$ and, using the above inequalities,

$$\|\bar{x} - y\| = \frac{1}{\|x - \bar{y}\|} \|x - (\bar{y} + \|x - \bar{y}\| y)\| \geq \frac{d}{\|x - \bar{y}\|} \geq 1 - \delta. \qquad \square$$

We recall a few important consequences of compactness.

THEOREM 1.45. *Let X be a normed vector space, K be a compact metric space, and $f \in C(K, X)$. Then f is uniformly continuous and bounded. If $X = \mathbb{R}$, then there are points $t_\pm$ in K such that $f(t_+) = \max_{t \in K} f(t)$ and $f(t_-) = \min_{t \in K} f(t)$.*

PROOF.[11] 1) Let $\varepsilon > 0$. Because f is continuous, for every $t \in K$ there is a radius $\delta_t > 0$ with $\|f(t) - f(s)\| < \varepsilon$ for all $s \in B(t, \delta_t)$. Since $K = \bigcup_{t \in K} B(t, \frac{1}{3}\delta_t)$ and K is compact, there are points $t_1, \dots, t_m \in K$ such that $K \subseteq B(t_1, \frac{1}{3}\delta_1) \cup \dots \cup B(t_m, \frac{1}{3}\delta_m)$, where $\delta_k := \delta_{t_k}$. Set $\delta = \min\{\frac{1}{3}\delta_1, \dots, \frac{1}{3}\delta_m\} > 0$. Take $s, t \in K$ with $\mathrm{d}(s,t) < \delta$. Then there are indices $k, l \in \{1, \dots, m\}$ such that $s \in B(t_k, \frac{1}{3}\delta_k)$ and $t \in B(t_l, \frac{1}{3}\delta_l)$, where we may assume that $\delta_k \geq \delta_l$. We thus obtain

$$\mathrm{d}(t_k, t_l) \leq \mathrm{d}(t_k, s) + \mathrm{d}(s,t) + \mathrm{d}(t, t_l) < \tfrac{1}{3}\delta_k + \delta + \tfrac{1}{3}\delta_l \leq \delta_k$$

so that $t_l \in B(t_k, \delta_k)$, and hence

$$\|f(s) - f(t)\| \leq \|f(s) - f(t_k)\| + \|f(t_k) - f(t_l)\| + \|f(t_l) - f(t)\| < 3\varepsilon.$$

[11] This proof was omitted in the lectures.

2) Suppose there were points t_n in K with $\|f(t_n)\| \to \infty$ as $n \to \infty$. Since K is compact, there is a subsequence $(t_{n_j})_j$ converging to some $t \in K$. The norms $\|f(t_{n_j})\|$ then tend to $\|f(t)\|$ as $j \to \infty$, due to the continuity of $t \mapsto \|f(t)\|$, which contradicts the limit $\|f(t_n)\| \to \infty$.

3) Let $X = \mathbb{R}$. By part 2), the supremum $z := \sup_{t\in K} f(t)$ belongs to $\mathbb{R}$. We can find elements r_n of K with $f(r_n) \to z$ as $n \to \infty$. The compactness of K again gives a subsequence $(r_{n_j})_j$ converging to some t_+ in K. Hence, $f(t_+) = \lim_{j\to\infty} f(r_{n_j}) = z$ since f is continuous. The minimum is treated in the same way. □

Theorem 1.42 shows that in an infinite dimensional Banach space a closed and bounded subset does not need to be compact. The next two results give stronger sufficient conditions for (relative) compactness of a set in ℓ^p or $C(K)$. Besides boundedness, one requires summability, resp. continuity, *uniformly* for elements in the set. (It can be seen that these conditions are in fact necessary.)

Proposition 1.46. *Let $p \in [1, \infty)$. A set $K \subseteq \ell^p$ is relatively compact if it is bounded and fulfills*

$$\lim_{N\to\infty} \sup_{(x_j)\in K} \sum_{j=N+1}^{\infty} |x_j|^p = 0.$$

Proof. Due to Corollary 1.39, it suffices to prove that K is totally bounded. Let $\varepsilon > 0$. By assumption, there is an index $N \in \mathbb{N}$ with

$$\sum_{j=N+1}^{\infty} |x_j|^p < \varepsilon^p \qquad \text{for all } \ (x_j) \in K.$$

For $x = (x_j) \in K$, we put $\hat{x} = (x_1, \dots, x_N) \in \mathbb{F}^N$. Since $|\hat{x}|_p \le \|x\|_p$, the set $\hat{K} = \{\hat{x} \,|\, x \in K\}$ is bounded in $\mathbb{F}^N$, and thus it is totally bounded by Example 1.41 and Corollary 1.39. So we obtain vectors $\hat{v}_1, \dots, \hat{v}_m \in \mathbb{F}^N$ such that for all $x \in K$ there is an index $l \in \{1, \dots, m\}$ with $|\hat{x} - \hat{v}_l|_p < \varepsilon$. Set $v_k = (\hat{v}_k, 0, \dots) \in \ell^p$ for all $k \in \{1, \dots, m\}$. The total boundedness of K now follows from

$$\|x - v_l\|_p^p = |\hat{x} - \hat{v}_l|_p^p + \sum_{j=N+1}^{\infty} |x_j|^p < 2\varepsilon^p. \qquad \square$$

We now establish one of the most important compactness results in analysis going back to *Arzela* and *Ascoli*.

Theorem 1.47. *Let K be a compact metric space and $F \subseteq C(K)$ be bounded and* equicontinuous*; i.e.,*

$$\forall\, \varepsilon > 0\ \exists\, \delta_\varepsilon > 0\ \forall\, f \in F\ \forall\, s,t \in K \text{ s.t. } \mathrm{d}(s,t) \le \delta_\varepsilon :\ |f(s) - f(t)| \le \varepsilon. \tag{1.3}$$

Then F is relatively compact. If F is also closed, it is compact.

Proof. 1) For each $\varepsilon_N := 1/N$ and $N \in \mathbb{N}$, we take the radius $\delta_N := \delta_{\varepsilon_N}$ from (1.3). Since K is compact, Corollary 1.39 gives points $t_{1,N}, \dots, t_{m_N,N} \in K$ with $K \subseteq B(t_{1,N}, \delta_N) \cup \dots \cup B(t_{m_N,N}, \delta_N)$. We relabel them as $\{s_l \,|\, l \in \mathbb{N}\} = \{t_{k,N} \,|\, k = 1, \dots, m_N; N \in \mathbb{N}\}$.

2) Take a sequence (f_n) in F. Since F is bounded, we have $\sup_{n\in\mathbb{N}} |f_n(s_1)| \le \sup_{f\in F} \|f\| < \infty$. By Bolzano-Weierstraß, there exists a subsequence $(f_{\nu_1(k)}(s_1))_k$ with a limit in $\mathbb{F}$. Similarly, we obtain a converging subsequence $(f_{\nu_2(k)}(s_2))_k$ of $(f_{\nu_1(k)}(s_2))_k$. Note that the points $f_{\nu_2(k)}(s_1)$ still converge in $\mathbb{F}$ as $k \to \infty$. We iterate this procedure and define the diagonal sequence $(f_{n_j})_j = (f_{\nu_j(j)})_j$. The values $f_{n_j}(s_l)$ have a limit as $j \to \infty$ for each fixed $l \in \mathbb{N}$, since $f_{n_j}(s_l) = f_{\nu_l(k_j)}(s_l)$ for all $j \ge l$ and certain indices $k_j \to \infty$ (depending on l).

3) Let $\varepsilon > 0$. Fix a number $N \in \mathbb{N}$ with $\varepsilon_N = 1/N \le \varepsilon$. Pick the radius $\delta_N = \delta_{\varepsilon_N} > 0$ from (1.3) and the points $t_{k,N}$ from step 1). Part 2) yields an index $J_\varepsilon \in \mathbb{N}$ such that $|f_{n_i}(t_{k,N}) - f_{n_j}(t_{k,N})| \le \varepsilon$ for all $i, j \ge J_\varepsilon$ and $k \in \{1, \dots, m_N\}$. Take $t \in K$. There is an integer $l \in \{1, \dots, m_N\}$ with $\mathrm{d}(t, t_{l,N}) < \delta_N$ by part 1). We can thus estimate

$$\begin{aligned}
&|f_{n_i}(t) - f_{n_j}(t)| \\
&\quad \le |f_{n_i}(t) - f_{n_i}(t_{l,N})| + |f_{n_i}(t_{l,N}) - f_{n_j}(t_{l,N})| + |f_{n_j}(t_{l,N}) - f_{n_j}(t)| \\
&\quad \le \varepsilon_N + \varepsilon + \varepsilon_N \le 3\varepsilon
\end{aligned}$$

for all $i, j \ge J_\varepsilon$ and all $t \in K$. Since J_ε does not depend on the argument t, we have shown that the subsequence $(f_{n_j})_j$ is Cauchy in $C(K)$. This space is complete by Example 1.4 so that $(f_{n_j})_j$ converges in $C(K)$. The relative compactness of F now follows from Corollary 1.39. The addendum is clear. □

In the next result we see that the equicontinuity (1.3) is consequence of a bit uniform extra regularity of the functions in F. We state this important fact in terms of a given sequence.

Corollary 1.48. *Let $\alpha \in (0, 1]$ and K be a compact metric space. Assume that the sequence (f_n) is bounded in $C(K)$ and uniformly Hölder continuous; i.e., there is a constant $c \ge 0$ such that*

$$|f_n(t)| \le c \quad \text{and} \quad |f_n(t) - f_n(s)| \le c\, \mathrm{d}(t, s)^\alpha \tag{1.4}$$

for all $t, s \in K$ and $n \in \mathbb{N}$. Then there exists a subsequence $(f_{n_j})_j$ with a limit f in $C(K)$ such that the function f also satisfies (1.4).

Proof. The first part follows from Theorem 1.47 with $F = \{f_n \,|\, n \in \mathbb{N}\}$, noting that (1.3) holds with $\delta_\varepsilon = (\varepsilon/c)^{1/\alpha}$. The last claim can be shown as in Example 1.14f). □

Condition (1.4) with $\alpha = 1$ is true for bounded sequences in $C^1([0, 1])$, for instance. We first illustrate the above concepts and then discuss in Example 1.49c) a typical application of Arzela–Ascoli.

Example 1.49. a) The sequence (f_n) in Example 1.41c) is not equicontinuous since $\left|f_n(\frac{3}{2}2^{-n}) - f_n(2^{-n})\right| = 1$ for all $n \in \mathbb{N}$, but $\frac{3}{2}2^{-n} - 2^{-n} \to 0$ as $n \to \infty$.

b) The analogue of Arzela–Ascoli in $C_b(\mathbb{R})$ or $C_0(\mathbb{R})$ is wrong. For instance, consider the sequence (f_n) in $C_b(\mathbb{R})$ given by $f_n(t) = 0$ for $|t-n| \geq 1/2$ and $f_n(t) = 1 - 2\,|t-n|$ for $|t-n| < 1/2$. It is bounded and equicontinuous, but $\|f_n - f_m\|_\infty = 1$ for $n \neq m$.

c) Let $k \in C([0,1]^2)$ and set $Tf(t) = \int_0^1 k(t,s)f(s)\,\mathrm{d}s$ for $t \in [0,1]$ and $f \in X = C([0,1])$. Then the set $F = \{g = Tf \mid f \in \overline{B}_X(0,1)\}$ is relatively compact in X.

Proof. Let $f \in \overline{B}_X(0,1)$ and $g = Tf$. Korollar 1.47 in Analysis 2 shows that g belongs to X. Since $\|Tf\|_\infty \leq \|k\|_\infty \|f\|_\infty \leq \|k\|_\infty$, the set F is bounded. Let $\varepsilon > 0$. By the uniformy continuity of k, we find a radius $\delta > 0$ such that $|k(t,s) - k(t',s)| \leq \varepsilon$ for all $t, t', s \in [0,1]$ with $|t'-t| \leq \delta$. Arzela–Ascoli now implies the assertion because

$$|g(t') - g(t)| \leq \int_0^1 |k(t,s) - k(t',s)|\,|f(s)|\,\mathrm{d}s \leq \varepsilon$$

for all $g \in F$ and $t, t' \in [0,1]$ with $|t'-t| \leq \delta$. □

In metric spaces which are not normed vector spaces, it is possible that the closed balls are compact also in the infinite dimensional situation. The most prominent theorem in this direction is due to *Montel* and deals with holomorphic functions, see Example 1.32.

Theorem 1.50. *Let $U \subseteq \mathbb{C}$ be open and $F \subseteq H(U)$ be* locally bounded*; i.e., for each point $z \in U$ there is a radius $r_z > 0$ such that $B_z := \overline{B}(z, r_z) \subseteq U$ and $\sup_{f\in F} \sup_{w \in B_z} |f(w)|$ is finite. Then F is relatively compact in $H(U)$. In particular, all closed balls in $H^\infty(U)$ are compact for the metric of $H(U)$.*

Proof. Let (f_n) be a sequence in F.

1) We first show that the sequence (f_n) is uniformly Lipschitz on certain balls. Take a point $z^0 \in U$ and set $r = r_{z^0}/2$. Let $w, z \in \overline{B}(z^0, r)$ and $n \in \mathbb{N}$. Cauchy's integral formula (see Theorem 2.8 in Analysis 4) then yields

$$\begin{aligned} f_n(w) - f_n(z) &= \frac{1}{2\pi \mathrm{i}} \int_{|\zeta - z^0| = 2r} f_n(\zeta)\Big(\frac{1}{\zeta - w} - \frac{1}{\zeta - z}\Big)\mathrm{d}\zeta \\ &= \frac{1}{2\pi \mathrm{i}} \int_{|\zeta - z^0| = 2r} f_n(\zeta)\, \frac{w - z}{(\zeta - w)(\zeta - z)}\,\mathrm{d}\zeta, \\ |f_n(w) - f_n(z)| &\leq \frac{4\pi r}{2\pi}\, \frac{|z-w|}{r^2} \max_{|\zeta - z^0| = 2r} |f_n(\zeta)| = k(z^0)\,|z-w| \end{aligned}$$

with $k(z^0) := 2r^{-1} \sup_{n\in\mathbb{N}} \sup_{\zeta \in B_{z^0}} |f_n(\zeta)| < \infty$.

2) As in Example 1.9 there are compact sets $K_j \subseteq K_{j+1} \subseteq U$ for $j \in \mathbb{N}$ whose union is equal to U. Since every K_j can be covered

with finitely many of the balls B_z, the restrictions $(f_n)|_{K_j}$ are bounded uniformly in $n \in \mathbb{N}$, for each fixed $j \in \mathbb{N}$.

Suppose the set $F_j = \{(f_n)|_{K_j} \,|\, n \in \mathbb{N}\}$ was not equicontinuous for some $j \in \mathbb{N}$. There thus exist points $w_n, z_n \in K_j$ and a number $\varepsilon_0 > 0$ such that $|w_n - z_n| \to 0$ as $n \to \infty$ and $|f_n(w_n) - f_n(z_n)| \geq \varepsilon_0$ for all $n \in \mathbb{N}$. Since K_j is compact, there are subsequences $(w_{n_l})_l$ and $(z_{n_l})_l$ with limits w^0 and z^0 in K_j, respectively. It follows $w^0 = z^0$. Take the radius r from part 1) for the point z^0. There is an index $L \in \mathbb{N}$ such that w_{n_l} and z_{n_l} belong to $\overline{B}(z^0, r)$ for all $l \geq L$. Step 1) then implies

$$0 < \varepsilon_0 \leq |f_{n_l}(w_{n_l}) - f_{n_l}(z_{n_l})| \leq k(z^0)\,|w_{n_l} - z_{n_l}| \longrightarrow 0, \qquad l \to \infty,$$

which is impossible. Hence, F_j is equicontinuous for each $j \in \mathbb{N}$.

3) Let $j = 1$. Theorem 1.47 now yields a subsequence $(f_{\nu_1(k)})_k$ which converges in $C(K_1)$ to a function $g^{\langle 1\rangle}$. Iteratively, for each $m \in \mathbb{N}$ one obtains subsequences $(f_{\nu_m(k)})_k$ of $(f_{\nu_{m-1}(k)})_k$ that converge on $C(K_m)$ to a function $g^{\langle m\rangle}$. Let $j < m$ in $\mathbb{N}$. The restriction of $g^{\langle m\rangle}$ to K_j coincides with $g^{\langle j\rangle}$ since the restrictions of $f_{\nu_m(k)}$ to K_j converge to both functions in $C(K_j)$. We can thus define a continuous map $g : U \to \mathbb{C}$ by setting $g(z) = g^{\langle m\rangle}(z)$ for $z \in K_m$. By construction, the diagonal sequence $(f_{n_m})_m = (f_{\nu_m(m)})_m$ tends to g in $H(U)$. Corollary 1.39 now says that F is relatively compact.

4) To prove the addendum, take a sequence (g_n) in a closed ball $\overline{B}(f, r)$ of $H^\infty(U)$. Because its uniform boundedness, it has a subsequence $(g_{n_j})_j$ with a limit g in $H(U)$. Moreover, g belongs to $\overline{B}(f, r)$ since for every $z \in U$ we have the estimate

$$|g(z) - f(z)| = \lim_{j\to\infty} |g_{n_j}(z) - f(z)| \leq r. \qquad \square$$

We next derive the amazing convergence theorem of *Vitali*. To this aim, we first recall a simple useful fact.

Lemma 1.51. *Let (x_n) be a sequence in a metric space M and $x \in M$. Then (x_n) tends to x if and only if each subsequence $(x_{n_j})_j$ has a subsequence with limit x.*

Proof. If (x_n) has the limit x, then the subsequence condition is clearly satisfied. Let this condition be true. Assume that (x_n) does not tend to x. Then there exists a number $\delta > 0$ and a subsequence such that $\mathrm{d}(x_{n_j}, x) \geq \delta$ for all $j \in \mathbb{N}$. But the assumption yields a subsequence of $(x_{n_j})_j$ converging to x, which is a contradiction. $\square$

Theorem 1.52. *Let $U \subseteq \mathbb{C}$ be open and pathwise connected, and let the set $A \subseteq U$ have an accumulation point in U. Assume that the sequence (f_n) in $H(U)$ is locally bounded and that it converges pointwise on A to a function $f_0 : A \to \mathbb{C}$. Then f_0 has an extension f in $H(U)$ and (f_n) tends to f in $H(U)$.*

Proof. The set $\{f_n \,|\, n \in \mathbb{N}\}$ is relatively compact in $H(U)$ by Theorem 1.50. Each subsequence $(f_{\nu(j)})_j$ thus has an accumulation point f^ν in $H(U)$. Because of the assumption, these functions coincide on A. Theorem 2.21 of Analysis 4 then yields that all functions $f^\nu =: f$ are equal on U, so that (f_n) tends to f due to Lemma 1.51. □

We add another concept that will be needed later in the course.

Definition 1.53. *A metric space is called* separable *if it contains a countable dense subset.*

In Exercise 6.2 simple properties of this notion are established; for instance, separability is preserved under homeomorphisms. Here we only discuss standard examples based on the following auxiliary fact.

Lemma 1.54. *Let X be a normed vector space and $Y \subseteq X$ be a countable subset such that* $\operatorname{lin} Y$ *is dense in X. Then X is separable.*

Proof. The set

$$\operatorname{lin}_{\mathbb{Q}} Y = \Big\{ y = \sum_{j=1}^{m} q_j\, y_j \,\Big|\, y_j \in Y, q_j \in \mathbb{Q} \ (\mathbb{Q} + \mathrm{i}\mathbb{Q} \text{ if } \mathbb{F} = \mathbb{C}),\ m \in \mathbb{N} \Big\}$$

is countable, since Y is countable. Let $x \in X$ and $\varepsilon > 0$. By assumption, there exists a vector $y = \sum_{j=1}^m a_j\, y_j$ in $\operatorname{lin} Y$ with $\|x - y\| \le \varepsilon$, where we may assume that all $y_j \in Y$ are non-zero. We then choose numbers $q_j \in \mathbb{Q}$ (or $q_j \in \mathbb{Q} + \mathrm{i}\mathbb{Q}$) with $|a_j - q_j| \le \varepsilon/(m\|y_j\|)$ and set $z = \sum_{j=1}^m q_j\, y_j$. The vector z then belongs to $\operatorname{lin}_{\mathbb{Q}} Y$ and satisfies

$$\|y - z\| \le \sum_{j=1}^{m} |a_j - q_j|\, \|y_j\| \le \varepsilon.$$

Hence, $\|x - z\| \le 2\varepsilon$. □

Example 1.55. a) The spaces $\ell^p, 1 \le p < \infty$, and c_0 are separable since $c_{00} = \operatorname{lin}\{e_k \,|\, k \in \mathbb{N}\}$ is dense in all of them by Proposition 1.31.

b) The space $C([0,1])$ is separable since $\operatorname{lin}\{p_n \,|\, n \in \mathbb{N}\}$ with $p_n(t) = t^n$ is dense in $C([0,1])$ by Weierstraß' Theorem 5.14 in Analysis 3.

c) Let $U \subseteq \mathbb{R}^m$ be open and $p \in [1,\infty)$. The space $L^p(U)$ is separable because Korollar VI.2.30 in [**El**] shows the density of the linear hull of the functions $\mathbb{1}_J$ for intervals $J = (a,b]$ with $a, b \in \mathbb{Q}^m$ and $\overline{J} \subseteq U$.

d) The space ℓ^∞ is not separable. In fact, the set Ω of $\{0,1\}$–valued sequences is uncountable and two different elements in Ω have distance 1. Suppose that the set $\{v_k \,|\, k \in \mathbb{N}\}$ was dense in ℓ^∞. Then Ω belongs to $\bigcup_{k\in\mathbb{N}} B(v_k, 1/4)$. As each ball $B(v_k, 1/4)$ contains at most one sequence $\omega \in \Omega$, the set Ω must be countable, which is wrong. ◊

CHAPTER 2

Continuous linear operators

2.1. Basic properties and examples of linear operators

The set of linear maps $T : X \to Y$ is designated by $L(X, Y)$. From Linear Algebra it is known that it is a vector space for the sum and scalar multiplication defined before Example 1.4. Recall that $T(0) = 0$. We usually write Tx instead of $T(x)$ and $ST \in L(X, Z)$ instead of $S \circ T$ for all $x \in X$, $T \in L(X, Y)$, $S \in L(Y, Z)$, and vector spaces Z. One often calls $T \in L(X, Y)$ an *operator*.

If $\dim X < \infty$ and $\dim Y < \infty$, each element T of $L(X, Y)$ can be represented by a matrix, and it is continuous. (See Beispiel 1.36 in Analysis 2 and the text preceding it.) However, in infinite dimensional spaces there are discontinuous linear maps.

Example 2.1. a) By $T(x_k) = (kx_k)$ we define a linear map $T : c_{00} \to c_{00}$. This operator is not continuous for any p–norm on c_{00}. Indeed, take the finite sequences $v_n = n^{-\frac{1}{2}} e_n$ for $n \in \mathbb{N}$, which satisfy $T(v_n) = n^{\frac{1}{2}} e_n$, $\|v_n\|_p = n^{-\frac{1}{2}} \to 0$, and $\|T(v_n)\|_p = n^{\frac{1}{2}} \to \infty$ as $n \to \infty$.

b) The map $Tf = f'$ is linear from $C^1([0, 1])$ to $C([0, 1])$, but not continuous if both spaces are endowed with the supremum norm. To check this claim, we consider the functions $f_n(t) = n^{-1/2} \sin(nt)$, and note that $\|f_n\|_\infty \le n^{-1/2}$ and $\|f_n'\|_\infty \ge |f_n'(0)| = n^{1/2}$.

On the other hand, T is continuous if we equip $C^1([0, 1])$ with the norm given by $\|f\|_{C^1} = \|f\|_\infty + \|f'\|_\infty$ and $C([0, 1])$ with $\|f\|_\infty$. ♢

We first characterize the (Lipschitz) continuity of linear operators by their *boundedness*.

Lemma 2.2. *Let X and Y be normed vector spaces and $T : X \to Y$ be linear. The following assertions are equivalent.*

a) T is Lipschitz continuous.

b) T is continuous.

c) T is continuous at $x = 0$.

d) T is bounded*; i.e, there is a constant $c > 0$ with $\|Tx\|_Y \le c\,\|x\|_X$ for every $x \in X$.*

Proof. The implications 'a) ⇒ b) ⇒ c)' are clear.

Let c) be true. Since also $T0 = 0$, there exists a radius $\delta > 0$ such that $\|Tz\| \le 1$ for all vectors $z \in \overline{B}(0, \delta)$. Let $x \in X \setminus \{0\}$, and set $z = \frac{\delta}{\|x\|} x \in \overline{B}(0, \delta)$. Because T is linear, we deduce $1 \ge \|Tz\| \ge \frac{\delta}{\|x\|} \|Tx\|$ and hence d) with $c = 1/\delta$. Of course, $x = 0$ fulfills assertion d).

Let d) be true. The linearity of T then yields $\|Tx - Tz\| = \|T(x - z)\| \le c\,\|x - z\|$ for all $x, z \in X$. $\square$

The above equivalence leads to the next definition.

DEFINITION 2.3. *For normed vector spaces X and Y, we set*

$$\mathcal{B}(X,Y) = \{T : X \to Y \,|\, T \text{ is linear and continuous}\}^{1}$$

and put $\mathcal{B}(X,X) = \mathcal{B}(X)$. The dual space $\mathcal{B}(X,\mathbb{F})$ *of X is denoted by $X^\star$. An element $x^\star \in X^\star$ is called linear functional on X, and one often writes $\langle x, x^\star\rangle_X = \langle x, x^\star\rangle$ instead of $x^\star(x)$.*

For maps $T : X \to Y$ the operator norm *is given by*

$$\|T\| = \inf\{c \ge 0 \,|\, \forall\, x \in X : \|Tx\|_Y \le c\,\|x\|_X\} \;\in\; [0,\infty].$$

The space $\mathcal{B}(X,Y)$ thus consists of all linear maps $T : X \to Y$ with $\|T\| < \infty$. Unless something else is said, we endow $\mathcal{B}(X,Y)$ with $\|T\|_{\mathcal{B}(X,Y)} := \|T\|$. We next show a few basic facts.

REMARK 2.4. Let X, Y, Z be normed vector spaces, $x \in X$, and $T \in L(X,Y)$. Then the following assertions are true.

a) The space $\mathcal{B}(X,Y)$ does not change if we replace the norms on X or Y by equivalent ones, though the norm $\|T\|$ of $T \in \mathcal{B}(X,Y)$ may depend on the choice of $\|\cdot\|_X$ or $\|\cdot\|_Y$.

b) The identity $I : X \to X$; $Ix = x$, has norm $\|I\| = 1$.

c) $\|T\| \overset{(1)}{=} \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} \overset{(2)}{=} \sup_{\|x\|\le 1} \|Tx\| \overset{(3)}{=} \sup_{\|x\|=1} \|Tx\| =: s.$[2]

d) $\|Tx\| \le \|T\|\,\|x\|$.

e) Let $T \in \mathcal{B}(X,Y)$ and $S \in \mathcal{B}(Y,Z)$. Then their product ST belongs to $\mathcal{B}(X,Z)$ with $\|ST\| \le \|S\|\,\|T\|$.

PROOF. Assertions a) and b) are clear.

c) Let $\|T\| < \infty$. For every $\varepsilon > 0$, Definition 2.3 yields $\|Tx\| \le (\|T\| + \varepsilon)\,\|x\|$. Taking the infimum over $\varepsilon > 0$, we obtain the relation '$\ge$' in (1). Similarly, one treats the case $\|T\| = \infty$. We clearly have '$\ge$' in (2) and (3) for both cases. The remaining inequality $\|T\| \le s$ is a consequence of the bound

$$\frac{\|Tx\|}{\|x\|} = \left\|T\big(\tfrac{1}{\|x\|}x\big)\right\| \le s \qquad \text{for} \quad x \neq 0.$$

Claim d) is true for $x = 0$, and follows for $x \neq 0$ from (1) in part c). Assertion d) yields $\|STx\| \le \|S\|\|Tx\| \le \|S\|\|T\|\|x\|$, and thus e). $\square$

PROPOSITION 2.5. *Let X and Y be normed vector spaces. Then $\mathcal{B}(X,Y)$ is a normed vector space with respect to the operator norm. If Y is a Banach space, then $\mathcal{B}(X,Y)$ is also a Banach space. In particular, $X^\star$ is a Banach space.*

[1]One also uses the notation $\mathcal{L}(X,Y)$ for this space.

[2]In contrast to the lectures, we allow for $\|T\| = \infty$ here.

Proof. Clearly, $\mathcal{B}(X,Y)$ is a vector space. If $\|T\| = 0$, then $Tx = 0$ for all $x \in X$ by Remark 2.4, and hence $T = 0$. Let $T, S \in \mathcal{B}(X,Y)$, $x \in X$, and $\alpha \in \mathbb{F}$. We then deduce from Remark 2.4 that

$$\|T + S\| = \sup_{\|x\|=1} \|(T+S)x\| \le \sup_{\|x\|=1} (\|Tx\| + \|Sx\|) \le \|T\| + \|S\|,$$
$$\|\alpha T\| = \sup_{\|x\|=1} \|\alpha Tx\| = \sup_{\|x\|=1} |\alpha|\, \|Tx\| = |\alpha| \sup_{\|x\|=1} \|Tx\| = |\alpha|\, \|T\|.$$

Thus, $\mathcal{B}(X,Y)$ is a normed vector space for the operator norm.

Assume that Y is a Banach space. Take a Cauchy sequence (T_n) in $\mathcal{B}(X,Y)$. Let $\varepsilon > 0$. There is an index $N_\varepsilon \in \mathbb{N}$ such that $\|T_n - T_m\| \le \varepsilon$ for all $n, m \ge N_\varepsilon$. Let $x \in X$. Since $T_n - T_m$ belongs to $\mathcal{B}(X,Y)$, Remark 2.4 yields

$$\|T_n x - T_m x\| \le \|T_n - T_m\|\, \|x\| \le \varepsilon \|x\|$$

for all $n, m \ge N_\varepsilon$. Hence, $(T_n x)$ is a Cauchy sequence in Y and possesses a unique limit $y =: Tx \in Y$. Let $\alpha, \beta \in \mathbb{F}$ and $x, z \in X$. We deduce from the linearity of T_n that

$$T(\alpha x + \beta z) = \lim_{n\to\infty} T_n(\alpha x + \beta z) = \lim_{n\to\infty} (\alpha T_n x + \beta T_n z) = \alpha Tx + \beta Tz.$$

Since (T_n) is Cauchy, there is a constant $c > 0$ with $\|T_n\| \le c$ for all n, whence $\|Tx\| = \lim_{n\to\infty} \|T_n x\| \le c\,\|x\|$. As a result, T is contained in $\mathcal{B}(X,Y)$. Let $n \ge N_\varepsilon$. The above displayed estimate implies

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

i.e., $\|T - T_n\| \le \varepsilon$ as required. □

We now discuss several important classes of bounded operators and compute their norms, starting with the most simple one. (See also the exercises.)

Example 2.6 (Multiplication operators). a) Let K be a compact metric space and $X = C(K)$. We fix a function m in $C(K)$, and define $Tf = mf$ for every $f \in X$.

1) The function Tf clearly belongs to X. We also have $T(\alpha f + \beta g) = \alpha mf + \beta mg = \alpha Tf + \beta Tg$ for all $f, g \in X$ and $\alpha, \beta \in \mathbb{F}$. The operator T is thus contained in $L(X)$.

2) For each $f \in X$ we estimate

$$\|Tf\|_\infty = \sup_{s\in K} |m(s)|\, |f(s)| \le \|m\|_\infty\, \|f\|_\infty,$$

so that T is an element of $\mathcal{B}(X)$ with norm $\|T\| \le \|m\|_\infty$.

3) To show equality, we take the map $f = \mathbb{1}$ in X. Since $\|\mathbb{1}\|_\infty = 1$, Remark 2.4 yields the lower bound $\|T\| \ge \|T\mathbb{1}\|_\infty = \|m\mathbb{1}\|_\infty = \|m\|_\infty$, and hence $\|T\| = \|m\|_\infty$.

b) Let $(S,\mathcal{A},\mu)$ be a σ–finite measure space, $1\le p\le\infty$, and $X=L^p(\mu)$. Fix an element $\hat m$ of $L^\infty(\mathbb{R}^m)$. For $\hat f=f+\mathcal{N}\in X$ we define $T\hat f=mf+\mathcal{N}$ for any representatives $f\in\hat f$ and $m\in\hat m$.[3]

1) Take other functions $f_1\in\hat f$ and $m_1\in\hat m$. Then there are null sets N' and N'' such that $f(s)=f_1(s)$ for all $s\in S\setminus N'$ and $m(s)=m_1(s)$ for all $s\in S\setminus N''$. We thus obtain $m(s)f(s)=m_1(s)f_1(s)$ for all s outside the null set $N'\cup N''$, so that the equivalence class $T\hat f$ does not depend on the representatives of $\hat f$ and $\hat m$.

2) The product mf is measurable. Pick a number $c>\|m\|_\infty$. We then have a null set N such that $|m(s)f(s)|\le c\,|f(s)|$ for all $s\in S\setminus N$. Hence, the map mf belongs to $\mathcal{L}^p(\mu)$ and satisfies $\|mf\|_p\le\|m\|_\infty\,\|f\|_p$. (Take the infimum over $c>\|m\|_\infty$.) We have thus shown that T maps X into X and $\|T\hat f\|_p=\|mf\|_p\le\|\hat m\|_\infty\,\|\hat f\|_p$.

3) For $f,g\in\mathcal{L}(\mathbb{R}^m)$ and $\alpha,\beta\in\mathbb{F}$ we compute

$$T(\alpha\hat f+\beta\hat g)=T(\alpha f+\beta g+\mathcal{N})=\alpha mf+\beta mg+\mathcal{N}=\alpha T\hat f+\beta T\hat g.$$

Consequently, T is an element of $\mathcal{B}(X)$ fulfilling $\|T\|\le\|\hat m\|_\infty$.

4) We finally claim that $\|T\|=\|\hat m\|_\infty$. This statement is true if $\hat m=0$. So let $\|\hat m\|_\infty=\|m\|_\infty>0$. By assumption, there are sets $S_n\in\mathcal{A}$ for $n\in\mathbb{N}$ of finite measure whose union is equal to S. Take $\varepsilon\in(0,\|m\|_\infty)$ and $n\in\mathbb{N}$. We put $A_{\varepsilon,n}=\{s\in S_n\mid |m(s)|\ge\|m\|_\infty-\varepsilon\}\in\mathcal{A}$. We can find an index $k\in\mathbb{N}$ with $\mu(A_{\varepsilon,k})\in(0,\infty)$ since $\{|m|\ge\|m\|_\infty-\varepsilon\}$ is not a null set. This fact allows us to define the map $f_\varepsilon=\|\mathbb{1}_{A_{\varepsilon,k}}\|_p^{-1}\mathbb{1}_{A_{\varepsilon,k}}$ belonging to $\mathcal{L}^p(\mathbb{R}^m)$ with norm $\|f_\varepsilon\|_p=1$. Using Remark 2.4, we infer

$$\|T\|\ge\|T\widehat{f_\varepsilon}\|_p=\|mf_\varepsilon\|_p=\frac{1}{\|\mathbb{1}_{A_{\varepsilon,k}}\|_p}\Big[\int_{A_{\varepsilon,k}}|m(s)|^p\,\mathrm{d}\mu(s)\Big]^{\frac1p}\ge\|m\|_\infty-\varepsilon,$$

where we let $p<\infty$. In the limit $\varepsilon\to0$ is follows that $\|T\|=\|\hat m\|_\infty$. The case $p=\infty$ is treated in the same way. ◊

We next look at a crucial class of bounded operators. More sophisticated examples for it are studied at the end of this section, for instance.

Example 2.7 (Integral operators). Let $X=C([0,1])$ and the *kernel* $k\in C([0,1]^2)$ be given. Let $f\in X$. We define

$$(Tf)(t)=\int_0^1 k(t,s)f(s)\,\mathrm{d}s,\qquad t\in[0,1].$$

1) In Example 1.49 we have seen that Tf belongs to X. Basic properties of the Riemann integral imply that $T:X\to X$ is linear.

2) Set $\kappa=\sup_{t\in[0,1]}\int_0^1|k(t,s)|\,\mathrm{d}s\le\|k\|_\infty$. We then calculate

$$\|Tf\|_\infty=\sup_{t\in[0,1]}\Big|\int_0^1 k(t,s)f(s)\,\mathrm{d}s\Big|\le\sup_{t\in[0,1]}\int_0^1|k(t,s)|\,|f(s)|\,\mathrm{d}s\le\kappa\|f\|_\infty.$$

[3] As an exception, here we explicitely take into acccount that $L^q=\mathcal{L}^q/\mathcal{N}$.

The operator T thus belongs to $\mathcal{B}(X)$ and satisfies $\|T\| \le \kappa$.

3) By continuity, there is a number $t_0 \in [0,1]$ with $\kappa = \int_0^1 |k(t_0,s)|\,\mathrm{d}s$. To show a lower bound for the norm, we introduce the functions

$$f_n : [0,1] \to \mathbb{F}; \quad f_n(s) = \frac{\bar{k}(t_0,s)}{|k(t_0,s)| + \frac{1}{n}},$$

for $n \in \mathbb{N}$. Since $f_n \in X$ and $\|f_n\|_\infty \le 1$, Remark 2.4 leads to

$$\|T\| \ge \|Tf_n\|_\infty \ge |Tf_n(t_0)| = \int_0^1 \frac{|k(t_0,s)|^2}{|k(t_0,s)| + \frac{1}{n}}\,\mathrm{d}s \;\longrightarrow \kappa$$

as $n \to \infty$, using e.g. Lebesgue's theorem. It follows that $\|T\| = \kappa$.

Note that for $k \ge 0$ one obtains this equality more easily as $\|T\| \ge \|T\mathbb{1}\|_\infty = \kappa$ in this case. ♢

The following type of linear maps will be studied in great detail in the penultimate chapter.

Example 2.8 (Linear functionals). a) Let $X = C([0,1])$. For a fixed $t_0 \in [0,1]$ we define the point evaluation $\varphi(f) = f(t_0)$ for all $f \in X$. It is clear that $\varphi : X \to \mathbb{F}$ is linear and that $|\varphi(f)| \le \|f\|_\infty$. Hence, φ belongs to X^* with norm $\|\varphi\| \le 1$. On the other hand, we have $\|\mathbb{1}\|_\infty = 1$ and thus $\|\varphi\| \ge |\varphi(\mathbb{1})| = 1$, implying $\|\varphi\| = 1$.

b) Let $X = L^p(\mu)$ for a measure space $(S,\mathcal{A},\mu)$ and $1 \le p \le \infty$, and let $g \in L^{p'}(\mu)$ be fixed. Hölder's inequality says that the integral $\varphi(f) = \int fg\,\mathrm{d}\mu \in \mathbb{F}$ exists for all $f \in X$ and that it is bounded by $|\varphi(f)| \le \|f\|_p \|g\|_{p'}$. Since linearity is clear, we see that φ is an element of $X^\star$ with $\|\varphi\| \le \|g\|_{p'}$. (Equality is shown in Proposition 5.1.)

c) On $X = C([0,1])$ with $\|\cdot\|_1$, the linear form $f \mapsto \varphi(f) = f(0)$ is not continuous. For instance, the functions f_n given by $f_n(t) = 1 - nt$ for $0 \le t < \frac{1}{n}$ and $f_n(t) = 0$ for $\frac{1}{n} \le t \le 1$ satisfy $\|f_n\|_1 = \frac{1}{2n} \to 0$ as $n \to \infty$, but $\varphi(f_n) = 1$ for all $n \in \mathbb{N}$. ♢

In the following examples, we encounter another fundamental difference between the finite and the infinite dimensional situation. See also Example 4.12 for related operators.

Example 2.9 (Shift operators). Let $X \in \{c_0, c, \ell^p \,|\, 1 \le p \le \infty\}$. The right and left shift operator on X are given by

$$Rx = (0, x_1, x_2, \dots) \qquad \text{and} \qquad Lx = (x_2, x_3, \dots)$$

for $x \in X$. Clearly, $R, L : X \to X$ are linear maps satisfying $\|Rx\|_p = \|x\|_p$ and $\|Lx\|_p \le \|x\|_p$ for all $x \in X$, as well as $\|Le_2\|_p = 1$. The operators R and L thus belong to $\mathcal{B}(X)$ with norm 1. We stress that $LR = I$, $RLx = (0, x_2, x_3, \dots)$, $Le_1 = 0$, and

- R is injective, but not surjective, and it has a left inverse, but no right inverse;

- L is surjective, but not injective, and it has a right inverse, but no left inverse.

Recall that in the case $\dim X < \infty$ the injectivity and the surjectivity of a map $T \in L(X)$ are equivalent, and that here a right or left inverse is automatically an inverse! ♢

We introduce several notions in the context of linear operators and list related observations.

DEFINITION 2.10. *Let $T : X \to Y$ be a linear map for normed vector spaces X and Y.*

a) We denote kernel *and* range *of T by $N(T) = \{x \in X \mid Tx = 0\}$ and $R(T) = T(X) = TX = \{y = Tx \mid x \in X\}$, respectively.*

b) An injective operator $T \in \mathcal{B}(X,Y)$ is called an embedding*, which is designated by $X \hookrightarrow Y$.*

c) A bijective operator $T \in \mathcal{B}(X,Y)$ having a continuous inverse T^{-1} is called isomorphism *or* invertible*. One then writes $X \simeq Y$.*

d) The map T is said to be isometric *if $\|Tx\| = \|x\|$ for all $x \in X$, and* contractive *if $\|T\| \le 1$.*

REMARK 2.11. a) Let X and Y be normed vector spaces and $J : X \to Y$ be an isomorphism. Let $x_n \in X$ and $y_n := Jx_n$ for $n \in \mathbb{N}$. Hence, $x_n = J^{-1}y_n$ and so the sequence (x_n) converges if and only if (y_n) converges. Proposition 1.12 and Theorem 1.37 then show that a set $C \subseteq X$ is closed [open, resp. compact] if and only if the image $D = JC \subseteq Y$ is closed [open, resp. compact]. Similarly, X is a Banach space if and only if Y is a Banach space.

b) The kernel $N(T)$ of a map $T \in \mathcal{B}(X,Y)$ is closed by Proposition 1.24. An isometry is contractive and injective, and a contraction is continuous. In Example 2.9, the right shift R is an isometry, and the left shift L has norm 1, but L is not an isometry.

c) Let $T \in \mathcal{B}(X,Y)$ be an isometry. Then its inverse $T^{-1} : R(T) \to X$ is linear and isometric. In fact, for $y = Tx$ in $R(T)$ we compute $\|T^{-1}y\| = \|x\| = \|Tx\| = \|y\|$.

d) Let X be a Banach space and let the operator $T \in \mathcal{B}(X,Y)$ satisfy the lower bound $\|Tx\| \ge c\,\|x\|$ for some $c > 0$ and all $x \in X$ (e.g., if T is isometric). Then its range $R(T)$ is closed in Y.

PROOF. Take a sequence $(y_n) = (Tx_n)$ in $R(T)$ with limit y in Y. The assumption yields $\|x_n - x_m\| \le c^{-1}\|y_n - y_m\| \to 0$ as $n, m \to \infty$. By completeness, there exists $x = \lim_{n\to\infty} x_n$. The continuity of T then implies that $y = Tx$ belongs to $R(T)$; i.e., $R(T)$ is closed. □

e) Let Y be a linear subspace of $(X, \|\cdot\|_X)$ with its own norm $\|\cdot\|_Y$. The identity $I : (Y, \|\cdot\|_Y) \to (X, \|\cdot\|_X)$ is continuous (and thus an embedding) if and only if $\|y\|_X = \|Iy\|_X \le c\,\|y\|_Y$ for all $y \in Y$ and a constant $c \ge 0$ if and only if $\|\cdot\|_Y$ is finer than $\|\cdot\|_X$. We have the examples $\ell^p \hookrightarrow \ell^q$ if $1 \le p \le q \le \infty$ and $C^1([0,1]) \hookrightarrow C([0,1])$. ♢

We illustrate the above concepts by several important operators.

Example 2.12. a) Let $\emptyset \neq B \in \mathcal{B}_m$ with $\lambda(B) < \infty$ such that for all $x \in B$ and $r > 0$ we have $\lambda(B \cap B(x,r)) > 0$. For $f \in C_b(B)$ set $Jf = f + \mathcal{N}$. This map is linear and bounded from $C_b(B)$ to $L^p(B)$ for $p \in [1,\infty]$, since $\|Jf\|_p = \|f\|_p \leq \lambda(B)^{1/p} \|f\|_\infty$ for all $f \in C_b(B)$. If $Jf = 0$, then $f = 0$ a.e.. Take any $x \in B$. The assumption gives $x_n \in B$ with $f(x_n) = 0$ and $x_n \to x$ as $n \to \infty$. It follows $f(x) = 0$ by continuity. As a result, $J : C_b(B) \to L^p(B)$ is an embedding.

b) We define a map J on $C([-1,1])$ by setting

$$Jf(t) = \begin{cases} 0, & |t| \geq 2, \\ (2+t)f(-2-t), & -2 < t < -1, \\ f(t), & |t| \leq 1, \\ (2-t)f(2-t), & 1 < t < 2, \end{cases}$$

for $f \in C([-1,1])$. Clearly, the map Jf is contained in $C_0(\mathbb{R})$ and J is linear and isometric. Hence, $J : C([-1,1]) \hookrightarrow C_0(\mathbb{R})$ is an isometric embedding. (One can also embed $C(K)$ into $C_0(\mathbb{R}^m)$ for any compact $K \subseteq \mathbb{R}^m$ using Tietze's extension theorem, see Satz B.1.5 in [**We**].)

c) Let $X = \{f \in C^1([0,1]) \,|\, f(0) = 0\}$ be endowed with the norm given by $\|f\| = \|f'\|_\infty$. Then the map $D : X \to C([0,1])$; $Df = f'$, is an isometric isomorphism with inverse defined by $D^{-1}g(t) = \int_0^t g(s)\,\mathrm{d}s$ for $t \in [0,1]$ and $g \in C([0,1])$. The Banach space structures of X and $C([0,1])$ are thus the 'same' by Remark 2.11. But, the isomorphism D destroys other properties such as non-negativity (e.g., $f(t) = t(1-t)$ is non-negative on $[0,1]$, in contrast to $Df(t) = f'(t) = 1 - 2t$).

d)[4] Each sequence $m \in \ell^\infty$ induces the multiplication operator $T_m : x \mapsto mx$ on ℓ^p for $p \in [1,\infty]$. Example 2.6b) with $(S,\mathcal{A},\mu) = (\mathbb{N},\mathcal{P}(\mathbb{N}),\zeta)$ implies that the mapping $\ell^\infty \to \mathcal{B}(\ell^p)$; $m \mapsto T_m$, is an isometry which is also linear. By the previous remark, ℓ^∞ is thus isomorphic to a closed subspace of $\mathcal{B}(\ell^p)$. From Example 1.55 and Exercise 6.2 we then infer that $\mathcal{B}(\ell^p)$ is not separable. ◊

The next simple extension lemma is used thoughout mathematics.

Lemma 2.13. *Let X be a normed vector space, Y be a Banach space, $D \subseteq X$ be a dense linear subspace (endowed with the norm of X), and $T_0 \in \mathcal{B}(D,Y)$. Then there exists exactly one extension $T \in \mathcal{B}(X,Y)$ of T_0; i.e., $T_0 x = Tx$ for all $x \in D$. We further have $\|T_0\| = \|T\|$, and T is isometric if T_0 is isometric.*

Proof. Let $x \in X$. Choose vectors $x_n \in D$ such that $x_n \to x$ in X as $n \to \infty$. Since $\|T_0 x_n - T_0 x_m\| \leq \|T_0\|\,\|x_n - x_m\|$, the sequence $(T_0 x_n)$ is Cauchy and thus converges to an element of Y denoted by Tx. Let also $(\widetilde{x}_n)$ in D tend to x. Because of $\|T_0 x_n - T_0 \widetilde{x}_n\| \leq \|T_0\| \|x_n - \widetilde{x}_n\| \to 0$

[4] This example was mentioned in the lectures at a different place.

as $n \to \infty$, the vector Tx indeed does not depend on the approximating sequence. It is clear that $Tx = T_0 x$ for $x \in D$ (take $x_n = x$).

Let $x, z \in D$ and $\alpha, \beta \in \mathbb{F}$. Pick $x_n, z_n \in D$ with $x_n \to x$ and $z_n \to z$ as $n \to \infty$. Because T_0 is continuous and linear, we obtain

$$T(\alpha x+\beta z) = \lim_{n\to\infty} T_0(\alpha x_n+\beta z_n) = \lim_{n\to\infty} (\alpha T_0 x_n+\beta T_0 z_n) = \alpha Tx+\beta Tz,$$

and hence $T : X \to Y$ is linear. Since

$$\|Tx\| = \lim_{n\to\infty} \|T_0 x_n\| \le \|T_0\| \lim_{n\to\infty} \|x_n\| = \|T_0\| \, \|x\| \,,$$

the operator T belongs to $\mathcal{B}(X,Y)$ with $\|T\| \le \|T_0\|$. (If T_0 is isometric, one sees that T is also isometric.) On the other hand, Remark 2.4 yields

$$\|T\| = \sup_{x\in X,\, \|x\|=1} \|Tx\| \ge \sup_{x\in D,\, \|x\|=1} \|T_0 x\| = \|T_0\| \,,$$

so that $\|T_0\| = \|T\|$.

Let $S \in \mathcal{B}(X,Y)$ satisfy $Sx = T_0 x$ for all $x \in D$. Let $z \in X$. Choose $x_n \in D$ with $x_n \to z$ as $n \to \infty$. The uniqueness assertion follows from

$$Sz = \lim_{n\to\infty} Sx_n = \lim_{n\to\infty} T_0 x_n = Tz. \qquad \square$$

Convolutions and Young's inequality. In this subsection we derive Young's important inequality for convolutions. Their relevance will become clear in Theorem 4.13 and Section 4.2. We first introduce and discuss the convolution $f * g$ of functions $f \in L^p(\mathbb{R}^m)$ and $g \in L^q(\mathbb{R}^m)$ for suitable $p, q \in [1,\infty]$. Note that the map

$$\varphi : \mathbb{R}^{2m} \to [0,\infty); \quad \varphi(x,y) = |f(x-y)g(y)|,$$

is measurable as a combination of measurable maps.

Step 1). Let $f, g \in L^1(\mathbb{R}^m)$. By means of Fubini's Theorem 3.29 in Analysis 3 and the transformation $z = x - y$, we derive

$$\begin{aligned}\int_{\mathbb{R}^{2m}} \varphi(x,y)\,\mathrm{d}(x,y) &= \int_{\mathbb{R}^m}\int_{\mathbb{R}^m} |f(x-y)|\,\mathrm{d}x\,|g(y)|\,\mathrm{d}y \\ &= \int_{\mathbb{R}^m}\int_{\mathbb{R}^m} |f(z)|\,\mathrm{d}z\,|g(y)|\,\mathrm{d}y = \|f\|_1\|g\|_1 < \infty\end{aligned}$$

and thus the integrability of φ on $\mathbb{R}^{2m}$. Therefore Fubini's theorem shows that the *convolution*

$$(f*g)(x) = \int_{\mathbb{R}^m} f(x-y)g(y)\,\mathrm{d}y \tag{2.1}$$

is defined in $\mathbb{F}$ for a.e. $x \in \mathbb{R}^m$ (where we set $f * g = 0$ on the null set) and that it belongs to $L^1(\mathbb{R}^m)$ with

$$\|f*g\|_1 \le \int_{\mathbb{R}^m}\int_{\mathbb{R}^m} |f(x-y)g(y)|\,\mathrm{d}y\,\mathrm{d}x = \|\varphi\|_1 = \|f\|_1\,\|g\|_1.$$

Step 2). Let $q \in [1, \infty)$, $f \in L^1(\mathbb{R}^m)$, and $g \in L^q(\mathbb{R}^m)$. Fubini's theorem again yields the existence of the integral

$$\psi(x) := \int_{\mathbb{R}^m} |f(x-y)g(y)| \,\mathrm{d}y = \int_{\mathbb{R}^m} |f(x-y)|^{\frac{1}{q'}} \, |f(x-y)|^{\frac{1}{q}} \, |g(y)| \,\mathrm{d}y$$

in $[0, \infty]$ for all $x \in \mathbb{R}^m$ and that ψ is measurable on $\mathbb{R}^m$. From Hölder's inequality, we then deduce

$$\begin{aligned} \psi(x)^q &\leq \left(\int_{\mathbb{R}^m} |f(x-y)| \,\mathrm{d}y \right)^{\frac{q}{q'}} \int_{\mathbb{R}^m} |f(x-y)| \, |g(y)|^q \,\mathrm{d}y \\ &= \|f\|_1^{q-1} \int_{\mathbb{R}^m} |f(x-y)| \, |g(y)|^q \,\mathrm{d}y, \end{aligned}$$

also using $q' = q/(q-1)$ and the transformation $z = x - y$. Step 1) now implies the estimate

$$\begin{aligned} \|\psi\|_q^q = \int_{\mathbb{R}^m} \psi^q \,\mathrm{d}x \leq \|f\|_1^{q-1} \, \| \, |f| * |g|^q \, \|_1 \leq \|f\|_1^{q-1} \, \|f\|_1 \int_{\mathbb{R}^m} |g|^q \,\mathrm{d}x \\ \leq \|f\|_1^q \, \|g\|_q^q \, . \end{aligned} \tag{2.2}$$

This time we cannot directly deduce the measurability of $f * g$ from Fubini's theorem since we integrated ψ^q instead of ψ. To deal with this problem, we use also Proposition 1.35 and compute

$$\|\psi\|_q^q \geq \int_{B(0,n)} \psi(x)^q \,\mathrm{d}x \geq \delta_n \int_{B(0,n)} \psi(x) \,\mathrm{d}x = \delta_n \int_{\mathbb{R}^{2m}} \mathbb{1}_{B(0,n)}(x) \varphi(x,y) \,\mathrm{d}(x,y)$$

for a constant $\delta_n > 0$ and every $n \in \mathbb{N}$. Therefore, the function given by $\varphi_n(x,y) = \mathbb{1}_{B(0,n)}(x)\varphi(x,y)$ is integrable on $\mathbb{R}^{2m}$. Fubini's theorem thus shows that $f * g$ is defined by (2.1) for a.e. $x \in B(0,n)$ and each $n \in \mathbb{N}$ (and hence for a.e. $x \in \mathbb{R}^m$) and that the map $\mathbb{1}_{B(0,n)} f * g$ is measurable on $\mathbb{R}^m$. Letting $n \to \infty$, we see that the pointwise limit $f * g$ is measurable on $\mathbb{R}^m$. Estimate (2.2) finally yields

$$\|f * g\|_q^q = \int_{\mathbb{R}^m} \left| \int_{\mathbb{R}^m} f(x-y) g(y) \,\mathrm{d}y \right|^q \mathrm{d}x \leq \int_{\mathbb{R}^m} \psi^q \,\mathrm{d}x \leq \|f\|_1^q \, \|g\|_q^q.$$

We have thus proved a part of the next result, see Theorem 4.33 in [**Br**] for the remaining cases. A different proof for the full statement is given at the end of Section 2.3.

THEOREM 2.14. *Let* $1 \leq p, q, r \leq \infty$ *with* $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. *Take* $f \in L^p(\mathbb{R}^m)$ *and* $g \in L^q(\mathbb{R}^m)$. *Then the convolution* $(f * g)(x)$ *in* (2.1) *is defined in* $\mathbb{F}$ *for a.e.* $x \in \mathbb{R}^m$ *(where we set* $f * g = 0$ *on the null set), and it gives a function in* $L^r(\mathbb{R}^m)$. *We further have* Young's inequality $\|f * g\|_r \leq \|f\|_p \, \|g\|_q$.

Let $p, q, r \in [1, \infty]$, $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$, and $g \in L^q(\mathbb{R}^m)$ be fixed. The above theorem yields the bounded linear operator $T : L^p(\mathbb{R}^m) \to L^r(\mathbb{R}^m)$; $f \mapsto f * g$, with norm $\|T\| \leq \|g\|_q$.

2.2. Standard constructions

One can construct new normed vector spaces out of given ones in various ways. The basic methods are treated below, see also Corollary 1.13. These results are standard tools in analysis.

A) Product spaces. We start with the simplest case. Let X and Y be normed vector spaces. The Cartesian product $X \times Y = \{(x,y) \,|\, x \in X, y \in Y\}$ is a normed vector space for each of the norms

$$\|(x,y)\|_p = \begin{cases} \max\{\|x\|_X, \|y\|_Y\}, & p = \infty, \\ (\|x\|_X^p + \|y\|_Y^p)^{1/p}, & p \in [1,\infty). \end{cases}$$

These norms are equivalent. We have $(x_n, y_n) \to (x, y)$ in $X \times Y$ if and only if $x_n \to x$ in X and $y_n \to y$ in Y, as $n \to \infty$. Moreover, $X \times Y$ is complete if X and Y are complete. These facts can be proved as in Analysis 2 for the case $\mathbb{R} = X = Y$. There are obvious modifications for finite products.

B) Direct sums. We now discuss how to decompose X into closed subspaces. Surprisingly, in a Banach space this procedure is equivalent to the Cartesian product, see Remark 2.17.

Definition 2.15. *Let X_1 and X_2 be* **closed** *linear subspaces of a normed vector space X such that $X_1 + X_2 = X$ and $X_1 \cap X_2 = \{0\}$. We then say that X is the* direct sum *of X_1 and X_2 and that X_2 is the* complement *of X_1. In this case we write $X = X_1 \oplus X_2$. Let Y be a vector space. A map $P \in L(Y)$ is called* projection *if $P^2 = P$.*

We first show that bounded projections yield direct sums.

Lemma 2.16. *Let X be a normed vector space and $P \in \mathcal{B}(X)$ be a projection. Then the operator $I - P \in \mathcal{B}(X)$ is also a projection. We have the equations*

$$R(P) = N(I-P) =: X_1, \quad N(P) = R(I-P) =: X_2, \quad X = X_1 \oplus X_2.$$

Moreover, P satisfies $\|P\| \ge 1$ if $P \ne 0$.

Proof. From $P = P^2$ we deduce $(I-P)^2 = I - 2P + P^2 = I - P$. If $y \in R(P)$, then there is a vector $x \in X$ with $y = Px$ and thus $(I-P)y = Px - P^2x = 0$; i.e. $y \in N(I-P)$. Conversely, if $(I-P)x = 0$, then $x = Px \in R(P)$. So we have shown the asserted equalities for X_1, which yield those for X_2 since $I - P$ is a projection and $I - (I-P) = P$. By Proposition 1.24 the subspaces X_1 and X_2 are closed as kernels of continuous maps. We can write $x = Px + (I-P)x \in X_1 + X_2$ for each $x \in X$. If $x \in X_1 \cap X_2$, then $Px = 0$ and thus $0 = x - Px = x$. Hence, $X = X_1 \oplus X_2$. The last assertion follows from $\|P\| = \|P^2\| \le \|P\|^2$. □

We next construct a projection for a given direct sum.

Remark 2.17. a) Let X be a normed vector space with $X = X_1 \oplus X_2$. For each $x \in X$ we then have unique vectors $x_1 \in X_1$ and $x_2 \in X_2$ with $x = x_1 + x_2$. Set $Px = x_1$. Then $P : X \to X$ is the unique linear projection with $R(P) = X_1$ and $N(P) = X_2$.

Proof. By assumption, for each $x \in X$ exists vectors $x_1 \in X_1$ and $x_2 \in X_2$ with $x = x_1 + x_2$. If also $\tilde{x}_k \in X_k$ satisfy $x = \tilde{x}_1 + \tilde{x}_2$, then the differences $x_1 - \tilde{x_1} = \tilde{x}_2 - x_2$ belong to $X_1 \cap X_2 = \{0\}$, so that the components $x_k \in X_k$ of x are unique.

The definition of P easily yields the identities $P^2 = P$, $R(P) = X_1$ and $N(P) = X_2$. Let $x, y \in X$ and $\alpha, \beta \in \mathbb{F}$. There are vectors $x_k, y_k \in X_k$ such that $x = x_1 + x_2$ and $y = y_1 + y_2$. We then compute

$$P(\alpha x + \beta y) = P((\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2)) = \alpha x_1 + \beta y_1 = \alpha P x + \beta P y,$$

so that P is linear.

Let also $Q \in L(X)$ satisfy $Q^2 = Q$, $R(Q) = X_1$ and $N(Q) = X_2$. Take $x \in X$ and write $x = x_1 + x_2$ as above. We then have $x_1 = Qy$ for some $y \in X$, and so $Qx = Qx_1 + Qx_2 = Q^2 y = Qy = x_1 = Px$. □

b) Let X be a Banach space. Proposition 4.32 says that the projection P in part a) is continuous and that $X_1 \oplus X_2 \cong X_1 \times X_2$. ♢

We illustrate the above concepts with simple examples.

Example 2.18. a) Let $X = \mathbb{R}^2$, $t \in \mathbb{R}$, and $P = \left(\begin{smallmatrix} 1 & t \\ 0 & 0 \end{smallmatrix}\right)$. Then P is a projection with $R(P) = \mathbb{R} \times \{0\}$, $N(P) = \{(-tr, r) \,|\, r \in \mathbb{R}\}$, and $\|P\| = 1 + |t|$ for $|\cdot|_1$.

b) Let $X = L^p(\mathbb{R})$, $p \in [1, \infty]$, and $Pf = \mathbb{1}_{\mathbb{R}_+} f$ for $f \in X$. Clearly, $\|Pf\|_p \le \|f\|_p$ and $P^2 = P$, so that $P \in \mathcal{B}(X)$ is a projection with $\|P\| = 1$. We further have $(I - P)f = \mathbb{1}_{(-\infty,0]} f$. To express the direct sum $X = R(P) \oplus N(P)$ more conveniently, we introduce the isometric isomorphism $J : R(P) \to L^p(\mathbb{R}_+)$; $Jf = f|_{\mathbb{R}_+}$, whose inverse is given by $J^{-1} g = g$ on $\mathbb{R}_+$ and $J^{-1} g = 0$ on $(-\infty, 0]$. On $\mathbb{R}_-$ one proceeds similarly. We can thus identify X with $L^p(\mathbb{R}_+) \oplus L^p(\mathbb{R}_-)$, considering $L^p(\mathbb{R}_\pm)$ as subspaces of $L^p(\mathbb{R})$ by extending functions by 0.

c) The closed subspace c_0 has no complement in ℓ^∞, see Satz IV.6.5 in [**We**]. ♢

C) Quotient spaces. Let X be a normed vector space, Y a linear subspace and

$$X/Y = \{\hat{x} = x + Y \,|\, x \in X\}$$

be the quotient space. The quotient map

$$Q : X \to X/Y, \quad Qx = \hat{x},$$

is linear and surjective with $N(Q) = Y$. (See Linear Algebra.) One sets $\operatorname{codim} Y = \dim X/Y$. We define the *quotient norm* q by

$$q(x) = \|\hat{x}\| = \|Qx\| := \inf_{y \in Y} \|x - y\| = \mathrm{d}(x, Y).$$

for $\hat{x} = x + Y \in X/Y$. If $\bar{x} + Y = x + Y$, then $\bar{x} - x$ belongs to Y and thus $\mathrm{d}(x, Y) = \mathrm{d}(\bar{x}, Y)$; i.e., $\|\hat{x}\|$ does not depend on the representative of $\hat{x}$. For $\alpha \neq 0$, we have

$$\|\alpha\hat{x}\| = \inf_{y\in Y} \|\alpha(x - \tfrac{1}{\alpha}y)\| = |\alpha| \inf_{z\in Y} \|x - z\| = |\alpha|\,\|\hat{x}\|.$$

Let $x_1, x_2 \in X$. Take $\varepsilon > 0$. There are $y_k \in Y$ mit $\|x_k - y_k\| \leq \|\hat{x}_k\| + \varepsilon$ for $k \in \{1, 2\}$. We then obtain

$$\|\hat{x_1} + \hat{x_2}\| = \inf_{y\in Y} \|x_1 + x_2 - y\| \leq \|x_1 + x_2 - (y_1 + y_2)\| \leq \|\hat{x}_1\| + \|\hat{x}_2\| + 2\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the quotient norm is a seminorm. Because of $\|Qx\| = \|\hat{x}\| \leq \|x\|$, the quotient map Q has norm $\|Q\| \leq 1$.[5]

Now, let Y be closed. If $\|\hat{x}\| = 0$ for some $\hat{x} \in X/Y$, then there exist $y_n \in Y$ with $\|x - y_n\| \to 0$ as $n \to \infty$. From the closedness of Y it follows that x belongs to Y, and hence $\hat{x} = 0$. So far we have established that X/Y is a normed vector space for the quotient norm. Moreover, for each $\delta \in (0, 1)$, Lemma 1.44 gives a vector $\bar{x} \in X$ with $\|\bar{x}\| = 1$ and

$$\|Q\| \geq \|Q\bar{x}\| = \inf_{y\in Y} \|\bar{x} - y\| \geq 1 - \delta.$$

Letting $\delta \to 1$, we deduce that $\|Q\| = 1$.

Proposition 2.19. *Let X be a normed vector space and Y be a linear subspace.*

a) Then X/Y is vector space with seminorm $q : x + Y \mapsto \mathrm{d}(x, Y)$. The map $Q : X \to X/Y$; $Qx = x + Y$, is linear and surjective with $N(Q) = Y$ and $\|Q\| \leq 1$.

b) Let Y be closed. Then q is a norm and $\|Q\| = 1$.

c) Let X be a Banach space and Y be closed. Then X/Y is complete.

Proof.[6] It remains to show the completeness of X/Y. Let $(\hat{x}_n)$ be a Cauchy sequence in X/Y. We find a subsequence such that

$$\|\hat{x}_m - \hat{x}_{n_k}\| \leq 2^{-k} \qquad \text{for all } \; m \geq n_k. \tag{2.3}$$

Hence, there are vectors $y_{n_k} \in Y$ with $\|x_{n_{k+1}} - x_{n_k} - y_{n_k}\| \leq 2 \cdot 2^{-k}$ for every $k \in \mathbb{N}$. Set $z_k = x_{n_{k+1}} - x_{n_k} - y_{n_k}$ and $v_N = x_{n_1} + \sum_{k=1}^N z_k$. Since X is a Banach space and $\sum_{k=1}^\infty \|z_k\| < \infty$, there exists the limit $x = \lim_{N\to\infty} v_N$ in X. (See Lemma 4.23.) We further have

$$v_N = x_{n_{N+1}} - \sum_{k=1}^N y_{n_k} \qquad \text{and} \qquad \sum_{k=1}^N y_{n_k} \in Y,$$

so that $\hat{v}_N = \hat{x}_{n_{N+1}}$. As Q is continuous, the difference $\hat{x}_{n_{N+1}} - \hat{x} = Q(v_N - x)$ tends to 0 in X/Y as $N \to \infty$. For $\varepsilon > 0$ we thus obtain an

[5] Here we use the operator norm also for semi-normed spaces.

[6] This proof was omitted in the lectures.

index $N = N_\varepsilon \in \mathbb{N}$ such that $2^{-N} \leq \varepsilon$ and $\|\hat{x} - \hat{x}_{n_N}\| \leq \varepsilon$. Using also inequality (2.3), we deduce

$$\|\hat{x} - \hat{x}_m\| \leq \|\hat{x} - \hat{x}_{n_N}\| + \|\hat{x}_{n_N} - \hat{x}_m\| \leq 2\varepsilon$$

for all $m \geq n_N$. □

We complement the examples of the preceding subsection.

Example 2.20. a) Let $X = Y \oplus Z$ for a Banach space X. Then the map $J : Z \to X/Y$; $Jz = \hat{z} = z + Y$, is linear and continuous. If $Jz = 0$, then $z \in Y$ and thus $z = 0$. Let $\hat{x} = x + Y \in X/Y$. There are vectors $y \in Y$ and $z \in Z$ with $x = y + z$, so that $\hat{x} = \hat{z} = Jz$. Hence, J is bijective. The continuity of J^{-1} follows from the open mapping theorem 4.28 below. As a result, $Z \simeq X/Y$ via J. In Example 2.18b) we thus obtain $L^p(\mathbb{R})/L^p(\mathbb{R}_+) \simeq L^p(\mathbb{R}_-)$ for $p \in [1, \infty]$.

b) The quotient construction is more general than the direct sum. For instance, ℓ^∞/c_0 exists, though c_0 has no complement in ℓ^∞ by Example 2.18c). ◊

D) Completion. On a Banach space one occasionally considers another weaker incomplete norm. One then wants to pass to a 'larger' Banach space with this norm. This is made precise in the following result. It possesses another much shorter proof which is indicated after Proposition 5.24, see also Korollar III.3.2 in [**We**]. We provide the more constructive proof below since it is useful in certain situations and it is similar to Cantor's construction of $\mathbb{R}$ out of $\mathbb{Q}$.

Proposition 2.21. *Let X be a normed vector space. Then there is a Banach space $\tilde{X}$ and a linear isometry $J : X \to \tilde{X}$ such that JX is dense in $\tilde{X}$. Any other Banach space with this property is isometrically isomorphic to $\tilde{X}$.*

Proof.[7] Let E be the vector space of all Cauchy sequences $v = (x_n)_{n\in\mathbb{N}}$ in X. Note that for $(x_n) \in E$ the sequence $(\|x_n\|)$ is Cauchy in $\mathbb{R}$ and thus the number $p(v) := \lim_{n\to\infty} \|x_n\|$ exists in $\mathbb{R}$. It is easy to check that p is a seminorm on E and that its kernel is given by the linear subspace $c_0(X)$ of all null sequences in X. We now define the vector space $\tilde{X} = E/c_0(X)$ and put $|||\tilde{v}||| = p(v)$ for any representative $v \in E$ of $\tilde{v} \in \tilde{X}$. If $w \in E$ is another representative of $\tilde{v}$, then $v - w \in c_0(X)$ and we thus obtain $p(v) \leq p(v - w) + p(w) = p(w)$ as well as $p(w) \leq p(v)$. Hence, $|||\tilde{v}|||$ is well defined and it gives a norm on $\tilde{X}$. We further introduce the map

$$J : X \to \tilde{X}; \quad x \longmapsto (x, x, \dots) + c_0(X),$$

which is linear and isometric. Let $\tilde{v} \in \tilde{X}$ and $\varepsilon > 0$. Choose a representative $v = (x_k) \in E$. There is an index $N = N_\varepsilon \in \mathbb{N}$ with $\|x_k - x_N\| \leq \varepsilon$

[7]This proof was only sketched in the lectures.

for all $k \geq N$. It follows $|||\tilde{v} - Jx_N||| = \lim_{k\to\infty} \|x_k - x_N\| \leq \varepsilon$, and hence the range of J is dense in $\tilde{X}$.

Let $(\tilde{v}_m)$ be a Cauchy sequence in $\tilde{X}$ with representives $v_m = (x_{m,j})_j \in E$. For each $m \in \mathbb{N}$, we can find an index j_m such that $j_{m+1} > j_m$ and

$$\|x_{m,j} - x_{m,j_m}\| \leq \frac{1}{m} \qquad \text{for all } \ j \geq j_m.$$

We define the diagonal sequence $w = (y_m)_m = (x_{m,j_m})_m$. The above inequality then yields

$$\begin{aligned}\|y_n - y_m\| &= |||Jy_n - Jy_m||| \leq |||Jy_n - \tilde{v}_n||| + |||\tilde{v}_n - \tilde{v}_m||| + |||\tilde{v}_m - Jy_m|||\\ &= \lim_{j\to\infty} \|x_{n,j_n} - x_{n,j}\| + |||\tilde{v}_n - \tilde{v}_m||| + \lim_{j\to\infty} \|x_{m,j} - x_{m,j_m}\|\\ &\leq \frac{1}{n} + |||\tilde{v}_n - \tilde{v}_m\ ||| + \frac{1}{m},\end{aligned}$$

so that $w \in E$. Let $\varepsilon > 0$. Using the above extimate and that $(\tilde{v}_m)$ is a Cauchy sequence in $\tilde{X}$, we find an index $N_\varepsilon \in \mathbb{N}$ such that $\|y_n - y_m\| \leq \varepsilon$ and $1/N_\varepsilon \leq \varepsilon$ for all $n, m \geq N_\varepsilon$. We can thus estimate

$$\begin{aligned}|||\tilde{w} - \tilde{v}_m||| &\leq |||\tilde{w} - Jy_m\ ||| + |||\ Jy_m - \tilde{v}_m|||\\ &= \lim_{n\to\infty} \|y_n - y_m\| + \lim_{j\to\infty} \|x_{m,j_m} - x_{m,j}\| \leq 2\varepsilon\end{aligned}$$

for all $m \geq N_\varepsilon$, so that $\tilde{X}$ is complete.

It remains to prove uniqueness. Let $\tilde{X}'$ be a Banach space and $J' : \ X \to \tilde{X}'$ be isometric and linear with dense range $J'X$ in $\tilde{X}'$. Remark 2.11 shows that the operator

$$T_0 = J \circ (J')^{-1} : \ J'X \to \tilde{X},$$

is well defined and is isometric. It has the dense range JX. Using Lemma 2.13 and the completeness of $\tilde{X}$, we can extend T_0 to an isometric linear map $T : \tilde{X}' \to \tilde{X}$ which still has a dense range. From Remark 2.11 and the completeness of $\tilde{X}'$, we conclude that the range of T is closed, and hence it is equal to $\tilde{X}$. Consequently, T is the required isometric isomorphism. □

Remark 2.22. Usually one identifies X with the subspace JX of $\tilde{X}$ (as one does with $\mathbb{Q}$ and $\mathbb{R}$). Let Y be a Banach space. Every $T \in \mathcal{B}(X, Y)$ can uniquely be extended to an operator $\tilde{T} \in \mathcal{B}(\tilde{X}, Y)$ by means of Lemma 2.13. Also, $\tilde{T}$ is isometric if T is isometric. ◊

Example 2.23. Let $p \in [1, \infty)$. The map $J : (C([0, 1]), \|\cdot\|_p) \to L^p(0, 1)$; $f \mapsto f + \mathcal{N}$, is isometric. Theorem 5.9 of Analysis yields that J has dense range. Using the above remark, we obtain a linear isometric map $\tilde{J} : (C([0, 1]), \|\cdot\|_p)^\sim \to L^p(0, 1)$ with dense range. By Remark 2.11, the range of $\tilde{J}$ is closed and thus $\tilde{J}$ is an isometric isomorphism. In this way one can view $L^p(0, 1)$ as the completion of $C([0, 1])$ with respect to the p–norm. ◊

E) Sum of Banach spaces. Let[8] X and Y be Banach spaces which are linear subspaces of a vector space Z that possesses a metric d which for addition and scalar multiplication of Z are continuous. (We call such a metric *compatible.*) We assume that the inclusion maps from X to Z and from Y to Z are continuous. We then define the sum

$$X + Y = \{z = x + y \,|\, x \in X,\ y \in Y\}$$

which is a linear subspace of Z. We can consider X and Y as linear subspaces of $X+Y$. A typical example is $L^p(\mu)+L^q(\mu)$ for $p, q \in [1, \infty]$ and a measure space $(S, \mathcal{A}, \mu)$, where we may take Z as the space of measurable functions modulo null functions, endowed with the metric describing local convergence in measure. This space will be used in Section 2.3.

We point out that the sum $X + Y$ does not need to be direct, i.e., for a given $z \in X + Y$ there may be many pairs $(x, y) \in X \times Y$ such that $z = x + y$. Moreover, the norm in X does not need to be finer or coarser than that of Y. We endow $X + Y$ with the sum norm

$$\|z\|_{X+Y} = \inf\{\|x\|_X + \|y\|_Y \,|\, z = x + y,\ x \in X,\ y \in Y\}$$

which turns out to be coarser those of X and of Y. Thus, $X + Y$ can serve as a space where we can compare the convergence in X with that in Y. We will further need the linear subspace $D = \{(u, -u) \,|\, u \in X \cap Y\}$ of $X \times Y$.

Proposition 2.24. *Let X and Y be Banach spaces which are linear subspaces of a vector space Z endowed with a compatible metric. We assume that the inclusion maps from X to Z and from Y to Z are continuous. Then $(X+Y, \|\cdot\|_{X+Y})$ is a Banach space which is isometrically isomorphic to the quotient space $(X \times Y)/D$, where $X \times Y$ is endowed with the norm $\|x\|_X + \|y\|_Y$. Moreover, $\|x\|_X \le \|x\|_{X+Y}$ for $x \in X$ and $\|y\|_Y \le \|y\|_{X+Y}$ for $y \in Y$.*

Proof. Let $z \in X + Y$ and $\alpha \in \mathbb{F}$. Note that $\|z\|_{X+Y}$ exists in $[0, \infty)$. If $\|z\|_{X+Y} = 0$, then there are $x_n \in X$ and $y_n \in Y$ such that $z = x_n + y_n$ for all $n \in \mathbb{N}$ and $\|x_n\|_X + \|y_n\|_Y \to 0$ as $n \to \infty$. By continuity, x_n and y_n both tend to 0 in Z, and so $z = 0$ since the metric is compatible. We further have

$$\|\alpha z\|_{X+Y} = \inf\{\|\alpha x\|_X + \|\alpha y\|_Y \,|\, z = x + y,\ x \in X,\ y \in Y\} = |\alpha|\, \|z\|_{X+Y}.$$

Let $z_1, z_2 \in X + Y$. For any $\varepsilon > 0$, we can choose $x_j \in X$ and $y_j \in Y$ such that $z_j = x_j + y_j$ and $\|x_j\|_X + \|y_j\|_Y \le \|z_j\|_{X+Y} + \varepsilon$ for $j \in \{1, 2\}$. Since $z_1 + z_2 = (x_1 + x_2) + (y_1 + y_2)$, we conclude

$$\|z_1 + z_2\|_{X+Y} \le \|x_1 + x_2\|_X + \|y_1 + y_2\|_Y \le \|z_1\|_{X+Y} + \|z_2\|_{X+Y} + 2\varepsilon.$$

[8] This subsection was not containd in the lectures.

As a result, $X+Y$ is a normed vector space. The last assertion is clear. We next show that

$$J : (X \times Y)/D \to X + Y; \quad J((x,y) + D) = x + y,$$

is a linear, isometric and surjective map and that D is closed in $X \times Y$. In view of Remark 2.11 and Proposition 2.19, these facts imply that J is an isometric isomorphism and that $X + Y$ is a Banach space.

First, let $v_n = (x_n, -x_n) \in D$ converge to v in $X \times Y$. Then x_n tends to some x in X and $-x_n$ to some y in Y. Since both sequences also converge in Z, we obtain $y = -x \in X \cap Y$ and $z \in D$; i.e., D is closed in $X \times Y$.

We next treat J. If $(x,y) + D = (x', y') + D$ in $(X \times Y)/D$, then $(x - x', y - y') \in D$ and thus $x - x' = y' - y$. This means that $J((x,y) + D) = x + y = x' + y' = J((x', y') + D)$, and J is in fact a map. It is clear that J is linear and surjective. Let $(x,y) \in X \times Y$ and set $E = (X \times Y)/D$. The operator J is isometric since

$$\begin{aligned}
\|(x,y) + D\|_E &= \inf\{\|(x+u, y-u)\|_{X\times Y} \,|\, u \in X \cap Y\} \\
&= \inf\{\|x+u\|_X + \|y-u\|_Y \,|\, u \in X \cap Y\} \\
&= \inf\{\|x'\|_X + \|y'\|_Y \,|\, x' \in X,\ y' \in Y,\ x' + y' = x + y\} \\
&= \|x+y\|_{X+Y}\,,
\end{aligned}$$

where we take $u = x' - x = y - y'$. □

Proposition 2.25. *Let X_j and Y_j (with $j \in \{0,1\}$) be Banach spaces which are linear subspaces of vector spaces V and W with compatible metrics, respectively. Moreover, these inclusion maps are continuous. Let $T_0 \in \mathcal{B}(X_0, Y_0)$ and $T_1 \in \mathcal{B}(X_1, Y_1)$ be operators such that $T_0 u = T_1 u =: Tu$ for all $u \in X_0 \cap X_1$. Then T has a unique extension $\widetilde{T} \in \mathcal{B}(X_0+X_1, Y_0+Y_1)$ with $\widetilde{T}x_j = T_j x_j$ for all $x_j \in X_j$ and $j \in \{0,1\}$.*

Proof. Let $x = x_0 + x_1$ for $x_j \in X_j$. We then define

$$\widetilde{T}x = T_0 x_0 + T_1 x_1 \in Y_0 + Y_1.$$

If $x = x_0' + x_1'$ for $x_j' \in X_j$, we obtain $u := x_0' - x_0 = x_1 - x_1' \in X_0 \cap X_1$. It follows that

$$T_0 x_0' + T_1 x_1' = T_0 x_0 + T_0 u + T_1 x_1 - T_1 u = \widetilde{T}x + Tu - Tu = \widetilde{T}x,$$

and thus $\widetilde{T} : X_0 + X_1 \to Y_0 + Y_1$ is a map. Clearly, $\widetilde{T}x_j = T_j x_j$ for all $x_j \in X_j$ and $j \in \{0,1\}$. Take any $x_j' \in X_j$ and $\alpha, \beta \in \mathbb{F}$. Set $x' = x_0' + x_1'$. We then compute

$$\begin{aligned}
\widetilde{T}(\alpha x + \beta x') &= T_0(\alpha x_0 + \beta x_0') + T_1(\alpha x_1 + \beta x_1') \\
&= \alpha(T_0 x_0 + T_1 x_1) + \beta(T_0 x_0' + T_1 x_1') = \alpha \widetilde{T}x + \beta \widetilde{T}x',
\end{aligned}$$

so that $\widetilde{T}$ is linear. Moreover,

$$\|\widetilde{T}x\|_{Y_0+Y_1} \le \|T_0 x_0\|_{Y_0} + \|T_1 x_1\|_{Y_1} \le \max\{\|T_0\|,\ \|T_1\|\}\,(\|x_0\| + \|y_1\|).$$

Taking the infimum over all decompositions $x = x_0 + x_1$ in $X_0 + X_1$, we derive that $\widetilde{T}$ is bounded. Let $S \in \mathcal{B}(X_0 + X_1, Y_0 + Y_1)$ be another extension of T_0 and T_1. Then $Sx = Sx_0 + Sx_1 = T_0x_1 + T_1x_1 = \widetilde{T}x$, and $\widetilde{T}$ is unique. □

2.3. The interpolation theorem of Riesz and Thorin

Interpolation[9] theory is an important branch of functional analysis which treats the following problem. Let X_j and Y_j (with $j \in \{0,1\}$) be Banach spaces which are linear subspaces of vector spaces W and Z with compatible metrics, respectively. Assume that $T_0 : X_0 \to Y_0$ and $T_1 : X_1 \to Y_1$ are bounded linear operators such that $T_0u = T_1u =: Tu$ for all $u \in X_0 \cap X_1$. Due to Paragraph 2.2E), we can extend T to a bounded linear operator $\widetilde{T} : X_0 + X_1 \to Y_0 + Y_1$ where the sum space

$$X + Y = \{z = x + y \,|\, x \in X,\ y \in Y\}$$

is endowed with the complete norm

$$\|z\|_{X+Y} = \inf\{\|x\|_X + \|y\|_Y \,|\, z = x + y,\ x \in X,\ y \in Y\}.$$

One now wants to find Banach spaces X between $X_0 \cap X_1$ and $X_0 + X_1$ and Y between $Y_0 \cap Y_1$ and $Y_0 + Y_1$ such that $\widetilde{T}$ can be restricted to a bounded linear map from X to Y which also extends T_0. We refer to the lecture notes [**Lu**] for an introduction to this area and its applications. Here we restrict ourselves to one of the seminal results in this subject due to Riesz and Thorin, which deals with L^p–spaces.

Let $(\Omega, \mathcal{A}, \mu)$ and $(\Lambda, \mathcal{B}, \nu)$ be σ-finite measure spaces and take $p_0, p_1, q_0, q_1 \in [1, \infty]$. Set $U = L^{p_0}(\mu) \cap L^{p_1}(\mu)$ and $V = L^{q_0}(\nu) \cap L^{q_1}(\nu)$. Take $\theta \in [0,1]$ and define

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} \quad \text{and} \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

Observe that every $p \in [p_0, p_1]$ if $p_0 \le p_1$ and every $p \in [p_1, p_0]$ if $p_0 \ge p_1$ can be written in this way. The exponent q between q_0 and q_1 is then fixed via $\theta \in [0,1]$. It is possible that the spaces $L^{p_0}(\mu)$ and $L^{p_1}(\mu)$ are not included in each other. We thus use the sum space $L^{p_0}(\mu) + L^{p_1}(\mu)$ to express that operators on $L^{p_0}(\mu)$ and on $L^{p_1}(\mu)$ are restrictions of common operator.

As seen in Analysis 3, Hölder's inequality shows that $U \subseteq L^p(\mu)$ and $V \subseteq L^q(\nu)$ with the norm bounds

$$\|f\|_p \le \|f\|_{p_0}^{1-\theta}\, \|f\|_{p_1}^{\theta} \quad \text{and} \quad \|g\|_q \le \|g\|_{q_0}^{1-\theta}\, \|g\|_{q_1}^{\theta} \tag{2.4}$$

for all $f \in U$ and $g \in V$. We recall from Theorem 5.9 of Analysis 3 that the space of simple functions

$$E(\mathcal{A}) = \mathrm{lin}\{\mathbb{1}_A \,|\, A \in \mathcal{A},\ \mu(A) < \infty\}$$

[9] This section was not containd in the lectures.

with a support of finite measure is dense in $L^r(\mu)$ if $r \in [1,\infty)$. The proof given there also shows that the space of simple functions

$$E_\infty(\mathcal{A}) = \mathrm{lin}\{\mathbb{1}_A \,|\, A \in \mathcal{A}\}$$

is dense in $L^\infty(\mu)$. Moreover, for $f \in U$ with $p_0 \neq p_1$ we can find a sequence (f_n) in $E(\mathcal{A})$ such that $f_n \to f$ in $L^{p_0}(\mu)$ and in $L^{p_1}(\mu)$, and thus also in $L^p(\mu)$ by (2.4).

We want to show that $L^p(\mu)$ with p as above is embedded into $L^{p_0}(\mu) + L^{p_1}(\mu)$. Let $p_0 \le p_1$. (The other case is treated analogously.) For $f \in L^p(\mu) \setminus \{0\}$, we set $\widetilde{f} = \|f\|_p^{-1} f$ so that $\|\widetilde{f}\|_p = 1$. Because of $\widetilde{f} \in L^p(\mu)$ the set $\{\widetilde{f} \ge 1\}$ has finite measure (if $p < \infty$), and thus the function $f_0 = \mathbb{1}_{\{\widetilde{f}\ge 1\}} \widetilde{f}$ belongs to $L^{p_0}(\mu)$, cf. Proposition 1.35. The function $f_1 = \mathbb{1}_{\{\widetilde{f}<1\}} \widetilde{f}$ is contained in $L^\infty(\mu)$ and hence in $L^{p_1}(\mu)$, see (2.4). The maps $\widetilde{f} = f_0 + f_1$ and f thus belong to $L^{p_0}(\mu) + L^{p_1}(\mu)$. Using $\|\widetilde{f}\|_p = 1$, we further compute

$$\begin{aligned}
\|f\|_{L^{p_0}(\mu)+L^{p_1}(\mu)} &= \|f\|_p \, \|\widetilde{f}\|_{L^{p_0}(\mu)+L^{p_1}(\mu)} \le \|f\|_p \,(\|f_0\|_{p_0} + \|f_1\|_{p_1}) \\
&= \|f\|_p \Big(\int_{\{\widetilde{f}\ge 1\}} |\widetilde{f}|^{p_0} \,\mathrm{d}\mu \Big)^{\frac{1}{p_0}} + \|f\|_p \Big(\int_{\{\widetilde{f}<1\}} |\widetilde{f}|^{p_1} \,\mathrm{d}\mu \Big)^{\frac{1}{p_1}} \\
&\le \|f\|_p \Big(\int_{\{\widetilde{f}\ge 1\}} |\widetilde{f}|^{p} \,\mathrm{d}\mu \Big)^{\frac{1}{p_0}} + \|f\|_p \Big(\int_{\{\widetilde{f}<1\}} |\widetilde{f}|^{p} \,\mathrm{d}\mu \Big)^{\frac{1}{p_1}} \\
&\le 2\, \|f\|_p \,,
\end{aligned}$$

so that $L^p(\mu) \hookrightarrow L^{p_0}(\mu) + L^{p_1}(\mu)$. Similarly, one verifies the embedding $L^q(\nu) \hookrightarrow L^{q_0}(\nu) + L^{q_1}(\nu)$.

Theorem 2.26 (Riesz–Thorin). *Let $(\Omega, \mathcal{A}, \mu)$ and $(\Lambda, \mathcal{B}, \nu)$ be σ-finite measure spaces, $\mathbb{F} = \mathbb{C}$, and $p_0, p_1, q_0, q_1 \in [1,\infty]$. Take $\theta \in [0,1]$ and define $p, q \in [1,\infty]$ via*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} \qquad \text{and} \qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

Assume there are operators $T_j \in \mathcal{B}(L^{p_j}(\mu), L^{q_j}(\nu))$ such that $T_0 u = T_1 u =: Tu$ for all $u \in U = L^{p_0}(\mu) \cap L^{p_1}(\mu)$. Then T has a unique extension $T_\theta \in \mathcal{B}(L^p(\mu), L^q(\nu))$. Moreover, T_θ is the restriction of $\widetilde{T} \in \mathcal{B}(L^{p_0}(\mu) + L^{p_1}(\mu), L^{q_0}(\nu) + L^{q_1}(\nu))$ which is the unique extension of T_0 and T_1 to this space (see Proposition 2.25), and we have

$$\|T_\theta\|_{\mathcal{B}(L^p(\mu),L^q(\nu))} \le \|T_0\|^{1-\theta}_{\mathcal{B}(L^{p_0}(\mu),L^{q_0}(\nu))} \, \|T_1\|^{\theta}_{\mathcal{B}(L^{p_1}(\mu),L^{q_1}(\nu))} \,.$$

We note that the theorem is also true for $\mathbb{F} = \mathbb{R}$ with an additional multiplicative constant in the estimate, see Satz II.4.2 in [**We**].

Proof. The result trivially holds if $\theta \in \{0,1\}$. So we can take $\theta \in (0,1)$. First let $p_0 = p_1$; i.e., $p_0 = p$. Let $f \in L^p(\mu)$. The assumptions then yield $Tf \in V = L^{q_0}(\nu) \cap L^{q_1}(\nu)$ and, using also (2.4),

$$\begin{aligned}\|Tf\|_q &\le \|T_0 f\|_{q_0}^{1-\theta} \, \|T_1 f\|_{q_1}^{\theta} \le \|T_0\|^{1-\theta} \, \|f\|_p^{1-\theta} \, \|T_1\|^{\theta} \, \|f\|_p^{\theta} \\ &= \|T_0\|^{1-\theta} \, \|T_1\|^{\theta} \, \|f\|_p\end{aligned}$$

so that the theorem has been shown in this case.

Next, let $p_0 \neq p_1$. In this case, we have $p < \infty$, and thus $E(\mathcal{A})$ is dense in $L^p(\mu)$. We consider simple functions

$$f = \sum_{j=1}^{m} a_j \, \mathbb{1}_{A_j} \in E(\mathcal{A}) \subseteq U = L^{p_0}(\mu) \cap L^{p_1}(\mu),$$

where we may assume that the sets A_j are pairwise disjoint and have finite measure. The assumptions again yield $Tf \in V$. By equation (5.5) we have

$$\|Tf\|_q = \sup_{g \in E(\mathcal{B}), \|g\|_{q'} \le 1} \Big| \int_\Lambda (Tf) \, g \, \mathrm{d}\nu \Big| \tag{2.5}$$

if $q' < \infty$. If $q' = \infty$ (i.e., $q = 1$ which is equivalent to $(q_0, q_1) = (1, 1)$), then (2.5) is valid with $E(\mathcal{B})$ replaced by $E_\infty(\mathcal{B})$. Further, take $g = \sum_{k=1}^n b_k \, \mathbb{1}_{B_k} \in E(\mathcal{B})$ with $\|g\|_{q'} \le 1$ where we may assume that the sets B_k are pairwise disjoint and have finite measure. (If $q' = \infty$, we allow for $\mu(B_k) = \infty$.) Let $z \in S := \{\zeta \in \mathbb{C} \mid \operatorname{Re} \zeta \in [0, 1]\}$. We then define the function $F(z)$ by

$$(F(z))(\omega) = |f(\omega)|^{p\left(\frac{1-z}{p_0} + \frac{z}{p_1}\right) - 1} f(\omega)$$

for $\omega \in \Omega$ with $f(\omega) \neq 0$, and by $(F(z))(\omega) = 0$ if $f(\omega) = 0$. For $q' < \infty$, we set

$$(G(z))(\lambda) = |g(\lambda)|^{q'\left(\frac{1-z}{q_0'} + \frac{z}{q_1'}\right) - 1} g(\lambda)$$

for $\lambda \in \Lambda$ with $g(\lambda) \neq 0$, and put $(G(z))(\lambda) = 0$ if $g(\lambda) = 0$. If $q' = \infty$, we simply take $G(z) = g$. We write $p(z) = p(\frac{1-z}{p_0} + \frac{z}{p_1}) - 1$ and $q(z) = q'(\frac{1-z}{q_0'} + \frac{z}{q_1'}) - 1$. Observe that $F(\theta) = f$ and $G(\theta) = g$, as well as $F(z) \in E(\mathcal{A})$, $G(z) \in E(\mathcal{B})$ if $q' < \infty$ and $G(z) \in E_\infty(\mathcal{B})$ if $q' = \infty$. The assumptions lead to $TF(z) \in V$. We further introduce the function

$$\varphi(z) = \int_\Lambda T(F(z)) \, G(z) \, \mathrm{d}\nu = \sum_{j=1}^{m} \sum_{k=1}^{n} a_j |a_j|^{p(z)} \, b_k |b_k|^{q(z)} \int_\Lambda (T \mathbb{1}_{A_j}) \, \mathbb{1}_{B_k} \, \mathrm{d}\nu$$

for $z \in S$. Observe that $\varphi \in C(S)$ is holomorphic on S°.

We want to apply the Three-Lines-Theorem, see Satz II.4.3 in [**We**], and thus check the estimates assumed in this result. Writing $z = s + \mathrm{i}t \in S$ with $s \in [0, 1]$ and $t \in \mathbb{R}$, we compute

$$|F(z)(\omega)| = |f(\omega)|^{p\left(\frac{1-s}{p_0} + \frac{s}{p_1}\right)}, \qquad |G(z)(\lambda)| = |g(\lambda)|^{q'\left(\frac{1-s}{q_0'} + \frac{s}{q_1'}\right)} \tag{2.6}$$

for all $s \in [0,1]$, $t \in \mathbb{R}$, $\omega \in \Omega$ and $\lambda \in \Lambda$. The right hand sides are bounded in $z \in S$ for fixed ω and λ. In the same way one sees that φ is bounded on S. The Three–Lines-Theorem then yields

$$\Big| \int_\Lambda (Tf) g \,\mathrm{d}\nu \Big| = |\varphi(\theta)| \le \Big(\sup_{t\in\mathbb{R}} |\varphi(\mathrm{i}t)| \Big)^{1-\theta} \Big(\sup_{t\in\mathbb{R}} |\varphi(1+\mathrm{i}t)| \Big)^{\theta} \tag{2.7}$$

Hölder's inequality, the assumptions and estimate (2.6) with $s = 0$ further imply

$$\begin{aligned}
|\varphi(\mathrm{i}t)| &\le \|T_0 F(\mathrm{i}t)\|_{q_0} \|G(\mathrm{i}t)\|_{q_0'} \le \|T_0\| \, \|F(\mathrm{i}t)\|_{p_0} \|G(\mathrm{i}t)\|_{q_0'} \\
&= \|T_0\| \Big(\int_\Omega |f|^{\frac{pp_0}{p_0}} \,\mathrm{d}\mu \Big)^{\frac{1}{p_0}} \Big(\int_\Lambda |g|^{\frac{q' q_0'}{q_0'}} \,\mathrm{d}\nu \Big)^{\frac{1}{q_0'}} \\
&= \|T_0\| \, \|f\|_p^{\frac{p}{p_0}} \|g\|_{q'}^{\frac{q'}{q_0'}} \le \|T_0\| \, \|f\|_p^{\frac{p}{p_0}},
\end{aligned}$$

where we also used that $\|g\|_{q'} \le 1$. Similarly one sees that

$$\begin{aligned}
|\varphi(1+\mathrm{i}t)| &\le \|T_1 F(1+\mathrm{i}t)\|_{q_1} \|G(1+it)\|_{q_1'} \\
&\le \|T_1\| \, \|F(1+\mathrm{i}t)\|_{p_1} \|G(1+it)\|_{q_1'} \\
&= \|T_1\| \Big(\int_\Omega |f|^{\frac{pp_1}{p_1}} \,\mathrm{d}\mu \Big)^{\frac{1}{p_1}} \Big(\int_\Lambda |g|^{\frac{q' q_1'}{q_1'}} \,\mathrm{d}\nu \Big)^{\frac{1}{q_1'}} \\
&= \|T_1\| \, \|f\|_p^{\frac{p}{p_1}} \|g\|_{q'}^{\frac{q'}{q_1'}} \le \|T_1\| \, \|f\|_p^{\frac{p}{p_1}}.
\end{aligned}$$

Formulas (2.5) and (2.7) now lead to

$$\|Tf\|_q \le \|T_0\|^{1-\theta} \|f\|_p^{p(1-\theta)/p_0} \|T_1\|^{\theta} \|f\|_p^{p\theta/p_1} = \|T_0\|^{1-\theta} \|T_1\|^{\theta} \|f\|_p \tag{2.8}$$

for all $f \in E(\mathcal{A})$. Let $f \in U$. As observed above the theorem, we can approximate f by $f_n \in E(\mathcal{A})$ in $L^{p_0}(\mu)$, in $L^{p_1}(\mu)$ and in $L^p(\mu)$. By the assumptions, also Tf_n tends to Tf in $L^{q_0}(\nu)$ and in $L^{q_1}(\nu)$, and hence in $L^q(\nu)$ due to (2.4). Inequality (2.8) thus holds for all $f \in U$. Lemma 2.13 then allows to extend T uniquely from its domain U to an operator $T_\theta \in \mathcal{B}(L^p(\mu), L^q(\mu))$ with norm less or equal $\|T_0\|^{1-\theta} \|T_1\|^{\theta}$. As observed before the theorem, we have $L^p(\mu) \hookrightarrow L^{p_0}(\mu) + L^{p_1}(\mu)$ and similarly for the range spaces. Hence, $Tf_n = \widetilde{T} f_n$ tends both to $T_\theta f$ and $\widetilde{T} f$ in $L^{q_0}(\nu) + L^{q_1}(\nu)$ so that T_θ is an restriction of $\widetilde{T}$. □

We next use the Riesz–Thorin theorem to give a different proof of Young's inequality for convolutions from Theorem 2.14.

1a) Recall the definition (2.1) of the convolution $f * g$ for $f, g \in L^1(\mathbb{R}^m)$ and that $\varphi(x,y) = |f(x-y)g(y)|$. There we have shown that

$$\|f * g\|_1 \le \|\varphi\|_1 = \|f\|_1 \, \|g\|_1.$$

1b) In a second step, we take $f \in L^1(\mathbb{R}^m)$ and $g \in L^\infty(\mathbb{R}^m)$. We compute

$$\int_{B(0,n)\times\mathbb{R}^d} \varphi(x,y)\,\mathrm{d}(x,y) \le \int_{B(0,n)} \int_{\mathbb{R}^m} |f(x-y)|\,\|g\|_\infty\,\mathrm{d}y\,\mathrm{d}x$$
$$= \|g\|_\infty \int_{B(0,n)} \int_{\mathbb{R}^m} |f(z)|\,\mathrm{d}z\,\mathrm{d}x = \lambda(B(0,n))\,\|f\|_1\,\|g\|_\infty.$$

Fubini's theorem now yields that $(f*g)(x)$ is defined for a.e. $x \in B(0,n)$ and gives a measurable function on $B(0,n)$. Letting $n \to \infty$, the same holds on $\mathbb{R}^m$. Replacing in the above estimate the integral over $x \in B(0,n)$ by a supremum in $x \in \mathbb{R}^m$, we further obtain

$$\|f * g\|_\infty \le \|f\|_1\,\|g\|_\infty.$$

1c) Fix any $f \in L^1(\mathbb{R}^m)$. We define $T_1 g = f * g$ for $g \in L^1(\mathbb{R}^m)$ and $T_\infty g = f * g$ for $g \in L^\infty(\mathbb{R}^m)$. We have shown that that $T_r \in \mathcal{B}(L^r(\mathbb{R}^m))$ with $\|T_r\| \le \|f\|_1$ for $r \in \{1,\infty\}$. We can now extend the convolution to an operator $Tg = f * g := f * (g_1 + g_\infty)$ for $g = g_1 + g_\infty \in L^1(\mathbb{R}^m) + L^\infty(\mathbb{R}^m)$. Let $q \in (1,\infty)$. Set $\theta = 1/q' \in (0,1)$, so that $\frac{1}{q} = \frac{1-\theta}{1} + \frac{\theta}{\infty}$. The Riesz–Thorin theorem allows us to restrict T to a bounded operator $T_q \in \mathcal{B}(L^q(\mathbb{R}^m))$ with $\|T_q\| \le \|f\|_1$. For all $f \in L^1(\mathbb{R}^m)$, $g \in L^q(\mathbb{R}^m)$ and $q \in [1,\infty]$, we have thus shown that $f * g \in L^q(\mathbb{R}^m)$ and

$$\|f * g\|_q \le \|f\|_1\,\|g\|_q.$$

2a) We fix $g \in L^q(\mathbb{R}^m)$ and $q \in [1,\infty]$, and vary f. For $f \in L^1(\mathbb{R}^m)$, step 1c) yields the bounded linear operator $S_1 : L^1(\mathbb{R}^m) \to L^q(\mathbb{R}^m)$; $S_1 f = f * g$, with norm $\|S_1\| \le \|g\|_q$. Let now $f \in L^{q'}(\mathbb{R}^m)$. Due to Hölder's estimate, the map $y \mapsto f(x-y)g(y)$ is integrable on $\mathbb{R}^m$ and

$$\Big|\int_{\mathbb{R}^m} f(x-y)g(y)\,\mathrm{d}y\Big| \le \|f(x-\cdot)\|_{q'}\,\|g\|_q = \|f\|_{q'}\,\|g\|_q$$

for each $x \in \mathbb{R}^m$. One sees as above that $f * g =: S_{q'} f$ is a measurable function and

$$\|S_{q'} f\|_\infty = \|f * g\|_\infty \le \|f\|_{q'}\,\|g\|_q\,;$$

i.e., $S_{q'}$ belongs to $\mathcal{B}(L^{q'}(\mathbb{R}^m), L^\infty(\mathbb{R}^m))$ with norm less or equal $\|g\|_q$. We can thus define the convolution $Sf = f * g$ for all $g \in L^q(\mathbb{R}^m)$, $f \in L^1(\mathbb{R}^m) + L^{q'}(\mathbb{R}^m)$ and $q \in [1,\infty]$.

2b) Finally, take $p \in [1,q']$ and $r \in [1,\infty]$ with $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Choose $\theta = q/p' \in [0,1]$. Observe that

$$\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{q'} \quad \text{and} \quad \frac{1}{r} = \frac{1-\theta}{q} + \frac{\theta}{\infty}.$$

By means of the Riesz–Thorin theorem we then restrict S to an operator $S_p \in \mathcal{B}(L^p(\mathbb{R}^m), L^r(\mathbb{R}^m))$ with norm less or equal $\|g\|_q$. In this way we have proved Theorem 2.14.

CHAPTER 3

Hilbert spaces

So far we can only treat lengths of vectors or their distance by means of the norm. In $\mathbb{F}^m$ one also uses orthogonality (or angles) in a crucial way. The relevant concepts are introduced below in our setting. We will see that the resulting 'Hilbert spaces' inherit much more structure of the finite dimensional case as a general Banach space.

3.1. Basic properties and orthogonality

DEFINITION 3.1. *A* scalar product *on a vector space X is a map $(\cdot|\cdot) : X^2 \to \mathbb{F}$ possessing the properties*

a) $(\alpha x + \beta z|y) = \alpha\,(x|y) + \beta\,(z|y)$,

b) $(x|y) = \overline{(y|x)}$,

c) $(x|x) \geq 0, \qquad (x|x) = 0 \iff x = 0$,

for all $x, y, z \in X$ and $\alpha, \beta \in \mathbb{F}$. The map is called a sesquilinear form *if it fulfills a) and b), as well as* positive definite *if c) is valid. The pair $(X, (\cdot|\cdot))$, or simply X, is said to be a* Pre-Hilbert space.[1] *We set $\|x\| = \sqrt{(x|x)}$, and call it a* Hilbert norm.

We start with several simple observations which will often be used.

REMARK 3.2. Let $(X, (\cdot|\cdot))$ be a Pre-Hilbert space.

a) Properties a) and b) in Definition 3.1 easily yield the relations $(0|y) = 0$, $(x|0) = 0$, $(x|x) \in \mathbb{R}$, and

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

for all $\alpha, \beta \in \mathbb{F}$ and $x, y, z \in X$.

b) We have the *Cauchy–Schwarz inequality*

$$|(x|y)| \leq \|x\|\,\|y\| \qquad \text{for all } x, y \in X. \tag{CS}$$

Here one obtains equality if and only if x and y are linearly dependent. See Linear Algebra or Satz V.1.2 in [**We**].

c) The Hilbert norm in Definition 3.1 is indeed a norm on X.

PROOF. Let $x, y \in X$ and $\alpha \in \mathbb{F}$. Definition 3.1 and (CS) yield

$$\|x\| = 0 \iff (x|x) = 0 \iff x = 0,$$

$$\|\alpha x\| = \sqrt{\alpha\bar{\alpha}\,(x|x)} = |\alpha|\,\|x\|,$$

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

[1] One also uses the notions 'inner product' and 'inner product space'.

$$= \|x\|^2 + 2\,\mathrm{Re}\,(x|y) + \|y\|^2 \tag{3.1}$$
$$\leq \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \qquad \square$$

d) The scalar product $(\cdot|\cdot) : X^2 \to \mathbb{F}$ is Lipschitz on each ball of X^2 and thus continuous.

PROOF. Using (CS), we estimate

$$\begin{aligned} |(x_1|y_1) - (x_2|y_2)| &\leq |(x_1 - x_2|y_1)| + |(x_2|y_1 - y_2)| \\ &\leq r\,\|x_1 - x_2\| + r\,\|y_1 - y_2\| \\ &\leq \sqrt{2}r\,\|(x_1, y_1) - (x_2, y_2)\| \end{aligned}$$

for all $x_k, y_k \in X$ with $\|(x_k, y_k)\| := (\|x_k\|^2 + \|y_k\|^2)^{\frac{1}{2}} \leq r$. $\square$

e) From (3.1) we deduce the *parallelogramm identity*

$$\begin{aligned} \|x+y\|^2 + \|x-y\|^2 &= \|x\|^2 + 2\,\mathrm{Re}\,(x|y) + \|y\|^2 + \|x\|^2 - 2\,\mathrm{Re}\,(x|y) + \|y\|^2 \\ &= 2\,\|x\|^2 + 2\,\|y\|^2 \end{aligned} \tag{3.2}$$

for all $x, y \in X$.

f) In view of Definition 3.1, a linear subspace of a Pre-Hilbert space is a Pre-Hilbert space with the restricted scalar product. $\Diamond$

For the deeper properties of $(X, (\cdot|\cdot))$ we need completeness.

DEFINITION 3.3. *Let* $(\cdot|\cdot)$ *be a scalar product on* X. *If the Hilbert norm* $\|\cdot\|$ *is complete, then* $(X, (\cdot|\cdot))$ *is called a* Hilbert space.

We discuss the basic examples of Hilbert spaces, see also Section 4.2 and the exercises.

EXAMPLE 3.4. a) On $X = \mathbb{F}^m$ we have the Euclidean scalar product

$$(x|y) = \sum_{k=1}^{m} x_k \bar{y}_k \qquad \text{with norm} \qquad \|x\|_2^2 = \sum_{k=1}^{m} |x_k|^2$$

for $x, y \in \mathbb{F}^m$. The pair $(\mathbb{F}^m, (\cdot|\cdot))$ is a Hilbert space by Example 1.4.

Note that $|\cdot|_p$ on $\mathbb{F}^m$ is not induced by a scalar product if $m \geq 2$ and $p \neq 2$, since $|e_1 + e_2|_p^2 + |e_1 - e_2|_p^2 = 2^{2/p} + 2^{2/p} \neq 2(|e_1|_p^2 + |e_2|_p^2) = 4$ contradicting (3.2).

b) Similarly, $X = \ell^2$ is a Hilbert space with the scalar product

$$(x|y) = \sum_{k=1}^{\infty} x_k \bar{y}_k \qquad \text{and the norm} \qquad \|x\|_2^2 = \sum_{k=1}^{\infty} |x_k|^2,$$

cf. Proposition 1.30. Note that the first series converges absolutely because of Hölder's inequality with $p = p' = 2$.

c) Let $(S, \mathcal{A}, \mu)$ be a measure space. Hölder's inequality with $p = p' = 2$ and basic properties of the integral imply that the space $X = L^2(\mu)$

possesses the scalar product

$$(f|g) = \int_S f\overline{g}\,\mathrm{d}\mu \qquad \text{with norm} \qquad \|f\|_2^2 = \int_S |f|^2\,\mathrm{d}\mu.$$

It is a Hilbert space by Theorem 5.5 in Analysis 3. ♢

In Pre-Hilbert spaces one can define angles, where we restrict ourselves to the angle $\pi/2$.

DEFINITION 3.5. *Two elements x and y of a Pre-Hilbert space X are called* orthogonal *if $(x|y) = 0$. Two non-empty subsets $A, B \subseteq X$ are called* orthogonal *if $(a|b) = 0$ for all $a \in A$ and $b \in B$. One then writes $x \perp y$ respectively $A \perp B$, and also $x \perp A$ instead of $\{x\} \perp A$. The* orthogonal complement *of A is given by*

$$A^\perp = \{x \in X \,|\, x \perp a \ \text{ for every } \ a \in A\}.$$

A projection $P \in L(X)$ is called orthogonal *if $R(P) \perp N(P)$.*

We discuss a few typical examples for orthogonal vectors.

EXAMPLE 3.6. a) Let $(S, \mathcal{A}, \mu)$ be a measure space. Functions $f, g \in L^2(\mu)$ with disjoint support (up to null sets) are orthogonal since then $(f|g) = \int_S f\overline{g}\,\mathrm{d}\mu = 0$.

b) Set $f_n(s) = \mathrm{e}^{\mathrm{i}ns}$ for $s \in [0, 2\pi]$ and $n \in \mathbb{Z}$. If $n \neq m$ the functions f_n and f_m are orthogonal in $L^2(0, 2\pi)$ because of

$$(f_n|f_m) = \int_0^{2\pi} \mathrm{e}^{\mathrm{i}ns}\mathrm{e}^{-\mathrm{i}ms}\,\mathrm{d}s = \frac{1}{\mathrm{i}(n-m)}\,\mathrm{e}^{\mathrm{i}(n-m)s}\Big|_0^{2\pi} = 0.$$

Here the orthogonality is caused by oscillations and not by a disjoint support as in a). (Compare $(1, 0) \perp (0, 1)$ and $(1, -1) \perp (1, 1)$ in $\mathbb{R}^2$.)

c) Let $f \in L^2(\mathbb{R})$ be even and $g \in L^2(\mathbb{R})$ be odd. Then $f \perp g$ since

$$\begin{aligned}(f|g) &= \int_{-\infty}^0 f(s)\overline{g}(s)\,\mathrm{d}s + \int_0^\infty f(s)\overline{g}(s)\,\mathrm{d}s \\ &= \int_0^\infty f(-t)\overline{g}(-t)\,\mathrm{d}t + \int_0^\infty f(s)\overline{g}(s)\,\mathrm{d}s = 0,\end{aligned}$$

where we substituted $s = -t$ in the first integral. ♢

We now collect various properties of orthogonality which are often employed in these lectures.

REMARK 3.7. Let X be a Pre-Hilbert space, $A, B \subseteq X$ be non-empty, and $x, y \in X$. The next assertions follow mostly from the above definitions.

a) We have $x \perp x$ if and only if $x = 0$; and thus $X^\perp = \{0\}$. Moreover, $x \perp y$ is equivalent to $y \perp x$. Observe that $\{0\}^\perp = X$.

b) Let $x \perp y$. Then equation (3.1) yields *Pythagoras'* identity

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

c) Part a) implies the relations $A \cap A^\perp \subseteq \{0\}$ and $A \subseteq (A^\perp)^\perp =: A^{\perp\perp}$.

d) It is easy to see that $A^\perp$ is a linear subspace of X. It is also closed. In fact, let (x_n) in $A^\perp$ tend to x in X. Since the scalar product is continuous (see Remark 3.2), we obtain $(x|a) = \lim_{n\to\infty} (x_n|a) = 0$ for all $a \in A$. Hence, $x \perp A$ and $A^\perp$ is closed.

e) If $A \subseteq B$, then $B^\perp \subseteq A^\perp$. As in d) one sees that $A^\perp = (\overline{\operatorname{lin} A})^\perp$.

f) If $(x|z) = (y|z)$ for some $x, y \in X$ and all z from a dense subset $D \subseteq X$, then $x = y$. Indeed, there are vectors z_n in D converging to $x - y$ in X. The continuity of the scalar product then yields

$$\|x-y\|^2 = (x-y|x-y) = \lim_{n\to\infty} (x-y|z_n) = 0. \qquad \Diamond$$

Many of the special properties of Hilbert spaces rely on the following *projection theorem.*

Theorem 3.8. *Let X be a Hilbert space and $Y \subseteq X$ be a closed linear subspace. Then there is a unique orthogonal projection $P \in \mathcal{B}(X)$ with $R(P) = Y$ and $N(P) = Y^\perp$. It satisfies $\|P\| = 1$ if $Y \neq \{0\}$ and*

$$\|x - Px\| = \inf_{y\in Y} \|x-y\| \qquad \text{for every } \; x \in X.$$

We have the decomposition $X = Y \oplus Y^\perp$ with $Y^{\perp\perp} = Y$ and $X/Y \cong Y^\perp$.

Given $x \in X$, the vector Px has the minimal distance to x within elements of Y, and the difference $x - Px \in N(P)$ is orthogonal to Y. Hilbert spaces thus inherit these basic geometric facts from $\mathbb{F}^m$.

Proof. 1) Due to Remark 3.7, $Y^\perp$ is a closed linear subspace of X and $Y \cap Y^\perp = \{0\}$. Let $x \in X$. To construct Px, we look for a vector y_x in Y satisfying $\|x - y_x\| = \inf_{y\in Y} \|x-y\| =: \delta$. There are elements y_n of Y with $\|x - y_n\| \to \delta$ as $n \to \infty$. From (3.2) and $\frac{1}{2}(y_n + y_m) \in Y$ we deduce the limit

$$\begin{aligned} 0 &\le \|\tfrac{1}{2}(y_n - y_m)\|^2 = \|\tfrac{1}{2}(y_n - x) - \tfrac{1}{2}(y_m - x)\|^2 \\ &= \tfrac{1}{2}\|y_n - x\|^2 + \tfrac{1}{2}\|y_m - x\|^2 - \|\tfrac{1}{2}(y_n + y_m) - x\|^2 \\ &\le \tfrac{1}{2}\|y_n - x\|^2 + \tfrac{1}{2}\|y_m - x\|^2 - \delta^2 \;\longrightarrow\; 0 \qquad \text{as } n, m \to \infty. \end{aligned}$$

Since X is complete and Y is closed, there exists the point $y_x = \lim_{n\to\infty} y_n$ in Y, and we obtain $\|y_x - x\| = \lim_{n\to\infty} \|y_n - x\| = \delta$.

2) Set $z_x := x - y_x$. We want to show that $z_x \perp Y$. Take $w \in Y \setminus \{0\}$ and put $\alpha = \|w\|^{-2}(z_x|w)$. The linear combination $\alpha w + y_x$ belongs to Y. We thus infer

$$\begin{aligned} \delta^2 &\le \|x - (\alpha w + y_x)\|^2 = \|z_x - \alpha w\|^2 \\ &= \|z_x\|^2 - 2\operatorname{Re} \overline{\alpha}(z_x|w) + |\alpha|^2 \|w\|^2 = \delta^2 - \frac{|(z_x|w)|^2}{\|w\|^2}, \end{aligned}$$

using (3.1) as well as the definition of z_x and α. It follows that $(z_x|w) = 0$ and thus $z_x \perp Y$. Taking into account $x = y_x + z_x$, we have shown the equality $Y + Y^\perp = X$, and hence $X = Y \oplus Y^\perp$.

3) Remark 2.17 gives a unique operator $P = P^2 \in L(X)$ with $R(P) = Y$ and $N(P) = Y^\perp$, where $Px = y_x$. Pythagoras further implies that $\|Px\|^2 \le \|y_x\|^2 + \|z_x\|^2 = \|x\|^2$ since $y_x \perp z_x$; i.e., $P \in \mathcal{B}(X)$ and $\|P\| \le 1$. Lemma 2.16 then yields $\|P\| = 1$ if $Y \neq \{0\}$.

The inclusion $Y \subseteq Y^{\perp\perp}$ follows from Remark 3.7. Let $x \perp Y^\perp$. We then compute $0 = (x|z_x) = (y_x + z_x|z_x) = \|z_x\|^2$ employing step 2). Hence, the vector $x = y_x$ belongs to Y so that $Y = Y^{\perp\perp}$. Example 2.20 further implies the isomorphy $Y^\perp \cong X/Y$. □

We illustrate the above theorem by a few basic examples.

Example 3.9. a) Let $X = \mathbb{F}^2$ and $Y = \mathbb{F} \times \{0\}$. Then $Y^\perp = \{0\} \times \mathbb{F}$ and $P = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, see Example 2.18.

b) Let $X = L^2(\mathbb{R})$ and $Y = \{f \in X \,|\, f = 0 \text{ a.e. on } \mathbb{R}_-\} \cong L^2(\mathbb{R}_+)$. With Theorem 3.8 and Example 2.18 we obtain the isomorphies $Y^\perp \cong X/Y \cong \{f \in X \,|\, f = 0 \text{ a.e. on } \mathbb{R}_+\} \cong L^2(\mathbb{R}_-)$, using $Pf = \mathbb{1}_{\mathbb{R}_+} f$.

c) Let $X = L^2(\mathbb{R})$ and $Y = \{f \in X \,|\, f(s) = f(-s) \text{ for a.e. } s \in \mathbb{R}\}$ be the set of even functions in X. As in Example 1.33 one sees that the linear subspace Y is closed in X. Let $f \in X$, and set $Pf(s) = \frac{1}{2}\,(f(s) + f(-s))$ for a.e. $s \in \mathbb{R}$. The map Pf belongs to Y, $\|Pf\|_2 \le \|f\|_2$, and we have $Pf = f$ if $f \in Y$. As a result, $P \in \mathcal{B}(X)$ is a projection with $\|P\| = 1$ and $R(P) = Y$. Its kernel $N(P) = \{f \in L^2(\mathbb{R}) \,|\, f(s) = -f(-s) \text{ for a.e. } s \in \mathbb{R}\}$ is the space of odd functions in X, so that the projection P is orthogonal by Example 3.6.

d) Let $A \neq \emptyset$ be a subset of a Hilbert space X. Set $Y = \overline{\text{lin}\, A}$. Remark 3.7 and Theorem 3.8 yield the identity $A^{\perp\perp} = Y^{\perp\perp} = Y$. ◊

Let X be a Hilbert space. For each fixed $y \in X$ we define the function

$$\Phi(y) : X \to \mathbb{F}; \quad x \mapsto (x|y)\,. \tag{3.3}$$

The map $\Phi(y)$ is linear and satisfies $|\Phi(y)(x)| \le \|x\|\,\|y\|$ by (CS) for all $x, y \in X$. Hence, $\Phi(y)$ is an element of $X^\star$ with $\|\Phi(y)\|_{X^\star} \le \|y\|_X$. The next important representation theorem by *F. Riesz* says that the resulting operator $\Phi_X = \Phi : X \to X^\star$ is isometric and bijective. This seemingly very abstract fact is a very powerful tool to solve (linear) partial differential equations.

Theorem 3.10. *Let X be a Hilbert space and define the map $\Phi_X = \Phi : X \to X^\star$ by (3.3). Then Φ is bijective, isometric and antilinear.*[2] *Let $x^\star \in X^\star$. Then $y = \Phi^{-1}(x^\star)$ is the unique element of X fulfilling $\langle x, x^\star \rangle = (x|y)$ for all $x \in X$, and we have $\|y\|_X = \|x^\star\|_{X^\star}$.*

[2] This means that $\Phi(\alpha x + \beta y) = \bar{\alpha}\Phi(y) + \bar{\beta}\Phi(y)$ for all $\alpha, \beta \in \mathbb{F}$ and $x, y \in X$.

Proof. Equation (3.3) and Remark 3.2 imply that Φ is antilinear. Let $y \in X \setminus \{0\}$, and set $x = \frac{1}{\|y\|} y$. Then $\|x\| = 1$, and we thus obtain

$$\|\Phi(y)\|_{X^\star} \geq |\langle x, \Phi(y)\rangle| = \frac{1}{\|y\|}\,(y|y) = \|y\|.$$

Hence, $\|\Phi(y)\|_{X^\star} = \|y\|_X$ and thus Φ is isometric and injective.

To show surjectivity, fix $\varphi \in X^\star \setminus \{0\}$. Then $Z := N(\varphi) \neq X$ is a closed linear subspace of X. Theorem 3.8 yields that $Z^\perp \neq \{0\}$. We take any $y_0 \in Z^\perp \setminus \{0\}$ and set $y_1 = \varphi(y_0)^{-1} y_0 \in Z^\perp \setminus \{0\}$. Let $x \in X$. We calculate $\varphi(x - \varphi(x)y_1) = \varphi(x) - \varphi(x)\varphi(y_1) = 0$ using that $\varphi(y_1) = 1$. As a result, $x - \varphi(x)y_1$ belongs to Z leading to

$$0 = (x - \varphi(x)y_1|y_1) = (x|y_1) - \varphi(x)\,\|y_1\|^2,$$
$$\varphi(x) = (x|\,\|y_1\|^{-2}y_1) \qquad \text{for all} \;\; x \in X.$$

We have shown that $\varphi \in R(\Phi)$ and so Φ is bijective. The other assertions easily follow. $\square$

Usually, one identifies a Hilbert space X with its dual $X^\star$; i.e., one omits the Riesz isomorphism Φ_X in the notation. We stress that this can only be done for one Hilbert space at the same time.

3.2. Orthonormal bases

In this section we extend the concept of the Euclidean basis in $\mathbb{F}^m$ to the setting of (separable) Hilbert spaces X. We only have to use series instead of finite sums if $\dim X = \infty$. Except for the definition, we restrict ourselves to countable bases to simplify the presentation a bit. This will lead to additional separability assumptions at the end of the section. Actually, it is not difficult to remove these restrictions, see Section V.4 in [**We**].

Definition 3.11. *Let X be a Pre-Hilbert space. A non-empty subset S of X is an* orthonormal system *if $\|v\| = 1$ and $(v|w) = 0$ for all $v, w \in S$ with $v \neq w$. Let B be a orthonormal system. It is called* orthonormal basis *if it is maximal; i.e., if S is another orthonormal system in X with $B \subseteq S$, then we already have $B = S$.*

We first state the most basic examples.

Example 3.12. a) The set $S = \{e_n \,|\, n \in \mathbb{N}\}$ is orthonormal in ℓ^2.

b) Let $X = L^2(0, 2\pi)$ with $\mathbb{F} = \mathbb{C}$. We put $v_k(s) = \frac{1}{\sqrt{2\pi}} e^{iks}$ for $s \in [0, 2\pi]$ and $k \in \mathbb{Z}$. The set $S = \{v_k \,|\, k \in \mathbb{Z}\}$ is orthonormal by Example 3.6 and $\|v_k\|_2^2 = \frac{1}{2\pi}\int_0^{2\pi} 1 \,\mathrm{d}s = 1$.

c) In $X = L^2(0, 2\pi)$ the set $S = \{\frac{1}{\sqrt{2\pi}}\mathbb{1}, \frac{1}{\sqrt{\pi}}\cos(n\cdot), \frac{1}{\sqrt{\pi}}\sin(n\cdot) \,|\, n \in \mathbb{N}\}$ is orthonormal, cf. Korollar 5.16 in Analysis 3. $\Diamond$

We recall the *Gram–Schmidt procedure* from Linear Algebra which allows us to construct a orthonormal system with the same linear span as a given linearly independent, at most countable subset. (See Satz V.4.2 of [**We**] for a proof.)

Lemma 3.13. *Let X be a Pre-Hilbert space and $\Sigma = \{x_j \,|\, j \in J\}$ be a linearly independent subset of X with $J = \{1, \dots, N\}$ or $J = \mathbb{N}$. Then there is an orthonormal system $S = \{v_j \,|\, j \in J\}$ with $\operatorname{lin}\Sigma = \operatorname{lin} S$, which is inductively given by $v_1 = \|x_1\|^{-1} x_1$ and*

$$y_{n+1} = x_{n+1} - \sum_{j=1}^{n} (x_{n+1}|v_j)\, v_j, \qquad v_{n+1} = \frac{1}{\|y_{n+1}\|}\, y_{n+1}.$$

In the next result we collect the most important properties of an orthonormal system S. In particular, it provides the formula for the orthogonal projection onto $\overline{\operatorname{lin} S}$. In the proof one uses arguments from Linear Algebra, as well as Pythagoras' identity to show convergence.

Proposition 3.14. *Let $S = \{v_n \,|\, n \in \mathbb{N}\}$ be an orthonormal system in a Hilbert space X and $x \in X$. Then the following assertions are true.*

a) $\sum_{n=1}^{\infty} |(x|v_n)|^2 \le \|x\|^2$. (Bessel's inequality)

b) The series $Px := \sum_{n=1}^{\infty} (x|v_n)\, v_n$ converges in X. It defines a map $P \in \mathcal{B}(X)$ which is the orthogonal projection onto $\overline{\operatorname{lin} S}$. We further have the identities $\|Px\|^2 = \sum_{n=1}^{\infty} |(x|v_n)|^2 \le \|x\|^2$ and $X = \overline{\operatorname{lin} S} \oplus S^{\perp}$.

c) Let numbers $\alpha_n \in \mathbb{F}$ satisfy $Px = \sum_{n=1}^{\infty} \alpha_n v_n$. Then these coefficients are given by $\alpha_n = (x|v_n)$ for all $n \in \mathbb{N}$.

Proof. Let $x \in X$ and $N, M \in \mathbb{N}$.

a) We set $x_N = x - \sum_{k=1}^{N} (x|v_k)\, v_k$. Orthonormality yields

$$(x_N|v_n) = (x|v_n) - \sum_{k=1}^{N} (x|v_k)\,(v_k|v_n) = (x|v_n) - (x|v_n) = 0$$

for all $n \in \{1, \dots, N\}$. Using also $(v_k|v_n) = 0$ for $k \neq n$, Pythagoras' formula and $\|v_k\| = 1$, we deduce the lower bound

$$\|x\|^2 = \|x_N\|^2 + \sum_{k=1}^{N} \|(x|v_k)v_k\|^2 = \|x_N\|^2 + \sum_{k=1}^{N} |(x|v_k)|^2 \ge \sum_{k=1}^{N} |(x|v_k)|^2.$$

Assertion a) follows taking the supremum over $N \in \mathbb{N}$.

b) Let $N > M$. As in a), we obtain

$$\Big\| \sum_{k=M}^{N} (x|v_k)\, v_k \Big\|^2 = \sum_{k=M}^{N} |(x|v_k)|^2 \longrightarrow 0 \qquad \text{as } M, N \to \infty,$$

since the sequence $(|(x|v_k)|^2)_k$ is summable by a). Because X is complete, there exists the series $Px := \sum_{k=1}^{\infty} (x|v_k)\, v_k$ in X. The map

$P : X \to X$ is linear. With $M = 1$ the above equality implies that

$$\|Px\|^2 = \lim_{N\to\infty} \Big\| \sum_{k=1}^{N} (x|v_k) v_k \Big\|^2 = \lim_{N\to\infty} \sum_{k=1}^{N} |(x|v_k)|^2$$
$$= \sum_{k=1}^{\infty} |(x|v_k)|^2 \le \|x\|^2, \tag{3.4}$$

where we also employed part a). The operator P thus belongs to $\mathcal{B}(X)$ with $\|P\| \le 1$. Since the scalar product is sesquilinear and continuous, we further obtain

$$P^2 x = \sum_{n=1}^{\infty} \Big(\sum_{m=1}^{\infty} (x|v_m) v_m \Big| v_n \Big) v_n = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} (x|v_m)\,(v_m|v_n)\, v_n$$
$$= \sum_{n=1}^{\infty} (x|v_n) v_n = Px;$$

i.e., P is a projection. Observe that $S \subseteq R(P) \subseteq \overline{\operatorname{lin} S}$. By Lemma 2.16, $R(P)$ is a closed linear subspace so that $R(P) = \overline{\operatorname{lin} S}$. Formula (3.4) implies that $y \in N(P)$ if and only if $y \perp v_n$ for every $n \in \mathbb{N}$ which is equivalent to $y \in S^\perp = \overline{\operatorname{lin} S}^\perp$, see Remark 3.7. Lemma 2.16 thus yields the last claim in statement b), too.

c) Let $Px = \sum_{n=1}^{\infty} \alpha_n v_n$ and $m \in \mathbb{N}$. By means of b), we compute

$$(x|v_m) = \sum_{n=1}^{\infty} (x|v_n)\,(v_n|v_m) = (Px|v_m) = \sum_{n=1}^{\infty} \alpha_n (v_n|v_m) = \alpha_m. \qquad \square$$

Let $\{v_n \,|\, n \in \mathbb{N}\}$ an orthonormal basis in a Hilbert space X. The next theorem says that we can write each vector $x \in X$ uniquely as a series of the basis vectors times the coefficients $(x|v_n)$. Moreover, the norm of x is equal to the ℓ^2–norm of the coefficients' sequence. We can thus work as in $\mathbb{F}^m$, except for additional limits. According to statement c), we need a density result to check that a given orthonormal system is a basis. A different way to construct a basis is given by Theorem 6.7.

Theorem 3.15. *Let $S = \{v_n \,|\, n \in \mathbb{N}\}$ be a orthonormal system in a Hilbert space X. Then the following assertions are equivalent.*

a) S is a orthonormal basis.
b) $S^\perp = \{0\}$.
c) $X = \overline{\operatorname{lin} S}$.
d) For all $x \in X$ we have $x = \sum_{n=1}^{\infty} (x|v_n)\, v_n$ with a limit in X.
e) For all $x, y \in X$ we have $(x|y) = \sum_{n=1}^{\infty} (x|v_n)\,(v_n|y)$.
f) For all $x \in X$ we have $\|x\|^2 = \sum_{n=1}^{\infty} |(x|v_n)|^2$. (Parseval's equality)

The coefficients in d) are uniquely determined.

Let X be separable with $\dim X = \infty$. Then X possesses a (countable) orthonormal basis, and for every orthonormal system S there exists an orthonormal basis B containing S.

Proof. 1) Let statement a) be true. Suppose there would exist a non-zero vector $y \in S^\perp$. Then the set $S' = S \cup \{\frac{1}{\|y\|} y\}$ is an orthonormal system, contradicting the maximality of S. So b) is shown.

Part b) and the projection theorem, more precisely Example 3.9d), yield that $\overline{\operatorname{lin} S} = S^{\perp\perp} = X$.

From Proposition 3.14 we deduce the implication 'c)⇒d)' and the first addendum.

Assertion d) implies e) taking the scalar product with y, and part f) follows from e) with $x = y$.

Let S fulfill property f). Suppose that S was not an orthonormal basis. There thus exists an orthonormal system $S' \supseteq S$ with $S' \neq S$, and hence a vector $x \in X$ with $\|x\| = 1$ and $x \perp S$. Statement f) would thus yield the wrong identity $0 = \sum_{n=1}^\infty |(x|v_n)|^2 = \|x\|^2 = 1$.

2) Let $\Sigma = \{x_n \,|\, n \in \mathbb{N}\}$ be dense in X. Put $y_1 = x_{n_1}$ where x_{n_1} is the first non-zero element x_n. Iteratively, we define y_{j+1} as the first vector x_n with $n > n_j$ which does not belong to the linear span of $y_1 = x_{n_1}, \dots, y_j = x_{n_j}$. By induction one sees that the set $\Gamma = \{y_j \,|\, j \in \mathbb{N}\}$ is linearly independent and has the span $D := \operatorname{lin} \Sigma$, which is dense in X. Using Lemma 3.13, out of Γ we construct an orthonormal system $S = \{w_j \,|\, j \in \mathbb{N}\}$ whose linear hull is equal to D. The implication 'c)⇒a)' then shows that S is a orthonormal basis.

3) Let X be separable and S be orthonormal system in X. Step 2) then yields an orthonormal basis B_0 of the space $(\operatorname{lin} S)^\perp$, which is separable as a subset of X by Exercise 6.2. Then $B = S \cup B_0$ is an orthonormal basis of X since $B^\perp = \{0\}$. □

Remark 3.16. Let $S = \{v_n \,|\, n \in \mathbb{N}\}$ be an orthonormal system in a Hilbert space X.

a) Theorem 3.15 implies that S is an orthonormal basis in $\overline{\operatorname{lin} S}$.

b) It can happen that $\sum_{n=1}^\infty \|(x|v_n)\, v_n\| = \sum_{n=1}^\infty |(x|v_n)| = \infty$ for some $x \in \overline{\operatorname{lin} S}$, see Example 3.17b); i.e, the series $x = \sum_n (x|v_n)\, v_n$ does not converge absolutely and the sequence $((x|v_n))_n$ of the coefficients does not belong to ℓ^1. (It is contained in ℓ^2 by Bessel's inequality in Proposition 3.14.)

However, the series always converges *unconditionally*: For every bijection $\pi : \mathbb{N} \to \mathbb{N}$ and each $x \in X$, the series $\sum_{n=1}^\infty (x|v_{\pi(n)})\, v_{\pi(n)}$ converges to x in X. See Satz V.4.8 in [**We**]. ♢

We discuss simple examples mainly based on Weierstraß' approximation theorem.

Example 3.17. a) Example 3.12 and Theorem 3.15 imply that the set $B = \{e_n \,|\, n \in \mathbb{N}\}$ is an orthonormal basis in ℓ^2 because its linear hull c_{00} is dense in ℓ^2 by Proposition 1.31.

b) Let $X = L^2(0, 2\pi)$ with $\mathbb{F} = \mathbb{C}$ and $v_n(t) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{int}$ for $n \in \mathbb{Z}$. Weierstraß' Theorem 5.14 in Analysis 3 implies that the linear hull

$\operatorname{lin}\{v_n \,|\, n \in \mathbb{Z}\}$ is dense in $C([0,2\pi])$, and hence in X by Example 2.12a). Using Example 3.12 and a variant of Theorem 3.15 for the index set $\mathbb{Z}$, we see that $\{v_n \,|\, n \in \mathbb{Z}\}$ is an orthonormal basis in X. Let $f \in X$. We define its *Fourier coefficients* by

$$c_n := (f|v_n) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(t)\mathrm{e}^{-\mathrm{i}nt}\,\mathrm{d}t.$$

Theorem 3.15 now yields the convergence of the *Fourier series*

$$f = \sum_{n=-\infty}^{\infty} c_n v_n$$

in $L^2(0,2\pi)$ and the identity

$$\int_0^{2\pi} |f(t)|^2\,\mathrm{d}t = \sum_{n\in\mathbb{Z}} \frac{1}{2\pi} \Big| \int_0^{2\pi} f(t)\mathrm{e}^{-\mathrm{i}nt}\,\mathrm{d}t \Big|^2,$$

cf. Theorem 5.15 in Analysis 3.

Let $f = \mathbb{1}_{[0,\pi]}$. We then we obtain $c_0 = \sqrt{\pi/2}$ and, for $n \neq 0$,

$$c_n = \frac{1}{\sqrt{2\pi}} \int_0^{\pi} \mathrm{e}^{-\mathrm{i}nt}\,\mathrm{d}t = -\frac{1}{\mathrm{i}n\sqrt{2\pi}}(e^{-\mathrm{i}n\pi} - 1) = \begin{cases} 0, & n \text{ even}, \\ \frac{\sqrt{2}}{\mathrm{i}n\sqrt{\pi}}, & n \text{ odd}. \end{cases}$$

Hence, the sequence (c_n) is not summable.

c) Similarly, the set $S = \{\frac{1}{\sqrt{2\pi}}\mathbb{1}, \frac{1}{\sqrt{\pi}}\cos(n\cdot), \frac{1}{\sqrt{\pi}}\sin(n\cdot); n \in \mathbb{N}\}$ is an orthonormal basis in $X = L^2(0,2\pi)$. See Korollar 5.16 in Analysis 3.

d) Let $X = L^2(-1,1)$ and $p_n(t) = t^n$ for $n \in \mathbb{N}_0$ and $t \in [-1,1]$. The set $D = \operatorname{lin}\{p_n \,|\, n \in \mathbb{N}\}$ of polynomials is dense in $C([-1,1])$ by Theorem 5.14 in Analysis 3, and thus in X due to Example 2.12. Lemma 3.13 gives us an orthonormal system $\{v_n \,|\, n \in \mathbb{N}\}$ (multiples of the so–called Legendre polynomials) with the dense linear hull D. Theorem 3.15 now shows that $\{v_n \,|\, n \in \mathbb{N}\}$ is an orthonormal basis. $\Diamond$

Each m–dimensional vector space is isomorphic to $\mathbb{F}^m$ after fixing a basis. This result is now extended to separable Hilbert spaces and ℓ^2.

THEOREM 3.18. *Every separable Hilbert space X with $\dim X = \infty$ is isometrically isomorphic to ℓ^2 via $J : X \to \ell^2$; $Jx = ((x|v_n))_n$, and $J^{-1}((\alpha_n)) = \sum_{n=1}^{\infty} \alpha_n v_n$, for any fixed orthonormal basis $\{v_n \,|\, n \in \mathbb{N}\}$.*

PROOF. The map J is isometric due to Parseval's equality in Theorem 3.15. Its linearity is clear. Let $(\alpha_n) \in \ell^2$. For $N > M$, Pythagoras' identity and $\|v_n\| = 1$ yield

$$\Big\| \sum_{n=M}^{N} \alpha_n v_n \Big\|^2 = \sum_{n=M}^{N} |\alpha_n|^2 \;\longrightarrow\; 0 \qquad \text{as } \; M,N \to \infty.$$

The series $x := \sum_n \alpha_n v_n$ thus converges since X is a Hilbert space. Proposition 3.14 now shows that $\alpha_n = (x|v_n)$ for all $n \in \mathbb{N}$ so that J is surjective. The theorem then follows from Remark 2.11. $\square$

CHAPTER 4

Two main theorems on bounded linear operators

We discuss two fundamental results of functional analysis, the principle of uniform boundedness and the open mapping theorem, which both rely on a corollary to Baire's theorem.

4.1. The principle of uniform boundedness and strong convergence

We start with a very helpful result by *Baire.*

THEOREM 4.1. *Let M be a complete metric space and $O_n \subseteq M$ be open and dense for each $n \in \mathbb{N}$. Then their intersection $D = \bigcap_{n\in\mathbb{N}} O_n$ is dense in M.*

PROOF. For every $x_0 \in M$ and $\delta > 0$ we must find a vector x in $B_0 \cap D$, where we put $B_0 = B(x_0, \delta)$. So let $x_0 \in M$ and $\delta > 0$. Since O_1 is open and dense, there is an element x_1 of $O_1 \cap B_0$ and a radius $\delta_1 \in (0, \frac{1}{2}\delta]$ with $\overline{B}(x_1, \delta_1) \subseteq O_1 \cap B_0$. Iteratively, one finds $x_n \in O_n \cap B_{n-1}$, $\delta_n \in (0, \frac{1}{2}\delta_{n-1}]$ and $B_n = B(x_n, \delta_n)$ such that

$$\overline{B_n} \subseteq O_n \cap B_{n-1} \subseteq O_n \cap (O_{n-1} \cap B_{n-2}) \subseteq \cdots \subseteq (O_n \cap \ldots \cap O_1) \cap B_0.$$

Since $\delta_m \le 2^{-m}\delta$, the vector x_n belongs to $B_m \subseteq B(x_m, 2^{-m}\delta)$ for all $n \ge m$. Hence, (x_n) is a Cauchy sequence. Its limit x is contained in each set $\overline{B_m}$, and thus in $D \cap B_0$ by the above inclusions. □

The set $\mathbb{R}^2 \setminus \mathbb{R}$ is open and dense in $\mathbb{R}^2$, for instance. A countable intersection D of open and dense sets is called *residual.* A property is *generic* if it is satisfied by all elements of such a set.

COROLLARY 4.2. *Let M be a complete metric space and $M = \bigcup_{n\in\mathbb{N}} A_n$ for closed subsets $A_n \subseteq M$. Then there exists an index $N \in \mathbb{N}$ with $A_N^\circ \ne \emptyset$.*

PROOF. Suppose that $A_n^\circ = \emptyset$ for all $n \in \mathbb{N}$. Then $O_n = M \setminus A_n$ is open and dense. Theorem 4.1 implies that $\bigcap_{n\in\mathbb{N}} O_n$ is dense in M. This fact contradicts the assumption since $\bigcap_n O_n = M \setminus \bigcup_n A_n = \emptyset$. □

EXAMPLE 4.3. One needs the completeness of M in the above corollary. For instance, take $(c_{00}, \|\cdot\|_p)$ with $p \in [1, \infty]$ and $A_n = \{x = (x_1, \ldots, x_n, 0, \ldots) \,|\, x_j \in \mathbb{F}\}$. These sets are closed for all $n \in \mathbb{N}$ by Lemma 1.43, and $\bigcup_n A_n = c_{00}$. But, each A_n has empty interior since for $x \in A_n$ the vectors $y_m = x + \frac{1}{m} e_{n+1} \notin A_n$ tend to x as $m \to \infty$. ♢

The following *principle of uniform boundedness* is one of the four most fundamental theorems of linear functional analysis treated in this course. It says that completeness provides uniformity for free in some cases. Using the full power of Baire's theorem instead of Corollary 4.2, one can considerably strengthen the result, see Section II.4 of [**Yo**].

THEOREM 4.4. *Let X be a Banach space, Y be a normed vector space, and $\mathcal{T} \subseteq \mathcal{B}(X,Y)$. If the set of operators $\mathcal{T}$ is* pointwise bounded *(i.e., $\forall\, x \in X\ \exists\, c_x \geq 0\ \forall T \in \mathcal{T} : \|Tx\| \leq c_x$), then $\mathcal{T}$ is* uniformly bounded *(i.e., $\exists\, c > 0\ \forall T \in \mathcal{T} : \|T\| \leq c$).*[1]

PROOF. We put $A_n = \{x \in X \,|\, \|Tx\| \leq n \text{ for all } T \in \mathcal{T}\}$. By asssumption, each vector $x \in X$ belongs to every set A_n with $n \geq c_x$; i.e., $\bigcup_{n\in\mathbb{N}} A_n = X$. Let (x_k) in A_n converge to x in X. We then have $\|Tx\| = \lim_{k\to\infty} \|Tx_k\| \leq n$ for all $T \in \mathcal{T}$ so that A_n is closed for each $n \in \mathbb{N}$. Corollary 4.2 now yields an index $N \in \mathbb{N}$, a point $x_0 \in A_N$ and a radius $\varepsilon > 0$ with $\overline{B}(x_0, \varepsilon) \subseteq A_N$. Let $z \in \overline{B}(0, \varepsilon)$. The vectors $x_0 \pm z$ then belong to $\overline{B}(x_0, \varepsilon) \subseteq A_N$, and hence

$$\begin{aligned}\|Tz\| &= \|T(\tfrac{1}{2}(z + x_0) + \tfrac{1}{2}(z - x_0))\| \leq \tfrac{1}{2}\,\|T(z + x_0)\| + \tfrac{1}{2}\,\|T(x_0 - z))\| \\ &\leq \tfrac{N}{2} + \tfrac{N}{2} = N.\end{aligned}$$

Finally, let $x \in X$ with $\|x\| \leq 1$. Set $z = \varepsilon x \in \overline{B}(0, \varepsilon)$. It follows $N \geq \|Tz\| = \varepsilon\,\|Tx\|$ and thus $\|T\| \leq \frac{N}{\varepsilon}$ for all $T \in \mathcal{T}$. □

The above result is often used in the next simpler version called *Banach–Steinhaus theorem.*

COROLLARY 4.5. *Let X be a Banach space, Y be a normed vector space, and T_n belong to $\mathcal{B}(X,Y)$ for every $n \in \mathbb{N}$. Assume that $(T_n x)_n$ converges in Y for each $x \in X$. Then $\sup_{n\in\mathbb{N}} \|T_n\| < \infty$.*

PROOF. The assumpotion shows that $c_x := \sup_n \|T_n x\|$ is finite for each $x \in X$, so that the result follows from Theorem 4.4. □

We again note that we really need completeness here.

EXAMPLE 4.6. Let $X = c_{00}$ and $Y = c_0$ be endowed with $\|\cdot\|_\infty$. Set $T_n x = (x_1, 2x_2, \dots, nx_n, 0, \dots)$ for $n \in \mathbb{N}$ and $x \in X$. Then T_n belongs to $\mathcal{B}(X,Y)$ with $\|T_n\| = n$ since $\|T_n x\|_\infty \leq n\,\|x\|_\infty$ and $\|T_n\| \geq \|T_n e_n\|_\infty = n$ for all $n \in \mathbb{N}$ and $x \in X$. Hence, the sequence (T_n) is unbounded. However, $(T_n x)_n$ converges for each $x \in X$ because there is an index $m = m_x \in \mathbb{N}$ with $x_k = 0$ if $k > m$ and thus $T_n x = (x_1, 2x_2, \dots, mx_m, 0, \dots)$ for all $n \geq m$. ◊

The principle of uniform boundedness is often used to establish the existence of interesting objects which are difficult to construct explicitly. As a typical example we look at pointwise divergent Fourier series.

[1] Uniform boundedness is equivalent to $\exists\, c > 0\ \forall\, T \in \mathcal{T},\ x \in \overline{B}(0,1) : \|Tx\| \leq c$.

EXAMPLE 4.7. We endow $X = \{f \in C([-\pi,\pi]) \,|\, f(-\pi) = f(\pi)\}$ with the supremum norm and let $f \in X$. As in Korollar 5.16 of Analysis 3 one can find coefficients $a_k, b_k \in \mathbb{F}$ such that the Fourier sum

$$S_n(f,t) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos(kt) + b_k \sin(kt)), \quad t \in [-\pi,\pi],$$

converges to f in $L^2(-\pi,\pi)$ as $n \to \infty$. We claim that there is a function $f \in X$ whose Fourier series diverges at $t = 0$.

PROOF. The fomulas for a_k and b_k from Analysis 3 and simple manipulations yield the representation

$$S_n(f,t) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(s+t) D_n(s)\,\mathrm{d}s$$

for $t \in [-\pi,\pi]$, see p. 146 of [**We**]. Here f has been extended to a 2π–periodic function on $\mathbb{R}$ and we use the 'Dirichlet kernel'

$$D_n(t) = \begin{cases} \frac{\sin(n+\frac{1}{2})t}{2\sin\frac{t}{2}}, & t \in [-\pi,\pi] \setminus \{0\}, \\ n + \frac{1}{2}, & t = 0. \end{cases}$$

In view of our claim, we define the map $\varphi : X \to \mathbb{F}$; $\varphi_n(f) = S_n(f,0)$, which clearly belongs to $X^\star$ for each $n \in \mathbb{N}$. Similar as in Example 2.7, one can compute

$$\|\varphi_n\| = \frac{1}{\pi} \int_{-\pi}^{\pi} |D_n(t)|\,\mathrm{d}t.$$

It is straightforward to show that this integral tends to ∞ as $n \to \infty$, see the proof of Satz IV.2.10 in [**We**]. Corollary 4.5 then implies that $\varphi_n(f)$ cannot converge for all $f \in X$, as asserted.[2] □

The convergence used in Corollary 4.5 plays an important role in analysis so that we discuss it a bit.

DEFINITION 4.8. *Let X and Y be normed vector spaces and $T_n, T \in \mathcal{B}(X,Y)$ for $n \in \mathbb{N}$. We say that (T_n)* converges strongly to T *if $T_n x \to Tx$ in Y as $n \to \infty$ for each $x \in X$. One then writes $T_n \xrightarrow{s} T$.*

Observe that it is not clear at the moment whether this type of convergence can be described by a metric, cf. Remark 5.35. Nevertheless its basic properties are easy to show.

REMARK 4.9. Let X and Y be normed vector spaces and T_n, T, S_n, and S belong to $\mathcal{B}(X,Y)$ for $n \in \mathbb{N}$. Then the following assertions hold.

a) If (T_n) tends to T and S strongly, then $S = T$ by the uniqueness of limits in Y.

b) Let (T_n) have the limit T in operator norm. Then (T_n) converges to T strongly since $\|T_n x - Tx\| \le \|T_n - T\|\,\|x\|$ for all $x \in X$. The

[2] A more or less concrete example of a pointwise divergent Fourier series is given in Section 18 of [**Ko**].

converse is wrong in general. As an example, consider the operators given by $P_n x = (x_1, \dots, x_n, 0, \dots)$ on $X = Y = \ell^2$ which converge strongly to I but $\|P_n - I\| \ge \|(P_n - I)e_{n+1}\|_2 = 1$ for all $n \in \mathbb{N}$.

c) Let the operators T_n tend strongly to T, S_n tend strongly to S, and $\alpha, \beta \in \mathbb{F}$. Then the vectors $(\alpha T_n + \beta S_n)x$ converge to $(\alpha T + \beta S)x$ as $n \to \infty$, so that $(\alpha T_n + \beta S_n)$ has the strong limit $\alpha T + \beta S$. ♢

The next result is an important tool in analysis. It allows to construct bounded linear operators as strong limits on a dense set of 'good' vectors, provided one has a uniform bound.

Lemma 4.10. *Let X be a normed vector space, Y be a Banach space, $T_n \in \mathcal{B}(X, Y)$ for all $n \in \mathbb{N}$, and $S \subseteq X$ be a subset whose span $D = \operatorname{lin} S$ is dense in X. Assume that $\sup_{n\in\mathbb{N}} \|T_n\| =: M < \infty$ and that $(T_n x)_n$ converges for every $x \in S$ as $n \to \infty$. Then there is a unique operator $T \in \mathcal{B}(X, Y)$ such that (T_n) converges strongly to T and $\|T\| \le \liminf_{n\to\infty} \|T_n\| =: M_0$. If $D = X$, these assertions are also true if Y is a normed vector space.*

Proof. By linearity, the limit $T_0 x := \lim_{n\to\infty} T_n x$ exists for every $x \in D$. As in the proof of Proposition 2.5 one checks that $T_0 : D \to Y$ is linear. Choose a subsequence with $M_0 = \lim_{j\to\infty} \|T_{n_j}\|$. We then obtain $\|T_0 x\| = \lim_{j\to\infty} \|T_{n_j} x\| \le M_0 \|x\|$ for each $x \in D$, and hence T_0 belongs to $\mathcal{B}(D, Y)$ with $\|T_0\| \le M_0$. So far we have not used that Y is Banach space, and the addendum is thus proved if $D = X$. In the general case, Lemma 2.13 yields a unique extension $T \in \mathcal{B}(X, Y)$ of T_0 with $\|T\| = \|T_0\| \le M_0$, since Y is a Banach space and D is dense.

Let $\varepsilon > 0$ and $x \in X$. Fix a vector $z \in D$ with $\|x - z\| \le \varepsilon$. Then there exists an index $N_\varepsilon \in \mathbb{N}$ such that $\|T_0 z - T_n z\| \le \varepsilon$ for all $n \ge N_\varepsilon$. We estimate

$$\begin{aligned}\|Tx - T_n x\| &\le \|T(x - z)\| + \|Tz - T_n z\| + \|T_n(z - x)\| \\ &\le M_0 \varepsilon + \varepsilon + M\varepsilon\end{aligned}$$

for $n \ge N_\varepsilon$, so that (T_n) tends strongly to T also if $D \ne X$. □

We first illustrate by standard examples that one cannot omit the uniform bound in the above lemma and that the operator norm is **not** 'continuous' for the strong limit in general.

Example 4.11. a) Let $X = Y = c_0$, $D = c_{00}$ and $T_n x = (x_1, 2x_2, \dots, nx_n, 0, \dots)$ for all $x \in c_0$ and $n \in \mathbb{N}$. As in Example 4.6 we see that $T_n \in \mathcal{B}(X)$ satisfies $\|T_n\| = n \to \infty$. The operators T_n thus do not have a strong limit by Corollary 4.5. However, for each $x \in c_{00}$ the vectors $T_n x$ tend to $(x_1, \dots, mx_m, 0, \dots)$ as $n \to \infty$ where $m = m_x \in \mathbb{N}$ is given by Example 4.6.

b) Let $X = Y = c_0$ and $T_n x = x_n e_n$ for all $x \in c_0$ and $n \in \mathbb{N}$. Since $\|T_n x\|_\infty = |x_n| \to 0$ as $n \to \infty$, the operators T_n have the strong limit

$T := 0$. However, $\|T_n e_n\|_\infty \geq 1$ and $\|T_n x\|_\infty \leq \|x\|_\infty$ for all $n \in \mathbb{N}$ and $x \in X$. Here the limit $\lim_{n\to\infty} \|T_n\| = 1$ exists, but it is strictly larger than $\|T\| = 0$, despite the strong convergence. $\Diamond$

We now look at two important examples of bounded operators which exhibit strong convergence, starting with *left translations*. Recall Example 2.9 for a discrete version.

Example 4.12. Let $X \in L^p(\mathbb{R})$ for some $1 \leq p < \infty$. For every $t \in \mathbb{R}$ we define $(T(t)f)(s) = f(s+t)$ for $s \in \mathbb{R}$ and $f \in X$. It is clear that $T(t) : X \to X$ is linear,

$$\|T(t)f\|_p^p = \int_{\mathbb{R}} |f(s+t)|^p \,\mathrm{d}s = \|f\|_p^p,$$

and $T(t)$ has the inverse $T(-t)$. Hence, $T(t) \in \mathcal{B}(X)$ is an isometric isomorphism. (This assertion can similarly be shown for $p = \infty$.) For all $b > a$ and $t \in \mathbb{R}$ we have

$$(T(t)\mathbb{1}_{[a,b]})(s) = \mathbb{1}_{[a,b]}(s+t) = \left\{\begin{matrix} 1, & a-t \leq s \leq b-t \\ 0, & \text{otherwise} \end{matrix}\right\} = \mathbb{1}_{[a-t,b-t]}(s)$$

for all $s \in \mathbb{R}$; i.e., $T(t)$ really is a left translation if $t \geq 0$.

Setting $f = t^{-1/p}\mathbb{1}_{[0,t]}$ for $t > 0$ we further obtain $\|f\|_p = 1$ and $\|T(t)f - f\|_p^p = t^{-1}\int_{-t}^{t} 1^p \,\mathrm{d}s = 2$. Therefore the map $\mathbb{R} \to \mathcal{B}(X)$; $t \mapsto T(t)$, is **not** continuous with respect to the operator norm. We claim that it is *strongly continuous*; i.e., the functions $\mathbb{R} \to X$; $t \mapsto T(t)f$, are continuous for every $f \in X$.

By Lemma 4.10, we only have to consider $f \in C_c(\mathbb{R})$ since $\|T(t)\| = 1$ for all $t \in \mathbb{R}$ and $C_c(\mathbb{R})$ is dense in X by Theorem 5.9 in Analysis 3. Let $t_0 \in \mathbb{R}$ and $t \in [t_0 - 1, t_0 + 1]$. Since $f \in C_c(\mathbb{R})$ is uniformly continuous, we derive

$$\|T(t)f - T(t_0)f\|_\infty = \sup_{s\in\mathbb{R}} |f(s+t) - f(s+t_0)| \longrightarrow 0 \qquad \text{as } t \to t_0.$$

There is a compact interval $J \subseteq \mathbb{R}$ with $\operatorname{supp}(T(t)f - T(t_0)f) \subseteq J$ for $t \in [t_0 - 1, t_0 + 1]$. We then obtain

$$\|T(t)f - T(t_0)f\|_p \leq \lambda(J)^{1/p} \|T(t)f - T(t_0)f\|_\infty \longrightarrow 0 \qquad \text{as } t \to t_0.$$

This result remains valid for $X = C_0(\mathbb{R})$ with an analogous proof. $\Diamond$

As a second example, we study **mollifiers** which are essential tools in analysis. Let $U \subseteq \mathbb{R}^m$ be open. We define the space of *test functions* on U by $C_c^\infty(U) = \{\varphi \in C^\infty(U) \mid \operatorname{supp}\varphi \text{ is compact}\}$. The map

$$\varphi_0(x) = \begin{cases} \mathrm{e}^{-\frac{1}{1-|x|_2^2}}, & |x|_2 < 1, \\ 0, & |x|_2 \geq 1, \end{cases}$$

belongs to $C_c^\infty(\mathbb{R}^m)$, for instance. Take any test function φ on $\mathbb{R}^m$ with support $\overline{B}(0,1)$ which is positive on $B(0,1)$. Set $k = \|\varphi\|_1^{-1}\varphi$

and $k_\varepsilon(x) = \varepsilon^{-m} k(\frac{1}{\varepsilon}\, x)$ for all $x \in \mathbb{R}$ and $\varepsilon > 0$. The function k_ε is contained in $C_c^\infty(\mathbb{R}^m)$ with $\operatorname{supp} k_\varepsilon = \overline{B}(0,\varepsilon)$, $k_\varepsilon \geq 0$, and

$$\|k_\varepsilon\|_1 = \int_{\mathbb{R}^m} \varepsilon^{-m} k(\tfrac{1}{\varepsilon}\, x)\,\mathrm{d}x = \int_{\mathbb{R}^m} k(y)\,\mathrm{d}y = 1$$

for all $\varepsilon > 0$, where we have used the transformation $y = \frac{1}{\varepsilon}\, x$.

Let $p \in [1,\infty]$. The space $L^p_{\mathrm{loc}}(U)$ contains all measurable maps $f : U \to \mathbb{F}$ (modulo null functions) such that $f|_K$ belongs to $L^p(K)$ for each compact set $K \subseteq U$. Proposition 1.35 yields the inclusion of the 'locally integrable functions' $L^1_{\mathrm{loc}}(U)$ in $L^p_{\mathrm{loc}}(U)$. A sequence (f_n) has a limit f in $L^p_{\mathrm{loc}}(U)$ if the restrictions to each compact set $K \subseteq U$ converge in $L^p(K)$. (As in Example 1.9 one can construct a distance on $L^p_{\mathrm{loc}}(U)$ corresponding to this convergence.)

The extension of f by 0 to $\mathbb{R}^m$ is denoted by $\tilde f$. Let $f : U \to \mathbb{F}$ be measurable and $\tilde f$ belong to $L^1_{\mathrm{loc}}(\mathbb{R}^m)$. We define the *mollifier* G_ε by

$$(G_\varepsilon f)(x) := (k_\varepsilon * \tilde f)(x) = \int_{\mathbb{R}^m} k_\varepsilon(x-y)\tilde f(y)\,\mathrm{d}y \qquad (4.1)$$

$$= \int_{U\cap\overline{B}(x,\varepsilon)} k_\varepsilon(x-y) f(y)\,\mathrm{d}y$$

for all $x \in U$ or $x \in \mathbb{R}^m$. (This integral exists since $\tilde f$ is integrable on $\overline{B}(x,\varepsilon)$ and k_ε is bounded.) The next result says that we can use the operators G_ε to approximate locally integrable functions by smooth ones. In particular, test functions are dense in $L^p(U)$ if $p < \infty$. As the proof indicates, mollifiers are often used in combination with cut-off arguments. We improve this proposition in the next section. Note that for $U = \mathbb{R}^m$ the number ε_0 is equal to ∞ since $\partial\mathbb{R}^m = \emptyset$.[3] This special case is considerably simpler.

PROPOSITION 4.13. *Let $U \subseteq \mathbb{R}^m$ be open, $f : U \to \mathbb{F}$ be measurable with $\tilde f \in L^1_{\mathrm{loc}}(\mathbb{R}^m)$, $\varepsilon > 0$, and $1 \leq p \leq \infty$. Define G_ε by (4.1). Then the following assertions hold.*

a) The map $G_\varepsilon f$ is an element of $C^\infty(\mathbb{R}^m)$. If there is a compact set $K \subseteq U$ with $f(x) = 0$ for a.e. $x \in U \setminus K$, then $G_\varepsilon f \in C_c^\infty(U)$ for all $\varepsilon \in (0,\varepsilon_0)$ with $\varepsilon_0 := \operatorname{dist}(K,\partial U)$.

b) The restriction G_ε to $L^p(U)$ belongs to $\mathcal{B}(L^p(U))$ with $\|G_\varepsilon\| \leq 1$. Let $1 \leq p < \infty$ and $f \in L^p(U)$. Then $G_\varepsilon f \to f$ in $L^p(U)$ as $\varepsilon \to 0$.

c) Let $1 \leq p < \infty$. Then $C_c^\infty(U)$ is dense in $L^p(U)$. More precisely, if $f \in L^p(U) \cap L^q(U)$ for some $1 \leq p, q < \infty$ then there are functions $f_n \in C_c^\infty(U)$ converging to f in $L^p(U)$ and in $L^q(U)$.

PROOF. a) Let $\varepsilon > 0$ and $\tilde f \in L^1_{\mathrm{loc}}(\mathbb{R}^m)$. Fix a point $x_0 \in \mathbb{R}^m$ and a radius $r > 0$. Let $x \in B(x_0,r)$ and $j \in \{1,\dots,m\}$. We then estimate

$$|\partial_{x_j} k_\varepsilon(x-y)\tilde f(y)| \leq \|\partial_{x_j} k_\varepsilon\|_\infty\, \mathbb{1}_{\overline{B}(x_0,r+\varepsilon)}(y)\,|\tilde f(y)| =: h(y)$$

[3] Concerning ε_0, the result presented in the lectures was slightly weaker.

for all $y \in \mathbb{R}^m$. The function h is integrable on $\mathbb{R}^m$. The differentiation theorem 3.18 of Analysis 3 now yields that $k_\varepsilon * f$ is partially differentiable at x_0. Iterating this argument, one concludes that $G_\varepsilon f$ belongs to $C^\infty(\mathbb{R}^m)$.

Let $K \subseteq U$ be compact such that $f(x) = 0$ for a.e. $x \in U \setminus K$. Since $K \cap \partial U = \emptyset$ and K is compact, the number $\varepsilon_0 = \operatorname{dist}(K, \partial U)$ is positive by Example 1.9. Take $\varepsilon \in (0, \varepsilon_0)$. The set $S_\varepsilon := K + \overline{B}(0,\varepsilon)$ then belongs to U. Let $z_n = x_n + y_n \in S_\varepsilon$ for $n \in \mathbb{N}$ with $x_n \in K$ and $y_n \in \overline{B}(0,\varepsilon)$. By compactness of both sets, (x_n) and (y_n) have subsequences with limits $x \in K$ and $y \in \overline{B}(0,\varepsilon)$, respectively; i.e., S_ε is compact. On the other hand, equation (4.1) yields $\operatorname{supp}(G_\varepsilon f) \subseteq S_\varepsilon$.

b) Let $f \in L^p(U)$. Young's inequality Theorem 2.14 implies that $G_\varepsilon f \in L^p(U)$ and $\|G_\varepsilon f\|_p \le \|k_\varepsilon\|_1 \|\tilde{f}\|_p = \|f\|_p$. As a result, G_ε induces an element of $\mathcal{B}(L^p(U))$ with $\|G_\varepsilon\| \le 1$ for all $\varepsilon > 0$.

Let $p \in [1,\infty)$. To show that $G_\varepsilon \to I$ strongly on $L^p(U)$, it suffices to consider $g \in C_c(U)$ due to Lemma 4.10 since $\|G_\varepsilon\| \le 1$ and $C_c(U)$ is dense in $L^p(U)$ by Theorem 5.9 in Analysis 3. Let $g \in C_c(U)$, $K := \operatorname{supp}(g)$, and $\varepsilon \in (0,\varepsilon_0)$. Again, $S_\varepsilon = K + \overline{B}(0,\varepsilon) \subseteq U$ is compact and $K, \operatorname{supp} G_\varepsilon g \subseteq S_\varepsilon$ for $\varepsilon \in (0,\varepsilon_0)$. We then derive

$$\begin{aligned}\sup_{x\in U} |G_\varepsilon g(x) - g(x)| &= \sup_{x\in S_\varepsilon} \Big| \int_{\mathbb{R}^m} k_\varepsilon(x-y)\tilde{g}(y)\,\mathrm{d}y - \int_{\mathbb{R}^m} k_\varepsilon(x-y)\,\mathrm{d}y\; g(x)\Big| \\ &\le \sup_{x\in S_\varepsilon} \int_{\bar{B}(x,\varepsilon)} k_\varepsilon(x-y)\,|\tilde{g}(y) - g(x)|\,\mathrm{d}y \\ &\le \|k_\varepsilon\|_1 \sup_{x\in S_\varepsilon, |x-y|\le\varepsilon} |\tilde{g}(y) - \tilde{g}(x)| \;\longrightarrow\; 0\end{aligned}$$

as $\varepsilon \to 0$, using that $\|k_\varepsilon\|_1 = 1$ and that $\tilde{g}$ is uniformly continuous. We fix some index $\overline{\varepsilon} \in (0,\varepsilon_0)$ and let $\varepsilon \in (0,\overline{\varepsilon}]$. The distance $\|G_\varepsilon g - g\|_p \le \lambda(S_{\overline{\varepsilon}})^{1/p}\,\|G_\varepsilon g - g\|_\infty$ then tends to 0 as $\varepsilon \to 0$.

c) Let $f \in L^p(U) \cap L^q(U)$ for some $1 \le p, q < \infty$. By Example 1.9 there are open and bounded sets U_n for $n \in \mathbb{N}$ with $\overline{U_n} \subseteq U$ whose union is equal U. Using assertion a), for each $n \in \mathbb{N}$ we can choose a number $\varepsilon_n \in (0, 1/n]$ such that the function $f_n = G_{\varepsilon_n}(\mathbb{1}_{U_n} f)$ belongs to $C_c^\infty(U)$. Since $\mathbb{1}_{U_n} f \to f$ pointwise as $n \to \infty$ and $|\mathbb{1}_{U_n} f| \le |f|$ for all n, Lebesgue's theorem yields the limits $\mathbb{1}_{U_n} f \to f$ in $L^p(U)$ and in $L^q(U)$. Employing $\|G_{\varepsilon_n}\| \le 1$ and part b), we then derive

$$\|G_{\varepsilon_n}(\mathbb{1}_{U_n} f) - f\|_r \le \|G_{\varepsilon_n}\|\,\|\mathbb{1}_{U_n} f - f\|_r + \|G_{\varepsilon_n} f - f\|_r \;\longrightarrow\; 0$$

as $n \to \infty$, where $r \in \{p, q\}$. □

4.2. Sobolev spaces

The classical (partial) derivative does not fit well to L^p spaces since it is defined via a pointwise limit. For a treatment of partial differential equations in an L^2 or L^p context one needs the more general concept of 'weak derivatives'. Here we restrict ourselves to basic results focusing on simple properties of the function spaces and a bit of calculus. The mollifiers from Proposition 4.13 are the crucial tool in this theory. For an introduction to this area we refer to the books [**Br**] or [**Do**], and also to the lecture notes [**ST**].

Throughout in this section, let $U \subseteq \mathbb{R}^m$ be open and non-empty. To motivate the definition below, we consider $f \in C^1(U)$ and $m \geq 2$. Take any test function $\varphi \in C_c^\infty(U)$. Let $\tilde{g} \in C^1(\mathbb{R}^m)$ be the 0–extension of the product $g := f\varphi \in C_c^1(U)$. There is a number $a > 0$ with $\operatorname{supp} \varphi \subseteq (-a,a)^m$. Set $C = (-a,a)^{m-1}$. Using the product rule and the fundamental theorem of calculus, we obtain

$$\begin{aligned}\int_U \varphi\, \partial_1 f \,\mathrm{d}x &= -\int_U f\partial_1\varphi \,\mathrm{d}x + \int_U \partial_1 g \,\mathrm{d}x \\ &= -\int_U f\partial_1\varphi \,\mathrm{d}x + \int_C \int_{-a}^a \partial_1 \tilde{g}(x_1, x') \,\mathrm{d}x_1 \,\mathrm{d}x' \\ &= -\int_U f\partial_1\varphi \,\mathrm{d}x + \int_C (\tilde{g}(a,x') - \tilde{g}(-a,x')) \,\mathrm{d}x' \\ &= -\int_U f\partial_1\varphi \,\mathrm{d}x, \end{aligned} \tag{4.2}$$

since $\tilde{g}(a,x') = \tilde{g}(-a,x') = 0$. Other partial derivatives can be treated in the same way. The following definition now relies on the observation that the right hand side of (4.2) is defined for all locally integrable f. It can thus serve as the definition of $\partial_1 f$ on the left hand side of (4.2).

Definition 4.14. *Let $U \subseteq \mathbb{R}^m$ be open, $f, g \in L^1_{\mathrm{loc}}(U)$, $j \in \{1, \ldots, m\}$, and $1 \leq p \leq \infty$. Assume that*

$$\int_U g\varphi \,\mathrm{d}x = -\int_U f\, \partial_j \varphi \,\mathrm{d}x \tag{4.3}$$

for all $\varphi \in C_c^\infty(U)$. Then $g =: \partial_j f$ is called weak derivative *of f. We write $D_j(U)$ for the space of such f. One further defines the* Sobolev spaces *(of first order) by*

$$W^{1,p}(U) = \{f \in L^p(U) \,|\, f \in D_j(U),\ \partial_j f \in L^p(U) \text{ for all } j \in \{1, \ldots, m\}\}$$

and endows them with

$$\|f\|_{1,p} = \begin{cases} \left(\|f\|_p^p \sum_{j=1}^m \|\partial_j f\|_p^p \right)^{\frac{1}{p}}, & 1 \leq p < \infty, \\ \max\left\{ \|f\|_\infty, \|\partial_1 f\|_\infty, \ldots, \|\partial_m f\|_\infty \right\}, & p = \infty. \end{cases}$$

As usually, the spaces $D_j(U)$ and $W^{1,p}(U)$ are spaces of equivalence classes modulo the space of null functions $\mathcal{N}$. The above definition can be extended to derivatives of higher order in a straigthforward way.

We first have to settle the basic question whether the weak derivative is uniquely determined by (4.3). This is done via the *fundamental lemma of calculus of variations.*

LEMMA 4.15. *Let $g \in L^1_{\mathrm{loc}}(U)$ satisfy*

$$\int_{\mathbb{R}^m} g\varphi \, \mathrm{d}x = 0 \qquad \text{for all } \varphi \in C_c^\infty(U).$$

Then $g = 0$ a.e.. Hence, a function $f \in D_j(U)$ has exactly one j–th weak derivative.

PROOF. Assume that $g \neq 0$ on a Borel set $B \subseteq U$ with $\lambda(B) > 0$. Theorem 1.25 of Analysis 3 yields a compact set $K \subseteq B$ with $\lambda(K) > 0$. We fix a positive number $\delta < \frac{1}{2}\operatorname{dist}(K, \partial U)$. The sum $S = K + \overline{B}(0,\delta) \subseteq U$ is also compact. Proposition 4.13 then shows that $\varphi = G_\delta \mathbb{1}_S$ belongs to $C_c^\infty(U)$. For all $x \in K$, we compute

$$\varphi(x) = \int_{B(x,\delta)} k_\delta(x-y)\, \mathbb{1}_S(y) \, \mathrm{d}y = \int_{B(x,\delta)} k_\delta(x-y) \, \mathrm{d}y = 1,$$

using the definition (4.1) of G_ε. Since $\varphi g \in L^1(U)$, the functions $G_\varepsilon(\varphi g)$ converge to φg in $L^1(U)$ as $\varepsilon \to 0$ by Proposition 4.13. There thus exist a nullset N and a subsequence $\varepsilon_j \to 0$ such that $(G_{\varepsilon_j}(\varphi g))(x) \to \varphi(x)g(x) = g(x) \neq 0$ as $j \to \infty$ for each $x \in K\backslash N$. Fix any $x \in K\backslash N$. For every $j \in \mathbb{N}$, we further deduce

$$(G_{\varepsilon_j}(\varphi g))(x) = \int_U k_{\varepsilon_j}(x-y)\varphi(y)g(y) \, \mathrm{d}y = 0$$

from the assumption, since the function $y \mapsto k_{\varepsilon_j}(x-y)\varphi(y)$ belongs to $C_c^\infty(U)$. This contradiction implies the assertions. □

We next collect simple properties of the spaces $D_j(U)$ and $W^{1,p}(U)$.

REMARK 4.16. Let $1 \le p \le \infty$ and $j \in \{1, \dots, m\}$.

a) Formula (4.2) yields the inclusion $C^1(U) + \mathcal{N} \subseteq D_j(U)$ and that weak and classical derivatives coincide for $f \in C^1(U)$. This fact justifies to use the same notation for both of them.

b) From Definition 4.14 one easily deduces that $D_j(U)$ is a vector space and $\partial_j : D_j(U) \to L^1_{\mathrm{loc}}(U)$ is linear.

c) It is straightforward to check that $(W^{1,p}(U), \|\cdot\|_{1,p})$ is a normed vector space. Moreover, a sequence (f_n) converges in $W^{1,p}(U)$ if and only if (f_n) and $(\partial_j f_n)$ converge in $L^p(U)$ for all $j \in \{1, \dots, m\}$.

d) The map

$$J : W^{1,p}(U) \to L^p(U)^{1+m}; \quad f \mapsto (f, \partial_1 f, \dots, \partial_m f),$$

is a linear isometry, where $L^p(U)^{1+m}$ is endowed with the norm $\big|(\|f\|_p, \|\partial_1 f\|_p, \dots, \|\partial_m f\|_p)\big|_p$. We see in the proof of Proposition 4.19 that $W^{1,p}(U)$ is isometrically isomorphic to a closed subspace of

$L^p(U)^{1+m}$. Since the p–norm and the 1–norm on $\mathbb{R}^{1+m}$ are equivalent, there are constants $C, c > 0$ with

$$c\Big(\|f\|_p + \sum_{j=1}^m \|\partial_j f\|_p\Big) \le \|u\|_{1,p} \le C\Big(\|f\|_p + \sum_{j=1}^m \|\partial_j f\|_p\Big)$$

for all $f \in W^{1,p}(U)$. $\Diamond$

The next lemma first gives a convergence result in $L^1_{\mathrm{loc}}(U)$ for weak derivatives which is analogous to that from Analysis 1 for uniform limits and the classical derivative. Second, it makes clear that the mollifiers fit perfectly well to weak derivatives except for some trouble near ∂U.

We use the following fact. Let $(S, \mathcal{A}, \mu)$ be a measure space and $p \in [1, \infty]$. As in Remark 3.2d), we deduce from Hölder's inequality

$$\text{the continuity of} \qquad L^p(\mu) \times L^{p'}(\mu) \to \mathbb{F}; \quad (f,g) \mapsto \int_S fg \,\mathrm{d}\mu. \tag{4.4}$$

Lemma 4.17. *Let $1 \le p \le \infty$ and $j \in \{1, \dots, m\}$.*

a) Let $f_n \in D_j(U)$ and $f, g \in L^1_{\mathrm{loc}}(U)$ such that $f_n \to f$ and $\partial_j f_n \to g$ in $L^1_{\mathrm{loc}}(U)$ as $n \to \infty$. Then $f \in D_j(U)$ and $\partial_j f = g$. If these limits exist in $L^p(U)$ for all j, then f belongs to $W^{1,p}(U)$.

b) Let $p < \infty$, $f \in D_j(U)$, $f, \partial_j f \in L^p_{\mathrm{loc}}(U)$, $x \in U$, and $0 < \varepsilon < \mathrm{d}(x, \partial U)$. We then have $\partial_j(G_\varepsilon f)(x) = G_\varepsilon(\partial_j f)(x)$. Moreover, $G_\varepsilon f$ tends to f and $\partial_j G_\varepsilon f$ to $\partial_j f$ in $L^p_{\mathrm{loc}}(U)$ as $\varepsilon \to 0$.

Proof. a) Let $\varphi \in C_c^\infty(U)$. By assumption a), the functions $\varphi\, \partial_j f_n$ converge to $\varphi\, g$ and $f_n\, \partial_j \varphi$ to $f\, \partial_j \varphi$ in $L^1(U)$. Using (4.4) and Definition 4.14, we compute

$$\int_U \varphi\, \partial_j f \,\mathrm{d}x = \lim_{n\to\infty} \int_U \varphi\, \partial_j f_n \,\mathrm{d}x = \lim_{n\to\infty} -\int_U f_n\, \partial_j \varphi \,\mathrm{d}x = -\int_U f\, \partial_j \varphi \,\mathrm{d}x.$$

The first claim in a) has been shown, the second one directly follows.

b) 1) Let $f \in D_j(U)$, $x \in U$, and $0 < \varepsilon < \mathrm{d}(x, \partial U) =: \delta$. Take $\eta \in (0, \frac{1}{2}(\delta - \varepsilon))$. Choose a point $x_0 \in \overline{B}(x, \eta)$. The function $y \mapsto \varphi_{\varepsilon,x}(y) := k_\varepsilon(x - y)$ belongs to $C_c^\infty(U)$ and has the support $\overline{B}(x, \varepsilon) \subseteq \overline{B}(x_0, \varepsilon + \eta) \subseteq U$. Observe that $|\partial_j \varphi_{\varepsilon,x}|$ is bounded by the integrable function $\|\partial_j k_\varepsilon\|_\infty \mathbb{1}_{\overline{B}(x_0,\varepsilon+\eta)}$. The differentiation theorem 3.18 of Analysis 3 and Definition 4.14 now imply

$$\begin{aligned}(\partial_j G_\varepsilon f)(x) &= \int_U \partial_{x_j} k_\varepsilon(x-y) f(y) \,\mathrm{d}y = -\int_U (\partial_j \varphi_{\varepsilon,x})(y) f(y) \,\mathrm{d}y \\ &= \int_U \varphi_{\varepsilon,x}(y)\, \partial_j f(y) \,\mathrm{d}y = (G_\varepsilon \partial_j f)(x).\end{aligned}$$

2) Let also $f, \partial_j f \in L^p_{\mathrm{loc}}(U)$. Fix a compact set $K \subseteq U$ and $\overline{\varepsilon} \in (0, \mathrm{dist}(K, \partial U))$. Take $\varepsilon \in (0, \overline{\varepsilon}]$ and $x \in K$. The integrand $y \mapsto k_\varepsilon(x -$

$y)f(y)$ in (4.1) then has support in the compact set $S = K + \overline{B}(0, \overline{\varepsilon}) \subseteq U$. From Proposition 4.13 and Example 2.6b) we deduce the limit

$$\mathbb{1}_K G_\varepsilon(f) = \mathbb{1}_K G_\varepsilon(\mathbb{1}_S f) \longrightarrow \mathbb{1}_K \mathbb{1}_S f = \mathbb{1}_K f, \qquad \varepsilon \to 0,$$

in $L^p(U)$ since $\mathbb{1}_S f \in L^p(U)$. This means that the restrictions of $G_\varepsilon f$ tend to $f|_K$ in $L^p(K)$. Step 1) also yields that $\mathbb{1}_K \partial_j G_\varepsilon(f) = \mathbb{1}_K G_\varepsilon(\partial_j f)$. As above we then infer that $\partial_j G_\varepsilon(f) \to \partial_j f$ in $L^p(K)$ as $\varepsilon \to 0$. Assertion b) is shown because K was arbitrary. □

The above convergence result is a crucial tool when extending properties from classical to weak derivatives, see Proposition 4.20 below. We first use it to compute weak derivatives and to show the completeness of Sobolev spaces.

Example 4.18. a) Let $f \in C(\mathbb{R})$ be such that $f_\pm := f_{|\mathbb{R}_\pm}$ belong to $C^1(\mathbb{R}_\pm)$. Then f is an element of $D_1(\mathbb{R})$ with weak derivative

$$\partial_1 f = \left\{ \begin{array}{ll} f_+' & \text{on } \mathbb{R}_+, \\ f_-' & \text{on } \mathbb{R}_- \end{array} \right\} =: g.$$

For $f(x) = |x|$, we thus obtain $\partial_1 f = \mathbb{1}_{\mathbb{R}_+} - \mathbb{1}_{\mathbb{R}_-}$.

Proof. For every $\varphi \in C_c^\infty(\mathbb{R})$, integration by parts yields

$$\begin{aligned} \int_{\mathbb{R}} f\varphi' \, \mathrm{d}s &= \int_{-\infty}^0 f_- \varphi' \, \mathrm{d}s + \int_0^\infty f_+ \varphi' \, \mathrm{d}s \\ &= -\int_{-\infty}^0 f_-' \varphi \, \mathrm{d}s + f_- \varphi \Big|_{-\infty}^0 - \int_0^\infty f_+' \varphi \, \mathrm{d}s + f_+ \varphi' \Big|_0^\infty \\ &= -\int_{\mathbb{R}} g\varphi \, \mathrm{d}s, \end{aligned}$$

since $f_+(0) = f_-(0)$ by the continuity of f. □

b) The function $f = \mathbb{1}_{\mathbb{R}_+}$ does not belong to $D_1(\mathbb{R})$.

Proof. Assume there would exist the weak derivative $g = \partial_1 f \in L^1_{\mathrm{loc}}(\mathbb{R})$. For every $\varphi \in C_c^\infty(\mathbb{R})$ we then obtain

$$\int_{\mathbb{R}} g\varphi \, \mathrm{d}s = -\int_{\mathbb{R}} \mathbb{1}_{\mathbb{R}_+} \varphi' \, \mathrm{d}s = -\int_0^\infty \varphi' \, \mathrm{d}s = \varphi(0).$$

Taking φ with $\operatorname{supp} \varphi \subseteq (0, \infty)$, we deduce from Lemma 4.15 that $g = 0$ on $(0, \infty)$. Similarly, it follows that $g = 0$ on $(-\infty, 0)$. Hence, $g = 0$ in $L^1_{\mathrm{loc}}(\mathbb{R})$ and so $\varphi(0) = 0$ for all $\varphi \in C_c^\infty(\mathbb{R})$ by the above identity in display. This statement is wrong. □

c) Let $U = B(0,1) \subseteq \mathbb{R}^m$, $p \in [1, \infty)$ and $\beta \in (1 - \frac{m}{p}, 1]$. Set $f(x) = |x|_2^\beta$ for $0 < |x|_2 < 1$ and $g_j(x) = \beta x_j |x|_2^{\beta - 2}$ for $0 < |x|_2 < 1$ and $j \in \{1, \dots, m\}$, as well as and $f(0) = g_j(0) = 0$. Then $f \in W^{1,p}(B(0,1))$ and $\partial_j f = g_j$. Observe that f is unbounded and has no continuous extension at $x = 0$ if $\beta < 0$ (which is admitted if $m > p$).

Proof. The functions $r \mapsto r^{\beta p} r^{m-1}$ and $r \mapsto r^{(\beta - 1)p} r^{m-1}$ are integrable on $(0,1)$ since $\beta > 1 - \frac{m}{p}$. Using polar coordinates, we infer that $f, g_j \in L^p(B(0,1))$ for $j \in \{1, \dots, m\}$, see Beispiel 5.3 in Analysis 3. We introduce the regularized functions $f_n(x) = (n^{-2} + |x|_2^2)^{\beta/2}$ for $n \in \mathbb{N}$ and $x \in B(0,1)$. Then $f_n \in C_b^1(B(0,1)) \hookrightarrow W^{1,p}(B(0,1))$ and $\partial_j f_n(x) = \beta x_j \, (n^{-2} + |x|_2^2)^{\frac{\beta}{2}-1}$. Observe that $f_n(x)$ and $\partial_j f_n(x)$ tend to $f(x)$ and $g_j(x)$ for $x \neq 0$ as $n \to \infty$, respectively. Moreover, $|f_n| \le \max\{|f|, 2^{\beta/2}\}$ and $|\partial_j f_n| \le |g_j|$ a.e. on $B(0,1)$ for all n and j. Dominated convergence then yields the limits $f_n \to f$ and $\partial_j f_n \to g_j$ in $L^p(B(0,1))$ as $n \to \infty$. The claim thus follows from Lemma 4.17. □

PROPOSITION 4.19. *Let $1 \le p \le \infty$ and $U \subseteq \mathbb{R}^m$ be open. Then $W^{1,p}(U)$ is a Banach space. It is separable if $1 \le p < \infty$. Moreover, $W^{1,2}(U) =: H^1(U)$ is a Hilbert space endowed with the scalar product*

$$(f|g)_{1,2} = \int_U f\overline{g} \, \mathrm{d}x + \sum_{j=1}^m \int_{\mathbb{R}^m} \partial_j f \, \overline{\partial_j g} \, \mathrm{d}x.$$

PROOF. Let (f_n) be a Cauchy sequence in $W^{1,p}(U)$. The sequences (f_n) and $(\partial_j f_n)$ thus are Cauchy in $L^p(U)$ for every $j \in \{1, \dots, m\}$, and hence have limits f and g_j in $L^p(U)$, respectively. Lemma 4.17 now implies that $f \in W^{1,p}(U)$ and $g_j = \partial_j f$ for all j; i.e., $W^{1,p}(U)$ is a Banach space. Using Remark 2.11, we then deduce from Remark 4.16d) that $W^{1,p}(U)$ is isometrically isomorphic to a closed subspace of $L^p(U)^{1+m}$, so that the separability for $p < \infty$ follows from Example 1.55 and Exercise 6.2. The last assertion is clear. □

Under suitable regularity and integrability assumptions, weak derivatives also satisfy the product and substitution rules. Here we only present the basic version of the product rule that is needed below.

PROPOSITION 4.20. *Let $p \in [1, \infty]$, $U \subseteq \mathbb{R}^m$ be open, $f \in W^{1,p}(U)$, and $g \in W^{1,p'}(U)$. Then the product fg belongs to $W^{1,1}(U)$ and satisfies $\partial_j(fg) = (\partial_j f)g + f\partial_j g$ for every $j \in \{1, \dots, m\}$.*

PROOF. Hölder's inequality implies that fg, $(\partial_j f)g$, and $f\partial_j g$ belong to $L^1(U)$ for all $j \in \{1, \dots, m\}$. Let $\varphi \in C_c^\infty(U)$ and $K := \operatorname{supp} \varphi$ Set $f_n = G_{1/n} f \in C^\infty(U) \cap L^p(U)$ and $g_n = G_{1/n} g \in C^\infty(U) \cap L^{p'}(U)$ for $n \in \mathbb{N}$. We let $1/n < \operatorname{dist}(K, \partial U)$.

1) Let $p \in (1, \infty)$ so that $p' \in (1, \infty)$. By Proposition 4.13 the functions f_n converge to f in $L^p(U)$ and g_n to g in $L^{p'}(U)$ as $n \to \infty$. Lemma 4.17 further yields the limits $\partial_j f_n \to \partial_j f$ and $\partial_j g_n \to \partial_j g$ in $L^p(K)$ as $n \to \infty$. Using (4.4), (4.2) and the product rule for C^1–functions, we now conclude that

$$\begin{aligned}\int_U fg \, \partial_j \varphi \, \mathrm{d}x &= \lim_{n\to\infty} \int_U f_n g_n \, \partial_j \varphi \, dx = -\lim_{n\to\infty} \int_U ((\partial_j f_n) g_n + f_n \partial_j g_n) \, \varphi \, \mathrm{d}x \\ &= -\int_U ((\partial_j f) g + f\partial_j g) \, \varphi \, \mathrm{d}x.\end{aligned}$$

Hence, fg has the weak derivative $\partial_j(fg) = g\partial_j f + f\partial_j g$.

2) Let $p = 1$. The above used convergence in $L^{p'}(U)$ may now fail since $p' = \infty$. The functions f_n and $\partial_j f_n$ still converge to f and $\partial_j f$ in $L^1(U)$. Passing to a subsequence, we can thus assume that they converge pointwise a.e. and that $|f_n|, |\partial_j f_n| \le h$ a.e. for some $h \in L^1(U)$ and all $n \in \mathbb{N}$ and $j \in \{1, \dots, m\}$. Proposition 4.13 and Lemma 4.17 imply that $\|g_n\|_\infty \le \|g\|_\infty$ and $\sup_{x\in K} |\partial_j g_n| \le \|\partial_j g\|_\infty$. Since $L^\infty(U) \subseteq L^1_{\text{loc}}(U)$, we deduce from Lemma 4.17 that after passing to subsequences (g_n) and $(\partial_j g_n)$ tend pointwise a.e. on K to g respectively $\partial_j g$. Based on dominated convergence we can now show the assertion as in step 1). The case $p = \infty$ is treated analogously. □

We now establish an important density result. In the proof we first use a cut-off argument to obtain a compact support and then perform a mollification. The cut-off must be chosen so that the extra terms caused by the $W^{1,p}$–norm vanish in the limit.

THEOREM 4.21. *The space $C_c^\infty(\mathbb{R}^m)$ is dense in $W^{1,p}(\mathbb{R}^m)$ for all $p \in [1, \infty)$.*

PROOF. 1) Let $f \in W^{1,p}(\mathbb{R}^m)$. Take a map $\phi \in C^\infty(\mathbb{R})$ with $0 \le \phi \le 1$, $\phi = 1$ on $[0, 1]$, and $\phi = 0$ on $[2, \infty)$. Set $\varphi_n(x) = \phi\left(\frac{1}{n}|x|_2\right)$ for $n \in \mathbb{N}$ and $x \in \mathbb{R}^m$. We then have $\varphi_n \in C_c^\infty(\mathbb{R}^m)$, $0 \le \varphi_n \le 1$ and $\|\partial_j \varphi_n\|_\infty \le \|\phi'\|_\infty \frac{1}{n}$ for all $n \in \mathbb{N}$, as well as $\varphi_n(x) \to 1$ for all $x \in \mathbb{R}^m$ as $n \to \infty$. So the functions $\varphi_n f$ converge to f in $L^p(\mathbb{R}^m)$ as $n \to \infty$ by Lebesgue's convergence theorem with majorant $|f|$. Let $j \in \{1, \dots, m\}$. Proposition 4.20 further implies that

$$\begin{aligned}\|\partial_j(\varphi_n f - f)\|_p &= \|(\varphi_n \partial_j f - \partial_j f) + (\partial_j \varphi_n) f\|_p \\ &\le \|\varphi_n \partial_j f - \partial_j f\|_p + \tfrac{1}{n}\|\phi'\|_\infty \|f\|_p\end{aligned}$$

Again by Lebesgue, the right hand side tends to 0 as $n \to \infty$. Given $\varepsilon > 0$, we can thus fix an index $N \in \mathbb{N}$ such that $\|\varphi_N f - f\|_{1,p} \le \varepsilon$.

2) Proposition 4.13 implies that the maps $G_{\frac{1}{n}}(\varphi_N f)$ belong to $C_c^\infty(\mathbb{R}^m)$ for all $n \in \mathbb{N}$ and that $G_{\frac{1}{n}}(\varphi_N f) \to \varphi_N f$ in $L^p(\mathbb{R}^m)$ as $n \to \infty$. From Lemma 4.17 we deduce that $\partial_j G_{\frac{1}{n}}(\varphi_N f) = G_{\frac{1}{n}} \partial_j(\varphi_N f)$ for all $\varepsilon > 0$, $n \in \mathbb{N}$, and $j \in \{1, \dots, m\}$, so that $\partial_j G_{\frac{1}{n}}(\varphi_N f)$ converges to $\partial_j(\varphi_N f)$ in $L^p(\mathbb{R}^m)$ as $n \to \infty$. For all sufficiently large n we finally estimate

$$\|G_{\frac{1}{n}}(\varphi_N f) - f\|_{1,p} \le \|G_{\frac{1}{n}}(\varphi_N f) - \varphi_N f\|_{1,p} + \|\varphi_N f - f\|_{1,p} \le 2\varepsilon. \quad \square$$

4.3. The open mapping theorem and invertibility

The invertibility of $T \in \mathcal{B}(X, Y)$ means that for each $y \in Y$ there is a unique solution $x = T^{-1}y$ of the equation $Tx = y$ which depends continuously on y. We start with a few simple properties of invertible operators and then establish the automatic continuity of T^{-1} in Banach spaces.

Lemma 4.22. *Let X, Y, and Z be normed vector spaces, $T \in \mathcal{B}(X,Y)$ and $S \in \mathcal{B}(Y,Z)$ be invertible. Then $ST \in \mathcal{B}(X,Z)$ is invertible with the inverse $T^{-1}S^{-1} \in \mathcal{B}(Z,X)$.*

Proof. The operators ST and $T^{-1}S^{-1}$ are continuous and linear. Moreover, $T^{-1}S^{-1}ST = I_X$ and $STT^{-1}S^{-1} = I_Z$. □

Lemma 4.23. *Let Z be a Banach space and $z_j \in Z$, $j \in \mathbb{N}_0$, satisfy $s := \sum_{j=0}^{\infty} \|z_j\| < \infty$. Then the partial sums $S_n = \sum_{j=0}^{n} z_j$ converge in Z as $n \to \infty$. Their limit is denoted by $\sum_{j=0}^{\infty} z_j$ and has norm less or equal s.*

Proof. Let $n > m$ in $\mathbb{N}_0$. Then $\|S_n - S_m\| \le \sum_{j=m+1}^{n} \|z_j\|$ tends to 0 as $n, m \to \infty$. Since Z is Banach space, the sequence (S_n) has a limit S with $\|S\| = \lim_{n\to\infty} \|S_n\| \le s$. □

We next show that small perturbations $T+S$ of an invertible operator T are again invertible and that the inverse is given by the *Neumann series*. The smallness condition below is sharp as shown by the example $T = I$ and $S = -I$. The basic idea is the formula $T+S = (I+ST^{-1})T$ which suggests to proceed as in the case of the geometric series.

Proposition 4.24. *Let X be Banach space, Y be a normed vector space and $T, S \in \mathcal{B}(X,Y)$. Assume that T is invertible and that $\|S\| < \|T^{-1}\|^{-1}$. Then $S+T$ is invertible and*

$$(S+T)^{-1} = \sum_{n=0}^{\infty} (-T^{-1}S)^n T^{-1} \qquad \text{(convergence in } \mathcal{B}(Y,X)\text{)},$$

$$\|(S+T)^{-1}\| \le \frac{\|T^{-1}\|}{1 - \|T^{-1}S\|}.$$

In particular, the set of invertible operators is open in $\mathcal{B}(X,Y)$.

Proof. We have $q := \|T^{-1}S\|_{\mathcal{B}(X)} < 1$ by assumption, and hence $\sum_{n=0}^{\infty} \|(T^{-1}S)^n\| \le 1/(1-q)$. Proposition 2.5 says that $\mathcal{B}(X)$ is a Banach space, and so Lemma 4.23 yields the convergence of

$$R := \sum_{n=0}^{\infty} (-T^{-1}S)^n T^{-1} = \sum_{n=0}^{\infty} T^{-1}(-ST^{-1})^n$$

in $\mathcal{B}(Y,X)$ with norm $\|R\| \le \|T^{-1}\|/(1-q)$. Moreover,

$$\begin{aligned} R(S+T) &= \sum_{n=0}^{\infty} (-T^{-1}S)^n (T^{-1}S + I) \\ &= -\sum_{j=1}^{\infty} (-T^{-1}S)^j + \sum_{n=0}^{\infty} (-T^{-1}S)^n = I, \\ (S+T)R &= \sum_{n=0}^{\infty} (ST^{-1} + I)(-ST^{-1})^n = I. \end{aligned}$$

□

The next example shows that in general the inverse of a bijective bounded operator is not continuous.

Example 4.25. Let $X = c_{00}$ with $\|\cdot\|_p$ and $Tx = (k^{-1}x_k)$. The operator T belongs to $\mathcal{B}(X)$ and is bijective with inverse $T^{-1}y = (ky_k)$. But $T^{-1} : c_{00} \to c_{00}$ is not continuous by Example 2.1. ♢

The following concept is useful in this context.

Definition 4.26. *Let M and M' be metric spaces. A map $f : M \to M'$ is called* open *if the image $f(O)$ is open in M' for each open $O \subseteq M$.*

Remark 4.27. a) A bijective map $f : M \to M'$ is open if and only if $f(O) = \{y \in M' \,|\, f^{-1}(y) \in O\} = (f^{-1})^{-1}(O)$ is open in M' for each open $O \subseteq M$ if and only if f^{-1} is continuous. (See Proposition 1.24.)

b) In Example 4.25 the maps $T, T^{-1} : c_{00} \to c_{00}$ are bijective and linear, T is continuous but not open, and T^{-1} is open but not continuous, due to part a). Observe that c_{00} is not a Banach space. ♢

We now prove the *open mapping theorem* which is the second fundamental principle of functional analysis. In Banach spaces it gives the continuity of an inverse for free. This is a very useful fact in many situations since often one does not have a formula for the inverse.

Theorem 4.28. *Let X and Y be Banach spaces and $T \in \mathcal{B}(X, Y)$ be surjective. Then T is open. If T is even bijective, then it is invertible.*

Proof. The second assertion is a consequence of the first one because of Remark 4.27. So let T be surjective. We abbreviate $U_r = B_X(0, r)$ and $V_r = B_Y(0, r)$ for every $r > 0$.

Claim A). There is a radius $\varepsilon > 0$ with $V_\varepsilon \subseteq TU_2$.

Assume that claim A) has been shown. Let $O \subseteq X$ be open, $x \in O$, and $y = Tx$ be an arbitrary element of TO. Then there is a number $r > 0$ with $B_X(x, r) \subseteq O$. From A) and the linearity of T we deduce

$$\begin{aligned} B_Y(y, \tfrac{\varepsilon r}{2}) &= y + \tfrac{r}{2} V_\varepsilon \subseteq Tx + \tfrac{r}{2} TU_2 = \{T(x + \tfrac{r}{2} z) \,|\, z \in U_2\} \\ &= TB_X(x, r) \subseteq TO; \end{aligned}$$

i.e., T is open.

Proof of A). The surjectivity of T yields $Y = \bigcup_{n=1}^\infty \overline{TU_n}$. Since Y is complete, Corollary 4.2 gives $N \in \mathbb{N}$, $y_0 \in Y$ and $r > 0$ such that

$$B_Y(y_0, r) \subseteq \overline{TU_N} = \overline{(2N)\, TU_{\frac{1}{2}}} = 2N\, \overline{TU_{\frac{1}{2}}},$$

where we also use the linearity of T and the characterization of closures from Proposition 1.17. Setting $z_0 = \frac{1}{2N} y_0$ and $\varepsilon = \frac{r}{2N}$, we deduce

$$B_Y(z_0, \varepsilon) = \tfrac{1}{2N}\{y_0 + w \,|\, w \in V_r\} = \tfrac{1}{2N} B_Y(y_0, r) \subseteq \overline{TU_{\frac{1}{2}}}$$

Observe that $TU_{\frac{1}{2}}$ is convex and $TU_{\frac{1}{2}} = -TU_{\frac{1}{2}}$. By approximation, these facts also hold for $\overline{TU_{\frac{1}{2}}}$, cf. Corollary 1.18. It follows

$$V_\varepsilon = B_Y(z_0, \varepsilon) - z_0 \subseteq \overline{TU_{\frac{1}{2}}} - \overline{TU_{\frac{1}{2}}} = \{x+y = 2(\tfrac{1}{2}x + \tfrac{1}{2}y) \,|\, x, y \in \overline{TU_{\frac{1}{2}}}\}$$
$$\subseteq 2\,\overline{TU_{\frac{1}{2}}} = \overline{TU_1}.$$

For later use, we note that the above inclusion yields

$$V_{\alpha\varepsilon} = \alpha V_\varepsilon \subseteq \alpha\,\overline{TU_1} = \overline{TU_\alpha} \qquad \text{for all } \alpha > 0. \tag{4.5}$$

Claim A) and thus the theorem then follow from the next assertion.

Claim B). We have $\overline{TU_1} \subseteq TU_2$.

Proof of B) Let $y \in \overline{TU_1}$. There is a vector $x_1 \in U_1$ with $\|y - Tx_1\| < \varepsilon/2$; i.e., $y - Tx_1 \in V_{\varepsilon/2}$ and so $y - Tx_1 \in \overline{TU_{1/2}}$ due to (4.5). Similarly, we obtain a point

$$x_2 \in U_{1/2} \qquad \text{with} \quad y - T(x_1 + x_2) = y - Tx_1 - Tx_2 \in V_{\frac{\varepsilon}{4}} \subseteq \overline{TU_{\frac{1}{4}}}\,.$$

Inductively, we find elements x_n of $U_{2^{1-n}}$ satisfying

$$y - T(x_1 + \cdots + x_n) \in V_{\varepsilon 2^{-n}} \tag{4.6}$$

for each $n \in \mathbb{N}$. Since X is a Banach space and $\sum_{n=1}^\infty \|x_n\| < 2$, by Lemma 4.23 there exists the limit $x := \sum_{n=1}^\infty x_n$ in U_2. Letting $n \to \infty$ in (4.6), we thus obtain

$$y = \sum_{n=1}^\infty Tx_n = Tx \;\in\; TU_2. \qquad \square$$

We collect important consequences of the open mapping theorem.

Corollary 4.29. *Let $\|\cdot\|$ and $|||\cdot|||$ be complete norms on the vector space X such that $\|x\| \le c\,|||x|||$ for some $c > 0$ and all $x \in X$. Then these norms are equivalent.*

Proof. The map $I : (X, |||\cdot|||) \to (X, \|\cdot\|)$ is continuous by assumption, and it is linear and bijective. Due to the completeness and Theorem 4.28, the map $I^{-1} : (X, \|\cdot\|) \to (X, |||\cdot|||)$ is also continuous. The assertion now follows from $I^{-1}x = x$. $\square$

We add a simple example showing that one needs completeness here.

Example 4.30. On $X = C^1([0,1])$ we have the complete norm $|||f||| = \|f\|_\infty + \|f'\|_\infty$ and the non complete norm $\|f\| = \|f\|_\infty \le |||f|||$, which are not equivalent. For instance, the functions $f_n(t) = \sin(nt)$ satisfy $\|f_n\|_\infty \le 1$, but $|||f_n||| \ge |f_n'(0)| = n$. $\Diamond$

The next result will be improved at the end of the next chapter.

Corollary 4.31. *Let X and Y be Banach spaces and $T \in \mathcal{B}(X, Y)$ be injective. The following assertions are equivalent.*

a) The operator $T^{-1} : R(T) \to X$ is continuous.

b) There is a constant $c > 0$ with $\|Tx\| \ge c\,\|x\|$ for every $x \in X$.

c) The range $R(T)$ is closed.

Proof. Let statement a) be true. For all $x \in X$ we then compute $\|x\| = \|T^{-1}Tx\| \leq \|T^{-1}\| \, \|Tx\|$, so that b) is shown.

The implication 'b)⇒c)' was proved in Remark 2.11.

Let c) hold. Corollary 1.13 then says that $(R(T), \|\cdot\|_Y)$ is Banach space. Since $T : X \to R(T)$ is bounded and bijective, assertion a) follows from Theorem 4.28. □

Proposition 4.32. *Let X be a Banach space and $X = Y \oplus Z$. Then the following assertions hold.*

a) $X \cong Y \times Z$.

b) The projection P with $R(P) = Y$ and $N(P) = Z$ is continuous (cf. Remark 2.17).

Proof. a) Since Y and Z are Banach spaces, their product $Y \times Z$ is a Banach space for the norm $\|(y,z)\| = \|y\| + \|z\|$, see Paragraph 2.2A). Further, the map $T : Y \times Z \to X$; $T(y,z) = y + z$, is linear, continuous (since $\|T(y,z)\| \leq \|y\| + \|z\| = \|(y,z)\|$) and bijective (since $X = Y \oplus Z$). Theorem 4.28 now yields part a).

b) Let $x = y + z \in Y \oplus Z$. Then $Px = y$. The operator T from a) satisfies $(y,z) = T^{-1}x$. Assertion a) now implies

$$\|Px\| = \|y\| \leq \|y\| + \|z\| = \|T^{-1}x\|_{Y\times Z} \leq \|T^{-1}\| \, \|x\|. \qquad \square$$

We have thus shown that in Banach spaces direct sums and Cartesian products are essentially the same.

CHAPTER 5

Duality

Let X be a normed vector space. We then have the Banach space $X^\star = \mathcal{B}(X, \mathbb{F})$ and the 'duality pairing'

$$X \times X^\star \to \mathbb{F}; \quad (x, x^\star) \mapsto x^\star(x) =: \langle x, x^\star \rangle = \langle x, x^\star \rangle_X. \tag{5.1}$$

This map is linear in both components, and it is continuous since $|\langle x, x^\star \rangle| \le \|x\|_X \, \|x^\star\|_{X^\star}$, cf. Remark 3.2d). To some extent, it replaces the scalar product in non-Hilbertian Banach spaces, though the quality of this replacement depends on the properties of the space.

We first determine the dual space $X^\star$ if $X = L^p(\mu)$ for $p \in [1, \infty)$, which is of course crucial to apply the theory below. In the main part of the chapter, we establish the third and fourth fundamental principle of functional analysis and discuss their consequences. In the last section we use duality theory for a better understanding of mapping properties of linear operators.

5.1. The duals of sequence and Lebesgue spaces

We first look at the simpler case of the sequence spaces. Set $X_p = \ell^p$ for $1 \le p < \infty$ and $X_\infty = c_0$ for $p = \infty$. By (1.2) the exponent $p' \in [1, \infty]$ is given by $\frac{1}{p} + \frac{1}{p'} = 1$. Given $y \in \ell^{p'}$, we define the function

$$\Phi_p(y) : \; X_p \to \mathbb{F}; \quad \Phi_p(y)(x) = \sum_{k=1}^{\infty} x_k y_k.$$

Hölder's inequality shows that this series converges absolutely and $|\Phi_p(y)(x)| \le \|x\|_p \, \|y\|_{p'}$. Since $\Phi_p(y)$ is linear, it is contained in $X_p^\star$ with $\|\Phi_p(y)\|_{X_p^\star} \le \|y\|_{p'}$. As a result, the mapping

$$\Phi_p : \ell^{p'} \to X_p^\star; \quad \langle x, \Phi_p(y) \rangle = \sum_{k=1}^{\infty} x_k y_k \qquad (\forall\, x \in X_p, \; y \in \ell^{p'}), \tag{5.2}$$

is contractive, and it is clearly linear. Since $\Phi_p(y)(e_k) = y_k$ for all $k \in \mathbb{N}$, the operator Φ_p is injective.

Proposition 5.1. *Equation (5.2) defines an isometric isomorphism $\Phi_p : \ell^{p'} \to X_p^\star$ for all $p \in [1, \infty]$. We thus obtain for each functional $x^\star \in X_p^\star$ exactly one sequence $y \in \ell^{p'}$ such that $\langle x, x^\star \rangle = \sum_k x_k y_k$ for all $x \in X_p$, where $\|y\|_{p'} = \|x^\star\|_{X_p^\star}$. Via this isomorphism,*

$$c_0^\star \cong \ell^1, \quad (\ell^1)^\star \cong \ell^\infty, \quad \textit{and} \quad (\ell^p)^\star \cong \ell^{p'} \;\; \textit{for} \;\; 1 < p < \infty.$$

Proof. In view of the above considerations it remains to show the surjectivity of Φ_p and that $\|\Phi_p y\|_{X_p^\star} \geq \|y\|_{p'}$ for all $y \in \ell^{p'}$. We restrict ourselves to the case $p = 1$, where $p' = \infty$. (The remainder can be proved similarly, see Theorem 5.4 and Exercise 11.3.)

Let $x^\star \in (\ell^1)^\star$. We define $y_k = x^\star(e_k)$ for every $k \in \mathbb{N}$, and set $y = (y_k)$. Since $|y_k| \leq \|x^\star\| \, \|e_k\|_1 = \|x^\star\|$ for all $k \in \mathbb{N}$, the sequence y belongs to ℓ^∞ and $\|y\|_\infty \leq \|x^\star\|$. Equation (5.2) also implies that $\Phi_1(y)(e_k) = y_k = x^\star(e_k)$. Since $x^\star$ and $\Phi_1(y)$ are linear, we arrive at $\Phi_1(y)(x) = x^\star(x)$ for all $x \in c_{00}$. Using continuity and the density of c_{00} in ℓ^1, see Proposition 1.31, we conclude that $\Phi_1(y) = x^\star$ and so Φ_1 is bijective. It also follows that $\|y\|_\infty \leq \|x^\star\| = \|\Phi_1 y\|$ for all $y \in \ell^\infty$. $\square$

We will see in Example 5.15 that the dual of ℓ^∞ is not isomorphic to ℓ^1. Observe that the right hand side in (5.2) does not depend on p and that it coincides with the scalar product in ℓ^2 except for the complex conjugation. The same is true in the next, more general case.

Let $(S, \mathcal{A}, \mu)$ be a measure space and $p \in [1, \infty)$. We define the map

$$\Phi_p(g) : L^p(\mu) \to \mathbb{F}; \quad \Phi_p(g)(f) = \int_S f g \, \mathrm{d}\mu. \tag{5.3}$$

for each fixed $g \in L^{p'}(\mu)$. Hölder's inequality yields that

$$\Phi_p(g) \in L^p(\mu)^\star \quad \text{and} \quad \|\Phi_p(g)\|_{X^\star} \leq \|g\|_{p'},$$

cf. Example 2.8. Consequently, $\Phi_p : L^{p'}(\mu) \to L^p(\mu)^\star$ is linear and contractive. To show that Φ_p is in fact an isometric isomorphism, we need further preparations.

A map $\nu : \mathcal{A} \to \mathbb{F}$ is called an $\mathbb{F}$-*valued measure*[1] if

$$\exists \sum_{n=1}^\infty \nu(A_n) = \nu\Big(\bigcup_{n=1}^\infty A_n\Big) \text{ in } \mathbb{F}$$

for all pairwise disjoint sets A_n in $\mathcal{A}$ for $n \in \mathbb{N}$. Since then $\nu(\emptyset) = \nu(\emptyset \cup \emptyset \cup \dots) = \sum_{n=1}^\infty \nu(\emptyset)$ in $\mathbb{F}$, we obtain that $\nu(\emptyset) = 0$. A measure μ in the sense of Paragraph 1.2C) is also called *positive measure*. (It is an $\mathbb{F}$–valued measure if and only if it is finite.) An $\mathbb{F}$-valued measure ν is called μ*–continuous* if one has $\nu(A) = 0$ for all $A \in \mathcal{A}$ with $\mu(A) = 0$. One then writes $\nu \ll \mu$.

Example 5.2. Let μ be a positive measure on $\mathcal{A}$ and $\rho \in L^1(\mu)$. For all $A \in \mathcal{A}$ we define

$$\nu(A) = \int_S \mathbb{1}_A \rho \, \mathrm{d}\mu = \int_A \rho \, \mathrm{d}\mu.$$

Then ν is an $\mathbb{F}$-valued measure on $\mathcal{A}$ with $\nu \ll \mu$. It is denoted by $\mathrm{d}\nu = \rho \, \mathrm{d}\mu$, and one calls ρ the *density* of ν with respect to μ.

[1]One usually says *signed measure* instead of '$\mathbb{R}$–valued measure'. The dual of $C(K)$ for a compact metric space can be identifieed with a space of $\mathbb{F}$–valued measures, see Satz II.2.5 in [**We**] or Appendix C of [**Co**].

We also assume that $\rho \geq 0$. In this case ν is a positive measure. Let $f : S \to \mathbb{F}$ be measurable with $f \geq 0$ or $\rho f \in L^1(\mu)$. We then obtain the equality $\int f \, \mathrm{d}\nu = \int f\rho \, \mathrm{d}\mu$, where $f \in L^1(\nu)$ in the second case.

PROOF. 1) Let $\rho \geq 0$. Then ν is a positive measure by Korollar 2.23 in Analysis 3, and it is finite since ρ is integrable. The definition of ν here means that characteristic functions f satisfy $\int f \, \mathrm{d}\nu = \int f\rho \, \mathrm{d}\mu$. By linearity this equation extends to simple functions f, and then by dominated or monotone convergence to integrable or non-negative ones.

2) A map $\rho \in L^1(\mu)$ can be written as $\rho = \rho_1 - \rho_2 + \mathrm{i}\rho_3 - \mathrm{i}\rho_4$ for integrable $\rho_j \geq 0$, Using part 1) one now easily checks that ν is an $\mathbb{F}$–valued measure. The definition of ν then implies its μ–continuity. □

The *Radon–Nikodym theorem* is a converse to the above example which is one of the fundamental results of measure theory: The seemingly very weak assumption of μ–continuity of ν already implies that ν has a density with respect to μ and is thus a very special $\mathbb{F}$–valued measure. We give a proof based on Riesz' Theorem 3.10, where we restrict ourselves to the main special case.

THEOREM 5.3. *Let $\mathcal{A}$ be a σ–algebra on S, μ be a σ–finite positive measure on $\mathcal{A}$, and ν be an $\mathbb{F}$–valued measure on $\mathcal{A}$ with $\nu \ll \mu$. Then there is a unique density $\omega \in L^1(\mu)$ such that $\mathrm{d}\nu = \omega \, \mathrm{d}\mu$. If ν is a positive measure, then $\omega \geq 0$.*

PROOF.[2] We show the theorem if μ and ν are positive and bounded. For the general case we refer to Satz VII.2.3 in [**El**] if $\mathbb{F} = \mathbb{R}$ and to Theorem C.7 in [**Co**] if $\mathbb{F} = \mathbb{C}$.

1) Consider the finite positive measure $\tau = \mu + \nu$ on $\mathcal{A}$. Observe that $\mu(A), \nu(A) \leq \tau(A)$ for all $A \in \mathcal{A}$. For simple functions $f = \sum_{j=1}^m a_k \mathbb{1}_{A_j}$ with an index $m \in \mathbb{N}$, values $a_j \in \mathbb{F}$ and sets $A_j \in \mathcal{A}$ (which are pairwise disjoint, without loss of generality), we deduce

$$\|f\|_{\mathcal{L}^2(\mu)}^2 = \sum_{j=1}^m |a_j|^2 \, \mu(A_j) \leq \sum_{j=1}^m |a_j|^2 \, \tau(A_j) = \|f\|_{\mathcal{L}^2(\tau)}^2.$$

Hölder's inequality further yields

$$\|f\|_{\mathcal{L}^1(\mu)} \leq \mu(S)^{1/2} \, \|f\|_{\mathcal{L}^2(\mu)} \leq \mu(S)^{1/2} \, \|f\|_{\mathcal{L}^2(\tau)},$$

see Proposition 1.35. If $\varphi = 0$ a.e. for τ, then also for μ. It thus follows $\|f\|_{L^1(\mu)} \leq \mu(S)^{1/2} \, \|f\|_{L^2(\tau)}$ for all simple functions f in $L^2(\tau)$. By approximation, each element f of $L^2(\tau)$ belongs to $L^1(\mu)$ and is bounded by $\|f\|_{L^1(\mu)} \leq \mu(S)^{1/2} \, \|f\|_{L^2(\tau)}$. We can then define the linear and continuous map

$$\varphi : L^2(\tau) \to \mathbb{F}; \quad \varphi(f) = \int_S f \, \mathrm{d}\mu.$$

[2]This proof was omitted in the lectures.

2) Theorem 3.10 now yields a function $\overline{g} \in L^2(\tau) \hookrightarrow L^1(\tau)$ such that

$$0 \le \mu(A) = \int_S \mathbb{1}_A \,\mathrm{d}\mu = \varphi(\mathbb{1}_A) = \int_S \mathbb{1}_A \overline{\overline{g}} \,\mathrm{d}\tau = \int_A g \,\mathrm{d}\tau$$

for all $A \in \mathcal{A}$; i.e, $\mathrm{d}\mu = g\,\mathrm{d}\tau$. To show that $g \ge 0$, we set $A_n = \{g \le -\frac{1}{n}\} \in \mathcal{A}$ for $n \in \mathbb{N}$. The above inequality implies that

$$0 \le \int_{A_n} g \,\mathrm{d}\tau \le -\frac{\tau(A_n)}{n},$$

and hence $\tau(A_n) = 0$ for all $n \in \mathbb{N}$. We deduce that $\{g < 0\} = \bigcup_n A_n$ is τ–null set, thus a μ–null set. Similarly, one sees that g is real–valued a.e.; and hence $g \ge 0$ a.e.. In the same way, one obtains a function $h \ge 0$ in $L^1(\tau)$ with $\mathrm{d}\nu = h\,\mathrm{d}\tau$.

3) Set $N = \{g = 0\}$. Since $\mu(N) = \int_N g\,\mathrm{d}\tau = 0$, the assumption $\nu \ll \mu$ yields that $\nu(N) = 0$. We now define the function

$$0 \le \omega(s) = \begin{cases} \frac{h(s)}{g(s)}, & s \in S \setminus N, \\ 0, & s \in N, \end{cases}$$

which is clearly measurable. For every $A \in \mathcal{A}$, we then compute

$$\nu(A) = \nu(A \cap N^c) = \int_{A\cap N^c} h \,\mathrm{d}\tau = \int_{A\cap N^c} \omega g \,\mathrm{d}\tau = \int_{A\cap N^c} \omega \,\mathrm{d}\mu = \int_A \omega \,\mathrm{d}\mu,$$

using step 2) and Example 5.2. This means that $\mathrm{d}\nu = \omega\,\mathrm{d}\mu$. Moreover, $\|\omega\|_1 = \int_S \omega\,\mathrm{d}\mu = \nu(S)$ is finite.

To show uniqueness, take another density $\widetilde{\omega} \in L^1(\mu)$ with $\nu(A) = \int_A \widetilde{\omega}\,\mathrm{d}\mu$ for all $A \in \mathcal{A}$. Arguing as for g, one sees that $\widetilde{\omega} \ge 0$. Set $B_n = \{\widetilde{\omega} \ge \omega + \frac{1}{n}\}$ for $n \in \mathbb{N}$. Because of

$$0 = \nu(B_n) - \nu(B_n) = \int_{B_n} (\omega - \widetilde{\omega})\,\mathrm{d}\mu \le -\frac{\mu(B_n)}{n},$$

we deduce the inequality $\widetilde{\omega} \le \omega$ μ–a.e. as in step 2). One analogously obtains $\omega \le \widetilde{\omega}$ μ–a.e., and so $\omega = \widetilde{\omega}$ in $L^1(\mu)$. □

We now represent the dual of $L^p(\mu)$ for $p \in [1,\infty)$ by $L^{p'}(\mu)$ via (5.3). This result is due to *F. Riesz* for $p>1$ and to Steinhaus for $p=1$.

THEOREM 5.4. *Let $1 \le p < \infty$ and $(S, \mathcal{A}, \mu)$ be a measure space which is σ–finite if $p = 1$. Then the map $\Phi_p : L^{p'}(\mu) \to L^p(\mu)^\star$ from (5.3) is an isometric isomorphism, and thus $L^p(\mu)^\star \cong L^{p'}(\mu)$ via*

$$\forall\, \varphi \in L^p(\mu)^\star \quad \exists!\, g \in L^{p'}(\mu) \quad \forall\, f \in L^p(\mu) : \ \langle f, \varphi\rangle_{L^p} = \int_S f g \,\mathrm{d}\mu.$$

PROOF. Set $X = L^p(\mu)$. It remains to prove that Φ_p is surjective and $\|\Phi_p(g)\|_{X^\star} \ge \|g\|_{p'}$ for all $g \in L^{p'}(\mu)$. We assume that $\mu(S) < \infty$ and that $1 < p < \infty$ and hence $p' = \frac{p}{p-1} \in (1,\infty)$. (The general case is treated in Satz VII.3.2 of [**El**] for $\mathbb{F} = \mathbb{R}$ and in Appendix B of [**Co**].)

1) Let $\varphi \in L^p(\mu)^\star$. Since $\mu(S) < \infty$, the characteristic function $\mathbb{1}_A$ belongs to $L^p(\mu)$ for each $A \in \mathcal{A}$. We define $\nu(A) = \varphi(\mathbb{1}_A)$.

Take sets $A_j \in \mathcal{A}$ with $A_j \cap A_k = \emptyset$ for all $j, k \in \mathbb{N}$ with $j \neq k$. We put $A = \bigcup_{j=1}^\infty A_j \in \mathcal{A}$ and $B_n = \bigcup_{j=1}^n A_j \in \mathcal{A}$ for all $n \in \mathbb{N}$. Observe that $\mathbb{1}_{B_n} \to \mathbb{1}_A$ pointwise as $n \to \infty$ and $0 \le \mathbb{1}_{B_n} \le \mathbb{1}_A \in L^p(\mu)$ for all $n \in \mathbb{N}$. Hence, $\mathbb{1}_{B_n}$ tends to $\mathbb{1}_A$ in $L^p(\mu)$ as $n \to \infty$ by the theorem of dominated convergence. We also have $\mathbb{1}_{B_n} = \mathbb{1}_{A_1} + \cdots + \mathbb{1}_{A_n}$ for all $n \in \mathbb{N}$ due to disjointness. Using the continuity and linearity of φ, we then conclude that

$$\nu(A) = \varphi(\mathbb{1}_A) = \lim_{n\to\infty} \varphi(\mathbb{1}_{B_n}) = \lim_{n\to\infty} \sum_{k=1}^n \varphi(\mathbb{1}_{A_k}) = \sum_{k=1}^\infty \nu(A_k).$$

As a result, ν is an $\mathbb{F}$–valued measure. If $\mu(A) = 0$, then $\mathbb{1}_A = 0$ in $L^p(\mu)$ and so $\nu(A) = \varphi(\mathbb{1}_A) = 0$; i.e., $\nu \ll \mu$.

The Radon–Nikodym Theorem 5.3 thus gives a map $g \in L^1(\mu)$ with

$$\varphi(\mathbb{1}_A) = \nu(A) = \int_S \mathbb{1}_A \, g \, \mathrm{d}\mu \qquad \text{for all } \; A \in \mathcal{A}.$$

The linearity of φ yields

$$\varphi(f) = \int_S f g \, \mathrm{d}\mu \quad \text{ for every simple function } \; f : S \to \mathbb{F}. \tag{5.4}$$

Since we only know that $g \in L^1(\mu)$, at first we can extend this equation to $f \in L^\infty(A)$ only. Indeed, by Satz 2.13 of Analysis 3 there are simple functions $f_n : S \to \mathbb{F}$ converging to f uniformly as $n \to \infty$. The integrals $\int f_n g \, \mathrm{d}\mu$ then tend to $\int f g \, \mathrm{d}\mu$ as $n \to \infty$ by (4.4). On the other hand, since $\mu(S) < \infty$ Proposition 1.35 implies that $f_n \to f$ in $L^p(\mu)$, and hence $\varphi(f_n) \to \varphi(f)$ as $n \to \infty$. As a result, (5.4) is satisfied by all functions f in $L^\infty(\mu)$.

2) To show $g \in L^{p'}(A)$, we set

$$h(s) = \begin{cases} 0, & g(s) = 0, \\ \frac{|g(s)|^{p'}}{g(s)}, & g(s) \neq 0. \end{cases}$$

Then h is measurable and

$$|h|^p = |g|^{p(p'-1)} = |g|^{p'} = gh.$$

Let $A_n = \{|g| \le n\} \in \mathcal{A}$ for $n \in \mathbb{N}$. The functions $\mathbb{1}_{A_n} g$ then belong to $L^\infty(\mu) \hookrightarrow L^{p'}(\mu)$ and $\mathbb{1}_{A_n} h$ to $L^\infty(\mu) \hookrightarrow L^p(\mu)$, again owing to $\mu(S) < \infty$. Employing (5.4) for $f = \mathbb{1}_{A_n} h$, we can now compute

$$\begin{aligned} \int_S \mathbb{1}_{A_n} |g|^{p'} \, \mathrm{d}\mu &= \int_S \mathbb{1}_{A_n} h g \, \mathrm{d}\mu = \varphi(\mathbb{1}_{A_n} h) \le \|\varphi\|_{X^\star} \, \|\mathbb{1}_{A_n} h\|_p \\ &= \|\varphi\| \Big[\int_S \mathbb{1}_{A_n} |h|^p \mathrm{d}\mu \Big]^{\frac{1}{p}} = \|\varphi\|_{X^\star} \Big[\int_S \mathbb{1}_{A_n} |g|^{p'} \mathrm{d}\mu \Big]^{\frac{1}{p}}, \end{aligned}$$

$$\Big[\int_S \mathbb{1}_{A_n}\, |g|^{p'}\, \mathrm{d}\mu\Big]^{\frac{1}{p'}} \le \|\varphi\|_{X^\star}.$$

As $n \to \infty$, Fatou's lemma yields that $g \in L^{p'}(\mu)$ and $\|g\|_{p'} \le \|\varphi\|_{X^\star}$.

3) Let $f \in L^p(\mu)$. There are simple functions f_n converging to f in $L^p(\mu)$ by Theorem 5.9 in Analysis 3. Using the continuity of φ, $g \in L^{p'}(\mu)$ and (4.4), we conclude that f fulfills (5.4). This means that $\Phi_p(g) = \varphi$ and Φ_p is surjective. The lower estimate $\|g\|_{p'} \le \|\varphi\|_{X^\star} = \|\Phi_p(g)\|_{X^\star}$ was shown in step 2). □

Usually one identifies $L^{p'}(\mu)$ with $L^p(\mu)^\star$ for $1 \le p < \infty$ and writes $\langle f, g\rangle = \int fg\,\mathrm{d}\mu$ for the duality, and analogously for the sequence spaces.

5.2. The extension theorem of Hahn-Banach

Many of the non-trivial properties of duality rely on the Hahn-Banach theorem proved below. By means of an extension process it produces 'tailor-made' continuous functionals. We start with the basic *order theoretic version* which will also yield a third variant of Hahn–Banach at the end of the section.

Let X be a vector space. A map $p : X \to \mathbb{R}$ is called *sublinear* if

$$p(\lambda x) = \lambda p(x) \quad \text{and} \quad p(x+y) \le p(x) + p(y)$$

for all $x, y \in X$ and $\lambda \ge 0$. A typical example is a seminorm.

Theorem 5.5. *Let X be a vector space with $\mathbb{F} = \mathbb{R}$, $p : X \to \mathbb{R}$ be sublinear, $Y \subseteq X$ be a linear subspace, and $\varphi_0 : Y \to \mathbb{R}$ be linear with $\varphi_0(y) \le p(y)$ for all $y \in Y$. Then there exists a linear map $\varphi : X \to \mathbb{R}$ satisfying $\varphi(y) = \varphi_0(y)$ for all $y \in Y$ and $\varphi(x) \le p(x)$ for all $x \in X$.*

Proof. 1) We define the set

$$\mathcal{M} = \{(Z,\psi) \,|\, Z \subseteq X \text{ is a linear subspace, } \psi : Z \to \mathbb{R} \text{ is linear with } Y \subseteq Z,\ \psi|_Y = \varphi_0,\ \psi \le p|_Z\},$$

which contains (Y, φ_0). On $\mathcal{M}$ we set $(Z,\psi) \preccurlyeq (Z',\psi')$ if $Z \subseteq Z'$ and $\psi'|_Z = \psi$. Straightforward calculations show that this relation is a partial order on $\mathcal{M}$ (i.e.; it is reflexive, antisymmetric and transitive).

Let $\mathcal{K}$ be a totally ordered non-empty subset of $\mathcal{M}$, which means that for $(Z,\psi), (Z',\psi') \in \mathcal{K}$ we have $(Z,\psi) \preccurlyeq (Z',\psi')$ or $(Z',\psi') \preccurlyeq (Z,\psi)$. We want to check that $U = \bigcup\{Z \,|\, (Z,\psi) \in \mathcal{K}\}$ is a linear subspace of X, that by setting $f(x) := \psi(x)$ for every $x \in U$ and any $(Z,\psi) \in \mathcal{K}$ with $x \in Z$ we define a linear map $f : U \to \mathbb{R}$, and that (U, f) belongs to $\mathcal{M}$ and is an upper bound for $\mathcal{K}$.

Take $x \in U$. There is a pair (Z,ψ) in $\mathcal{K}$ with $x \in Z$. If x also belongs to $\hat{Z}$ for some $(\hat{Z},\hat{\psi}) \in \mathcal{K}$, then we have $(\hat{Z},\hat{\psi}) \preccurlyeq (Z,\psi)$, for instance, and thus $\hat{\psi}(x) = \psi(x)$. Hence, $f : U \to \mathbb{R}$ is a well defined map. Next, pick $y \in U$ and $\alpha, \beta \in \mathbb{R}$. Choose $(Z',\psi') \in \mathcal{K}$ with $y \in Z'$ and, say,

$(Z,\psi) \preccurlyeq (Z',\psi')$. The linear combination $\alpha x + \beta y$ is then an element of $Z' \subseteq U$ and the linearity of ψ' yields

$$f(\alpha x + \beta y) = \psi'(\alpha x + \beta y) = \alpha\psi'(x) + \beta\psi'(y) = \alpha f(x) + \beta f(y).$$

Therefore f is a linear map on the subspace U. Similarly one checks the properties $f|_Y = \varphi_0$ and $f \le p|_U$ so that the pair (U,f) is an element of $\mathcal{M}$. By construction, we obtain $(Z,\psi) \preccurlyeq (U,f)$ for all $(Z,\psi) \in \mathcal{K}$.

Zorn's Lemma (see Theorem 1.2.7 of [**DS**]) now gives a maximal element (V,φ) in $\mathcal{M}$. We next show that $V = X$; i.e., the linear form φ has the required properties.

2) We suppose that $V \neq X$ and fix a vector $x_0 \in X \setminus V$. We use the linear subspace $\tilde{V} = V + \operatorname{lin}\{x_0\}$. Since $V \cap \operatorname{lin}\{x_0\} = \{0\}$, for each $x \in \tilde{V}$ we have unique elements $v \in V$ and $t \in \mathbb{R}$ with $x = v + tx_0$. We will construct a linear map $\tilde{\varphi} : \tilde{V} \to \mathbb{R}$ with $(V,\varphi) \preccurlyeq (\tilde{V},\tilde{\varphi}) \in \mathcal{M}$. Since $(V,\varphi) \neq (\tilde{V},\tilde{\varphi})$, this fact contradicts the maximality of (V,φ), so that $V = X$ and the theorem is shown. Let $u,w \in V$. We compute

$$\varphi(u) + \varphi(w) = \varphi(u+w) \le p(u+w) \le p(u+x_0) + p(w-x_0),$$
$$\varphi(w) - p(w-x_0) \le p(u+x_0) - \varphi(u).$$

There thus exists a number

$$\alpha \in \Big[\sup_{w\in V}(\varphi(w) - p(w-x_0)),\ \inf_{u\in V}(p(u+x_0) - \varphi(u))\Big].$$

We next introduce the linear map $\tilde{\varphi} : \tilde{V} \to \mathbb{R}$; $\tilde{\varphi}(x) = \varphi(v) + \alpha t$, using the decomposition $x = v + tx_0 \in \tilde{V}$.

For $y \in Y \subseteq V$, the definitions yield $\tilde{\varphi}(y) = \varphi(y) = \varphi_0(y)$. To show $\tilde{\varphi} \le p|_{\tilde{V}}$, take $x = v + tx_0 \in \tilde{V}$. For $t = 0$ and $x = v$, we have $\tilde{\varphi}(v) = \varphi(v) \le p(v) = p(v)$. By means of the definition of α and the sublinearity of p, we estimate

$$\tilde{\varphi}(x) = \varphi(v) + t\alpha \le \varphi(v) + t\big(p(\tfrac{1}{t}v + x_0) - \varphi(\tfrac{1}{t}v)\big) = p(v + tx_0) = p(x)$$

for $t > 0$ and inserting $u = \frac{1}{t}v$. For $t < 0$, we similarly compute

$$\tilde{\varphi}(x) = \varphi(v) + t\alpha \le \varphi(v) + t\big(\varphi(-\tfrac{1}{t}v) - p(-\tfrac{1}{t}v - x_0)\big) = p(v+tx_0) = p(x)$$

with $w = -\frac{1}{t}v$. As a result, we have shown that $\tilde{\varphi}(x) \le p(x)$ for $x \in \tilde{V}$, and hence $(\tilde{V},\tilde{\varphi})$ belongs to $\mathcal{M}$ as needed. $\square$

As a preparation for later results we describe how to pass from $\mathbb{C}$–linear functionals to $\mathbb{R}$–linear ones, and vice versa.

Lemma 5.6. *Let X be a normed vector space with $\mathbb{F} = \mathbb{C}$.*

a) Let $x^\star \in X^\star$. Set $\xi^\star(x) = \operatorname{Re} x^\star(x)$ for all $x \in X$. Then the map $\xi^\star : X \to \mathbb{R}$ is $\mathbb{R}$–linear with $\|\xi^\star\| := \sup_{\|x\|\le 1} |\xi^\star(x)| = \|x^\star\|$.

b) Let $\xi^\star : X \to \mathbb{R}$ be continuous and $\mathbb{R}$–linear. Set $x^\star(x) = \xi^\star(x) - \mathrm{i}\xi^\star(\mathrm{i}x)$ for all $x \in X$. Then the functional $x^\star$ belongs to $X^\star$ and satisfies $\|x^\star\| = \|\xi^\star\|$ and $\operatorname{Re} x^\star = \xi^\star$.

Proof. The $\mathbb{R}$–linearity of $\xi^\star$ in part a) can be shown in a straightforward way. Define $x^\star$ as in assertion b). Take $x \in X$ and $\alpha, \beta \in \mathbb{R}$. Using the $\mathbb{R}$–linearity of $\xi^\star$, we then compute

$$\begin{aligned} x^\star((\alpha + \mathrm{i}\beta)x) &= \xi^\star((\alpha x + \mathrm{i}\beta x) - \mathrm{i}\xi^\star((\mathrm{i}\alpha x - \beta x) \\ &= \alpha\xi^\star(x) + \beta\xi^\star(\mathrm{i}x) - \mathrm{i}\alpha\xi^\star(\mathrm{i}x) + \mathrm{i}\beta\xi^\star(x) \\ &= (\alpha + \mathrm{i}\beta)x^\star(x). \end{aligned}$$

Since its additivity is clear, the functional $x^\star$ in b) thus is $\mathbb{C}$–linear. In part b) we also have the identity $\xi^\star = \operatorname{Re} x^\star$. It implies the inequality $\|\xi^\star\| \le \|x^\star\|$ in statements a) and b).

To show the converse, pick $x \in X$ with $\|x\| = 1$. We set $\alpha = 1$ if $x^\star(x) = 0$ and $\alpha = x^\star(x)/\,|x^\star(x)|$ otherwise. Since $\|\alpha^{-1}x\| = 1$, the $\mathbb{C}$–linearity of $x^\star$ yields

$$0 \le |x^\star(x)| = x^\star(\tfrac{1}{\alpha}x) = \xi^\star(\tfrac{1}{\alpha}x) \le \|\xi^\star\|.$$

Taking the supremum over x with $\|x\| = 1$, we obtain $\|x^\star\| \le \|\xi^\star\|$ in a) and b). □

We can now establish the main version of the *Hahn–Banach theorem.* It allows to extend every bounded linear functional on a subspace to the full normed vector space X keeping its norm. Such a result is wrong for operators in general. For instance, an extension $P \in \mathcal{B}(\ell^\infty, c_0)$ of the identity $I : c_0 \to c_0$ would be a bounded projection onto c_0 in ℓ^∞, which does not exist by Example 2.18.

In Hilbert spaces even an operator version of the Hahn–Banach theorem follows from the projection Theorem 3.8. Let X and Z be Hilbert spaces, $Y \subseteq X$ be a linear subspace, and T_0 belong $\mathcal{B}(Y, Z)$. By Lemma 2.13, the operator T_0 has a linear extension T_1 to $\overline{Y}$ with the same norm. Let P be the orthogonal projection onto $\overline{Y}$. Then the operator $T = T_1 P \in \mathcal{B}(X, Z)$ extends T_0 and has the same norm.

We also point out that the proof of Hahn–Banach is highly nonconconstructive. Nevertheless under certain assumptions on X^* one obtains uniqueness in the next result, see Exercise 12.1.

Theorem 5.7. *Let X be a normed vector space, $Y \subseteq X$ be a linear subspace (endowed with the norm of X), and $y^\star \in Y^\star$. Then there exists a functional $x^\star \in X^\star$ such that $\langle y, x^\star \rangle = \langle y, y^\star \rangle$ for all $y \in Y$ and $\|x^\star\| = \|y^\star\|$.*

Proof. 1) Let $\mathbb{F} = \mathbb{R}$. Set $p(x) = \|y^\star\|\,\|x\|$ for all $x \in X$. The map p is sublinear and $y^\star(y) \le p(y)$ for all $y \in Y$. Theorem 5.5 yields a linear functional $x^\star : X \to \mathbb{R}$ with $x^\star|_Y = y^\star$ and $x^\star(x) \le p(x)$ for all $x \in X$. We further have

$$-x^\star(x) = x^\star(-x) \le p(-x) = p(x)$$

so that $|x^\star(x)| \le p(x) = \|y^\star\|\,\|x\|$ for all $x \in X$. As a result, $x^\star$ belongs to $X^\star$ with $\|x^\star\| \le \|y^\star\|$. The equality $\|x^\star\| = \|y^\star\|$ now follows from the estimate

$$\|x^\star\| \ge \sup_{\|y\|=1, y\in Y} |\langle y, x^\star\rangle| = \sup_{\|y\|=1, y\in Y} |\langle y, y^\star\rangle| = \|y^\star\|.$$

2) Let $\mathbb{F} = \mathbb{C}$. We consider X as a normed vector space $X_\mathbb{R}$ over $\mathbb{R}$ by restricting the scalar multiplication to real scalars; i.e., to the map $\mathbb{R}\times X \to X$; $(\alpha, x) \mapsto \alpha x$. Lemma 5.6a) first shows that the real part $\eta^\star = \operatorname{Re} y^\star$ belongs to $Y_\mathbb{R}^\star$ and $\|\eta^\star\| = \|y^\star\|$. Due to step 1), the functional $\eta^\star$ then has an extension $\xi^\star \in (X_\mathbb{R})^\star$ with $\|\xi^\star\| = \|\eta^\star\| = \|y^\star\|$. From Lemma 5.6b) we finally obtain a map $x^\star \in X^\star$ satisfying $\|x^\star\| = \|\xi^\star\| = \|y^\star\|$ and

$$\begin{aligned} x^\star(y) &= \xi^\star(y) - \mathrm{i}\xi^\star(\mathrm{i}y) = \operatorname{Re} y^\star(y) - \mathrm{i}\operatorname{Re} y^\star(\mathrm{i}y) \\ &= \operatorname{Re} y^\star(y) - \mathrm{i}\operatorname{Re}(\mathrm{i}y^\star(y)) = y^\star(y) \end{aligned}$$

for all $y \in Y$, where we used the $\mathbb{C}$–linearity of $y^\star$. □

Example 5.8. Let $Y = c \subseteq X = \ell^\infty$ and $y^\star(y) = \lim_{n\to\infty} y_n$ for $y \in Y$. Clearly, $y^\star$ belongs to $Y^\star$ and has norm 1. The Hahn–Banach theorem yields an extension $x^\star \in (\ell^\infty)^\star$ of $y^\star$ with norm 1. Note that $x^\star(y) = y^\star(y) = 0$ for $y \in c_0$. The functional $x^\star$ cannot be represented by a sequence $z \in \ell^1$ as in (5.2) since otherwise it would follow both $z \neq 0$ and

$$0 = \langle e_n, x^\star\rangle_{\ell^\infty} = \sum_{j=1}^\infty \delta_{nj} z_j = z_n \qquad \text{for all } \; n \in \mathbb{N}. \qquad \Diamond$$

The next result allows to distinguish between a closed linear subspace Y and a vector $x_0 \notin Y$. This fact will lead to the main corollaries of the Hahn–Banach theorem below.

Proposition 5.9. *Let X be a normed vector space, $Y \subsetneqq X$ be a closed linear subspace, and $x_0 \in X\setminus Y$. Then there exists a functional $x^\star \in X^\star$ such that $x^\star(y) = 0$ for all $y \in Y$, $x^\star(x_0) = \mathrm{d}(x_0, Y) = \inf_{y\in Y}\|x_0 - y\| > 0$, and $\|x^\star\| = 1$.*

Proof. We define the linear subspace $Z = Y + \operatorname{lin}\{x_0\}$ of X and the linear map $z^\star : Z \to \mathbb{F}$ by $z^\star(y + tx_0) = t\,\mathrm{d}(x_0, Y)$ for all $y \in Y$ and $t \in \mathbb{F}$, using that $Y \cap \operatorname{lin}\{x_0\} = \{0\}$. Clearly, $z^\star|_Y = 0$ and $z^\star(x_0) = \mathrm{d}(x_0, Y)$. We further compute

$$\|z^\star\| = \sup_{\|y+tx_0\|\le 1} |t| \inf_{\tilde{y}\in Y} \|x_0 - \tilde{y}\| \le \sup_{\|y+tx_0\|\le 1} \|tx_0 + y\| \le 1,$$

where we have chosen $\tilde{y} = -\frac{1}{t}y$ assuming that $t \neq 0$ without loss of generality. Recall that $\mathrm{d}(x_0, Y) > 0$ by Example 1.9 since Y is closed. Take vectors $y_n \in Y$ with $\|x_0 - y_n\| \to \mathrm{d}(x_0, Y)$. The properties of $z^\star$ now yield the limit

$$\|z^\star\| \ge \left| \left\langle \frac{1}{\|x_0 - y_n\|}(x_0 - y_n), z^\star \right\rangle \right| = \frac{\mathrm{d}(x_0, Y) - 0}{\|x_0 - y_n\|} \;\longrightarrow\; 1$$

as $n \to \infty$, so that $\|z^\star\| = 1$. The Hahn–Banach extension of $z^\star$ is the required map $x^\star \in X^\star$. □

The next easy consequence is used very often in analysis.

COROLLARY 5.10. *Let X be a normed vector space, $x, x_1, x_2 \in X$, and $D_\star \subseteq X^\star$ be dense. Then the following assertions hold.*

a) Let $x \neq 0$. Then there exists a functional $x^\star \in X^\star$ with $\langle x, x^\star\rangle = \|x\|$ and $\|x^\star\| = 1$.

b) Let $x_1 \neq x_2$. Then there is a map $x^\star \in X^\star$ with $x^\star(x_1) \neq x^\star(x_2)$.

c) $\|x\| = \max_{x^\star \in X^\star, \|x^\star\|_{X^\star} \le 1} |\langle x, x^\star\rangle| = \sup_{x^\star \in D_\star, \|x^\star\|_{X^\star} \le 1} |\langle x, x^\star\rangle|$.

PROOF. Assertion a) is a consequence of Proposition 5.9 with $Y = \{0\}$, and a) implies b) taking $x = x_1 - x_2$. For x in X we have $\sup_{\|x^\star\| \le 1} |\langle x, x^\star\rangle| \le \|x\|$, and so assertion c) follows from a) and an approximation argument. □

Interestingly, one has a precise formula for the functional in part a) above in some cases.

EXAMPLE 5.11. Let $1 \le p < \infty$ and $(S, \mathcal{A}, \mu)$ be a measure space, which is σ–finite if $p = 1$. Given $f \in L^p(\mu) \setminus \{0\}$, we set $g = \|f\|_p^{1-p} \bar{f}\, |f|^{p-2}\, \mathbb{1}_{\{f \neq 0\}}$ (i.e., $g = \frac{1}{\|f\|_2} \bar{f}$ for $p = 2$). One can then check that $g \in L^{p'}(\mu)$ with $\|g\|_{p'} = 1$ and $\langle f, g\rangle = \|f\|_p$. Let D be a dense subset of $L^{p'}(\mu)$. Corollary 5.10 and Theorem 5.4 further yield

$$\|f\|_p = \sup_{g \in D, \|g\|_{p'} \le 1} \Big| \int_S f g \,\mathrm{d}\mu \Big|. \tag{5.5}$$

We next combine the Hahn-Banach theorem with the principle of uniform boundedness.

COROLLARY 5.12. *Let X be a normed vector space and $M \subseteq X$. Then set M is bounded in X if and only if the sets $x^\star(M)$ are bounded in $\mathbb{F}$ for each $x^\star \in X^\star$.*

PROOF. The implication "⇒" is clear. To show the converse, set $T_x(x^\star) = \langle x, x^\star\rangle$ for each fixed $x \in M$ and all $x^\star \in X^\star$. By assumption, the functionals $T_x \in \mathcal{B}(X^\star, \mathbb{F})$ are pointwise bounded by $|T_x(x^\star)| \le c(x^\star) := \sup_{x \in M} |x^\star(x)| < \infty$ for all $x^\star \in X^\star$ and $x \in M$. Since $X^\star$ is a Banach space, Theorem 4.4 yields a constant C such that

$$C \ge \|T_x\| = \sup_{\|x^\star\| \le 1} |\langle x, x^\star\rangle| = \|x\|$$

for all $x \in M$, where we have used Corollary 5.10c). □

Density is often checked by means of the following result.

COROLLARY 5.13. *Let X be a normed vector space and $Y \subseteq X$ be a linear subspace. Then Y is not dense in X if and only if there exists a functional $x^\star \in X^\star \setminus \{0\}$ with $\langle y, x^\star\rangle = 0$ for all $y \in Y$.*

Proof. The implication "⇒" follows from Proposition 5.9. If Y is dense and $x^\star \in X^\star$ vanishes on Y, then $x^\star$ must be 0 by continuity. □

The next consequence of Hahn–Banach says that $X^\star$ is at least as large as X in terms of separability.

Corollary 5.14. *Let X be a normed vector space and $X^\star$ be separable. Then X is separable.*

Proof. By assumption and Exercise 6.2, we have a dense subset $\{x_n^\star \,|\, n \in \mathbb{N}\}$ in $\partial B_{X^\star}(0,1)$. There are vectors $y_n \in X$ with $\|y_n\| = 1$ and $|\langle y_n, x_n^\star\rangle| \geq \frac{1}{2}$ for every $n \in \mathbb{N}$. Set $Y = \operatorname{lin}\{y_n \,|\, n \in \mathbb{N}\}$. Suppose that $\overline{Y} \neq X$. Corollary 5.13 then yields a functional $x^\star \in X^\star$ such that $\|x^\star\| = 1$ and $\langle y, x^\star\rangle = 0$ for all $y \in Y$. There exists an index $j \in \mathbb{N}$ with $\|x^\star - x_j^\star\| \leq \frac{1}{4}$. We then deduce the contradiction

$$\frac{1}{2} \leq \left|\langle y_j, x_j^\star\rangle\right| = \left|\langle y_j, x_j^\star - x^\star\rangle\right| \leq \frac{1}{4}. \qquad \square$$

Example 5.15. The spaces c_0 and $\ell^1 = c_0^\star$ are separable, whereas $\ell^\infty = (\ell^1)^\star$ is not separable, see Example 1.55. (Since separability is preserved under isomorphisms by Exercise 6.2, we can omit here the isomorphisms from Proposition 5.1.) The above result implies that also $(\ell^\infty)^\star$ is not separable. In particular, ℓ^1 cannot be isomorphic to $(\ell^\infty)^\star$, cf. Proposition 5.1. ◊

So far we do not know whether there are non-zero bounded linear maps between two normed vector spaces $X \neq Y$. (If $X = Y$, the identity belongs to $\mathcal{B}(X)$.) We now can at least construct operators of *finite rank*.

Example 5.16. Let X and Y be normed vector spaces with $\dim X \geq n$, the vectors $x_1, \dots, x_n \in X$ be linearly independent, and $y_1, \dots, y_n \in Y$. For each $k \in \{1, \dots, n\}$, we put $Z_k = \operatorname{lin}\{x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n\}$ if $n \geq 2$ and $Z_1 = \{0\}$ if $n = 1$. Proposition 5.9 provides us with a functional $x_k^\star \in X^\star$ such that $x_k^\star|_{Z_k} = 0$ and $x_k^\star(x_k) = 1$; i.e., $\langle x_j, x_k^\star\rangle = \delta_{jk}$ for $j, k \in \{1, \dots, n\}$. We now define

$$Tx = \sum_{k=1}^{n} \langle x, x_k^\star\rangle\, y_k \qquad \in \operatorname{lin}\{y_1, \dots, y_n\} =: Y_0$$

for all $x \in X$. Clearly, $T \in \mathcal{B}(X, Y)$ with $\|T\| \leq \sum_{k=1}^n \|x_k^\star\| \|y_k\|$. Since also $Tx_j = y_j$, we have $R(T) = Y_0$. If $X = Y$ and $x_k = y_k$ for all k, we further deduce the identity

$$T^2 x = \sum_{j=1}^{n} \sum_{k=1}^{n} \langle x, x_k^\star\rangle \langle x_k, x_j^\star\rangle x_j = Tx. \qquad ◊$$

We now show that closed subspaces of finite dimension or co-dimension have a complement, which is needed in Spectral Theory.

Proposition 5.17. *For a normed vector space X the following assertions hold.*

a) Let $U \subseteq X$ be a finite dimensional linear subspace. Then U is closed and there exists a closed linear subspace Z with $X = U \oplus Z$.

b) Let Y be a closed linear subspace with finite codimension $\dim X/Y$. Then there exists a closed linear subspace V with $\dim V = \dim X/Y$ and $X = Y \oplus V$.

Proof. a) Let $\{x_1, \dots, x_n\}$ be a basis of U. We define T as in Example 5.16 with $x_k = y_k$. Then $T \in \mathcal{B}(X)$ is a projection with range U and thus Lemma 2.16 implies assertion a).

b) Let $Q : X \to X/Y$; $Qx = x + Y$, be the quotient map (see Proposition 2.19) and let $B = \{b_1, \dots, b_n\}$ be a basis of X/Y. Since Q is surjective, there are vectors $x_k \in X$ with $Qx_k = b_k$. Set $V = \operatorname{lin}\{x_1, \dots, x_n\}$.

If $\sum_{k=1}^n \alpha_k x_k = 0$ for some $\alpha_k \in \mathbb{F}$, then $\sum_{k=1}^n \alpha_k b_k = Q0 = 0$, and hence $\alpha_k = 0$ for all $k \in \{1, \dots, n\}$ because B is a basis. The set $\{x_1, \dots, x_n\}$ is thus linearly independent and $\dim V = \dim X/Y$. By part a), the space V is closed.

If $x \in Y \cap V$, then $x = \sum_{k=1}^n \beta_k x_k$ for some $\beta_k \in \mathbb{F}$ since $x \in V$. From $x \in Y$ we infer $Qx = 0$ which yields $0 = \sum_{k=1}^n \beta_k b_k$ and thus $\beta_k = 0$ for all $k \in \{1, \dots, n\}$. As a result, $x = 0$. Take $x \in X$. There are coefficients $\alpha_k \in \mathbb{F}$ with $Qx = \alpha_1 b_1 + \cdots + \alpha_n b_n$. Set $v = \alpha_1 x_1 + \cdots + \alpha_n x_n \in V$. We then obtain $Q(x - v) = \sum_{k=1}^n \alpha_k (b_k - Qx_k) = 0$, and hence $x - v \in Y$ and $x = x - v + v \in Y + V$. It follows $X = Y \oplus V$. $\square$

Geometric version of Hahn–Banach. Let X be a normed vector space, $A, B \subseteq X$, $A \cap B = \emptyset$ and $A, B \neq \emptyset$. A functional $x^\star \in X^\star$ *separates* the sets A and B if

$$\forall\, a \in A,\, b \in B : \qquad \operatorname{Re}\langle a, x^\star\rangle < \operatorname{Re}\langle b, x^\star\rangle,$$

and it *separates* A and B *strictly* if

$$s_A := \sup_{a \in A} \operatorname{Re}\langle a, x^\star\rangle < i_B := \inf_{b \in B} \operatorname{Re}\langle b, x^\star\rangle.$$

Observe that $N(x^\star)$ is a closed linear subspace with codimension 1 (cf. Linear Algebra). Hence, $H_0 = x_0 + N(x^\star)$ is a closed affine hyperplane. Let $\mathbb{F} = \mathbb{R}$, $x^\star$ separate A and B, and $x_0 \in X$ satisfy $\gamma := \langle x_0, x^\star\rangle \in [s_A, i_B]$. Then A and B are contained in the different halfspaces $H_\pm := \{x \in X \,|\, x^\star(x) \gtreqless \gamma\}$ separated by H_0, since $x^\star|_A \le s_A \le x^\star|_{H_0} = \gamma \le i_B \le x^\star|_B$.

Let $A \subseteq X$. As a crucial tool we define the *Minkowski functional*

$$p_A : X \to [0, \infty]; \quad p_A(x) = \inf\{\lambda > 0 \,|\, \tfrac{1}{\lambda} x \in A\},$$

where $\inf \emptyset = \infty$. Note that $p_{B(0,1)}(x) = \|x\|$. We show several of the basic properties of this map below, using the following observations.

REMARK 5.18. If $A, B \subseteq X$ are convex and $\alpha \in \mathbb{F}$, then also αA and $A+B$ are convex. To check the second assertion, take $a_k \in A$, $b_k \in B$, and $t \in [0,1]$ for $k \in \{1,2\}$. We then have $t(a_1+b_1)+(1-t)(a_2+b_2) = (ta_1+(1-t)a_2)+(tb_1+(1-t)b_2)$ which belongs to $A+B$ by the convexity of A and B. The first assertion is shown similarly. $\Diamond$

LEMMA 5.19. *Let X be a normed vector space and $A \subseteq X$ convex with $0 \in A^\circ$. (There thus exists a radius $\delta > 0$ with $\overline{B}(0,\delta) \subseteq A$.) Then the following assertions hold.*
a) We have $p_A(x) \leq \frac{1}{\delta}\|x\|$ for all $x \in X$.
b) The map p_A is sublinear.
c) If A is also open, we have $A = p_A^{-1}([0,1))$.

PROOF. a) The assumption implies that the vector $\frac{\delta}{\|x\|}x$ belongs to A for each $x \in X \setminus \{0\}$, so that $p_A(x) \leq \frac{1}{\delta}\|x\|$.
b) Let $t > 0$, $x, y \in X$, and $\varepsilon > 0$. The definition of p_A yields the equalities $p_A(0x) = 0p_A(x)$ and

$$p_A(tx) = \inf\{\lambda > 0 \mid \tfrac{t}{\lambda}x \in A\} = \inf\{t\mu > 0 \mid \tfrac{1}{\mu}x \in A\} = t\,p_A(x).$$

Further, there are numbers $0 < \lambda \leq p_A(x)+\varepsilon$ and $0 < \mu \leq p_A(y)+\varepsilon$ with $\frac{1}{\lambda}x, \frac{1}{\mu}y \in A$. Since A is convex, the vector

$$\frac{1}{\lambda+\mu}(x+y) = \frac{\lambda}{\lambda+\mu}\frac{1}{\lambda}x + \frac{\mu}{\lambda+\mu}\frac{1}{\mu}y$$

is contained in A, so that $p_A(x+y) \leq \lambda+\mu \leq p_A(x)+p_A(y)+2\varepsilon$. In the limit $\varepsilon \to 0$, we deduce assertion b).
c) Let A be open. First, take an element $x \in X$ with $p_A(x) < 1$. Then there exists a number $\lambda \in (0,1)$ with $\frac{1}{\lambda}x \in A$. The convexity of A yields that $x = \lambda\frac{1}{\lambda}x + (1-\lambda)0 \in A$. Conversely, pick x with $p_A(x) \geq 1$. The product $\frac{1}{\lambda}x$ is then contained in $X \setminus A$ for all $\lambda < 1$. Since $X \setminus A$ is closed, we obtain $x \in X \setminus A$ letting $\lambda \to 1$. $\square$

We now establish the *separation theorems* which are geometric versions of Hahn–Banach. Simple examples in $\mathbb{R}^2$ show that one needs convexity for separation and also compactness for strict separation.

THEOREM 5.20. *Let X be a normed vector space, $A, B \subseteq X$ be convex and non-empty, and $A \cap B = \emptyset$. The following assertions hold.*
a) Let A and B be open. We have $x^\star \in X^\star$ separating A and B.
b) Let A be closed and B be compact. There is a functional $x^\star \in X^\star$ separating A and B strictly.

PROOF. We let $\mathbb{F} = \mathbb{R}$. The case $\mathbb{F} = \mathbb{C}$ then follows by Lemma 5.6.
a) Let A and B be open. Fix a vector $x_0 \in A - B$ and put

$$C = A - B - x_0 = \bigcup_{b \in B} A - b - x_0.$$

The set C is open because of Proposition 1.15, it is convex by Remark 5.18, 0 belongs to C, and $y_0 := -x_0 \notin C$ (since $0 \notin A - B$). Thanks to Lemma 5.19, the Minkowski functional p_C is thus sublinear and satisfies $p_C(y_0) \geq 1$. We define $y^\star(ty_0) = tp_C(y_0)$ for all $t \in \mathbb{R}$. The map $y^\star : \operatorname{lin}\{y_0\} \to \mathbb{R}$ is linear and $y^\star(y) \leq p_C(y)$ for all $y = ty_0$. (If $t < 0$, we have $p_C(y) \geq 0 \geq tp_C(y_0)$.) Theorem 5.5 now gives a linear functional $x^\star : X \to \mathbb{R}$ with $x^\star(x) \leq p_C(x)$ for all $x \in X$ and $x^\star(y_0) = y^\star(y_0) = p_C(y_0) \geq 1$. For $x \in X$, Lemma 5.19a) implies that

$$|x^\star(x)| = \max\{x^\star(x), x^\star(-x)\} \leq \max\{p_C(x), p_C(-x)\} \leq \tfrac{1}{\delta}\, \|x\|$$

for some $\delta > 0$, and hence $x^\star$ belongs to $X^\star$. Let $a \in A$ and $b \in B$. Then the vector $x = a - b - x_0$ is an element of C. Since $y_0 = -x_0$, Lemma 5.19c) yields

$$1 > p_C(x) \geq \langle x, x^\star\rangle = \langle a, x^\star\rangle - \langle b, x^\star\rangle + \langle y_0, x^\star\rangle.$$

Using $\langle y_0, x^\star\rangle \geq 1$, we deduce that $\langle a, x^\star\rangle < \langle b, x^\star\rangle$, and thus part a).

b) Let A be closed and B be compact. The number $\varepsilon := \frac{1}{3}\operatorname{dist}(A, B) > 0$ is positive by Example 1.9. The sets $A_\varepsilon = A + B(0,\varepsilon) = \bigcup_{a\in A} B(a,\varepsilon)$ and $B_\varepsilon = B + B(0,\varepsilon)$ are thus disjoint and, as above, open and convex. From step a) we obtain a functional $x^\star \in X^\star$ satisfying

$$\langle a + x, x^\star\rangle < \langle b + y, x^\star\rangle$$

for all $a \in A$, $b \in B$ and $x, y \in B(0,\varepsilon)$. For $y = 0$ and $x = \pm\varepsilon z$ with $z \in B(0,1)$, it follows $\varepsilon\,|\langle z, x^\star\rangle| < \langle b - a, x^\star\rangle$. Taking the supremum over $z \in B(0,1)$, we derive the inequality $0 < \varepsilon\,\|x^\star\| \leq \langle b - a, x^\star\rangle$ for all $a \in A$ and $b \in B$ which implies assertion b). □

A typical application of the above result is given in Theorem 5.36. We first discuss another consequence based on new concepts which we will used in Section 5.4 to describe the mapping properties of operators.

Let X be a normed vector space, $A \subseteq X$ and $B_\star \subseteq X^\star$ be non-empty. The *annihilators* of A and $B_\star$ are defined by

$$\begin{aligned} A^\perp &= \{x^\star \in X^\star \,|\, \forall\, a \in A \ \text{ we have } \ \langle a, x^\star\rangle = 0\} \subseteq X^\star, \\ {}^\perp B_\star &= \{x \in X \,|\, \forall\, b^\star \in B_\star \ \text{ we have } \ \langle x, b^\star\rangle = 0\} \subseteq X. \end{aligned} \tag{5.6}$$

In view of Riesz' Theorem 3.10, in a Hilbert space $X \cong X^\star$ these two sets are isomorphic to the orthogonal complement. We first collect their simple properties which follow from the corollaries to Hahn–Banach.

Remark 5.21. Let X be a normed space, $A \subseteq X$, and $B_\star \subseteq X^\star$.

a) As in Remark 3.7, but now using (5.1), one verifies that $A^\perp$ and ${}^\perp B_\star$ are closed linear subspaces of $X^\star$ and X, respectively, and that $\overline{\operatorname{lin} A} \subseteq {}^\perp(A^\perp)$, $(\overline{\operatorname{lin} A})^\perp = A^\perp$, and ${}^\perp(\overline{\operatorname{lin} B_\star}) = {}^\perp B_\star$.

b) Corollary 5.13 shows that $A^\perp = \{0\}$ if and only if $\overline{\operatorname{lin} A} = X$. From Corollary 5.10 we deduce that $A^\perp = X^\star$ if and only if $A = \{0\}$.

c) By definition, we have ${}^{\perp}B_\star = X$ if and only if $B_\star = \{0\}$. Due to Corollary 5.10, $\overline{\operatorname{lin} B_\star} = X^\star$ implies that ${}^{\perp}B_\star = \{0\}$.

d) The converse implication in part c) is wrong in general. For instance, let $X = \ell^1$ and $B_\star = \{e_n \,|\, n \in \mathbb{N}\} \subseteq X^\star = \ell^\infty$. We then have $\overline{\operatorname{lin} B_\star} = c_0$, but ${}^{\perp}B_\star = \{0\}$ since all $y \in \ell^1$ satisfy $\langle y, e_n\rangle_{\ell^1} = y_n$. ♢

In Hilbert spaces the next result follows from the projection theorem, see Example 3.9d). We now need the separation theorem to show it.

Proposition 5.22. *Let X be a normed vector space X and be $A \subseteq X$ non-empty. We then have $\overline{\operatorname{lin} A} = {}^{\perp}(A^\perp)$.*

Proof. Remark 5.21 yields the inclusion $\overline{\operatorname{lin} A} \subseteq {}^{\perp}(A^\perp)$. Suppose there was a vector x_0 in ${}^{\perp}(A^\perp) \setminus \overline{\operatorname{lin} A}$. Theorem 5.20b) with $B = \{x_0\}$ then gives a separating functional $x^\star \in X^\star$ satisfying

$$s := \sup_{x \in \overline{\operatorname{lin} A}} \operatorname{Re}\langle x, x^\star\rangle < \operatorname{Re}\langle x_0, x^\star\rangle.$$

Suppose $r := \operatorname{Re}\langle x, x^\star\rangle \neq 0$ for some $x \in \overline{\operatorname{lin} A}$. For $t \in \mathbb{R}$ we then infer

$$\operatorname{Re}\langle tx, x^\star\rangle = tr \to \infty \begin{cases} \text{as } \ t \to \infty, & \text{if } \ r > 0, \\ \text{as } \ t \to -\infty, & \text{if } \ r < 0. \end{cases}$$

This contradicts the above estimate for s, and thus $\operatorname{Re}\langle x, x^\star\rangle = 0$ for all $x \in \overline{\operatorname{lin} A}$. Hence, $s = 0$. Using it instead of t, we similarly obtain $\operatorname{Im}\langle x, x^\star\rangle = 0$ for all $x \in \overline{\operatorname{lin} A}$ so that $x^\star$ belongs to $A^\perp$. We arrive at the contradiction $0 = s < \operatorname{Re}\langle x_0, x^\star\rangle = 0$ since $x_0 \in {}^{\perp}(A^\perp)$. □

On a dual space one can interchange the order of the annihilators, but one does not have an analogue of the above result in general. As an example, let $X = \ell^1$ and $B_\star = \{e_n \,|\, n \in \mathbb{N}\} \subseteq X^\star = \ell^\infty$. In Remark 5.21d), we have seen that $\overline{\operatorname{lin} B_\star} = c_0$ and that ${}^{\perp}B_\star = \{0\}$, hence $({}^{\perp}B_\star)^\perp = X^\star = \ell^\infty \neq \overline{\operatorname{lin} B_\star}$. See Theorem 4.7 in [**Ru**] for more information.

Proposition 5.23. *Let X be a normed vector space and $Y \subseteq X$ be a closed linear subspace. Then the maps*

$$T : X^\star/Y^\perp \to Y^\star; \quad T(x^\star + Y^\perp) = x^\star|_Y,$$
$$S : (X/Y)^\star \to Y^\perp; \quad S\varphi = \varphi \circ Q,$$

are isometric isomorphisms, where $Q : X \to X/Y$; $Qx = x + Y$.

See Exercise 12.3 for a proof. This result is used in Spectral Theory. In a Hilbert space X it is a part of the projection theorem 3.8.

5.3. Reflexivity and weak convergence

Let X be a normed vector space. The *bidual* of X is $X^{\star\star} := (X^\star)^\star$. For each $x \in X$ we define the map

$$J_X(x) : X^\star \to \mathbb{F}; \quad \langle x^\star, J_X(x)\rangle_{X^\star} = \langle x, x^\star\rangle_X. \tag{5.7}$$

Clearly, $J_X(x)$ is linear in $x^\star$ and we have $|J_X(x)(x^\star)| \le \|x\|\,\|x^\star\|$, so that $J_X(x)$ belongs to $X^{\star\star}$. Moreover, the operator J_X is linear in x, and by means of Corollary 5.10 and equation (5.7) we obtain

$$\|x\|_X = \sup_{\|x^\star\|\le 1} |\langle x, x^\star\rangle| = \|J_X(x)\|_{X^{\star\star}}.$$

We state these obeservations in a proposition.

Proposition 5.24. *Let X be a normed vector space. Then equation (5.7) defines a linear isometry $J_X : X \to X^{\star\star}$.* □

This result leads to a quick proof of the existence part of Proposition 2.21: The closure $\tilde{X}$ of the range $J_X(X)$ in $X^{\star\star}$ is a Banach space; i.e., X is isometrically isomorphic to a dense subspace of a Banach space. We next introduce an interesting class of Banach spaces.

Definition 5.25. *A normed vector space X is called* reflexive *if the isometry J_X from (5.7) is surjective.*

Remark 5.26. a) Let X be reflexive. Then $X \cong X^{\star\star}$ since also J_X^{-1} is isometric, and hence X is a Banach space. However, there are non-reflexive Banach spaces which are isomorphic to their biduals (with an isomorphism different from J_X), see Example 1.d.2 in [**LT**].

b) If X reflexive and $B_\star \subseteq X^\star$, then $B_\star^\perp = J_X({}^\perp B_\star)$ by (5.7). Here (and in similar points below) reflexive Banach spaces share some properties of Hilbert spaces, which are not true in a general Banach space.

c) In Corollary 5.51 we will show that reflexivity is preserved under isomorphisms.

d) One usually identifies the space X with the range $R(J_X)$ in $X^{\star\star}$ and a reflexive space X with its bidual $X^{\star\star}$. ◊

We next discuss the basic exampels of reflexive spaces.

Example 5.27. a) Hilbert spaces X are reflexive.

Proof. For a Hilbert space Z we have the antilinear bijection $\Phi_Z : Z \to Z^\star$; $\langle v, \Phi_Z(z)\rangle_Z = (v|z)_Z$, for all $v \in Z$ (see Theorem 3.10). It is straightforward to check that the dual space $X^\star$ of X is a Hilbert space equipped with the scalar product $(x^\star|y^\star)_{X^\star} := (\Phi_X^{-1} y^\star | \Phi_X^{-1} x^\star)_X$ for $x^\star, y^\star \in X^\star$. Take any $x^{\star\star} \in X^{\star\star}$. Set $x^\star = \Phi_{X^\star}^{-1} x^{\star\star} \in X^\star$ and $x = \Phi_X^{-1} x^\star \in X$. Using the above definitions, we compute

$$\langle y^\star, x^{\star\star}\rangle_{X^\star} = (y^\star|x^\star)_{X^\star} = (\Phi_X^{-1}x^\star|\Phi_X^{-1}y^\star)_X = \langle x, y^\star\rangle_X$$

for every $y^\star \in X^\star$; i.e., $J_X(x) = x^{\star\star}$ as asserted. □

b) Let $1 < p < \infty$ and $(S, \mathcal{A}, \mu)$ be a measure space. Then $X = L^p(\mu)$ is reflexive. (An example is $X = \ell^p$.)

Proof. Let $r \in (1, \infty)$. Theorem 5.4 yields the isomorphism $\Phi_r : L^{r'}(\mu) \to L^r(\mu)^\star$; $\langle \varphi, \Phi_r(\psi)\rangle_{L^r} = \int \varphi\psi \,\mathrm{d}\mu$, for all $\varphi \in L^r(\mu)$ and $\psi \in L^{r'}(\mu)$. Take $\phi \in X^{\star\star}$. The map $\phi \circ \Phi_p$ then belongs to $L^{p'}(\mu)^\star$. Since $p'' = p$, we have a function $f \in L^p(\mu)$ satisfying $\Phi_{p'}(f) = \phi \circ \Phi_p$. Let $\Phi_p(g)$ with $g \in L^{p'}(\mu)$ be an arbitrary element of $X^\star$. We now calculate

$$\langle \Phi_p(g), \phi\rangle_{X^\star} = \phi(\Phi_p(g)) = \langle g, \Phi_{p'}(f)\rangle_{L^{p'}} = \int_S gf \,\mathrm{d}\mu = \langle f, \Phi_p(g)\rangle_{L^p}.$$

Hence, $\phi = J_X(f)$ as asserted. □

c) The space c_0 is not reflexive. Indeed, its bidual $c_0^{\star\star}$ is isomorphic to ℓ^∞ by Proposition 5.1. From Example 1.55 we know that c_0 is separable and ℓ^∞ not, so that they cannot be isomorphic. Remark 5.26a) thus yields the claim. ♢

In the above example and also below, we use that separability is preserved under isomorphisms by Exercise 6.2. We show permanence properties of reflexivity needed later on.

Proposition 5.28. *Let X be a normed vector space.*

a) Let X be reflexive and Y be a closed linear subspace of X. Then $(Y, \|\cdot\|_X)$ is reflexive.

b) The space X is reflexive if and only if $X^\star$ is reflexive.

c) Let X be reflexive. Then is X separable if and only if $X^\star$ separable.

Proof. a) Let $Y \subseteq X$ be a closed linear subspace. Take $y^{\star\star} \in Y^{\star\star}$. For each $x^\star \in X^\star$, the restriction $x^\star|_Y$ belongs to $Y^\star$ with $\|x^\star|_Y\|_{Y^\star} \le \|x^\star\|_{X^\star}$. We define the linear map $x^{\star\star} : X^\star \to \mathbb{F}$ by $x^{\star\star}(x^\star) = \langle x^\star|_Y, y^{\star\star}\rangle_{Y^\star}$ for all $x^\star \in X^\star$. As $|x^{\star\star}(x^\star)| \le \|x^\star\|_{X^\star}\|y^{\star\star}\|_{Y^{\star\star}}$, the functional $x^{\star\star}$ is an element of $X^{\star\star}$. By the reflexivity of X, there exists a vector $y \in X$ such that

$$\langle x^\star|_Y, y^{\star\star}\rangle_{Y^\star} = \langle x^\star, x^{\star\star}\rangle_{X^\star} = \langle y, x^\star\rangle_X \qquad \text{for all } \; x^\star \in X^\star.$$

Suppose that $y \notin Y$. Since Y is closed, Proposition 5.9 yields a map $\tilde{x}^\star \in X^\star$ satisfying $\tilde{x}^\star|_Y = 0$ and $\langle y, \tilde{x}^\star\rangle_X \neq 0$. This fact contradicts the above equation in display, and y is thus contained in Y.

Take any $y^\star \in Y^\star$. Let $x^\star \in X^\star$ be a Hahn–Banach extension of $y^\star$. We then obtain $\langle y, y^\star\rangle_Y = \langle y, x^\star\rangle_X = \langle y^\star, y^{\star\star}\rangle_{Y^\star}$ by the above considerations, and therefore $J_Y(y) = y^{\star\star}$.

b) Let X be reflexive. Take a functional $x^{\star\star\star} \in X^{\star\star\star}$. We set $x^\star(x) = \langle J_X(x), x^{\star\star\star}\rangle_{X^{\star\star}}$ for all $x \in X$. Clearly, $x^\star$ belongs to $X^\star$. Let $x^{\star\star} \in X^{\star\star}$. By assumption, there is a vector $x \in X$ with $x^{\star\star} = J_X(x)$. It follows $\langle x^\star, x^{\star\star}\rangle_{X^\star} = \langle x, x^\star\rangle_{X^\star} = \langle x^{\star\star}, x^{\star\star\star}\rangle_{X^{\star\star}}$. Hence, $X^\star$ is reflexive.

Conversely, assume that X is not reflexive. Proposition 5.9 then yields a map $x^{\star\star\star} \in X^{\star\star\star} \setminus \{0\}$ satisfying $\langle J_X(x), x^{\star\star\star}\rangle_{X^{\star\star}} = 0$ for all

$x \in X$. Suppose that $X^\star$ was reflexive. There thus exists a functional $x^\star \in X^\star$ with $x^{\star\star\star} = J_{X^\star}(x^\star)$. We infer that

$$0 = \langle J_X(x), J_{X^\star}(x^\star)\rangle_{X^{\star\star}} = \langle x^\star, J_X(x)\rangle_{X^\star} = \langle x, x^\star\rangle_X \qquad \text{for all } \ x \in X,$$

which means that $x^\star = 0$, contradicting $x^{\star\star\star} \neq 0$.

c) The implication "$\Leftarrow$" was shown in Corollary 5.14. If X is separable, then $X^{\star\star} \cong X$ is also separable. Hence, $X^\star$ is separable by Corollary 5.14. □

We can now treat further main examples.

EXAMPLE 5.29. a) The space $X = \ell^1$ is not reflexive, because it is separable and its dual $X^\star = \ell^\infty$ is not. Since c_0 is not reflexive by Example 5.27 and it is a closed subspace of ℓ^∞, also the space ℓ^∞ fails to be reflexive.

b) The spaces $C([0,1])$, $L^\infty(0,1)$, and $L^1(0,1)$ are not reflexive by Exercise 13.2.

c) Let $U \subseteq \mathbb{R}^m$ be open and $p \in (1,\infty)$. By Remark 4.16d), the Sobolev space $W^{1,p}(U)$ is isomorphic to a closed subspace F of $L^p(U)^{1+m}$. As in Example 5.27 one sees that $L^p(U)^{1+m}$ is reflexive, and hence also F by Proposition 5.28. Corollary 5.51 thus shows that $W^{1,p}(U)$ is reflexive. ◊

In order to obtain fundamental compactness results below, we introduce new convergence concepts.

DEFINITION 5.30. *Let X be a normed vector space.*

a) *A sequence (x_n) in X* converges weakly *to $x \in X$ if*

$$\forall\, x^\star \in X^\star : \quad \langle x_n, x^\star\rangle \to \langle x, x^\star\rangle \quad \text{as} \ \ n \to \infty.$$

b) *A sequence $(x_n^\star)$ in $X^\star$* converges weakly* *to $x^\star \in X^\star$ if*

$$\forall\, x \in X : \quad \langle x, x_n^\star\rangle \to \langle x, x^\star\rangle \quad \text{as} \ \ n \to \infty.$$

We then write $x_n \rightharpoonup x$ or $x_n \overset{\sigma}{\to} x$ or $\sigma\text{–}\lim_{n\to\infty} x_n = x$, respectively, $x_n^\star \overset{\star}{\rightharpoonup} x^\star$ or $x_n^\star \overset{\sigma^\star}{\longrightarrow} x^\star$ or $\sigma^\star\text{–}\lim_{n\to\infty} x_n^\star = x^\star$. One often replaces here the letter 'σ' by 'w'.

We first collect simple, but important properties of weak and weak* convergence, see also the exercises.

REMARK 5.31. Let X be a normed vector space, $x_n, x, y \in X$, $x_n^\star, x^\star \in X^\star$, and $n \to \infty$.

a) The weak* convergence in $X^\star = \mathcal{B}(X,\mathbb{F})$ is just the strong convergence of a sequence of operators in $\mathcal{B}(X,\mathbb{F})$ as discussed in Section 4.1.

Let the sequence $(\langle x, x_n^\star\rangle)_n$ be Cauchy in $\mathbb{F}$ for each $x \in X$. Lemma 4.10 then yields a functional $x^\star \in X^\star$ satisfying $x_n^\star \overset{\sigma^\star}{\longrightarrow} x^\star$ as $n \to \infty$. In this sense, $X^\star$ is 'weakly* sequentially complete'.

b) Weak and weak* convergence are linear in view of Definition 5.30.

c) If $x_n \to x$ in the norm of X, then $x_n \xrightarrow{\sigma} x$ (since $|\langle x_n - x, x^\star\rangle| \le \|x_n - x\|\,\|x^\star\|$). If $x_n^\star \to x^\star$ in the norm of $X^\star$, then $x_n^\star \xrightarrow{\sigma^\star} x^\star$. In $X = \mathbb{F}^m$ weak or weak* convergence are equivalent to componentwise convergence (take $x = e_k$ or $x^\star = e_k$), and thus to convergence in norm.

d) The weak and weak* limits are unique.

Proof. Let $x_n \xrightarrow{\sigma} x$ and $x_n \xrightarrow{\sigma} y$ with $x \neq y$. Corollary 5.10 then yields a functional $x^\star \in X^\star$ satisfying $\langle x, x^\star\rangle \neq \langle y, x^\star\rangle$. This is impossible since $\langle x_n, x^\star\rangle$ converges to both $\langle x, x^\star\rangle$ and $\langle y, x^\star\rangle$. The second assertion follows from part a). □

e) For $x \in X$ we have $\langle x, x^\star\rangle_X = \langle x^\star, J_X(x)\rangle_{X^\star}$ so that the σ-convergence on $X^\star$ implies the $\sigma^\star$-convergence. If X is reflexive, then the two types of convergence on $X^\star$ coincide. Reflexive spaces thus are 'weakly sequentially complete' due to statement a). ♢

For weak or weak* convergence several properties can fail which one might hope to be true. These examples are of great importance. See also the exercises.

Remark 5.32. Let $n \to \infty$.

a) For weakly or weakly* convergent sequences each subsequence may diverge in norm. The norm of a weak or weak* limit may be strictly smaller than the limes inferior of the norms of the sequence.

Proof. In $X = \ell^2$ we have $\langle e_n, x\rangle = x_n \to 0$ for all $x \in \ell^2$. This means that $e_n \xrightarrow{\sigma} 0$ and $e_n \xrightarrow{\sigma^\star} 0$. But each subsequence of (e_n) diverges in ℓ^2 since $\|e_n - e_m\|_2 = 2^{1/2}$ for all $n \neq m$. Moreover, $\|e_n\| = 1$ and the weak limit 0 has a strictly smaller norm. □

b) A weakly* convergent sequence in a non-reflexive space may not possess a weakly converging subsequence.

Proof. In $X^\star = \ell^1 = c_0^\star$ we have $\langle x, e_n\rangle_{c_0} = x_n \to 0$ for each $x \in c_0$ so that $e_n \xrightarrow{\sigma^\star} 0$. Take any subsequence $(e_{n_j})_j$. For $k \in \mathbb{N}$, we set $y_k = (-1)^j$ if $k = n_j$ and $y_k = 0$ otherwise, and put $y = (y_j) \in \ell^\infty = (\ell^1)^\star$. Then $\langle e_{n_j}, y\rangle_{\ell^1} = y_{n_j}$ does not converge as $j \to \infty$. □

c) There are (non-reflexive) spaces which are not 'weakly sequentially complete'.

Proof. Let $X = c_0$ and $v_n = e_1 + \cdots + e_n \in c_0 \subseteq \ell^\infty$. For each $y \in \ell^1$, we have

$$\langle v_n, y\rangle_{c_0} = \langle y, v_n\rangle_{\ell^1} = \sum_{k=1}^{n} y_k \;\longrightarrow\; \sum_{k=1}^{\infty} y_k = \langle y, \mathbb{1}\rangle_{\ell^1} \quad \text{as} \quad n \to \infty.$$

This means that $v_n \xrightarrow{\sigma^\star} \mathbb{1}$ in ℓ^∞ and that $(\langle v_n, y\rangle_{c_0})_n$ is a Cauchy sequence in $\mathbb{F}$ for every $y \in \ell^1$. If (v_n) had a weak limit x in c_0, then x would also be the weak* limit of (v_n) in ℓ^∞ and thus $x = \mathbb{1}$ by the uniqueness of weak* limits, which is impossible. □

We state a useful characterization of weak and weak* convergence.

Proposition 5.33. *Let X be a normed vector space, $x_n, x \in X$ and $x_n^\star, x^\star \in X^\star$ for $n \in \mathbb{N}$, $D_\star \subseteq X^\star$ with $\overline{\operatorname{lin} D_\star} = X^\star$, and $D \subseteq X$ with $\overline{\operatorname{lin} D} = X$. Then the following equivalences are true.*

a) $x_n \xrightarrow{\sigma} x$ as $n \to \infty$ if and only if $\sup_n \|x_n\|_X < \infty$ and $\langle x_n, y^\star\rangle \to \langle x, y^\star\rangle$ as $n \to \infty$ for all $y^\star \in D_\star$.

b) Let X be complete. Then $x_n^\star \xrightarrow{\sigma^\star} x^\star$ as $n \to \infty$ if and only if $\sup_n \|x_n^\star\|_{X^\star} < \infty$ and $\langle y, x_n^\star\rangle \to \langle y, x^\star\rangle$ as $n \to \infty$ for all $y \in D$.

If a) is valid, then $\|x\| \le \varliminf_{n\to\infty} \|x_n\|$; and if b) holds, then $\|x^\star\| \le \varliminf_{n\to\infty} \|x_n^\star\|$. Moreover, in equivalence b) the implication '$\Leftarrow$' is true for all normed vector spaces X.

Proof. For the implications '$\Rightarrow$' we use Corollary 4.5, whereas '$\Leftarrow$' and the addendum follow from Lemma 4.10. In part a) one has to apply these results to the vectors $J_X(x_n) \in X^{\star\star} = \mathcal{B}(X^\star, \mathbb{F})$. □

In the following examples we use the above characterization to describe weak convergence in sequence and Lebesgue spaces quite well, whose duals were described in Proposition 5.1 and Theorem 5.4.

Example 5.34. a) Let $X = c_0$ oder $X = \ell^p$ for $1 < p < \infty$, (v_n) in X be bounded, and $x \in X$. Proposition 5.33 with $D_\star = \{e_n \,|\, n \in \mathbb{N}\}$ then yields the equivalence

$$v_n \xrightarrow{\sigma} x \iff \forall\, k \in \mathbb{N}: \quad (v_n)_k = \langle v_n, e_k\rangle \to \langle x, e_k\rangle = x_k$$

as $n \to \infty$, since $c_{00} = \operatorname{lin} D_\star$ is dense in $X^\star$. (On the right-hand side one has 'componentwise convergence'.)

For sequences in ℓ^1 or ℓ^∞ one obtains an analogous result for the $\sigma^\star$-convergence taking $D = \{e_n \,|\, n \in \mathbb{N}\}$ in c_0, respectively ℓ^1.

b) The implication '$\Leftarrow$' fails in a) for $p = 1$. In fact, the sequence (e_n) converges componentwise to 0, but it is bounded and diverges weakly in ℓ^1 by the proof of Remark 5.32b).

c) The assumption of boundedness cannot be omitted in Proposition 5.33. For instance, the sequence (ne_n) in ℓ^2 converges compenentwise to 0, but is unbounded and thus cannot converge weakly.

d) Let $(S, \mathcal{A}, \mu)$ be a measure space, $X = L^p(\mu)$, $1 < p < \infty$, and (f_n) be bounded in X. We then deduce the equivalence

$$f_n \xrightarrow{\sigma} f \iff \int_A f_n \,\mathrm{d}\mu \to \int_A f \,\mathrm{d}\mu \quad \text{for all } A \in \mathcal{A} \quad \text{with } \mu(A) < \infty$$

as $n \to \infty$ from Proposition 5.33 and Theorem 5.9 of Analysis 3 (saying that the subspace of simple functions is dense in $L^{p'}(\mu)$).

The weak* convergence in $L^\infty(\mu)$ can be characterized analogously if the measure space is σ–finite.

e) The sequence given by $f_n(s) = \sin(ns)$ tends weakly to 0 in $L^2(0,1)$, though $\|f_n\|_2^2 \to \frac{1}{2}$ as $n \to \infty$. This weak limit can be checked using part d) or integrating by parts. For each $\varphi \in C_c^\infty(0,1)$ we have

$$|\langle f_n, \varphi\rangle| = \Big|\frac{1}{n}\int_0^1 \cos(ns)\varphi'(s)\,\mathrm{d}s\Big| \le \frac{\|\varphi'\|_\infty}{n} \longrightarrow 0$$

as $n \to \infty$, and thus $f_n \xrightarrow{\sigma} f$ by Proposition 5.33 and the density result Proposition 4.13. Here oscillations lead to weak convergence.[3] ♢

Weak convergence in a bounded subset is often given by a metric.

REMARK 5.35. If $X^\star$ is separable, then the weak sequential convergence in a bounded subset M of X is given by the metric

$$\mathrm{d}(x,y) = \sum_{j=1}^\infty 2^{-j} \frac{\big|\langle x-y, x_j^\star\rangle\big|}{1+\big|\langle x-y, x_j^\star\rangle\big|}$$

for $x, y \in M$, where $D_\star = \{x_j^\star \,|\, j \in \mathbb{N}\}$ is dense in $X^\star$. This result follows from Propositions 5.33 and 1.8 using the seminorms $p_j(x) = \big|\langle x, x_j^\star\rangle\big|$. Observe that for each $x \in X \setminus \{0\}$ there is an index $k \in \mathbb{N}$ such that $\langle x, x_k^\star\rangle \neq 0$ due to Corollary 5.10 and the density of $D_\star$. Similarly, if X is separable then the weak* sequential convergence in a bounded subset of $X^\star$ is given by an analogous metric. ♢

The next theorem by *Mazur* says that closed convex sets are 'weakly sequentially closed'. It is a consequence of the separation theorem and often used in combination with Theorem 5.40.

THEOREM 5.36. *Let X be a normed vector space, $C \subseteq X$ be closed and convex, and let (x_n) in C converge weakly to some $x \in X$. Then x belongs to C. Moreover, a sequence (y_N) of convex combinations of the vectors $\{x_n \,|\, n \ge N\}$ tends to x in norm as $N \to \infty$.*

PROOF. 1) Suppose that x was not contained in C. Theorem 5.20b) with $A = C$ and $B = \{x\}$ then gives a functional $x^\star \in X^\star$ with $\sup_n \mathrm{Re}\langle x_n, x^\star\rangle < \mathrm{Re}\langle x, x^\star\rangle$. But this inequality cannot hold since $\langle x_n, x^\star\rangle$ converges to $\langle x, x^\star\rangle$ as $n \to \infty$. Hence, x belongs to C.

2) Let $N \in \mathbb{N}$. It is straightforward to check that the set

$$C_N = \Big\{ y = \sum_{j=N}^m t_j x_j \,\Big|\, m \in \mathbb{N},\ m \ge N,\ t_j \ge 0,\ t_N + \cdots + t_m = 1 \Big\}$$

is convex, and it contains all x_n for $n \ge N$. Its closure $\overline{C_N}$ is also convex by Corollary 1.18. Step 1) then shows that x belongs to every $\overline{C_N}$. So we can choose points $y_N \in C_N$ with $\|x - y_N\| \le \frac{1}{N}$ for each $N \in \mathbb{N}$. □

We note that one needs convexity in Mazur's theorem and that it may fail for the weak* convergence.

[3]This example was omitted in the lectures.

Remark 5.37. a) The vectors e_n in the closed unit sphere S of ℓ^2 converge weakly in ℓ^2 to $0 \notin S$.

b) The elements e_n tend weakly* to 0 in $X^\star = \ell^1$, and they belong to the closed affine subspace $A = \{y \in \ell^1 \mid \langle y, \mathbb{1}\rangle = 1\}$, but $0 \notin A$. ◊

We now prove a simplified version of the *Banach–Alaoglu theorem*. It extends the Bolzano–Weierstraß theorem to Banach spaces which are adjoints of a separable space and says that the balls $\overline{B}_{X^\star}(0,r)$ are 'weakly* sequentially compact' (instead being compact as in finite dimensions). This fact is the fourth of the fundamental principles of linear functional analysis. At the end of this section we give one of its many applications.

Theorem 5.38. *Let X be a separable normed vector space. Let $(x_n^\star)$ be a bounded sequence in $X^\star$. Then there is a functional $x^\star \in X^\star$ and a subsequence $(x_{n_j}^\star)_j$ converging weakly* to $x^\star$ as $j \to \infty$, where $\|x^\star\| \le \underline{\lim}_{n\to\infty}\|x_n^\star\|$. Hence, each sequence in a ball $\overline{B}_{X^\star}(0,r)$ has a weak* accumulation point.*

Proof. There are vectors whose norms $\|x_{n_l}^\star\|$ tend to $\underline{\lim}_{n\to\infty}\|x_n^\star\|$ as $l \to \infty$. We replace $x_n^\star$ by this subsequence without relabelling it.

Let $\{x_k \mid k \in \mathbb{N}\}$ be dense in X. Since $(\langle x_1, x_n^\star\rangle)_{n\in\mathbb{N}}$ is bounded in $\mathbb{F}$, there exists a subsequence $(\langle x_1, x_{\nu_1(j)}^\star\rangle)_j$ with a limit in $\mathbb{F}$. Since also $(\langle x_2, x_{\nu_1(j)}^\star\rangle)_j$ is bounded, there is a converging subsequence $(\langle x_2, x_{\nu_2(j)}^\star\rangle)_j$. For each $k \in \mathbb{N}$ we iteratively obtain subsequences $(x_{\nu_k(j)}^\star)_j$ of $(x_{\nu_{k-1}(j)}^\star)_j$ such that $(\langle x_k, x_{\nu_k(j)}^\star\rangle)_j$ converges. We define $y_m^\star = x_{\nu_m(m)}^\star$ for each $m \in \mathbb{N}$. Then the sequence $(\langle x_k, y_m^\star\rangle)_m$ has a limit for every $k \in \mathbb{N}$ since for each $m \ge k$ there is an index $j_m = j_m(k) \ge m$ with $\nu_m(m) = \nu_k(j_m)$. Since $\{x_k \mid k \in \mathbb{N}\}$ is dense and the subsequence $(y_m^\star)$ of $(x_n^\star)$ is also bounded, Proposition 5.33 implies the assertion. □

Remark 5.39. a) The above version of the theorem of Banach–Alaoglu can fail if X is not separable. As the simplest example we consider $X = \ell^\infty$, where $X^\star$ is already a rather unpleasant space, cf. Theorem IV.5.1 in [**DS**]. The maps $\varphi_n : X \to \mathbb{F}$; $\varphi_n(x) = x_n$, belong to $X^\star$ with $\|\varphi_n\| = 1$ for all $n \in \mathbb{N}$. Take any subsequence $(\varphi_{n_j})_j$. As in Remark 5.32b) we find a sequence $x \in \ell^\infty$ such that $\langle x, \varphi_{n_j}\rangle = x_{n_j}$ diverges and hence (φ_n) has no weakly* convergent subsequence.[4]

b) The theorem of Banach–Alaoglu can fail for the weak convergence. For instance, look at $e_n \in \ell^1 = X$ for $n \in \mathbb{N}$. Take any subsequence $(e_{n_j})_j$ and choose as above an element y in $\ell^\infty = (\ell^1)^\star$ such that $(y_{n_j})_j$ diverges. Then $\langle e_{n_k}, y\rangle = y_{n_k}$ does not converge. ◊

---

[4]In Theorem V.3.1 in [**Co**] one can find the full version of the Banach–Alaoglu theorem without the separability assumption.

In a reflexive space we can derive a version of the *Banach-Alaoglu theorem* with weak convergence, cf. Remark 5.31. A standard trick even allows us to get rid of the separability assumption,

Theorem 5.40. *Let X be a reflexive Banach space. Let (x_n) be a bounded sequence in X. Then there is a vector $x \in X$ and a subsequence $(x_{n_j})_j$ converging weakly to x as $j \to \infty$, where $\|x\| \le \underline{\lim}_{n\to\infty} \|x_n\|$.*

Proof. Let $Y = \overline{\text{lin}}\{x_n \,|\, n \in \mathbb{N}\}$ be endowed with the norm of X. By Proposition 5.28, the space Y is reflexive, and it is separable by Lemma 1.54. Proposition 5.28 then shows that $Y^\star$ is separable. Theorem 5.38 now yields a subsequence $(J_Y(x_{n_j}))_j$ of $(J_Y(x_n))_n$ with weak* limit $y^{\star\star}$ in $Y^{\star\star}$. Since Y is reflexive, there exists a vector $x \in Y$ such that $J_Y(x) = y^{\star\star}$ and

$$\langle x_{n_j}, y^\star\rangle_Y = \langle y^\star, J_Y(x_{n_j})\rangle_{Y^\star} \longrightarrow \langle y^\star, J_Y(x)\rangle_{Y^\star} = \langle x, y^\star\rangle_Y$$

for all $y^\star \in Y^\star$, as $j \to \infty$. We further obtain

$$\|x\| = \|J_Y(x)\| \le \underline{\lim}_{n\to\infty} \|J_Y(x_n)\| = \underline{\lim}_{n\to\infty} \|x_n\|,$$

using also Proposition 5.24. Let $x^\star \in X^\star$. The restriction $x^\star|_Y$ then belongs to $Y^\star$. As a result,

$$\langle x_{n_j}, x^\star\rangle_X = \langle x_{n_j}, x^\star|_Y\rangle_Y \longrightarrow \langle x, x^\star|_Y\rangle_Y = \langle x, x^\star\rangle_X,$$

which means that $x_{n_j} \xrightarrow{\sigma} x$ as $j \to \infty$. □

We next use the Banach-Alaoglu theorem and results about Sobolev spaces to solve a basic problem about static electric fields.

Example 5.41. Let $D \subseteq \mathbb{R}^3$ open and bounded with a C^1–boundary. The trace theorem shows that the mapping $W^{1,2}(D) \cap C(\overline{D}) \to L^2(\partial D, \sigma)$; $u \mapsto u|_{\partial D}$, has a unique continuous linear extension $\text{tr} : W^{1,2}(D) \to L^2(\partial D, \sigma)$, where σ is the surface measure from Analysis 3. The kernel of tr is equal to the closure $W_0^{1,2}(D)$ of the test functions in $W^{1,2}(D)$. (See Theorem 3.33 in [**ST**].) We fix a map $g \in C^{1/2}(\partial D)$ and define the closed affine subspace

$$A = \{u \in W^{1,2}(D) \mid \text{tr}\, u = g\}$$

of $W^{1,2}(D)$. We are looking for the potential u of the electric field $E = \nabla u$ in the vacuum D which is generated by the charge density g at the boundary. (We ignore the physical units and related constants.) A general principle in physics says that this potential $u \in A$ has to minimize the 'electrical energy'

$$\varphi(u) := \int_D |\nabla u|^2 \, dx$$

among all functions in A. We now show that such a minimizer exists.[5]

[5]One can also prove its uniqueness.

Proof. A more sopisticated version of the trace theorem yields a function $u_0 \in A$ so that this set is non-empty. (See Theorem 2.5.7 in [**Ne**].) There thus exists the number $\mu := \inf_{u\in A} \varphi(u) \ge 0$. We fix a sequence (u_n) in A with $\varphi(u_n) \to \mu$ as $n \to \infty$. The functions $\tilde u_n := u_n - u_0$ belong to $W_0^{1,2}(D) = N(\mathrm{tr})$. On $W_0^{1,2}(D)$ the quantity $|||v||| = \varphi(v)^{1/2}$ defines a norm which is equivalent to $\|\cdot\|_{1,2}$ as discussed at the end of Secton V.3 in [**We**]. We thus obtain the inequalities

$$\begin{aligned}\|u_n\|_{1,2} &\le \|\tilde u_n\|_{1,2} + \|u_0\|_{1,2} \le c\,|||\tilde u_n|||+\|u_0\|_{1,2}\\ &\le c(\varphi(u_n)^{1/2} + \varphi(u_0)^{1/2})\|u_0\|_{1,2}\end{aligned}$$

for all n; i.e., (u_n) is bounded in $W^{1,2}(D)$. This space is reflexive by Example 5.29. Theorem 5.40 thus yields a a subsequence $(u_{n_j})_j$ with a weak limit u in $W^{1,2}(D)$. Due to Exercise 13.4, also the continuous images $\mathrm{tr}\, u_{n_j} = g$ converge to $\mathrm{tr}\, u$ as $j \to \infty$ so that $\mathrm{tr}\, u = g$ and u is contained in A. We recall the isometric map $J : W^{1,2(D)} \to L^2(D)^4$; $v \mapsto (v, \partial_1 v, \partial_2, v, \partial_3 v)$ from Remark 4.16d). Using again Exercise 13.4, we then deduce that the partial derivatives $(\partial_k u_{n_j})_j$ tend weakly in $L^2(D)$ to $\partial_k u$. Proposition 5.33 now imples that

$$\varphi(u) = \sum_{k=1}^{3} \|\partial_k u\|_2^2 \le \liminf_{j\to\infty} \sum_{k=1}^{3} \|\partial_k u_{n_j}\|_2^2 \le \mu,$$

and hence $\varphi(u) = \mu$ by the definition of μ.[6] □

5.4. Adjoint operators

In this section we introduce the notions which allow to connect duality theory with linear operators.

Let X and Y be normed vector spaces and $T \in \mathcal{B}(X, Y)$. For each $y^\star \in Y^\star$ we define a map $\varphi_{y^\star} : X \to \mathbb{F}$ by setting $\varphi_{y^\star}(x) = \langle Tx, y^\star\rangle_Y$. It is clear that $\varphi_{y^\star}$ is linear in $x \in X$ and that $|\varphi_{y^\star}(x)| \le \|T\|\,\|y^\star\|$ if $\|x\| \le 1$. So we obtain $\varphi_{y^\star} \in X^\star$ with $\|\varphi_{y^\star}\| \le \|T\|\,\|y^\star\|$. Observe that $\varphi_{y^\star}$ is uniquely determined by $y^\star$ for a given T. We now introduce $T^\star y^\star := \varphi_{y^\star} \in X^\star$. we have thus defined a map

$$T^\star : Y^\star \to X^\star;\quad \langle x, T^\star y^\star\rangle_X = \langle Tx, y^\star\rangle_Y \quad (\forall\, x \in X,\ y^\star \in Y^\star), \tag{5.8}$$

which is called the *adjoint* of T.

If X and Y are Pre-Hilbert spaces, analogously we introduce the *Hilbert space adjoint* T' of T by

$$T' : Y \to X;\quad (x|T'y)_X = (Tx|y)_Y \quad (\forall\, x \in X,\ y \in Y). \tag{5.9}$$

Let X and Y be Hilbert spaces. We then obtain $T' = \Phi_X^{-1} T^\star \Phi_Y$ for the Riesz isomorphisms from Theorem 3.10.

[6] One can check that all partial derivatives $\partial_k u$ belong to $W^{1,2}_{\mathrm{loc}}(D)$ and that $\Delta u = (\partial_{11} + \partial_{22} + \partial_{33})u = 0$ on D, see Theorems 8.2.4 and 8.3.1 in [**Ev**].

Proposition 5.42. *Let X, Y and Z be normed vector spaces, $S,T \in \mathcal{B}(X,Y)$, $R \in \mathcal{B}(Y,Z)$, and $\alpha \in \mathbb{F}$. The following assertions hold.*

a) $T^\star \in \mathcal{B}(Y^\star, X^\star)$ with $\|T^\star\| = \|T\|$.

b) $(T+S)^\star = T^\star + S^\star$, $(\alpha T)^\star = \alpha T^\star$, and $(RT)^\star = T^\star R^\star$.

The analogous assertions (with $(\alpha T)' = \overline{\alpha} T'$) and $T = (T')' =: T''$ are true for Pre-Hilbert spaces and the Hilbert space adjoints.

Proof. Let $\alpha, \beta \in \mathbb{F}$, $x \in X$, $y^\star, u^\star \in Y^\star$ and $z^\star \in Z^\star$. For a), we compute

$$\begin{aligned}\langle x, T^\star(\alpha y^\star + \beta u^\star)\rangle &= \langle Tx, \alpha y^\star + \beta u^\star\rangle = \alpha\langle Tx, y^\star\rangle + \beta\langle Tx, u^\star\rangle \\ &= \alpha\langle x, T^\star y^\star\rangle + \beta\langle x, T^\star u^\star\rangle = \langle x, \alpha T^\star y^\star + \beta T^\star u^\star\rangle.\end{aligned}$$

Since $x \in X$ is arbitrary, this means that $T^\star(\alpha y^\star + \beta u\star) = \alpha T^\star y^\star + \beta T^\star u^\star$ and thus $T^\star$ is linear. Moreover, Corollary 5.10 yields

$$\begin{aligned}\|T\| &= \sup_{\|x\|\le 1} \|Tx\| = \sup_{\|x\|\le 1, \|y^\star\|\le 1} |\langle Tx, y^\star\rangle| = \sup_{\|x\|\le 1, \|y^\star\|\le 1} |\langle x, T^\star y^\star\rangle| \\ &= \sup_{\|y^\star\|\le 1} \|T^\star y^\star\| = \|T^\star\|,\end{aligned}$$

and assertion a) is shown. We further calculate

$$\langle x, (RT)^\star z^\star\rangle = \langle RTx, z^\star\rangle = \langle Tx, R^\star z^\star\rangle = \langle x, T^\star R^\star z^\star\rangle$$

so that $(RT)^\star = T^\star R^\star$. The remaing parts of b) and the Hilbert space variants of a) and b) are shown similarly. In the Hilbert setting, we finally compute

$$(Tx|y) = (x|T'y) = \overline{(T'y|x)} = \overline{(y|T''x)} = (T''x|y)$$

for all $y \in X$; i.e., $T = T''$. (Note that T'' exists due to a).) ☐

In view of the above result, each operator $T \in \mathcal{B}(X,Y)$ possesses its *bi-adjoint* $T^{\star\star} := (T^\star)^\star \in \mathcal{B}(X^{\star\star}, Y^{\star\star})$ with $\|T\| = \|T^{\star\star}\|$. We introduce important concepts in the Hilbert space setting.

Definition 5.43. *Let X and Y be Pre-Hilbert spaces and $T \in \mathcal{B}(X,Y)$. The operator T is called* unitary *if $T'T = I_X$ and $TT' = I_Y$ (i.e., it exists $T^{-1} = T'$). Let $X = Y$. Then T is called* self-adjoint *if $(Tx|y) = (x|Ty)$ for all $x, y \in X$ (i.e., $T = T'$).*

We compute adjoints for basic classes of operators, cf. the exercises.

Example 5.44. a) For $X = \mathbb{F}^m$ and a matrix $T = [a_{kl}]$, the adjoint $T^\star$ is given by $[a_{lk}]$ and T' by $[\overline{a_{lk}}]$. (See Linear Algebra.)

b) On $X = c_0$ or $X = \ell^p$ with $1 \le p < \infty$ (and also on ℓ^∞) we consider the shift operators $Lx = (x_{n+1})_n$ and $Rx = (0, x_1, x_2, \dots)$. Take $x \in X$ and $y \in X^\star$. Using Proposition 5.1, we calculate

$$\langle x, L^\star y\rangle = \langle Lx, y\rangle = \sum_{k=1}^\infty x_{k+1} y_k = \sum_{n=2}^\infty x_n y_{n-1} = \langle x, Ry\rangle.$$

Since x and y are arbitrary, we first obtain $L^\star y = Ry$ and then $L^\star = R$. Similarly one derives that $R^\star = L$ and, for $p = 2$, $L' = R$ and $R' = L$.

c) Let $X = L^2(\mathbb{R})$ and $(T(t)f)(s) = f(s+t)$ for $f \in X$ and $s, t \in \mathbb{R}$, see Example 4.12. For $f, g \in X$ we compute

$$(T(t)f|g) = \int_{\mathbb{R}} f(s+t)\overline{g(s)}\,\mathrm{d}s = \int_{\mathbb{R}} f(\tau)\overline{g(\tau - t)}\,\mathrm{d}\tau = (f|T(-t)g).$$

As in part b), it follows that $T(t)' = T(-t) = T(t)^{-1}$ and hence $T(t)$ is unitary. Analogously one sees that $T(t)^\star = T(-t)$ on $L^p(\mathbb{R})$ for $p \in [1, \infty)$ and $t \in \mathbb{R}$, employing Theorem 5.4.

d) Let $A \in \mathcal{B}_m$, $1 \le p < \infty$, and $k : A \times A \to \mathbb{F}$ be measurable with

$$\kappa_p = \left(\int_A \left(\int_A |k(x,y)|^{p'} \,\mathrm{d}y \right)^{p/p'} \mathrm{d}x \right)^{1/p} < \infty, \qquad \text{if } \; p > 1,$$

$$\kappa_1 = \operatorname*{ess\,sup}_{y \in A} \int_A |k(x,y)|\,\mathrm{d}x < \infty, \qquad \text{if } \; p = 1.$$

For $p = 2$ this means that $k \in L^2(A \times A)$. Let $p \in (1, \infty)$, $f \in L^p(A)$ and $g \in L^{p'}(A)$. The function $(x, y) \mapsto \varphi(x, y) = k(x, y)f(y)g(x)$ is measurable on $A \times A$ as a product of measurable functions. Using Fubini's Theorem 3.29 in Analysis 3 and Hölder's inequality (first in the y– and then in the x–integral), we deduce

$$\begin{aligned}
\int_{A\times A} |\varphi|\,\mathrm{d}(x,y) &= \int_A \int_A |k(x,y)|\,|f(y)|\,\mathrm{d}y\,|g(x)|\,\mathrm{d}x \\
&\le \int_A \Big(\int_A |k(x,y)|^{p'}\,\mathrm{d}y \Big)^{\frac{1}{p'}} \Big(\int_A |f(y)|^p\,\mathrm{d}y \Big)^{\frac{1}{p}} |g(x)|\,\mathrm{d}x \\
&\le \kappa_p\,\|f\|_p \Big(\int_A |g(x)|^{p'}\,\mathrm{d}x \Big)^{\frac{1}{p'}} = \kappa_p\,\|f\|_p\,\|g\|_{p'} < \infty.
\end{aligned}$$

Since φ is integrable on $A \times A$, by Fubini's theorem the integral

$$h_g(x) := \int_A k(x,y)\,f(y)\,g(x)\,\mathrm{d}y = g(x)\int_A k(x,y)\,f(y)\,\mathrm{d}y$$

exists for $x \notin N_{fg}$ and a null set N_{fg}, and h_g is measurable on A (after setting $h_g(x) = 0$ for all $x \in N_{fg}$). We now take $g_n = \mathbb{1}_{A\cap B(0,n)}$ for every $n \in \mathbb{N}$ and define the null set $N_f = \bigcup_n N_{fg_n}$. In this way we see that the function

$$Tf(x) := \begin{cases} \int_A k(x,y)\,f(y)\,\mathrm{d}y, & x \in A \setminus N_f, \\ 0, & x \in N_f, \end{cases} \tag{5.10}$$

exists and that it is measurable. Note that this definition does not depend on the representative of f.

Using again Hölder's inequality in the y–integral, we further estimate

$$\|Tf\|_p^p = \int_A \Big| \int_A k(x,y)\,f(y)\,\mathrm{d}y \Big|^p \,\mathrm{d}x$$

$$\leq \int_A \Big(\int_A |k(x,y)|^{p'} \,\mathrm{d}y \Big)^{\frac{p}{p'}} \Big(\int_A |f(y)|^p \,\mathrm{d}y \Big)^{\frac{p}{p}} \,\mathrm{d}x$$
$$= \kappa_p^p \, \|f\|_p^p \,,$$

so that Tf belongs to $L^p(A)$. The linearity of T is clear, and hence T is an element of $\mathcal{B}(L^p(A))$ with norm less or equal κ_p.

Since φ is integrable, Fubini's theorem finally implies

$$\langle Tf, g \rangle = \int_A \int_A k(x,y) f(y) \,\mathrm{d}y \, g(x) \,\mathrm{d}x = \int_A \int_A f(y) k(x,y) g(x) \,\mathrm{d}x \,\mathrm{d}y$$
$$= \int_A f(y) \int_A k(x,y) g(x) \,\mathrm{d}x \,\mathrm{d}y = \langle f, T^\star g \rangle.$$

As before, for each $g \in L^{p'}(A)$ his equality means that

$$T^\star g(y) = \int_A k(x,y) g(x) \, dx \qquad \text{for a.e. } y \in A. \tag{5.11}$$

For $p = 2$ one derives analogously

$$T' g(y) = \int_A \overline{k(x,y)} g(x) \,\mathrm{d}x$$

for every $g \in L^2(A)$ and a.e. $y \in A$. Hence, T is self adjoint if $k(x,y) = \overline{k(y,x)}$ for a.a. $x, y \in A$ (which is equivalent to 'for a.a. $y \in A$ we have $k(x,y) = \overline{k(y,x)}$ for a.a. $x \in A$' by Korollar 3.25 in Analysis 3).

Let $p = 1$. We first deduce from Fubini's theorem the inequality

$$\tau := \int_{A \times A} |k(x,y) f(y)| \,\mathrm{d}(x,y) = \int_A \int_A |k(x,y)| \,\mathrm{d}x \, |f(y)| \,\mathrm{d}y \leq \kappa_1 \, \|f\|_1$$

for all $f \in L^1(A)$. The function $(x,y) \mapsto k(x,y) f(y)$ thus is integrable on A^2. By Fubini's theorem, formula (5.10) now defines a function Tf in $L^1(A)$ satisfying $\|Tf\|_1 \leq \tau \leq \kappa_1 \|f\|_1$ for each $f \in L^1(A)$. As a result, T belongs to $\mathcal{B}(L^1(A))$ with $\|T\| \leq \kappa_1$. As above one can also check that $T^* \in \mathcal{B}(L^\infty(A))$ is given as in (5.11).[7] $\Diamond$

We now relate $T^{\star\star}$ to T also in non-Hilbertian Banach spaces.

PROPOSITION 5.45. *For normed vector spaces X and Y, the following assertions hold.*

a) For $T \in \mathcal{B}(X,Y)$ we have $T^{\star\star} \circ J_X = J_Y \circ T$.

b) If Y is reflexive, then $T = J_Y^{-1} T^{\star\star} J_X$.

PROOF. Let $x \in X$ and $y^\star \in Y^\star$. Using (5.8) and (5.7), we compute

$$\langle y^\star, T^{\star\star} J_X(x) \rangle_{Y^\star} = \langle T^\star y^\star, J_X(x) \rangle_{X^\star} = \langle x, T^\star y^\star \rangle_X = \langle Tx, y^\star \rangle_Y$$
$$= \langle y^\star, J_Y(Tx) \rangle_{Y^\star}.$$

These equalities yields assertion a) which implies assertion b). $\square$

One usually identifies T and $T^{\star\star}$ if X and Y are reflexive.

[7] In the lectures we have only sketched the case $p > 1$.

Range and kernel of bounded linear operators. Let $T \in \mathcal{B}(X,Y)$. Then the kernel $N(T)$ is closed, but $R(T)$ need not to be closed. An example is the Volterra operator given by

$$Tf(t) = \int_0^t f(s)\,\mathrm{d}s$$

for $f \in X = C([0,1])$ and $t \in [0,1]$. It is straightforward to check that T belongs to $\mathcal{B}(X)$ with $R(T) = \{g \in C^1([0,1]) \,|\, g(0) = 0\}$ This space is different from $\overline{R(T)} = \{g \in X \,|\, g(0) = 0\}$.

Leaving aside this difficulty for a moment, we now describe range and kernel of T by means of its adjoint and the annihilators from (5.6).

Proposition 5.46. *Let X and Y be normed vector spaces and $T \in \mathcal{B}(X,Y)$. Then the following assertions hold.*

a) $R(T)^\perp = N(T^\star)$.

b) $\overline{R(T)} = {}^\perp N(T^\star)$. Thus, $R(T)$ is dense if and only if $T^\star$ is injective.

c) $N(T) = {}^\perp R(T^\star)$. Hence, T is injective if $R(T^\star)$ is dense.

d) $\overline{R(T^\star)} \subseteq N(T)^\perp$.

e) Let X be reflexive. Then $\overline{R(T^\star)} = N(T)^\perp$. Hence, $R(T^\star)$ is dense if and only if T is injective.

Proof. a) Let $y^\star \in Y^\star$. The functional $y^\star$ belongs to $R(T)^\perp$ if and only if for all $x \in X$ we have $0 = \langle Tx, y^\star\rangle = \langle x, T^\star y^\star\rangle$, which is equivalent to $y^\star \in N(T^\star)$.

b) Proposition 5.22 and part a) show that $\overline{R(T)} = {}^\perp(R(T)^\perp) = {}^\perp N(T^\star)$. The addendum now follows from Remark 5.21.

c) Let $x \in X$. Due Corollary 5.10, the vector x is contained in $N(T)$ if and only if for all $y^\star \in Y^\star$ we have $0 = \langle Tx, y^\star\rangle = \langle x, T^\star y^\star\rangle$, which is equivalent to $x \in {}^\perp R(T^\star)$. Remark 5.21 yields the second part.

d) Proposition 5.22 and step a) imply that $\overline{R(T^\star)} = {}^\perp(R(T^\star)^\perp) = {}^\perp N(T^{\star\star})$. We thus have to prove the inclusion ${}^\perp N(T^{\star\star}) \subseteq N(T)^\perp$. Let $y^\star \in {}^\perp N(T^{\star\star})$ and take any $x \in N(T)$. Proposition 5.45 then yields the equality $T^{\star\star} J_X x = J_Y T x = 0$ so that $J_X x$ is an element of $N(T^{\star\star})$. We now infer $\langle x, y^\star\rangle = \langle y^\star, J_X x\rangle = 0$; i.e., $y^\star$ belongs to $N(T)^\perp$.

e) Let X be reflexive. It remains to show that $N(T)^\perp \subseteq {}^\perp N(T^{\star\star})$ in view of step d). So let $y^\star \in N(T)^\perp$ and take any $x^{\star\star} \in N(T^{\star\star})$. Because X is reflexive, there exists a vector $x \in X$ with $J_X x = x^{\star\star}$. Proposition 5.45 yields that $J_Y T x = T^{\star\star} J_X x = 0$. Since J_Y is injective by Proposition 5.24, the vector x is contained in $N(T)$, and hence $\langle y^\star, x^{\star\star}\rangle = \langle x, y^\star\rangle = 0$ so that $y^\star$ is an element of ${}^\perp N(T^{\star\star})$. □

Remark 5.47. In parts c) and d) of the above result, the converse implication and inclusion, respectively, do not hold in general. In fact, let $X = c_0$ and $T = I - L$ for the left shift L. If $Tx = 0$ then $x_n = x_{n+1}$ for all $n \in \mathbb{N}$ and thus $N(T) = \{0\}$. We have $T^\star = I - R$ by Example 5.44. In Example 1.25 of [**ST**] it is shown that the range of $I - R$ in $X^\star = \ell^1$ is not dense, hence $\overline{R(T^\star)} \neq X^\star = N(T)^\perp$. ◊

We next state an easy consequence of Proposition 5.46b). The lecture Spectral Theory will elaborate on this and the following results.

Corollary 5.48. *Let X and Y be normed vector spaces, $y \in Y$, and let $T \in \mathcal{B}(X,Y)$ have* closed range. *Then the equation $Tx = y$ has a solution $x_0 \in X$ if and only if we have $\langle y, y^\star \rangle = 0$ for all $y^\star \in N(T^\star)$. Every other solution is given by $x = x_0 + z$ for any $z \in N(T)$. Hence, T is surjective if and only if $R(T)$ is closed and $T^\star$ is injective.*

Using also the results of Chapter 4, we can now characterize the invertibility of T by injectivity properties of T and $T^\star$. We stress that usually injectivity is easier to check than surjectivity, but of course one must know the adjoint to apply the result.

Corollary 5.49. *Let X and Y be Banach spaces and $T \in \mathcal{B}(X,Y)$. The operator T is invertible if and only if*
a) $T^\star$ injective and
b) there is a constant $c > 0$ such that $\|Tx\| \geq c\,\|x\|$ for all $x \in X$.

Proof. Statement b) clearly implies the injectivity of T. From Corollary 4.31 we then deduce that b) is true if and only if T is injective and $R(T)$ is closed. Corollary 5.48 thus yields that the validity of a) and b) is equivalent to the bijectivity of T, and hence to its invertibility by Theorem 4.28. □

We can now prove that invertibility is preserved when taking adjoints. In contrast, by Proposition 5.46 injectivity and the density of the range are exchanged when passing to $T^\star$ at least in reflexive spaces.

Theorem 5.50. *Let X and Y be Banach spaces and $T \in \mathcal{B}(X,Y)$. The operator T is invertible if and only if $T^\star \in \mathcal{B}(Y^\star, X^\star)$ is invertible. In this case we have $(T^{-1})^\star = (T^\star)^{-1}$.*

Proof. 1) Let T be invertible. Since $I_X = T^{-1}T$, we obtain $I_{X^\star} = (I_X)^\star = T^\star (T^{-1})^\star$ by Proposition 5.42. Similarly, it follows $I_{Y^\star} = (T^{-1})^\star T^\star$. Hence, $T^\star$ has the inverse $(T^{-1})^\star \in \mathcal{B}(X^\star, Y^\star)$.

2) Let $T^\star$ be invertible, and thus injective. By step 1), $T^{\star\star}$ is invertible. Let $x \in X$. Propositions 5.24 and 5.45 imply the lower bound

$$\begin{aligned} \|x\| &= \|J_X x\| = \|(T^{\star\star})^{-1} T^{\star\star} J_X x\| \leq \|(T^{\star\star})^{-1}\|\, \|T^{\star\star} J_X(x)\| \\ &= \|(T^{\star\star})^{-1}\|\, \|J_Y T x\| = \|(T^{\star\star})^{-1}\|\, \|Tx\| \end{aligned}$$

From Corollary 5.49 we now deduce the invertibility of T. □

The next corollary was already stated in Remark 5.26, but only used in examples.

Corollary 5.51. *Let X be reflexive and $\Phi : X \to Y$ be an isomophism. Then Y is reflexive.*

Proof. Let $y^{\star\star} \in Y^{\star\star}$. Set $y = \Phi J_X^{-1}(\Phi^{\star\star})^{-1}y^{\star\star} \in Y$, using the easy part of the above theorem. Take $y^\star \in Y^\star$. By means of (5.8) and (5.7), we compute

$$\begin{aligned}\langle y, y^\star\rangle_Y &= \langle \Phi J_X^{-1}(\Phi^{\star\star})^{-1}y^{\star\star}, y^\star\rangle_Y = \langle J_X^{-1}(\Phi^{\star\star})^{-1}y^{\star\star}, \Phi^\star y^\star\rangle_X \\ &= \langle \Phi^\star y^\star, (\Phi^{\star\star})^{-1}y^{\star\star}\rangle_{X^\star} = \langle y^\star, \Phi^{\star\star}(\Phi^{\star\star})^{-1}y^{\star\star}\rangle_{Y^\star} = \langle y^\star, y^{\star\star}\rangle_{Y^\star},\end{aligned}$$

so that $J_Y y = y^{\star\star}$. □

We add a couple of results in Hilbert spaces which are used in the next chapter. The first one says that unitary operators preserve the full structure of Hilbert spaces.

Proposition 5.52. *Let X and Y be Hilbert spaces and $T \in \mathcal{B}(X,Y)$. Then equivalences are true.*

a) T is a isometry if and only if we have $(Tx|Tz)_Y = (x|z)_X$ for all $x, z \in X$.

b) T is unitary if and only if T is bijective and isometric if and only if T is bijective and preserves the scalar product.

Proof. a) The implication '⇐' is shown by setting $x = z$. To verify '⇒', take $\alpha \in \{1, \mathrm{i}\}$ and $x, z \in X$. Using (3.1) the isometry of T and $|\alpha| = 1$, we calculate

$$\begin{aligned}\|Tx + \alpha Tz)\|^2 &= \|Tx\|^2 + 2\,\mathrm{Re}\,(Tx|\alpha Tz) + \|\alpha Tz\|^2 \\ &= \|x\|^2 + 2\,\mathrm{Re}\,\bar{\alpha}\,(Tx|Tz) + \|z\|^2, \\ \|T(x+\alpha z)\|^2 &= \|x + \alpha z\|^2 = \|x\|^2 + 2\,\mathrm{Re}\,\bar{\alpha}\,(x|z) + \|z\|^2.\end{aligned}$$

It follows $\mathrm{Re}\,\bar{\alpha}\,(Tx|Tz) = \mathrm{Re}\,\bar{\alpha}\,(x|z)$ and thus assertion a).

b) The second equivalence is a consequence of step a). To show the first equivalence, take $x, z \in X$. If T is unitary, we obtain $(Tx|Tz) = (x|T'Tz) = (x|z)$ so that T is isometric by assertion a). If T is isometric, part a) yields $(T'Tx|z) = (Tx|Tz) = (x|z)$. Since $z \in X$ is arbitrary, we conclude that $T'Tx = x$ for all $x \in X$ and hence $T'T = I$. Now, the bijectivity of T implies that $T' = T^{-1}$. □

In the next two results we see that the numerical range $\{(Tx|x) \,|\, x \in \partial B(0,1)\}$ plays an important role for self-adjoint operators.

Proposition 5.53. *Let X be a Hilbert space with $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{B}(X)$. Then T is self-adjoint if and only if $(Tx|x) \in \mathbb{R}$ for every $x \in X$.*

Proof. The implication '⇒' follows from $(Tx|x) = (x|Tx) = \overline{(Tx|x)}$. To show the implication '⇐', take $\alpha \in \{1, -\mathrm{i}\}$ and $x, y \in X$. By means of $|\alpha| = 1$ and the assumption, we calculate

$$\begin{aligned}a &:= (T(x+\alpha y)|x+\alpha y) = (Tx|x) + \bar{\alpha}\,(Tx|y) + \alpha\,(Ty|x) + (Ty|y)\,, \\ \bar{a} &= (Tx|x) + \alpha\,(y|Tx) + \bar{\alpha}\,(x|Ty) + (Ty|y)\,.\end{aligned}$$

Since $a = \overline{a}$ by assumption, we obtain

$$(Tx|y) + (Ty|x) = (y|Tx) + (x|Ty) \qquad (\text{taking} \ \ \alpha = 1),$$
$$\mathrm{i}\,(Tx|y) - \mathrm{i}\,(Ty|x) = -\mathrm{i}\,(y|Tx) + \mathrm{i}\,(x|Ty) \qquad (\text{taking} \ \ \alpha = -\mathrm{i}).$$

These equations imply that $(Tx|y) = (x|Ty)$; i.e., T is self adjoint. □

PROPOSITION 5.54. *A self-adjoint operator $T \in \mathcal{B}(X)$ on a Hilbert space X satisfies*

$$\|T\| = \sup_{\|x\| \le 1} |(Tx|x)| =: M.$$

PROOF. The inequality "$\ge$" is clear. Let $x, y \in \partial B(0,1)$. Employing $T' = T$, we compute

$$(T(x+y)|x+y) - (T(x-y)|x-y)$$
$$= 2\,(Tx|y) + 2\,(Ty|x) = 2\,(Tx|y) + 2\overline{(Tx|y)} = 4\,\mathrm{Re}\,(Tx|y)\,.$$

Observe that $|(Tz|z)| \le M\|z\|^2$ for all $z \in X$. The above equation and (3.2) thus yield

$$4\,\mathrm{Re}\,(Tx|y) \le M\|x+y\|^2 + M\|x-y\|^2 = 2M(\|x\|^2 + \|y\|^2) = 4M.$$

If $Tx \neq 0$, we can replace y by $\tilde{y} := \|Tx\|^{-1}\,Tx$ in this inequality. We then obtain $\|Tx\| \le M$ for all $x \in X$ with $\|x\| = 1$. (This fact trivially holds if $Tx = 0$.) As a result, $\|T\| \le M$. □

We illustrate the above results by matrix examples.

REMARK 5.55. The non self-adjoint matrix $T = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ satisfies $(Tx|x) = 0$ for all $x \in \mathbb{R}^2$, so that we need $\mathbb{F} = \mathbb{C}$ in the implication '$\Leftarrow$' of Proposition 5.53 and the self-adjointness in Proposition 5.54. For $X = \mathbb{C}^2$, consider the non self-adjoint matrix $T = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$. Then $|(Tx|x)| \le \frac{1}{2}\,|x|_2^2$ for all $x \in \mathbb{C}^2$, but $\|T\| \ge |Te_2|_2 = 1$, which again shows that we need the self-adjointness in Proposition 5.54. ◊

PROPOSITION 5.56. *Let X be a Hilbert space and $P = P^2 \in \mathcal{B}(X)$ be orthogonal. We then have $\|P\| = 1$ (if $P \neq 0$), $P = P'$, and $(Px|x) = \|Px\|^2 \ge 0$ for all $x \in X$.*

PROOF. The first assertion was shown in Theorem 3.8. For $x, y \in X$ te vector $y - Py$ belongs to $N(P) \perp R(P)$ and thus

$$(Px|y) = (Px|Py + (I-P)y) = (Px|Py)\,,$$

and similarly $(x|Py) = (Px|Py)$. Therefore, $P = P'$. This fact further yields that $(Px|x) = (PPx|x) = \|Px\|^2$. □

In fact the above properties of a projection on a Hilbert space are all equivalent, see Satz V.5.9 in [**We**].

CHAPTER 6

The spectral theorem for compact self-adjoint operators

In the lecture Spectral Theory one establishes a diagonalization theorem for self-adjoint operators generalizing the corresponding result for hermitian matrices. Moreover, many of the properties related to the Jordan normal form will be extended to so-called compact operators. These theorems play a crucial role in mathematics and its applications. Their quite sophisticated proofs are out of reach in the present course. However, for compact *and* self-adjoint maps we can show them by our means below. Fortunately, this special case is sufficient for many examples. We first discuss the basic properties of compact operators.

6.1. Compact operators

Throughout X, Y and Z are Banach spaces. We study operators which provide compactness, using concepts from Section 1.3 freely.

Definition 6.1. *A linear map $T : X \to Y$ is called* compact *if $T\overline{B}(0,1)$ is relatively compact in Y. The set of all compact operators is denoted by $\mathcal{B}_0(X,Y)$. We put $\mathcal{B}_0(X) = \mathcal{B}_0(X,X)$.*

We start with some simple observations.

Remark 6.2. a) Let T be compact. Then the set $T\overline{B}(0,1)$ is bounded so that T is continuous; i.e., $\mathcal{B}_0(X,Y) \subseteq \mathcal{B}(X,Y)$.

b) The space of *operators of finite rank* is defined by

$$\mathcal{B}_{00}(X,Y) = \{T \in \mathcal{B}(X,Y) \,\big|\, \dim TX < \infty\},$$

cf. Example 5.16. For T in $\mathcal{B}_{00}(X,Y)$, the set $T\overline{B}(0,1)$ is relatively compact by Example 1.41, and hence $\mathcal{B}_{00}(X,Y) \subseteq \mathcal{B}_0(X,Y)$.

c) The identity $I : X \to X$ is compact if and only if $\overline{B}(0,1)$ is compact in X which is equivalent to $\dim X < \infty$ by Theorem 1.42.

d) For $T \in L(X,Y)$ the following assertions are equivalent.

(i) T is compact.
(ii) T maps bounded sets in X into relatively compact sets in Y.
(iii) For every bounded sequence $(x_n)_n$ in X there exists a convergent subsequence $(Tx_{n_j})_j$ in Y.

Proof. Let statement (i) be true. Take a bounded sequence (x_n) in X. Set $r = \sup_n \|x_n\|$. The images Tx_n then belong to the relatively compact set $T\overline{B}(0,r) = rT\overline{B}(0,1)$. Corollary 1.39 now implies assertion (iii). This corollary also yields the implication '(iii)⇒(ii)', whereas '(ii)⇒(i)' is clear. □

The next result yields the very useful fact that $\mathcal{B}_0(X,Y)$ is a closed 'two-sided ideal' in $\mathcal{B}(X,Y)$.

PROPOSITION 6.3. *The set $\mathcal{B}_0(X,Y)$ is a closed linear subspace of $\mathcal{B}(X,Y)$. Let $T \in \mathcal{B}(X,Y)$ and $S \in \mathcal{B}(Y,Z)$. If one of the operators T or S is compact, then ST is compact.*

PROOF. Take vectors $x_k \in X$, $k \in \mathbb{N}$, satisfying $c := \sup_k \|x_k\| < \infty$.

1) Let $T, R \in \mathcal{B}_0(X,Y)$ and $\alpha \in \mathbb{F}$. We then have converging subsequences $(Tx_{k_j})_j$ and $(Rx_{k_{j_l}})_l$. Hence also $(\alpha T x_{k_j})_j$ and $((T+R)x_{k_{j_l}})_l$ have a limit, so that $\mathcal{B}_0(X,Y)$ is a vector space.

2) Let $T_n \in \mathcal{B}_0(X,Y)$ tend in $\mathcal{B}(X,Y)$ to an operator T as $n \to \infty$. As in step 1), for each $n \in \mathbb{N}$ we find a subsequence $(x_{\nu_n(j)})_j$ of $(x_{\nu_{n-1}(j)})_j$ such that $(T_n x_{\nu_n(j)})_j$ converges. Set $u_m = x_{\nu_m(m)}$ for $m \in \mathbb{N}$. By construction, for every $n \in \mathbb{N}$ the sequence $(T_n u_m)_m$ has a limit. Let $\varepsilon > 0$. Fix an index $N = N_\varepsilon \in \mathbb{N}$ such that $\|T_N - T\| \le \varepsilon$. Take $m, k \ge N$ in $\mathbb{N}$. We then estimate

$$\begin{aligned}\|Tu_m - Tu_k\| &\le \|(T - T_N)u_m\| + \|T_N(u_m - u_k)\| + \|(T_N - T)u_k\| \\ &\le 2\varepsilon c + \|T_N(u_m - u_k)\|.\end{aligned}$$

Therefore (Tu_m) is a Cauchy sequence and thus has a limit. We have shown that T is compact and so $\mathcal{B}_0(X,Y)$ is closed.

3) Let $S \in \mathcal{B}_0(X,Y)$. Since $(Tx_k)_k$ is bounded, there is a converging subsequence $(STx_{k_j})_j$; i.e., ST is compact. Let $T \in \mathcal{B}_0(X,Y)$. We then have a subsequence $(Tx_{k_j})_j$ with a limit, and thus $(STx_{k_j})_j$ converges. Again, ST is compact. □

In the next examples we first note that strong limits may lose compactness. The typical examples of compact operators are integral operators on bounded sets. To ensure compactness on unbounded sets, the kernels have to decay at infinity sufficiently fast.

EXAMPLE 6.4. a) Strong limits of compact operators may fail to be compact. Consider, e.g., $X = \ell^2$ and $T_n x = (x_1, \dots, x_n, 0, 0, \dots)$ for all $x \in X$ and $n \in \mathbb{N}$. We have seen in Remark 4.9 that $T_n \to I$ strongly as $n \to \infty$. By Remark 6.2, each T_n belongs $\mathcal{B}_{00}(X) \subseteq \mathcal{B}_0(X)$ but I is not compact.

b) Let $X \in \{C([0,1]), L^p([0,1]) \,|\, 1 \le p \le \infty\}$, $k \in C([0,1]^2)$, and

$$Tf(t) = \int_0^1 k(t,\tau) f(\tau)\,\mathrm{d}\tau$$

for $f \in X$ and $t \in [0,1]$. As in Examples 1.49 and 2.7 one checks that T belongs to $\mathcal{B}_0(X, C([0,1]))$, using the Arzela–Ascoli Theorem. The map $J : C([0,1]) \to L^p(0,1)$; $f \mapsto f + \mathcal{N}$, is linear and bounded by Example 2.12. Proposition 6.3 now shows that $S = JT$ belongs to $\mathcal{B}_0(X, C([0,1]))$. Note that S is given as T if one ignores null functions.

c) Let $A \in \mathcal{B}_m$, $E = L^2(A)$ and $k \in L^2(A \times A)$. For $f \in E$, we set

$$Tf(x) = \int_A k(x,y) f(y)\,\mathrm{d}y, \qquad x \in A.$$

By Example 5.44, this defines an operator $T \in \mathcal{B}(E)$. We claim that T is compact.

Proof. We first extend k by 0 to a function $\tilde{k}$ in $L^2(\mathbb{R}^{2m})$, and analogously for f. As above, this kernel induces an operator $\tilde{T}$ in $\mathcal{B}(L^2(\mathbb{R}^m))$. For $f \in E$ we have $\tilde{T}\tilde{f} = Tf$ on A. The compactness of T thus follows from that of $\tilde{T}$. Hence, we restrict ourselves to the case $A = \mathbb{R}^m$ and drop the tilde.

Theorem 5.9 of Analysis 3 yields maps $k_n \in C_c(\mathbb{R}^{2m})$ that tend to k in E. Let T_n be the corresponding integral operators in $\mathcal{B}(E)$. There is a closed ball $B_n \subseteq \mathbb{R}^m$ such that $\operatorname{supp} k_n \subseteq B_n \times B_n$. We then have

$$T_n f(x) = \begin{cases} 0, & x \in \mathbb{R}^m \setminus B_n, \\ \int_{B_n} k_n(x,y) f(y)\,\mathrm{d}y, & x \in B_n. \end{cases}$$

for $f \in E$ and $n \in \mathbb{N}$. Let $R_n f = f|_{B_n}$. Fix $n \in \mathbb{N}$ and take a bounded sequence (f_k) in E. Arguing as in part b), one finds a subsequence such that the functions $R_n T_n f_{k_j}$ have a limit g in $C(B_n)$ as $j \to \infty$. Since B_n has finite measure, this sequence converges also in $L^2(B_n)$. The maps $T_n f_{k_j}$ thus tend in E to the 0–extension $\tilde{g}$ of g as $j \to \infty$, and so T_n is compact. Since $T - T_n$ is an integral operator on E with kernel $k - k_n \in L^2(\mathbb{R}^{2m})$, Example 5.44 shows the bound $\|T - T_n\| \le \|k - k_n\|_2$ for all $n \in \mathbb{N}$. The operators T_n thus converge to T in $\mathcal{B}(E)$ so that T is compact by Proposition 6.3. □

d) Let $X = L^2(\mathbb{R})$. For $f \in X$, we define

$$Tf(t) = \int_{\mathbb{R}} \mathrm{e}^{-|t-s|} f(s)\,\mathrm{d}s, \qquad t \in \mathbb{R}.$$

Theorem 2.14 yields that $T \in \mathcal{B}(X)$. The map T is not compact.[1]

Proof. For $f_n = \mathbb{1}_{[n,n+1]}$ and $n > m$ in $\mathbb{N}$ we compute $\|f_n\|_2 = 1$ and

$$\begin{aligned} \|Tf_n - Tf_m\|_2^2 &\ge \int_{n+1}^{n+2} \left| \int_n^{n+1} \mathrm{e}^{s-t}\,\mathrm{d}s - \int_m^{m+1} \mathrm{e}^{s-t}\,\mathrm{d}s \right|^2 \mathrm{d}t \\ &= \int_{n+1}^{n+2} \mathrm{e}^{-2t} \big(\mathrm{e}^{n+1} - \mathrm{e}^n - e^{m+1} + \mathrm{e}^m\big)^2\,\mathrm{d}t \\ &\ge \tfrac{1}{2}(\mathrm{e}^{-2n-2} - \mathrm{e}^{-2n-4})(\mathrm{e}^{n+1} - 2\mathrm{e}^n)^2 \\ &= \tfrac{1}{2}(\mathrm{e}^{-2} - \mathrm{e}^{-4})(\mathrm{e} - 2)^2 > 0. \end{aligned}$$

Hence, (Tf_n) has no converging subsequence. □

[1] This proof and the next example were omitted in the lectures.

e) The right shift R on c_0 or ℓ^p with $p \in [1,\infty]$ is not compact since the sequence $(Re_n) = (e_{n+1})$ has no converging subsequence (cf. Example 2.9). Similarly, one shows the non-compactness of the left shift L or of the translation $T(t)$, $t \in \mathbb{R}$, on $L^p(\mathbb{R})$, see Example 4.12.$\Diamond$

We show that a compact operator improves weak to strong limits.[2]

PROPOSITION 6.5. *Let $T \in \mathcal{B}_0(X,Y)$ and (x_n) tend weakly to x in X. Then the images Tx_n converge to Tx in Y as $n \to \infty$.*

PROOF. We have $\langle Tx_n - Tx, y^\star\rangle = \langle x_n - x, T^\star y^\star\rangle \to 0$ as $n \to \infty$ for each $y^\star \in Y^\star$ and hence $Tx_n \stackrel{\sigma}{\to} Tx$. Take any subsequence $(x_{n_j})_j$. It is bounded by Proposition 5.33. The compactness of T thus yields a subsubsequence $(Tx_{n_{j_l}})_l$ converging to some y in Y. There thus exist the weak limits $Tx_{n_{j_l}} \stackrel{\sigma}{\to} Tx$ and $Tx_{n_{j_l}} \stackrel{\sigma}{\to} y$ as $l \to \infty$, and hence $y = Tx$. Lemma 1.51 now implies that (Tx_n) tends to Tx in Y. $\square$

The following theorem by *Schauder* nicely connects duality with compactness. It will be used in Spectral Theory.

THEOREM 6.6. *An operator $T \in \mathcal{B}(X,Y)$ is compact if and only if its adjoint $T^\star \in \mathcal{B}(Y^\star, X^\star)$ is compact.*

PROOF. 1) Let T be compact and take $y_n^\star \in Y^\star$ with $\sup_{n\in\mathbb{N}} \|y_n^\star\| =: c < \infty$. The set $K = \overline{T\overline{B}_X(0,1)}$ is a compact metric space for the distance induced by $\|\cdot\|_Y$. We use the restrictions $f_n = y_n^\star|_K \in C(K)$ for $n \in \mathbb{N}$. Putting $c_1 := \max_{y\in K} \|y\| < \infty$, we obtain the bound

$$\|f_n\|_\infty = \max_{y\in K} |\langle y, y_n^\star\rangle| \le cc_1$$

for every $n \in \mathbb{N}$. Moreover, $(f_n)_{n\in\mathbb{N}}$ is equicontinuous since

$$|f_n(y) - f_n(z)| = |\langle y - z, y_n^\star\rangle| \le c\,\|y - z\|$$

for all $n \in \mathbb{N}$ and $y, z \in K$. Arzela–Ascoli's Theorem 1.47 thus yields a subsequence $(f_{n_j})_j$ with a limit in $C(K)$. We deduce that

$$\begin{aligned}\|T^\star y_{n_j}^\star - T^\star y_{n_l}^\star\|_{X^\star} &= \sup_{\|x\|\le 1} \left|\langle x, T^\star(y_{n_j}^\star - y_{n_l}^\star)\rangle\right| = \sup_{\|x\|\le 1} \left|\langle Tx, y_{n_j}^\star - y_{n_l}^\star\rangle\right| \\ &= \sup_{y\in K} |f_{n_j}(y) - f_{n_l}(y)|\end{aligned}$$

tends to 0 as $j, l \to \infty$. This means that $(T^\star y_{n_j}^\star)_j$ converges and so $T^\star$ is compact.

2) Let $T^\star$ be compact. By step 1), the bi-adjoint $T^{\star\star}$ is compact. Let $J_X : X \to X^{\star\star}$ be the isometry from Proposition 5.24. Proposition 5.45 says that $T^{\star\star}J_X = J_Y T$, and hence $J_Y T$ is compact by Proposition 6.3. If (x_n) is bounded in X, we thus obtain a converging subsequence $(J_Y T x_{n_j})_j$ which is Cauchy. Since J_Y is isometric, also $(Tx_{n_j})_j$ is Cauchy and thus has a limit; i.e., T is compact. $\square$

[2]The following two proofs were omitted in the lectures.

6.2. The spectral theorem

In this section we extend the diagonalization theorem for hermitian matrices to compact self-adjoint operators in a Hilbert space. We first introduce a few basic concepts and state simple facts.

Let $T \in \mathcal{B}(X)$. A number $\lambda \in \mathbb{F}$ is called *eigenvalue* of T with *eigenvector* $v \in X$ if $v \neq 0$ and $Tv = \lambda v$. We write

$$E_\lambda = N(\lambda I - T)$$

for the correspoding *eigenspace*. It is a closed linear subspace of X. (These and related concepts are intensively treated in Spectral Theory.)

Let now X be a Hilbert space. We denote by P_λ the orthogonal projection onto E_λ for an eigenvalue λ of $T \in \mathcal{B}(X)$, see Theorem 3.8. Assume that T is self-adjoint. Let $\lambda \neq \mu$ be eigenvalues of T with eigenvectors v and w, respectively. We then obtain

$$\lambda \in \mathbb{R}, \qquad \text{since} \;\; \lambda \|v\|^2 = (\lambda v|v) = (Tv|v) = (v|Tv) = \overline{\lambda}\, \|v\|^2; \tag{6.1}$$

$$\begin{aligned} E_\lambda \perp E_\mu\,, \quad \text{since} \;\; (\lambda-\mu)\,(v|w) &= (\lambda v|w) - (v|\mu w) \\ &= (Tv|w) - (v|Tw) = 0, \end{aligned} \tag{6.2}$$

using the self-adjointness and in (6.2) also that μ is real. Since $T = T'$ and X is reflexive, Proposition 5.46 yields

$$R(T)^\perp = N(T) \qquad \text{and} \qquad \overline{R(T)} = N(T)^\perp. \tag{6.3}$$

The following basic version of the *spectral theorem* says that a compact and self-adjoint operator largely behaves like a hermitian matrix. It possesses at most countably many eigenvalues and the non-zero ones have finite dimensional eigenspaces. The eigenvectors yield an orthonormal basis $\tilde{B}$ of X, in which T becomes an infinite diagonal matrix. The diagonal elements are the eigenvalues listed according to their multiplicity. See the comments below the theorem. Equivalently, one can write the image Tx as the series (6.4) over the basis vectors where T is represented by its eigenvalues. Since the non-zero eigenvalues form a null sequence (if there are infinitely many), one can often neglect all but finitely many and thus work on a finite dimensional subspace spanned by eigenvectors.

Theorem 6.7. *Let X be a Hilbert space with $\dim X = \infty$, and $T \in \mathcal{B}(X)$ be compact and self-adjoint. Then T has at most countably many eigenvalues. The pairwise different non-zero eigenvalues of T are denoted by λ_j for $j \in J$ and an index set $J \in \{\emptyset, \mathbb{N}, \{1,\dots,N\} \,\big|\, N \in \mathbb{N}\}$. We write $P_j = P_{\lambda_j}$ and $E_j = E_{\lambda_j}$ for $j \in J$, as well as P_0 for the orthogonal projection onto $N(T)$. Let $\lambda_0 = 0$ and $J_0 = J \cup \{0\}$. Moreover, the following assertions hold.*

a) $\nu_j := \dim E_j < \infty$ for $j \in J$, as well as $TP_j = P_jT$ and $P_jP_k = P_kP_j = 0$ for $j \neq k$ in J_0.

b) Let $J = \mathbb{N}$. Then $\lambda_j \to 0$ as $j \to \infty$.

c) Let $J \neq \mathbb{N}$. *Then* $\dim N(T) = \infty$.

d) $\|T\| = \sup_{j \in J_0} |\lambda_j|$.

e) $T = \sum_{j \in J} \lambda_j P_j$ *with convergence in* $\mathcal{B}(X)$.

f) We write $(\mu_k)_{k \in K}$ *with* $K \subseteq \mathbb{N}$ *for the eigenvalues* λ_j *repeated* ν_j *times for each* $j \in J$. *The space* $N(T)^\perp = \overline{R(T)}$ *has an orthonormal basis* B *consisting of eigenvectors* b_k, $k \in K$. *For each* $x \in X$ *we obtain*

$$Tx = \sum_{k \in K} \mu_k \, (x|b_k) b_k \, . \tag{6.4}$$

We can extend the basis B of $N(T)^\perp$ in part f) to an orthonormal basis $\tilde{B} = \{\tilde{b}_i | i \in I\}$ of X by means of Theorem 3.15, assuming[3] that X is separable if $\dim N(T) = \infty$ and taking an index set $I \subseteq \mathbb{Z}$. In the possibly two-sided sequence $(\tilde{b}_i)_{i \in I}$ we start with the basis vectors for $N(T)$. We obtain a sequence $(\tilde{\mu}_i)_{i \in I}$ by adding $\dim N(T)$–many zeros before the eigenvalues $(\mu_k)_{k \in K}$. Theorem 3.18 provides the isometric isomorphism

$$J : X \to \ell^2(I); \quad Jx = ((x|\tilde{b}_i))_{i \in I} \, ,$$

with inverse $J^{-1}(\alpha_i)_i = \sum_i \alpha_i \tilde{b}_i$. It is then easy to see that T is isomorphic to the inifinite diagonal matrix

$$JTJ^{-1} = \begin{pmatrix} \ddots & 0 & 0 \\ 0 & \tilde{\mu}_i & 0 \\ 0 & 0 & \ddots \end{pmatrix}$$

acting on the sequence space $\ell^2(I) = \{(\alpha_i))_{i \in I} \mid \sum_{i \in I} |\alpha_i|^2 < \infty\}$.

Proof of Theorem 6.7. 1) Let λ be an eigenvalue of T and v be a corresponding eigenvector. Note that λ is real by (6.1) and that the image $Tv = \lambda v$ belongs to E_λ. For $y \in E_\lambda^\perp$ we have $(Ty|v) = (y|Tv) = (y|\lambda v) = 0$, and hence T leaves also $E_\lambda^\perp$ invariant. We then compute

$$TP_\lambda x = P_\lambda T P_\lambda x = P_\lambda T(P_\lambda - I)x + P_\lambda Tx = P_\lambda Tx$$

for $x \in X$, using that $P_\lambda = I$ on E_λ and $P_\lambda = 0$ on $E_\lambda^\perp = (P_\lambda - I)X$. Formula (6.2) easily yields the last part of assertion a).

Let $\dim E_\lambda = \infty$. By Lemma 3.13, we obtain an orthonormal system $\{v_n \,|\, n \in \mathbb{N}\}$ in E_λ. Compactness provides a converging subsequence $(Tv_{n_k})_k$. Pythagoras' formula and orthonormality now imply

$$\|Tv_{n_k} - Tv_{n_l}\|^2 = \|\lambda v_{n_k} - \lambda v_{n_l}\|^2 = \lambda^2 \left(\|v_{n_k}\|^2 + \|v_{n_l}\|^2\right) = 2\lambda^2.$$

Since the left-hand side tends to 0 as $k, l \to \infty$, the eigenvalue λ has to be 0. So statement a) is shown.

2) The main step is the *claim*: $\|T\|$ or $-\|T\|$ is an eigenvalue of T.

If $T = 0$, then the claim and the theorem with $J = \emptyset$ are true. So let $T \neq 0$. Since $T = T'$, Proposition 5.54 provides vectors $x_n \in X$ with $\|x_n\| = 1$ such that the numbers $|(Tx_n|x_n)|$ tend to $\|T\|$ as $n \to \infty$.

[3] As noted in Section 3.2 this assumption can be avoided.

Note that $(Tx_n|x_n)$ is real by Proposition 5.53. There thus exists a subsequence $((Tx_{n_l}|x_{n_l}))_l$ with a limit λ in $\mathbb{R}$ and $|\lambda| = \|T\| \neq 0$. From formula (3.1) we then deduce

$$\begin{aligned}\|Tx_{n_l} - \lambda x_{n_l}\|^2 &= \|Tx_{n_l}\|^2 - 2\,\mathrm{Re}(Tx_{n_l}|\lambda x_{n_l}) + \lambda^2 \|x_{n_l}\|^2 \\ &\leq 2\lambda^2\|x_{n_l}\|^2 - 2\lambda(Tx_{n_l}|x_{n_l}) = 2\lambda^2 - 2\lambda(Tx_{n_l}|x_{n_l}) \to 0\end{aligned}$$

as $l \to \infty$, and hence $(Tx_{n_l} - \lambda x_{n_l})_l$ is a null sequence. Since T is compact, another subsequence $(Tx_{n_{l_k}})_k$ tends to a vector y in X. This fact yields the convergence

$$x_{n_{l_k}} = \tfrac{1}{\lambda}\big((\lambda I - T)x_{n_{l_k}} + Tx_{n_{l_k}}\big) \longrightarrow \lambda^{-1}y$$

as $k \to \infty$. In particular, y is nonzero as $\|x_{n_{l_k}}\| = 1$. Using the continuity of T, we further infer

$$y = \lim_{k\to\infty} Tx_{n_{l_k}} = \lambda^{-1}Ty,$$

so that $\lambda =: \lambda_1$ is an eigenvalue and the claim is established.

3) We next iterate step 2, starting from the the closed subspaces $X_1 := E_1$ and $X_2 := X_1^\perp$ of X. By part 1), the operator T leaves invariant X_2 and we can thus define the restriction $T_2 := T|_{X_2} \in \mathcal{B}(X_2)$. It is straightforward to check that T_2 is again self-adjoint and compact. If $T_2 \neq 0$, we set $J = \{1\}$ and stop the iteration. Otherwise, T_2 has an eigenvalue $\lambda_2 \neq 0$ with $|\lambda_2| = \|T_2\| \leq \|T\| = |\lambda_1|$ due to the claim in 2). Observe that $\lambda_2 \neq \lambda_1$ since $X_2 \cap E_1 = \{0\}$ and eigenvectors of T_2 are also eigenvectors of T. Statement (6.2) thus shows that all eigenvectors of T for λ_2 belong to X_2; i.e., $N(\lambda_2 I - T_2) = N(\lambda_2 I - T) = E_2$.

We now iterate this procedure obtaining the restriction $T_j = T|_{X_j}$ to the orthogonal complement $X_j = (E_1 \oplus \cdots \oplus E_{j-1})^\perp$. If $T_{N+1} = 0$ for some $N \in \mathbb{N}$, we stop and take $J = \{1, \ldots, N\}$. With arguments as below one can finish the proof in this simpler case.

We focus on the other alternative that $J = \mathbb{N}$. Here we obtain a sequence $(\lambda_j)_{j\in\mathbb{N}}$ of non-zero eigenvalues with $\|T_j\| = |\lambda_j| \geq |\lambda_{j+1}|$ for all $j \in \mathbb{N}$. In this case the absolute values $|\lambda_j|$ tend to a number $\alpha \geq 0$ as $j \to \infty$. Moreover, assertion d) has been shown.

4) To check that $\alpha = 0$, we take a unit vector $v_n \in E_n$ for each $n \in \mathbb{N}$. The vectors are pairwise orthogonal by (6.2). Compactness yields a converging subsequence $(Tv_{n_l})_l$. Employing Pythagoras' identity, we calculate

$$\|Tv_{n_l} - Tv_{n_k}\|^2 = \|\lambda_{n_l}v_{n_l} - \lambda_{n_k}v_{n_k}\|^2 = \lambda_{n_l}^2\|v_{n_l}\|^2 + \lambda_{n_k}^2\|v_{n_k}\|^2 = \lambda_{n_l}^2 + \lambda_{n_k}^2.$$

as $k, l \to \infty$, the left-hand side tends to 0 and the right-hand side to $2\alpha^2$ so that $\alpha = 0$ and statement b) is true.

Let $x \in X$ and $n \in \mathbb{N}$. Set $Q_n = P_1 + \cdots + P_n$. Since the projections P_j onto E_n are orthogonal to each other and have the kernels $E_n^\perp$,

Q_n is a projection with range X_n and kernel X_{n+1} so that it is also orthogonal. Using the equations $\lambda_j P_j = TP_j = P_j T$, we compute

$$Tx - \sum_{j=1}^{n} \lambda_j P_j x = Tx - \sum_{j=1}^{n} TP_j x = T(I - Q_n)x.$$

Assertion b) then implies that

$$\Big\| T - \sum_{j=1}^{n} \lambda_j P_j \Big\| \le \|T_{n+1}\| \, \|I - Q_n\| = |\lambda_{n+1}| \;\longrightarrow\; 0$$

as $n \to \infty$ and hence part e) is valid.

5) The Gram–Schmidt Lemma 3.13 yields a orthonormal basis of eigenvectors in each eigenspace $E_j = P_j X$ whose union B forms a orthonormal system because of (6.2). The subspace $L = \operatorname{lin}\{P_j x \,|\, x \in X, \, j \in \mathbb{N}\}$ is contained in $R(T)$ and $R(T)$ in $\overline{L}$ because of assertion e), so that $\overline{R(T)} = \overline{L}$. Theorem 3.15 now implies that B is a orthonormal basis of $\overline{R(T)}$. The operator T does not have another non-zero eigenvalue, since its eigenvector would belong to $R(T)$ and be orthogonal to B by (6.2). Assertion e) and Theorem 3.15 yield

$$Tx = \sum_{k \in K} (Tx|b_k) b_k = \sum_{k \in K} \mu_k \, (x|b_k) b_k$$

and the other parts of assertion f), using also 6.3. Finally, if $J \neq \mathbb{N}$, formula 6.3 implies statement c) since the spaces E_j are finite dimensional and $\dim X = \infty$. □

We next sketch a standard application of the above result to boundary value problems for ordinary differential equations. This class of problems is the source for many orthonormal bases used in mathematics and other sciences. We follow Section II.6 of [**Co**].

Example 6.8. Let $q \in C([0,1], \mathbb{R})$ and $\alpha_j, \beta_j \in \mathbb{R}$ with $\alpha_j^2 + \beta_j^2 > 0$ for $j \in \{0,1\}$. In $X = L^2(0,1)$ we define the linear map $A : D(A) \to X$ by

$$Au = -u'' + qu,$$
$$D(A) = \{u \in W^{2,2}(0,1) \,|\, \alpha_j u(j) + \beta_j u'(j) = 0 \;\text{ for } j \in \{0,1\}\}.$$

(Note that the boundary conditions are included in the domain.) Here we write u' instead of $\partial_1 u$, $W^{2,2}(0,1)$ is the space of $u \in W^{1,2}(0,1)$ with $u' \in W^{1,2}(0,1)$, and we use that $W^{1,2}(0,1) \hookrightarrow C([0,1])$, cf. Theorem 3.11 in [**ST**]. We *assume* that A is injective.

As in Satz 4.17 of Analysis 4 one can check that for each $f \in X$ there is a unique function $u \in D(A)$ fulfilling $Au = f$ and given by

$$u(t) = \int_0^1 g(t,s) f(s) \, \mathrm{d}s =: (Tf)(t), \qquad t \in [0,1],$$

for a continuous map $g : [0,1]^2 \to \mathbb{R}$. Lemma II.6.8 of [**Co**] shows that $g(t,s) = g(s,t)$ for all $t, s \in [0,1]$. The linear operator $T : X \to X$ is compact and self-adjoint by Examples 5.44 and 6.4. It satisfies $R(T) = D(A)$, $ATf = f$ for $f \in X$, and $TAu = u$ for $u \in D(A)$, see Theorem II.6.9 in [**Co**]. Finally, for all eigenvalues $\lambda \neq 0$ of T, the eigenspace $N(\lambda I - T)$ is one-dimensional by Lemma II.6.11 of [**Co**].

The Spectral Theorem 6.7 and some calculations then provide eigenvalues $\mu_n \in \mathbb{R} \setminus \{0\}$ with $|\mu_n| \to \infty$ as $n \in \mathbb{N}$ and an orthonormal basis of X consisting of eigenvectors $v_n \in D(A) \setminus \{0\}$ with $Av_n = \mu_n v_n$ for $n \in \mathbb{N}$, Let $f \in X$. We obtain the Fredholm alternative:

a) Let $\mu \neq \mu_n$ for all $n \in \mathbb{N}$. Then there is a unique solution $u \in D(A)$ of the equation $Au - \mu u = f$.

b) Let $\mu = \mu_n$ for some $n \in \mathbb{N}$. Then there is solution $u \in D(A)$ of $Au - \mu u = f$ if and only if $(f|v_n) = 0$. Each other solution is of the form $u + \alpha v_n$ for some $\alpha \in \mathbb{F}$.

See Theorem II.6.12 in [**Co**]. $\Diamond$

Bibliography

[Br] H. Brézis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Springer–Verlag, 2011.

[Co] J.B. Conway, *A Course in Functional Analysis.* 2nd edition. Springer–Verlag, 1990.

[Do] M. Dobrowolski, *Angewandte Funktionalanalysis.* Springer–Verlag, 2005.

[DS] N. Dunford and J.T. Schwartz, *Linear Operators Part I: General Theory.* John Wiley & Sons, 1957.

[El] J. Elstrodt, *Maß- und Integrationstheorie.* 5te korrigierte Auflage. Springer–Verlag, 2007.

[Ev] L.C. Evans, *Partial Differential Equations.* American Math. Society, 1998.

[Ko] T.W. Körner, *Fourier Analysis.* Cambridge University Press, 1989.

[LT] J. Lindenstraus and L. Tzafriri, *Classical Banach Spaces 1: Sequence spaces.* Springer–Verlag, 1996.

[Lu] A. Lunardi, *Interpolation Theory.* 2nd Edition. Edizioni della Normale, Pisa, 2009.

[Ne] J. Nečas, *Direct Methods in the Theory of Elliptic Equations.* Translated from the 1967 French original. Springer–Verlag, 2012.

[Ru] W. Rudin, *Functional Analysis.* 2nd edition. McGraw–Hill, 1991.

[ST] R. Schnaubelt, *Spectral Theory.* Class Notes. Karlsruhe, 2015.

[We] D. Werner, *Funktionalanalysis.* 5te erweiterte Auflage. Springer–Verlag, 2005.

[Yo] K. Yosida, *Functional Analysis.* Reprint of the 1980 Edition, Springer–Verlag, 1995.