Lecture Notes

Functional Analysis (2014/15)

Roland Schnaubelt

These lecture notes are based on my course from winter semester 2014/15. Compared to the notes from three years ago, several details and very few subjects have been changed. I mostly kept the contents of the results discussed in the lectures, but the numbering has been shifted in some parts. Typically, the proofs and calculations in the notes are a bit shorter than those given in class. Moreover, the drawings and many additional, mostly oral remarks from the lectures are omitted here. On the other hand, in the notes I have added a few results (e.g., the Riesz–Thorin theorem) and a couple of proofs for peripheral statements not presented during the course.

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, December 30, 2019 Roland Schnaubelt

Contents

Chapter 1

Banach spaces

1.1 Basic properties of Banach and metric spaces

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

Definition 1.1. *A* seminorm *on X is a mapping $p : X \to \mathbb{R}_+$ 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 in addition

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

holds for all $x \in X$, then p is called a norm. *One usually 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.

Definition 1.2. *Let $\|\cdot\|$ be a seminorm on a vector space X. A sequence $(x_n)_{n\in\mathbb{N}} = (x_n)$ 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 called a Banach space *if each Cauchy sequence in $(X, \|\cdot\|)$ converges in X. Then one also calls $\|\cdot\|$ or $(X, \|\cdot\|)$* complete.

Remark 1.3. Let $\|\cdot\|$ be a seminorm on a vector space X and (x_n) be a sequence in X. The following results are shown as in Analysis 2, see also Remark 1.7.

a) $|\|x\| - \|y\|| \leq \|x - y\|$ for all $x, y \in X$.

b) The vector 0 has the norm 0.

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

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

e) Limits are linear: If $x_n \to x$ and $y_n \to y$ in X as $n \to \infty$ and $\alpha, \beta \in \mathbb{F}$, then $\alpha x_n + \beta y_n$ converges 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 $x = y$. ♢

Let X be a vector space and $M \neq \emptyset$ be a set. For functions $f, g : M \to X$ and numbers $\alpha, \beta \in \mathbb{F}$ one defines the functions

$$f + g : M \to X;\ (f+g)(t) = f(t) + g(t) \quad \text{and} \quad \alpha f : M \to X;\ (\alpha f)(t) = \alpha f(t).$$

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

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

Let $\alpha, \beta \in \mathbb{R}$. We then write $f \geq \alpha$ ($f > \alpha$, respectively) if $f(t) \geq \alpha$ ($f(t) > \alpha$, respectively) for all $t \in M$. Similarly one defines $\alpha \leq f \leq \beta$, $f \leq g$, and so on.

Example 1.4. a) $X = \mathbb{F}^d$ is a Banach space with respect to the norms

$$|x|_p = \Big(\sum_{k=1}^{d} |x_k|^p \Big)^{\frac{1}{p}}, \quad 1 \leq p < \infty, \quad \text{and} \quad |x|_\infty = \max\{|x_k| \mid k = 1, \ldots, d\}$$

where $x = (x_1, \ldots, x_d) \in \mathbb{F}^d$, see Analysis 1+2 or the proof of Proposition 1.30.

b) $X = C([0,1]) = \{f : [0,1] \to \mathbb{F} \mid f \text{ is continuous}\}$ with the supremum norm

$$\|f\|_\infty := \sup_{t \in [0,1]} |f(t)| \quad (= \max_{t \in [0,1]} |f(t)|).$$

is a Banach space. We always endow X with this norm, unless something else is specified. In fact, it is clear that X is a vector space and that $\|f\|_\infty = 0$ implies $f = 0$. We further compute

$$\|\alpha f\|_\infty = \sup_{t \in [0,1]} |\alpha|\ |f(t)| = |\alpha| \sup_{t \in [0,1]} |f(t)| = |\alpha|\ \|f\|_\infty,$$
$$\|f + g\|_\infty = \sup_{t \in [0,1]} |f(t) + g(t)| \leq \sup_{t \in [0,1]} (|f(t)| + |g(t)|) \leq \|f\|_\infty + \|g\|_\infty,$$

for all $f, g \in X$ and $\alpha \in \mathbb{F}$, so that X is a normed vector space. Let (f_n) be a Cauchy sequence in X; i.e., for all $\varepsilon > 0$ there is an index $N_\varepsilon \in \mathbb{N}$ with

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

for all $n, m \geq N_\varepsilon$ and $t \in [0,1]$. We stress that N_ε does not depend on t. By this estimate, $(f_n(t))_{n \in \mathbb{N}}$ is a Cauchy sequence in $\mathbb{F}$. Since $\mathbb{F}$ is complete, there exists $f(t) := \lim_{n \to \infty} f_n(t)$ in $\mathbb{F}$ for each $t \in [0,1]$. Let $t \in [0,1]$ and $\varepsilon > 0$ be given. Take N_ε from above and $n \geq N_\varepsilon$. We then deduce

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

Since this inequality holds for all $t \in [0,1]$, it follows that $\|f - f_n\|_\infty \leq \varepsilon$ for all $n \geq N_\varepsilon$. Hence, f_n converges *uniformly* (in t) to f, and so f is continuous by a convergence theorem from Analysis 1+2. Summing up, we have shown that f_n converges in X to $f \in X$, as required.

c) We further introduce $\|f\|_1 = \int_0^1 |f(t)|\, dt$ for $f \in X = C([0,1])$. The number $\|f - g\|_1$ corresponds to area between the graphs of f and g (if $\mathbb{F} = \mathbb{R}$), i.e., $\|f - g\|_1$ is the distance in the mean; whereas $\|f - g\|_\infty$ corresponds to the maximal vertical distance between the graphs. We claim that $\|\cdot\|_1$ is an **incomplete** norm on X.

In fact, let $\alpha \in \mathbb{F}$ and $f, g \in X$. Clearly, $\|f\|_1 \geq 0$. If $f \neq 0$, then there are $0 \leq a < b \leq 1$ and $\delta > 0$ such that $|f(t)| \geq \delta$ for all $t \in [a,b]$, since f is continuous. It then follows

$$\|f\|_1 \geq \int_a^b |f(t)|\, dt \geq (b-a)\delta > 0.$$

We also obtain

$$\|\alpha f\|_1 = \int_0^1 |\alpha|\,|f(t)|\,dt = |\alpha|\,\|f\|_1,$$
$$\|f+g\|_1 = \int_0^1 |f(t)+g(t)|\,dt \le \int_0^1 (|f(t)|+|g(t)|)\,dt = \|f\|_1 + \|g\|_1.$$

As a result, $\|\cdot\|_1$ is a norm on X. Next, consider the functions given by

$$f_n(t) = \begin{cases} 1, & \frac{1}{2} \le t \le 1, \\ nt - \frac{n}{2} + 1, & \frac{1}{2} - \frac{1}{n} \le t \le \frac{1}{2}, \\ 0, & 0 \le t \le \frac{1}{2} - \frac{1}{n}, \end{cases}$$

for $n \in \mathbb{N}$ with $n \ge 3$. For $m \ge n \ge 3$ we have

$$\|f_n - f_m\|_1 = \int_{\frac{1}{2}-\frac{1}{n}}^{\frac{1}{2}} |f_n(t) - f_m(t)|\,dt \le \frac{1}{n} \longrightarrow 0$$

as $n \to \infty$, so that (f_n) is a Cauchy sequence with respect to $\|\cdot\|_1$. We suppose that (f_n) would converge to some $f \in X$ with respect to $\|\cdot\|_1$. Fix any $a \in (0, \frac{1}{2})$ and take $n \in \mathbb{N}$ with $\frac{1}{2} - \frac{1}{n} \ge a$. We then obtain that

$$0 \le \int_0^a |f(t)|\,dt = \int_0^a |f(t) - f_n(t)|\,dt \le \|f_n - f\|_1 \longrightarrow 0$$

as $n \to \infty$, i.e., $\int_0^a |f(t)|\,dt = 0$. The continuity of f implies as above that $f = 0$ on $[0,a]$ for every $a < \frac{1}{2}$, and hence $f(1/2) = 0$. On the other hand,

$$0 \le \int_{\frac{1}{2}}^1 |f(t) - 1|\,dt = \int_{\frac{1}{2}}^1 |f(t) - f_n(t)|\,dt \le \|f - f_n\|_1 \longrightarrow 0$$

as $n \to \infty$. As a consequence, f is equal to 1 on $[\frac{1}{2}, 1]$, which is a contradiction.

d) Let $X = C(\mathbb{R})$ and $a < b$ in $\mathbb{R}$. We see as in b) that $p(f) = \sup_{t\in[a,b]} |f(t)|$ 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. ◊

Remark. The vector space $X = C([0,1])$ is infinite dimensional since the functions $p_n \in X$, $n \in \mathbb{N}$, given by $p_n(t) = t^n$ are linearly independent. If one interprets a function $0 \le u \in X$ as a mass density, then $\|u\|_1$ is the total mass and $\|u\|_\infty$ is the maximal density. ◊

Before discussing further examples, we want to investigate various 'topological' concepts in a more general framework without vector space structure.

Definition 1.5. *A* distance *d on a set $M \ne \emptyset$ is a map $d : M \times M \to \mathbb{R}_+$ satisfying*

a) $d(x,y) = 0 \iff x = y$ *(definiteness),*

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

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

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

$$\forall\,\varepsilon > 0 \;\; \exists\, N_\varepsilon \in \mathbb{N} \;\; \forall\, n \ge N_\varepsilon : \quad d(x, x_n) \le \varepsilon, \tag{1.1}$$

and it is a Cauchy sequence *if*

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

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

Example 1.6. a) Let X be a normed vector space and $\emptyset \neq M \subseteq X$. Set $d(x,y) = \|x-y\|$ for $x,y \in M$. Then Definition 1.5a) follows from Definition 1.1c), 1.5b) from 1.1a) with $\alpha = -1$, and 1.5c) from 1.1b). As a result, (M,d) is a metric space (but not a vector space, in general).

b) Let $N \subseteq M$ and d be a metric on M. Then $d_N(x,y) = d(x,y)$ for $x,y \in N$ defines the *subspace metric* d_N on N. One often writes d instead of d_N.

c) Let $M \neq \emptyset$ be any set. One defines the *discrete metric* on M by setting $d(x,x) = 0$ and $d(x,y) = 1$ 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 $x_n \to x$ as $n \to \infty$ in the discrete metric if and only if there is an $m \in \mathbb{N}$ such that $x_n = x$ for all $n \geq m$. ◊

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. (In fact, we have $d(x_n,x_m) \leq d(x_n,x) + d(x,x_m) \leq 2\varepsilon$ for each $\varepsilon > 0$ and all $n,m \geq N_\varepsilon$ with N_ε from (1.1).)

b) If $x_n \to x$ and $x_n \to y$ in M for some $x,y \in X$, then $x = y$. (In fact, we have $d(x,y) \leq d(x,x_n) + d(x_n,y) \leq 2\varepsilon$ for each $\varepsilon > 0$ and all $n \geq \max\{N_\varepsilon(x), N_\varepsilon(y)\}$, where $N_\varepsilon(x)$ and $N_\varepsilon(y)$ are given by (1.1) for x and y, respectively.)

c) If (x_n) converges or is Cauchy, then it is *bounded*, i.e., there is a $z \in M$ such that $\sup_n d(x_n,z) < \infty$. (In fact, there is an $N \in \mathbb{N}$ such that $d(x_n,x_N) \leq 1$ for all $n \geq N$, and hence $d(x_n,x_N) \leq 1 + \max\{d(x_j,x_N) \,|\, j = 1,\dots,N\}$ 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 a $k \in \mathbb{N}$ with $p_k(x) > 0$. Then*

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

defines a metric on X such that $d(x_n,x) \to 0$ as $n \to \infty$ if and only if $p_j(x_n - x) \to 0$ as $n \to \infty$ for each $j \in \mathbb{N}$.

Proof. Note that the function $\varphi(t) = t/(1+t)$ increases strictly on $\mathbb{R}_+$, that $\varphi(0) = 0$ and that $\varphi(t) \in (0,1)$ for $t > 0$. In particular, the series in the statement converges in $\mathbb{R}_+$. Let $x,y,z,x_n \in X$ for $n \in \mathbb{N}$. We have $d(x,y) = 0$ if and only if $p_j(x-y) = 0$ for all $j \in \mathbb{N}$ which is equivalent to $x = y$ by the assumptions. Moreover, the identity $d(x,y) = d(y,x)$ holds since $p_j(x-y) = p_j(y-x)$ for each $j \in \mathbb{N}$. Using the monotonicity of φ, we further estimate

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

Thus, d is a metric on X.

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

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

so that $p_j(x-x_n) \to 0$ as $n \to \infty$.

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

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

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

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

for all $n \ge N_\varepsilon$. □

Example 1.9. a) Let $X = C(\mathbb{R})$ and $p_j(f) = \max_{|t|\le j} |f(t)|$ for $j \in \mathbb{N}$. Since each p_j is a seminorm and $p_j(f) = 0$ for all $j \in \mathbb{N}$ yields $f = 0$, Proposition 1.8 shows that X possesses the metric

$$d(f, g) := \sum_{j=1}^{\infty} 2^{-j} \frac{\max_{|t|\le j} |f(t) - g(t)|}{1 + \max_{|t|\le j} |f(t) - g(t)|},$$

and that $d(f_n, f) \to 0$ as $n \to \infty$ if and only if $\max_{|t|\le j} |f_n(t) - f(t)| \to 0$ as $n \to \infty$ for each $j \in \mathbb{N}$. In particular, (X, d) is endowed with the 'uniform convergence on compact sets' as introduced in Funktionentheorie.

Moreover, (X, d) is complete. In fact, let (f_n) be Cauchy in (X, d). As in the proof of Proposition 1.8 we see that

$$\forall\, \varepsilon \in (0, 1/2),\ j \in \mathbb{N}\ \ \exists\, N_{\varepsilon,j} \in \mathbb{N}\ \ \forall\, n, m \ge N_{\varepsilon,j} : \quad \max_{|t|\le j} |f_n(t) - f_m(t)| \le \varepsilon.$$

Since $C([-j, j])$ is Banach space for the supremum norm, for each $j \in \mathbb{N}$ there is a function $g^{(j)} \in C([-j, j])$ such that the restrictions $f_n|[-j, j]$ converge to $g^{(j)}$ in $C([-j, j])$. This convergence implies that $g^{(j)}(t) = g^{(k)}(t)$ whenever $t \in [-j, j] \subset [-k, k]$. We can thus define a function $f \in C(\mathbb{R})$ by setting $f(t) := g^{(j)}(t)$ for any $j \in \mathbb{N}$ with $|t| \le j$. By construction, we have $p_j(f_n - f) \to 0$ for each $j \in \mathbb{N}$, and hence $d(f_n, f) \to 0$ as $n \to \infty$.

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 $f_n(t) = |t|$ if $|t| \le n$ and $f_n(t) = n$ otherwise. Then f_n converges in (X, d) to the function f given by $f(t) = |t|$. 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 $f \notin Y$. ◊

In a metric space (M, d) we write $B(x, r) = \{y \in X \,|\, d(x, y) < r\}$ for the *open ball* with center $x \in M$ and radius $r > 0$, and $\overline{B}(x, r) = \{y \in X \,|\, d(x, y) \le r\}$ for the *closed ball*. 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 $r_x > 0$ such that $B(x, r_x) \subseteq O$. Moreover, $\emptyset$ is open. Let $x_0 \in M$. A subset $N \subseteq M$ is called a* neighborhood *of x_0 and x_0 is an* interior point *of N if there is a radius $r_0 > 0$ such that $B(x_0, r_0) \subseteq N$. Finally, a subset $A \subseteq M$ is called* closed *if $M \setminus A$ is open.*

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

a) The ball $B(x, r)$ is open. In fact, for any given $y \in B(x, r)$ we set $\rho = r - d(x, y) > 0$. Let $z \in B(y, \rho)$. We then have

$$d(z, x) \le d(z, y) + d(y, x) < \rho + d(x, y) = r.$$

This estimate yields $z \in B(x,r)$, and hence $B(y,\rho) \subseteq B(x,r)$, as required.

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

$$R = d(x,y) \leq d(x,z) + d(z,y) < d(x,z) + R - r,$$

and thus $d(x,z) > r$. This fact implies that $B(y, R-r) \subset M \setminus \bar{B}(x,r)$, so that $M \setminus \bar{B}(x,r)$ is open, and hence $\bar{B}(x,r)$ 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 $x \in M$.

d) In the discrete metric of Example 1.6 we have $\{x\} = B(x,1)$ for each $x \in M$, and hence $\{x\}$ is open. ◊

Proposition 1.12. *In a metric space M the following equivalences hold.*

(a) $A \subseteq M$ is closed $\Leftrightarrow$ If $x_n \in A$ and $x_n \to x$ in M as $n \to \infty$, then $x \in A$.

(b) $O \subseteq M$ is open $\Leftrightarrow$ For no $x \in O$ there exists (x_n) in $M \setminus O$ converging to x.

Proof. By definition, $O \subseteq M$ is open if and only if for each $x \in O$ there is an $r > 0$ such that for all $y \in M$ with $d(x,y) < r$ it follows that $y \in O$. 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. □

Corollary 1.13. *Let M be a complete metric space and $A \subseteq M$. If A is closed, then A is complete for the subspace metric d_A. In particular, if X is a Banach space and Y is closed linear subspace, then Y is a Banach space for the restriction $\|\cdot\|_Y$ of the norm $\|\cdot\|_X$ to Y.*

Proof. Let (x_n) be a Cauchy sequence in A with respect to d. Because of the completeness of M, there exists an $x \in M$ with $x_n \to x$ in (M,d) as $n \to \infty$. Since A is closed, x belongs to A; and thus $x_n \to x$ in (A, d_A). □

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

a) Let $Y = \{f \in X \,|\, f(0) = 0\}$. Clearly, Y is a linear subspace. Let $f_n \in Y$ converge to f in X. It follows that $0 = f_n(0) \to f(0)$ as $n \to \infty$, hence $f(0) = 0$ and $f \in Y$. Therefore Y is closed and Y is a Banach space with respect to $\|\cdot\|_\infty$.

b) The set $O = \{f \in X \,|\, f > 0\}$ is open in X. In fact, let $f \in O$. Then $\min_{t\in[0,1]} f(t) =: \delta > 0$. Let $g \in X$ with $\|f - g\|_\infty < \delta$. We now deduce that $g(t) = f(t) + g(t) - f(t) \leq \delta - \|f-g\|_\infty > 0$ for all $t \in [0,1]$, so that $g \in Y$. This means that $B(f,\delta) \subset O$, and hence O is open.

c) The set $C = \{f \in X \,|\, f(0) > 0 \text{ and } f(1) \geq 0\}$ is neither open nor closed in X. In fact, the functions $f_n = \frac{1}{n}\mathbb{1}$ belong to C and converge in X to $0 \neq C$, so that C is not closed. Moreover, the functions g_n given by $g_n(t) = 1 - (1 + \frac{1}{n})t$ do not belong to C, but they have the limit $g(t) = 1 - t$ in X which is contained in C, so that C is not open.

d) Let $\ell > 0$. The set $L = \{f \in X \,|\; |f(t) - f(s)| \leq \ell\,|t-s| \;(\forall\, t,s \in [0,1])\}$ of functions with Lipschitz constant less or equal ℓ is closed in X. In fact, even if functions $f_n \in L$ converge pointwise to some f as $n \to \infty$, we conclude

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

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

On the other hand, the set $D = \{f \in C^1([-1,1]) \,|\, \|f'\|_\infty \leq 1\}$ is not closed in $C([-1,1])$. In fact, the functions $f_n \in D$ given by $f_n(t) = (t^2 + \frac{1}{n})^{1/2}$ for $n \in \mathbb{N}$ converge uniformly on $[-1,1]$ to the continuous function $t \mapsto f(t) = |t|$ which is not differentiable at 0. ◊

Proposition 1.15. *In a metric space M 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. Let be $\mathcal{C}$ be a collection of open sets $O \subset M$ and let $x \in V := \bigcup_{O\in\mathcal{C}} O$. Then there is a set $O' \in \mathcal{C}$ containing x. Since O' is open, we have an $r > 0$ such that $B(x,r) \subset O' \subset V$. Therefore V is open. Let $O_1, \cdots, O_n \subset M$ be open and $x \in D := O_1 \cap \cdots \cap O_n$. Again, there are $r_j > 0$ such that $B(x, r_j) \subset O_j$ for each $j \in \{1, \ldots, n\}$. Setting $\rho := \min\{r_1, \ldots, r_n\} > 0$, we arrive at $B(x,\rho) \subset D$, so that D is open. Assertion b) follows from a) by taking complements. □

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 $\mathbb{R}$.

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

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

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

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 $\overline{N} = M$.*

The above concepts can be characterized in various ways.

Proposition 1.17. *Let M be a metric space M and $N \subseteq M$. We then have:*

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

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

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

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) $\overline{N} = \{x \in M \,|\, \exists\, (x_n) \subseteq N \text{ with } x_n \to x \text{ as } n \to \infty\} =: N_3$.

c) (i) ∂N is closed.

(ii) $\partial N = \{x \in M \,|\, \exists\, x_n \in N,\ y_n \notin N \text{ with } x_n \to x,\ y_n \to x \text{ as } n \to \infty\}$.

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. a) The inclusion $N^\circ \subset N$ follows from the definition of N°, and N° is open by Proposition 1.15. If $N^\circ \subset O \subset N$ for an open set O, then $O \subset N^\circ$ due to the definition of N°, so that $O = N^\circ$. We have thus shown (i) which implies (ii). In (iii), we deduce $N_1 \subset N^\circ$ from the definition of N°, and $N^\circ \subset N_2$ 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). Assertion (iii) in a) implies that $N_3 = M \setminus (M \setminus N)^\circ$, so that N_3 is closed. Clearly, $N \subseteq N_3$. From Proposition 1.12 we infer $N_3 \subseteq \overline{N}$. Now, (i) yields (iv).

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

Corollary 1.18. *Let X be a normed vector space.*
a) If $Y \subset X$ is a linear subspace, then $\overline{Y}$ is also a linear subspace.
b) If $Y \subset X$ is convex, then $\overline{Y}$ is also convex.

Proof. a) Let $x, y \in \overline{Y}$ and $\alpha, \beta \in \mathbb{F}$. By Proposition 1.17 there are $x_n, y_n \in Y$ with $x_n \to x$ and $y_n \to y$ as $n \to \infty$. Clearly, $z_n := \alpha x_n + \beta y_n \in Y$ and

$$\|z_n - (\alpha x + \beta y)\| \le |\alpha| \, \|x_n - x\| + |\beta| \, \|y_n - y\| \;\longrightarrow\; 0$$

as $n \to \infty$. Proposition 1.17 yields assertion a). Part b) is shown similarly. □

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

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

Example 1.19. a) Let $\mathbb{F} = \mathbb{R}$. The set P of polynomials is dense in $X = C([0,1])$ by Weierstraß' approximation theorem from Analysis 3. Since $P \subset C^k([0,1])$, also the subspace $C^k([0,1])$ is dense in $C([0,1])$ for every $k \in \mathbb{N}$.

b) Let X be a normed vector space, $x \in X$, and $r > 0$. We then have $\overline{B(x,r)} = \overline{B}(x,r)$ and $\partial B(x,r) = \partial \overline{B}(x,r) = \{y \in X \mid \|x - y\| = r\} =: S(x,r)$. In fact, Proposition 1.17 and Example 1.11 show that $B(x,r) \dot\cup \partial B(x,r) = \overline{B(x,r)} \subset \overline{B}(x,r)$. Take $y \in S(x,r) = \overline{B}(x,r) \setminus B(x,r)$. Then $y_n = y - \frac{1}{n}(y - x) \in B(x,r)$ and $z_n = y + \frac{1}{n}(y - x) \notin \overline{B}(x,r)$ for all $n \in \mathbb{N}$, and we have $y_n \to y$ and $z_n \to y$ as $n \to \infty$. Consequently, $y \in \partial B(x,r)$ and $y \in \partial \overline{B}(x,r)$ due to Proposition 1.17. These facts imply the assertions.

c) Let M contain at least two elements and let d be the discrete metric on M from Example 1.6. We then have $B(x,1) = \{x\}$ and $\overline{B}(x,1) = M$, whence $B(x,1) = \overline{B(x,1)} \neq \overline{B}(x,1)$.

d) Let $X = C([0,1])$ and $N = \{f \in X \mid f \ge 0\}$. Then N is closed in X, $N^\circ = \{f \in X \mid f > 0\} =: O$, and $\partial N = \{f \in N \mid \exists\, t \in [0,1] \text{ with } f(t) = 0\} =: R$.
Proof. Let $f_n \in N$ converge to some f in X. Then $f(t) = \lim_{n\to\infty} f_n(t) \ge 0$ for each $t \in [0,1]$, and so $f \in N$. Thus, N is closed. We have $N = O \dot\cup R$. The set O is open by Example 1.14. If $f \in R$, then $f_n := f - \frac{1}{n}\mathbb{1} \notin N$ and $f_n \to f$ in X as $n \to \infty$. Hence, $R \cap N^\circ = \emptyset$. Proposition 1.17 now yields $O = N^\circ$ and $R = \partial N$. □

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

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

in X is given by

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

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.
Proof. As in Example 1.14 we see that $C_0(0,1)$ is closed in X. Clearly, $C_c(0,1) \subset C_0(0,1)$, so that $\overline{C_c(0,1)} \subset C_0(0,1)$ by Proposition 1.17. 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 $n \ge 2$. Then $f_n := \varphi_n f \in C_c(0,1)$ and

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

and the right hand side tends to 0 as $n \to \infty$ since $f \in C_0(0,1)$. As a result, $C_0(0,1) \subset \overline{C_c(0,1)}$ and the assertion is proved. □

f) The sets L and D from Example 1.14 have no interior points since for each $\varepsilon > 0$ there is a continuous, but not Lipschitz continuous f with $\|f\|_\infty < \varepsilon$. ◊

Remark 1.20. Let (M, d) be a metric space and $N \subseteq M$ be endowed with the subspace metric d_N. Then the open balls in (N, d_N) are given by

$$B_N(x,r) = \{y \in N \mid r > d_N(x,y) = d(x,y)\} = B(x,r) \cap N.$$

A set $C \subseteq N$ is called *relatively open (closed)* if it is open (closed) in (N, d_N). Similarly, one introduces the relative closure etc.. For instance, the open unit ball in $N = \mathbb{R}^2_+$ for $|\cdot|_2$ is given by $B_N(0,1) = \{(x,y) \in \mathbb{R}^2 \mid x, y \geq 0,\ x^2 + y^2 < 1\}$. The definitions further imply the following facts.

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

b) N is open (closed) in M if and only if openness (closedness) in N and M coincide.

(In the proof of a), for an open S one may set $S' = \bigcup_{x \in S} B_M(x, r_x)$ where $(B_M(x, r_x) \cap N) \subseteq S$.) ♢

Definition 1.21. *Let (M, d) and (M', d') be metric spaces, $f : M \to M'$, $x_0 \in M$, and $y_0 \in M'$. We write $y_0 = \lim_{x \to x_0} f(x)$ if $f(x_n) \to y_0$ in M' as $n \to \infty$ for* **every** *sequence $(x_n) \subseteq M$ with $x_n \to x_0$ in M as $n \to \infty$. The map f is called* continuous at x_0 *if $f(x_0) = \lim_{x \to x_0} f(x)$. If f is continuous at every $x_0 \in M$, then it is called* continuous *(on M). We write $C(M, M') := \{f : M \to M' \mid f$ is continuous$\}$, and put $C(M) := C(M, \mathbb{F})$, where $\mathbb{F}$ is endowed with the usual metric. 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 preserves convergence in the sense that: If $x_n \to x_0$, then $f(x_n) \to f(x_0)$ as $n \to \infty$.

Remark 1.22. Let (M_1, d_1) and (M_2, d_2) be metric spaces. On the product space $M = M_1 \times M_2$, we can define a metric by setting $d((x,y),(u,v)) = d_1(x,u) + d_2(y,v)$. A sequence (x_n, y_n) in M converges with respect to d to $(x,y) \in M$ if and only if $x_n \to x$ in M_1 and $y_n \to y$ in M_2 as $n \to \infty$. These facts can be shown as for $M_1 = M_2 = \mathbb{R}$ in Analysis 2. ♢

Example 1.23. a) Every distance $d : M \times M \to \mathbb{R}$ is continuous, where $M \times M$ is endowed with the metric from Remark 1.22. In fact, let $(x_n, y_n) \to (x,y)$ in $M \times M$ as $n \to \infty$. Using

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

we deduce that $|d(x_n, y_n) - d(x,y)|$ tends to 0 as $n \to \infty$.

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 $r \geq 0$. Set $(F(g))(t) = \varphi(g(t))$ for all $t \in [0,1]$ and $g \in X$. Clearly, $F(g) \in X$. Let $g_n \to g$ in X as $n \to \infty$. Then $R := \sup_{n \in \mathbb{N}} \|g_n\|_\infty < \infty$ and $\|g\|_\infty \leq R$. We thus deduce

$$\|\varphi(g_n) - \varphi(g)\|_\infty = \sup_{t \in [0,1]} |\varphi(g_n(t)) - \varphi(g(t))| \leq \sup_{t \in [0,1]} L_R\, |g_n(t) - g(t)| = L_R\, \|g_n - g\|_\infty,$$

so that $F : X \to X$ is continuous. (Similarly, one sees that F is Lipschitz on every ball of X.) ♢

One can characterize continuity in terms of open or closed sets.

Proposition 1.24. *Let (M,d), (M',d') and (M'',d'') be metric spaces, $x_0 \in M$, $f : M \to M'$ and $g : M' \to M''$. It then holds:*

a) If f is continuous at x_0 and g is continuous at $f(x_0)$, then also $h = g \circ f : M \to M''$ is continuous at x_0.

b) The following assertions are equivalent.

(i) f is continuous at x_0.

(ii) $\forall \varepsilon > 0 \ \exists \delta > 0 : \ \forall x \in M$ with $d(x,x_0) < \delta$ we have $d(f(x), f(x_0)) < \varepsilon$.

(iii) If V is a neighborhoood of $f(x_0)$ in M', then $f^{-1}(V)$ is a neighborhood of x_0 in M.

c) The following assertions are equivalent.

(i) 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. a) follows from Definition 1.21 as in Analysis 1+2.

b) (i) $\Rightarrow$ (iii): If (iii) were false, then there would exist a neighborhood V of $f(x_0)$ in M' and a sequence (x_n) which does not belong to $f^{-1}(V)$ and converges to x_0 in M. Hence, $f(x_n) \notin V$ for all $n \in \mathbb{N}$ so that $f(x_n)$ cannot converge to $f(x_0)$. This fact contradicts (i).
(iii) $\Rightarrow$ (ii): Set $V = B(f(x_0), \varepsilon)$ for any given $\varepsilon > 0$. By (iii) there exists a $\delta > 0$ such that $B(x_0, \delta) \subseteq f^{-1}(V)$. Thus, for every $x \in B(x_0, \delta)$ we have that $f(x) \in V = B(f(x_0), \varepsilon)$, and so (ii) holds.
(ii) $\Rightarrow$ (i) can be deduced from Definition 1.21 as in Analysis 1+2.

c) (i) $\Rightarrow$ (iii): Let $A \subset M'$ be closed. Take $x_n \in f^{-1}(A)$ with $x_n \to x$ as $n \to \infty$. Then $A \ni f(x_n) \to f(x)$ by (i). Since A is closed, $f(x)$ belongs to A; i.e., $x \in f^{-1}(A)$. Proposition 1.12 now implies that $f^{-1}(A)$ is closed.
(iii) $\Rightarrow$ (ii) follows by taking complements.
(ii) $\Rightarrow$ (i): Let $x \in M$. Take a neighborhood V of $f(x)$ in M'. Hence, there is an $r > 0$ such that $B(x,r) \subseteq V$. Due to (ii), $f^{-1}(B(f(x), r))$ is open in M. As a result $f^{-1}(V)$ is a neighborhood of x in M, so that f is continuous at x by part b). □

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

Example 1.26. Let $X = C([0,1])$ and $\varphi : X \to \mathbb{F}$; $\varphi(g) = g(0)$. If $g_n \to g$ in X, then $g_n(0) \to g(0)$ as $n \to \infty$. Thus φ is continuous. So, if $B \subseteq \mathbb{F}$ is open (closed), then $\varphi^{-1}(B) = \{g \in X \,|\, g(0) \in B\}$ is open (closed) in X. ◊

Quite often one has several norms on a vector space. The next definition is concerned with concepts that allow to compare norms. These concepts can then again be characterized by means of sequences and open or closed sets.

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

$$c\,|||x||| \le \|x\| \le C\,|||x|||$$

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

Proposition 1.28. *Let $\|\cdot\|$ and $|||\cdot|||$ be norms on a vector space X. Then the following assertions are equivalent.*

a) $\|\cdot\|$ is coarser than $|||\cdot|||$.

b) If a sequence $(x_n) \subseteq X$ converges for $|||\cdot|||$, then it converges for $\|\cdot\|$.

c) If a set $A \subseteq X$ is closed for $\|\cdot\|$, then it is closed for $|||\cdot|||$.

d) If a set $O \subseteq X$ is open for $\|\cdot\|$, then it is open for $|||\cdot|||$.

In this case the limits in b) are equal. Moreover, the norms are equivalent if and only if one of the implications in b), c) or d) becomes an equivalence. Finally, let the norms be equivalent. Then $(X, \|\cdot\|)$ is complete if and only if $(X, |||\cdot|||)$ is complete.

Proof. The implication 'a) $\Rightarrow$ b)' and the equality of the limits are a consequence of the estimate $\|x_n - x\| \le C\,|||x_n - x|||$. To show 'b) $\Rightarrow$ c)', let A be closed for $\|\cdot\|$ and let $x_n \in A$ converge to some $x \in X$ for $|||\cdot|||$. By b), x_n also converge for $\|\cdot\|$ so that $x \in A$; i.e., A is closed for $|||\cdot|||$. The implication 'c) $\Rightarrow$ d)' follows by taking complements. Let statement d) hold. The ball $B_{\|\cdot\|}(0,1)$ is thus open for $|||\cdot|||$, so that we can find an $r > 0$ with $B_{|||\cdot|||}(0,r) \subseteq B_{\|\cdot\|}(0,1)$. This means that $|||y||| < r$ implies that $\|y\| < 1$. Let $x \in X$. Since the vector $y := r(2\,|||x|||)^{-1}\,x$ has the triple norm $r/2 < r$, we conclude that $1 > \|y\| = r(2\,|||x|||)^{-1}\,\|x\|$ which yields a).

The assertions concerning equivalence follow from the above results. One similarly shows the claim about completeness. □

Example 1.29. a) On a finite dimensional vector space X all norms are equivalent. (See Analysis 2.) On $X = \mathbb{F}^n$ we have the more precise result

$$|x|_\infty \le |x|_p \le |x|_1 \le n^{1/q'}\,|x|_q$$

for all $x \in \mathbb{F}^n$ and $p, q \in [1,\infty]$, where $\frac{1}{q} + \frac{1}{q'} = 1$. (See Analysis 2 or the next section.)

b) Let $X = C([0,1])$ and $0 < w \in X$. Observe that $\delta := \min_{t\in[0,1]} w(t) > 0$. Then the norms $\|\cdot\|_\infty$ and $\|f\|_w := \max_{0\le t\le 1} w(t)\,|f(t)|$ on X are equivalent since $\delta\,\|f\|_\infty \le \|f\|_w \le \|w\|_\infty\,\|f\|_\infty$ for all $f \in X$.

c) Let $X = C([0,1])$. Since $\|f\|_1 \le \|f\|_\infty$ for $f \in X$ the supremum-norm is finer that 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 since only the first one is complete. This fact can directly be seen using the functions f_n given by $f_n(t) = 1 - nt$ for $0 \le t \le \frac{1}{n}$ and $f_n(t) = 0$ for $\frac{1}{n} \le t \le 1$, since $\|f_n\|_\infty = 1$ and $\|f_n\|_1 = \frac{1}{2n}$ for all $n \in \mathbb{N}$.

The 1-norm and the sup-norm on $C_c(\mathbb{R}_+) = \{f \in C(\mathbb{R}_+) \,|\, \operatorname{supp} f \text{ is bounded}\}$ are not comparable. Indeed, using functions as f_n one sees that $\|\cdot\|_1$ cannot be finer than $\|\cdot\|_\infty$. Conversely, take functions $g_n \in C_c(\mathbb{R}_+)$ with $0 \le g_n \le$ and $g_n = 1$ on $[0,n]$ for $n \in \mathbb{N}$. Then $\|g_n\|_\infty = 1$ and $\|g_n\|_n \ge n$ for all n, so that $\|\cdot\|_\infty$ cannot be finer than $\|\cdot\|_1$. ◊

1.2 More examples of Banach spaces

A) Sequence spaces

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

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

Moreover, for all $v_n, x \in s$, we have $d(v_n, x) \to 0$ as $n \to \infty$ if and only $(v_n)_j \to x_j$ for each $j \in \mathbb{N}$, as $n \to \infty$. These assertions follow from Proposition 1.8 with $p_j(x) = |x_j|$. As in Example 1.9 one shows the completeness of (s, d). For $x \in s$, one defines the supremum norm by

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

and introduces by

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

the spaces of bounded, converging, null and finite sequences, repectively. For $1 \le p < \infty$ and $x \in s$, we further define

$$\|x\|_p^p := \sum_{j=1}^\infty |x_j|^p \ \in [0,\infty] \qquad \text{and the space} \qquad \ell^p = \{x \in s \,|\, \|x\|_p < \infty\}.$$

Observe that

$$|x_k| \le \|x\|_p \qquad \text{for all } \ k \in \mathbb{N},\, 1 \le p \le \infty,\, x = (x_j) \in \ell^p. \tag{1.2}$$

We also set

$$p' = \begin{cases} \infty, & p = 1, \\ \frac{p}{p-1}, & 1 < p < \infty, \\ 1, & p = \infty. \end{cases} \tag{1.3}$$

We note that

$$\frac{1}{p} + \frac{1}{p'} = 1; \quad p'' = p; \quad p' = 2 \iff p = 2; \quad p \in (1,2) \iff p' \in (2,\infty). \tag{1.4}$$

Proposition 1.30. *Let $p \in [1,\infty]$, $x, y \in \ell^p$ and $z \in \ell^{p'}$. We then have:*

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

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

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

Proof. Assertions a) and b) are straightforward to verify for $p = 1$ and $p = \infty$, so that we only consider $p, p' \in (1,\infty)$ in a) and b).

a) Let $x \in \ell^p$ and $z \in \ell^{p'}$. We recall Young's inequality

$$ab \le \frac{a^p}{p} + \frac{b^{p'}}{p'} \qquad \text{for all } \ a, b > 0 \text{ and } p \in (1,\infty) \tag{1.5}$$

from Analysis 1+2. If $\|x\|_p = 0$ or $\|z\|_{p'} = 0$, then $x_j = 0$ or $z_j = 0$ for all $j \in \mathbb{N}$ by (1.2), and a) holds. Otherwise, we fix a $j \in \mathbb{N}$ and set $a = |x_j|/\|x\|_p$ and $b = |z_j|/\|z\|_{p'}$. Estimate (1.5) then yields

$$\frac{|x_j z_j|}{\|x\|_p \|z\|_{p'}} \le \frac{|x_j|^p}{p\,\|x\|_p^p} + \frac{|z_j|^{p'}}{p'\,\|z\|_{p'}^{p'}}$$

Since both terms on the right are summable, we have $xz \in \ell^1$ and

$$\frac{1}{\|x\|_p \, \|z\|_{p'}} \sum_{j=1}^{\infty} |x_j z_j| \le \frac{1}{p\,\|x\|_p^p} \sum_{j=1}^{\infty} |x_j|^p + \frac{1}{p'\,\|z\|_{p'}^{p'}} \sum_{j=1}^{\infty} |z_j|^{p'} = 1,$$

using also (1.4). Thus, a) holds.

b) Let $x, y \in \ell^p$. Observe that $|x_k + y_k|^p \le 2^p(|x_k|^p + |y_k|^p)$ for all $k \in \mathbb{N}$, so that $x + y \in \ell^p$. By part a) and (1.3), the sequence $w = (|x_k + y_k|^{p-1})_k$ belongs to $\ell^{p'}$ since $w_k^{p'} = |x_k + y_k|^{(p-1)p/(p-1)} = |x_k + y_k|^p$ is summable. From a) we thus infer

$$\begin{aligned}
\|x+y\|_p^p &= \sum_{k=1}^{\infty} |x_k + y_k|\, |x_k + y_k|^{p-1} \\
&\le \sum_{k=1}^{\infty} |x_k|\, |x_k + y_k|^{p-1} + \sum_{k=1}^{\infty} |y_k|\, |x_k + y_k|^{p-1} \\
&\le \|x\|_p \left(\sum_{k=1}^{\infty} |x_k + y_k|^{(p-1)\frac{p}{p-1}} \right)^{\frac{p-1}{p}} + \|y\|_p \left(\sum_{k=1}^{\infty} |x_k + y_k|^{(p-1)\frac{p}{p-1}} \right)^{\frac{p-1}{p}} \\
&= \|x+y\|_p^{p-1} (\|x\|_p + \|y\|_p).
\end{aligned}$$

Minkowski's inequality now follows.

c) The assertions on ℓ^∞, c and c_0 were shown in the exercises. So let $p \in [1, \infty)$, $x, y \in \ell^p$ and $\alpha \in \mathbb{F}$. Clearly, $\alpha x \in \ell^p$ and $\|\alpha x\|_p = |\alpha|\, \|x\|_p$. If $\|x\|_p = 0$, then $x_j = 0$ for all j by (1.2). Hence, ℓ^p is a normed vector space in view of b). Let $v_n = (v_n)_{k\in\mathbb{N}}$ be a Cauchy sequence in ℓ^p for each $n \in \mathbb{N}$. This means that for every $\varepsilon > 0$ there is an $N_\varepsilon \in \mathbb{N}$ such that

$$|(v_n)_k - (v_m)_k| \le \|v_n - v_m\|_p \le \varepsilon$$

for all $n, m \ge N_\varepsilon$ and each $k \in \mathbb{N}$. So there exists $x_k = \lim_{n\to\infty} (v_n)_k$ in $\mathbb{F}$. Set $x = (x_k)_{k\in\mathbb{N}}$. Using the second of the above inequalities, we obtain

$$\sum_{k=1}^{K} |(v_n)_k - (v_m)_k|^p \le \|v_n - v_m\|_p^p \le \varepsilon^p$$

for all $n, m \ge N_\varepsilon$ and $K \in \mathbb{N}$. Letting $m \to \infty$, it follows

$$\sum_{k=1}^{K} |(v_n)_k - x_k|^p \le \varepsilon^p.$$

Taking the supremum over $K \in \mathbb{N}$, we deduce

$$\sum_{k=1}^{\infty} |(v_n)_k - x_k|^p \le \varepsilon^p$$

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

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

$$c_{00} \underset{\neq}{\subsetneq} \ell^1 \underset{\neq}{\subsetneq} \ell^p \underset{\neq}{\subsetneq} \ell^q \underset{\neq}{\subsetneq} c_0 \underset{\neq}{\subsetneq} \ell^\infty \quad \textit{and} \quad \|x\|_\infty \le \|x\|_q \le \|x\|_p \le \|x\|_1,$$

$$\overline{c_{00}}^{\ell^p} = \ell^p \quad \textit{for} \ \ 1 \le p < \infty \qquad \textit{and} \qquad \overline{c_{00}}^{\|\cdot\|_\infty} = c_0.$$

Proof. It is clear that $c_{00} \subsetneqq \ell^1$ and $\ell^q \subseteq c_0 \subsetneqq \ell^\infty$ for all $q < \infty$. Set $z_k = k^{-\frac{1}{p}}$. Then $(z_k) \notin \ell^p$, but $(z_k) \in \ell^q \cap c_0$ for all $1 \le p < q < \infty$. It further holds $\|x\|_\infty = |x_N| \le \|x\|_q$ for all $x \in \ell^q \subseteq c_0$ and some $N = N(x) \in \mathbb{N}$, using also (1.2). Let $x \in \ell^p \setminus \{0\}$ and $1 \le p < q < \infty$, and set $y = \frac{1}{\|x\|_p} x$. Then $\|y\|_p = 1$ and $|y_k| \le 1$ by (1.2), so that $|y_k|^q \le |y_k|^p$ for all $k \in \mathbb{N}$. Summing over k, we obtain $\|y\|_q^q \le \|y\|_p^p = 1$. Thus, $1 \ge \|y\|_q = \|x\|_q / \|x\|_p$, whence $\|x\|_q \le \|x\|_p$ and $\ell^p \subseteq \ell^q$.

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

B) L^p spaces

We have already discussed several Banach spaces with (variants of) the supremum norm, partly in the exercises: $C([0,1])$, $C_0(0,1)$, $C_b(\mathbb{R})$, $C_0(\mathbb{R})$, and $C^1([0,1])$. There are various natural combinations and generalizations, e.g., to compact $K \subseteq \mathbb{R}^d$ or open $U \subseteq \mathbb{R}^d$ or to higher derivatives. We add two related classes of supnorm spaces.

Example 1.32. Let $D \subseteq \mathbb{C}$ be open. Then $H^\infty(D) = \{f : D \to \mathbb{C} \,|\, f \text{ bounded and holomorphic}\}$ is a closed subspace of $C_b(D)$ for the supremum norm, and thus a Banach space, since the uniform limit of holomorphic functions is holomorphic due to a convergence theorem by Weierstraß from Funktionentheorie.

Moreover, $A(D) = C(\overline{D}, \mathbb{C}) \cap H^\infty(D)$ is a closed subspace of $H^\infty(D)$, and for bounded D the maximum principle yields $\|f\|_\infty = \sup_{z \in \partial D} |f(z)|$ for $f \in A(D)$. ♢

We look for a Banach space of functions with respect to the norm $\|f\|_p = (\int |f|^p \, d\mu)^{1/p}$, where $p \in [1, \infty)$. In the following we briefly repeat the necessary material from Analysis 3 without proofs. Actually some of the statements below are more general than in Analysis 3 since they are formulated for every measure space, but they can be shown in the same way, see also [Rud87] or [Wer06].

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

a) $\emptyset \in \mathcal{A}$.

b) If $A \in \mathcal{A}$, then $S \setminus A \in \mathcal{A}$.

c) If $A_k \in \mathcal{A}$ for all $k \in \mathbb{N}$, then $\bigcup_{k \in \mathbb{N}} A_k \in \mathcal{A}$.

Note that then S belongs to $\mathcal{A}$ and $\bigcap_k A_k \in \mathcal{A}$ if $A_k \in \mathcal{A}$ for all $k \in \mathbb{N}$. We give a few basic examples for σ–algebras.

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

b) Let $A \subset S$. Then $\{\emptyset, A, S \setminus A, S\}$ is a σ–algebra over S.

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

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

and is called *Borel σ–algebra* on M. It contains all open and closed sets, all countable unions and intersections of closed or open sets, and so on. We write $\mathcal{B}_d$ instead of $\mathcal{B}(\mathbb{R}^d)$ and endow $\mathbb{C}^d$ with $\mathcal{B}_{2d}$. For $M = \mathbb{R}^d$ one can replace $\mathcal{O}$ by the collection of all compact sets or of all intervals. We stress that $\mathcal{B}_d$ is a strict subset of $\mathcal{P}(\mathbb{R}^d)$. For $B \in \mathcal{B}_d$ one has $\mathcal{B}(B) = \{A \in \mathcal{B}_d \,|\, A \subset B\} = \{A' \cap B \,|\, A' \in \mathcal{B}_d\}$.

d) On $[0, \infty]$ one considers the metric discussed in the exercises for which also $(a, \infty]$ for $a \in [0, \infty)$ is an open set. One obtains $\mathcal{B}([0, \infty]) \supset \mathcal{B}(\mathbb{R}_+)$.

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

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

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

Let $A, B, A_k \in \mathcal{A}$ for $k \in \mathbb{N}$ with $A \subset B$ and μ be a measure on $\mathcal{A}$. We then obtain $\mu(A) \le \mu(B)$, and for $\mu(A) < \infty$ we have $\mu(B\setminus A) = \mu(B) - \mu(A)$. Moreover, $\mu(\bigcup_k A_k) \le \sum_{k=1}^{\infty} \mu(A_k)$.

Continuing the above examples, we present a few measure spaces.

a) 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 $\mathcal{P}(S)$.

b) For $A \subset \mathbb{N}$, we define the *counting measure* $\zeta(A)$ as the number of elements of A. It is a σ–finite measure on $\mathcal{P}(\mathbb{N})$.

c) On $\mathcal{B}_d$ there is exactly one measure $\lambda = \lambda_d$ such that $\lambda(J) = \ell_1 \cdot \ldots \cdot \ell_d$ for any interval J with side lengths ℓ_j. It is σ–finite and called *Lebesgue measure*. Let $B \in \mathcal{B}_d$. The restriction $\lambda_B = \lambda$ of λ to $\mathcal{B}(B)$ is a σ–finite measure, also called Lebesgue measure. Unless otherwise specified, we endow Borel sets in $\mathbb{R}^d$ with $\mathcal{B}(B)$ and λ.

Let $\mathcal{A}$ and $\mathcal{A}'$ be σ–algebras on S and S', respectively. A map $f : S \to S'$ is called *measurable* if $f^{-1}(A) \in \mathcal{A}$ for every $A \in \mathcal{A}'$. We collect the basic properties of measurable maps.

a) Let $A \subset S$. The characteristic function $\mathbb{1}_A : S \to \mathbb{F}$ is measurable if and only if $A \in \mathcal{A}$. Linear combinations of measurable characteristic functions are called *simple functions*.

b) Compositions of measurable functions are measurable.

c) Continuous functions are measurable for the respective Borel σ–algebras.

d) A function with values in $\mathbb{F}^d$ is measurable if and only if the real and imaginary parts of all its components are measurable.

e) By $\tilde{\mathbb{F}}$ we denote temporarily either $\mathbb{F}$ or $[0,\infty]$ endowed with the Borel σ–algebras. Let $f, g, f_n : S \to \tilde{\mathbb{F}}$ be measurable for $n \in \mathbb{N}$ (for any σ–algebra on S) and let $\alpha \in \tilde{\mathbb{F}}$. Then the functions αf, $f + g$, fg, $|f|$, and (if it exists) $\lim_{n\to\infty} f_n : S \to \tilde{\mathbb{F}};\ t \mapsto \lim_{n\to\infty} f_n(t)$ are measurable (where we set $0 \cdot \infty = 0$). If $\tilde{\mathbb{F}} \neq \mathbb{C}$, then also the functions $\sup_n f_n$, $\inf_n f_n$, $\limsup_n f_n$ and $\liminf_n f_n$ are measurable (provided they exist if $\mathbb{F} = \mathbb{R}$). In particular, we obtain the measurability of the positive part $f_+ = \max\{0, f\}$ and the negative part $f_- = \max\{0, -f\}$ of a real–valued measurable function f.

Let $(S, \mathcal{A}, \mu)$ be a measure space and $f : S \to [0,\infty]$ be measurable. Then f can be approximated monotonically by simple functions $f_n : S \to [0,\infty)$. Using this fact, one can then define an *integral* $\int_S f\, d\mu$ which belongs to $[0,\infty]$.

It is monotone, additive and positive homogeneous in f. The function f is called *integrable* if the integral is finite.

A measurable function $f : S \to \mathbb{F}$ is called *integrable* if the positive measurable maps $(\operatorname{Re} f)_\pm$ and $(\operatorname{Im} f)_\pm$ are integrable, and then one defines the *integral* of f by

$$\int f\, d\mu = \int_S f(s)\, d\mu(s) := \int_S (\operatorname{Re} f)_+\, d\mu - \int_S (\operatorname{Re} f)_-\, d\mu$$

$$+\,i\int_S (\operatorname{Im} f)_+\,d\mu - i\int_S (\operatorname{Im} f)_-\,d\mu.$$

One writes dx instead of $d\lambda$ or $d\lambda(x)$ for the Lebesgue measure. The integral has the following properties.

a) A measurable function f is integrable if and only if the (positive and measurable) function $|f|$ is integrable. In this case, one has $|\int f\,d\mu| \le \int |f|\,d\mu$.

b) The space of integrable functions $f : S \to \mathbb{F}$ is a vector space. The integral is linear and (if $\mathbb{F} = \mathbb{R}$) monotone in f.

c) If $f : S \to [0,\infty]$ is measurable, the integral is equal to 0 if and only if the set $\{s \in S \,|\, f(s) \neq 0\} \in \mathcal{A}$ has measure 0.

For $p \in [1,\infty)$ and measurable $f : S \to \mathbb{F}$, we define $\|f\|_p^p := \int_S |f|^p\,d\mu$ and

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

One can show Minkowski's inequality as in Proposition 1.30 so that $\|\cdot\|_p$ is a seminorm. We list again a few basic examples.

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

b) We have $\mathcal{L}^p(\mathbb{N},\mathcal{P}(\mathbb{N}),\zeta) = \ell^p$ for $p \in [1,\infty)$ and $\int_{\mathbb{N}} f\,d\zeta = \sum_{k=1}^\infty f(k)$.

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

$$\int_B f\,dx = \int_B f\,d\lambda_B = \int_{\mathbb{R}^d} \widetilde{f}\,dx.$$

The crucial results for the Lebesgue integral are the following *convergence theorems*, where $(S,\mathcal{A},\mu)$ is a measure space and $p \in [1,\infty)$.

a) *Fatou's lemma*: For measurable $f_n : S \to [0,\infty]$, $n \in \mathbb{N}$, we have

$$\int_S \liminf_{n\to\infty} f_n\,d\mu \le \liminf_{n\to\infty} \int_S f_n\,d\mu.$$

b) *Monotone convergence*: Let $f_n : S \to [0,\infty]$ and $f_n \le f_{n+1}$ for all $n \in \mathbb{N}$. Then the following limits exist in $[0,\infty]$ and satisfy

$$\int_S \lim_{n\to\infty} f_n\,d\mu = \lim_{n\to\infty} \int_S f_n\,d\mu.$$

c) *Dominated convergence, Lebesgue*: Let $f_n, g \in \mathcal{L}^p(\mu)$ such that $|f_n(s)| \le g(s)$ for all $n \in \mathbb{N}$ and $f_n(s) \to f(s)$ as $n \to \infty$, for every $s \in S$. Then $f \in \mathcal{L}^p(\mu)$ and $\|f_n - f\|_p \to 0$ as $n \to \infty$. If $p = 1$, we also have

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

d) *Riesz–Fischer*: Let (f_n) be a Cauchy sequence in $\mathcal{L}^p(\mu)$. Then there are a subsequence $(f_{n_k})_k$, a set $N \in \mathcal{A}$ with $\mu(N) = 0$ and $f, g \in \mathcal{L}^p(\mu)$ such that $|f_{n_k}(s)| \le g(s)$ for all $k \in \mathbb{N}$ and $f_{n_k}(s) \to f(s)$ as $k \to \infty$, for every $s \in S \setminus N$, and such that $\|f_n - f\|_p$ as $n \to \infty$.

It would be nice if $\mathcal{L}^p(\mu)$ was a Banach space, but unfortunately $\|f\|_p = 0$ does not imply $f = 0$, but only that the set $\{s \in S \,|\, f(s) \neq 0\}$ has measure 0. This fact leads us 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.)*. For instance, the Riesz–Fischer theorem says that the subsequence f_{n_k} converges pointwise a.e. to f. In the above convergence theorems, one replace pointwise properties by their counterparts holding a.e..

We note that a countable union of null sets is a null set and that $M \in \mathcal{A}$ is a null set if it is contained in a null set. Hyperplanes or each countable subset are null sets in $\mathbb{R}^d$ for the Lebesgue measure. Uncountable unions of null sets can have even infinite measure; e.g., $\mathbb{R}$ is the union of all singletons $\{x\}$. We further introduce the set of *null functions*

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

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

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

We further 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 $f \in \hat{f}$. It can be seen that these definitions do not depend on the choice of the representative. By means of the Riesz–Fischer theorem, one can show that $(L^p(\mu), \|\cdot\|_p)$ is a Banach space.

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

The set of simple functions in $L^p(\mu)$ is dense in $L^p(\mu)$ if $p \in [1, \infty)$, see e.g. Satz IV.7.6 in [Wer06].

Example 1.33. Let $B \in \mathcal{B}_d$ and $p \in [1, \infty)$. The set $P = \{f \in L^p(B) \,|\, f \ge 0\}$ is closed in $L^p(B)$ and contains no interior point.

Proof. Let $f_n \in P$ converge to f in X. Choose representatives g_n of f_n and g of f. The Riesz–Fischer theorem gives a subsequence and a null set N (possibly depending on the representatives) such that $g_{n_j}(x) \to g(x)$ for all $x \notin N$ as $j \to \infty$. On the other hand, there are null sets N_j such that $g_{n_j}(x) \ge 0$ for all $x \notin N_j$ and $j \in \mathbb{N}$. Therefore $g(x) \ge 0$ for all x not contained in the nullset $\bigcup_j N_j \cup N$. This means that $f \in P$, and so P is closed.

Let $f \in P$. There is a constant $c \ge 0$ such that the set $B_0 = \{x \in B \,|\, 0 \le f(x) \le c\}$ is not a null set (otherwise f would be equal to ∞ a.e.). Let $(C_j^n)_j$ be a countable disjoint covering of $\mathbb{R}^d$ with cubes of volume $1/n$. Since $\lambda(B_0) > 0$, for each $n \in N$ there is an index j_n such that $A_n = B_0 \cap C_{j_n}^n$ has measure in $(0, 1/n]$. Set $f_n = f - (c+1)\mathbb{1}_{A_n}$. Then f_n does not belong to P and $\|f - f_n\|_p \le (c+1)\lambda(A_n)^{1/p}$ tends to 0 as $n \to \infty$; i.e., f is not an interrior point of P. □

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

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

For a measurable function $f : S \to \mathbb{F}$ we define the *essential supremum norm* 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. One can see that for the Lebesgue measure on $B \in \mathcal{B}_d$ 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 $r > 0$. We also set

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

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

Proof. Let $f_k \in \mathcal{L}^\infty(\mu)$ and $\alpha_k \in \mathbb{F}$ for $k = 1, 2$. Then there are $c_k \geq 0$ and null sets N_k such that $|f_k(s)| \leq c_k$ for all $s \in S \setminus N_k$. We thus obtain $|\alpha_1 f_1(s) + \alpha_2 f_2(s)| \leq |\alpha_1|\, c_1 + |\alpha_2|\, c_2$ for $s \notin N_1 \cup N_2 = N$, where N is a null set. Therefore, $\alpha_1 f_1 + \alpha_2 f_2 \in \mathcal{L}^\infty(\mu)$, and so $\mathcal{L}^\infty(\mu)$ is a vector space. Moreover, $\|f_1 + f_2\|_\infty \leq \|f_1\|_\infty + \|f_2\|_\infty$. It is clear that $\|\alpha_1 f_1\|_\infty = |\alpha_1|\, \|f_1\|_\infty$. Observe that $\mathcal{N}$ is a linear subspace of $\mathcal{L}^\infty(\mu)$ and that $\|f_1\|_\infty = \|f_2\|_\infty$ if $f_1 = f_2$ a.e.. As result, $L^\infty(\mu)$ 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 functions $f_n \in \hat{f}_n$. For every $j \in \mathbb{N}$ there is a $k(j) \in \mathbb{N}$ such that $\|f_n - f_m\|_\infty \leq \frac{1}{j}$ for all $n, m \geq k(j)$. Hence, the set

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

is a null set for these n, m, j. Define the null set $N := \bigcup_{n,m \geq k(j), j \in \mathbb{N}} N_{n,m,j}$. For $s \in S \setminus N$, we then obtain

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

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 f is measurable. Let $\varepsilon > 0$ and take $j \geq 1/\varepsilon$. It follows that

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

if $n \geq k(j)$. Hence, $\hat{f} \in L^\infty(\mu)$ and $\hat{f}_n \to \hat{f}$ in $L^\infty(\mu)$ as $n \to \infty$. □

Proposition 1.35 (Hölder)**.** *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 $\|fg\|_1 \leq \|f\|_p\, \|g\|_{p'}$.

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

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

where $r = q/p > 1$ and $r' = q/(q-p)$ by (1.3).

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

1.3 Compactness and separability

Compactness is one of the most important concepts in analysis, as one could already see in Analysis 1+2. 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 set $\mathcal{C}$ of open sets $O \subseteq M$ with $K \subseteq \bigcup\{O \mid O \in \mathcal{C}\}$) 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 $m \in \mathbb{N}$ and $x_1, \dots, x_m \in M$ such that $K \subseteq B(x_1, \varepsilon) \cup \dots \cup B(x_m, \varepsilon)$.*

In the definition of total boundedness, one can equivalently require that the points x_j belong to K. In other words, if K is totally bounded in M, then it is also totally bounded in the metric space (K, d_K). To show this, take $\varepsilon > 0$ and the points $x_1, \cdots, x_m$ 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 $B(x_j, \varepsilon/2) \subset B(y_j, \varepsilon)$, and hence also the balls $B(y_j, \varepsilon)$ cover K.

If K is totally bounded, then K is bounded since $K \subseteq B(x_1, r)$ with $r = \varepsilon + \max_{j \in \{1,\dots,m\}} d(x_1, x_j)$ and the points x_j from part d) of the above definition.

As in Analysis 1 one sees that a sequence (x_n) 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; i.e., for each $r > 0$ and $m \in \mathbb{N}$ there is an $n \geq m$ with $x_n \in B(x, r)$ (or equivalently, in each $B(x, r)$ there are infinitely many x_n).

Theorem 1.37. *A subset K of a metric space M is compact if and only if it sequentially compact.*

Proof. (1) Let K be compact and suppose it is 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 an $r_y > 0$ such that $B(y, r_y)$ contains only finitely many x_n. Since $K \subseteq \bigcup_{y \in K} B(y, r_y)$ and K is compact, there are $y_1, \dots, y_m \in K$ such that $K \subseteq B(y_1, r_{y_1}) \cup \dots \cup B(y_m, r_{y_m}) =: B$. But this inclusion cannot hold because B only contains finitely many x_n; i.e., K must be sequentially compact.

(2a) **Claim:** Let K be a subset of a metric space M such that each sequence in K has an accumulation point in $\overline{K}$. Then K is totally bounded.

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

(2b) 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 part (2a), for each $n \in N$ there are finitely many balls of radius $1/n$ covering K. For every $n \in \mathbb{N}$, we thus find a ball B_n or radius $1/n$ such that $B_n \cap K$ is not covered by finitely many sets from $\mathcal{C}$. Since 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 $B(\widehat{x}, \varepsilon) \subset \widehat{O}$ for some $\varepsilon > 0$. We then obtain an index $N \in \mathbb{N}$ such that $d(x_N, \widehat{x}) < \varepsilon/2$ and $N \geq 2/\varepsilon$. Each $x \in B_N$ thus satisfies

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

i.e., $B_N \subset B(\widehat{x}, \varepsilon) \subset \widehat{O}$. This contradiction implies that K is compact. $\square$

Corollary 1.38. *Let M be a complete metric space and $N \subseteq M$. Then the following assertions are equivalent.*

a) N is relatively compact.

b) Each sequence in N has an accumulation point in $\overline{N}$.

c) N is totally bounded.

Proof. The implication "a)⇒b)" follows from Theorem 1.37 applied to $\overline{N}$, whereas "b)⇒c)" was shown in the proof of this theorem.

Let N be totally bounded, and take a sequence (x_n) in $\overline{N}$. There are $\widetilde{x}_n \in N$ such that $d(x_n, \widetilde{x}_n) \le 1/n$ for every $n \in \mathbb{N}$. 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 ν_1 such that $\widetilde{x}_{\nu_1(k)} \in B_{j_1}^1$ for all $k \in \mathbb{N}$. 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 ν_2 of ν_1 such that $\widetilde{x}_{\nu_2(k)} \in B_{j_2}^2$ for all $k \in \mathbb{N}$. By induction, for every $m \in \mathbb{N}$ we obtain a subsequence ν_m of ν_{m-1} and a ball $B_{j_m}^m$ of radius $1/m$ such that $\widetilde{x}_{\nu_m(k)} \in B_{j_m}^m$ for all $k \in \mathbb{N}$. We now define the diagonal sequence $n_m = \nu_m(m)$ for $m \in \mathbb{N}$. Note that $\widetilde{x}_{n_m}, \widetilde{x}_{n_p} \in B_{j_m}^m$ for $p \ge m$ since n_p and n_m belong to ν_m. Take $\varepsilon > 0$. Choose $m \in \mathbb{N}$ with $m, n_m \ge 1/\varepsilon$. It follows that

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

for all $p \ge m$. Because M is complete, the Cauchy sequence $(x_{n_m})_m$ has the limit $x \in \overline{N}$. Theorem 1.37 now implies that $\overline{N}$ is compact; i.e., a) holds. □

Corollary 1.39. *Let K be a compact subset of a metric space M. Then K is closed and bounded.*

Proof. The boundedness follows from Corollary 1.38. Let $x_n \in K$ converge to some $x \in M$. Since K is sequentially compact, there is a subsequence $(x_{n_k})_k$ and a $y \in K$ with $x_{n_k} \to y$ as $k \to \infty$. Hence, $x = y \in K$ and K is closed. □

Example 1.40. a) A subset K of a finite dimensional normed vector space X is compact if and only if K is closed and bounded. (See Analysis 2.) In particular, $\overline{B}(0,1)$ is compact in finite dimensions.

b) Let $X = \ell^p$, $1 \le p \le \infty$, and $e_n = (0, \dots, 0, 1, 0, \dots) \in \overline{B}(0,1)$ be the n–th unit vector. We then have $\|e_n - e_m\|_p = 2^{\frac{1}{p}}$ if $n \ne m$, 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 $X = C([0,1])$ and, for $n \in \mathbb{N}$,

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

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

The next theorem shows that the above simple examples are typical.

Theorem 1.41. *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 1.40 a). 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 1.42 and $X \neq U_1$ because of $\dim X = \infty$. Lemma 1.43 thus yields an $x_2 \in X$ 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 Lemma 1.43 gives an $x_3 \in X$ 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 $(x_n) \subseteq B$ mit $\|x_n - x_m\| \geq \frac{1}{2}$ for all $n \neq m$, so that B is not sequentially compact. □

Lemma 1.42. *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 $\bar{y} = (y_1, \dots, y_m) \in \mathbb{F}^m$ be the (uniquely determined) vector of coefficients with respect to B for a given $y \in Y$. Set $|||\bar{y}||| := \|y\|$. This gives a norm on $\mathbb{F}^m$, which is equivalent to $|\cdot|_2$ by Analysis 2, and thus it is complete. Let $v_n \in Y$ converge to some x in X. Then $(\bar{v}_n)$ is Cauchy in $\mathbb{F}^m$, and so there is a vector $\bar{z} = (z_1, \dots, z_m) \in \mathbb{F}^m$ such that $\bar{v}_n \to \bar{z}$ in $|||\cdot|||$ as $n \to \infty$. Set $z = z_1 b_1 + \cdots + z_m b_m \in Y$. We then deduce that $v_n \to z$ in X and thus $x = z \in Y$; i.e., Y is closed. □

Lemma 1.43 (F. Riesz)**.** *Let X be a normed vector space, $Y \neq X$ be a closed linear subspace, and $\delta \in (0, 1)$. Then there exists an $\bar{x} \in X$ 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 $\bar{x} = \frac{1}{\|x - \bar{y}\|}(x - \bar{y})$. We then obtain $\|\bar{x}\| = 1$ and, using the above inequalities,

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

for all $y \in Y$. □

We note some important consequences of compactness essentially known from Analysis 2.

Theorem 1.44. *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 $t_\pm \in K$ such that $f(t_+) = \max_{t \in K} f(t)$ and $f(t_-) = \min_{t \in K} f(t)$.*

Proof. 1) Let $\varepsilon > 0$. Since $f \in C(K, X)$, for every $t \in K$ there is a $\delta_t > 0$ such that $\|f(t) - f(s)\| < \varepsilon$ for all $s \in B(t, \delta_t) \subseteq K$. Since $K = \bigcup_{t \in K} B(t, \frac{1}{3}\delta_t)$ and K is compact, there are $t_1, \dots, t_m \in K$ such that $K \subseteq B(t_1, \frac{1}{3}\delta_1) \cup \cdots \cup B(t_m, \frac{1}{3}\delta_m)$, where $\delta_k := \delta_{t_k}$. Set $\delta = \min\{\frac{1}{3}\delta_1, \dots, \frac{1}{3}\delta_m\} > 0$. Take any $s, t \in K$ with $d(s, t) < \delta$. 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

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

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

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

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

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

Theorem 1.41 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 in ℓ^p and $C(K)$. (It can be seen that the conditions are in fact necessary.)

Proposition 1.45. *Let $p \in [1,\infty)$. A set $K \subseteq \ell^p$ is relatively compact if it is bounded and*

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

Proof. Due to Corollary 1.38, it suffices to prove that K is totally bounded. So let $\varepsilon > 0$. By the assumption, there is an $N \in \mathbb{N}$ such that

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

For $x = (x_j) \in K$, 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 1.40 and Corollary 1.38. So we obtain vectors $\hat{v}_1, \dots, \hat{v}_m \in \mathbb{F}^m$ 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

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

Theorem 1.46 (Arzela–Ascoli)**.** *Let K be a compact metric space and $F \subseteq C(K)$ be bounded and* equicontinuous*, i.e.,*

$$\forall\, \varepsilon > 0 \;\exists\, \delta_\varepsilon > 0 \;\forall\, f \in F \;\forall\, s,t \in K \text{ with } d(s,t) \le \delta_\varepsilon : \quad |f(s) - f(t)| \le \varepsilon. \tag{1.6}$$

Then $\overline{F}$ is compact. If F is also closed, it is compact.

Proof. 1) For every $\varepsilon_N := 1/N$ and $N \in \mathbb{N}$, we have $\delta_N := \delta_{\varepsilon_N}$ from (1.6). Since K is compact, Corollary 1.38 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 write $\{t_{k,N} : k = 1, \dots, m_N; N \in \mathbb{N}\} = \{s_l : l \in \mathbb{N}\}$.

2) Take a sequence (f_n) in F. Since F is bounded, we have $\sup_{n\in\mathbb{N}} |f_n(s_1)| \le \sup_{f\in F} \|f\| < \infty$. So there exists a converging subsequence $(f_{\nu_1(k)}(s_1))_k$ in $\mathbb{F}$. Similarly, we obtain a subsequence ν_2 of ν_1 such that $(f_{\nu_2(k)}(s_2))_k$ converges. Note that $f_{\nu_2(k)}(s_1)$ still converges in $\mathbb{F}$ as $k \to \infty$. We iterate this procedure and set $f_{n_j} = f_{\nu_j(j)}$ for all $j \in \mathbb{N}$. Then $f_{n_j}(s_l)$ converges 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 a sequence $k_j \to \infty$ (depending on l).

3) Let $\varepsilon > 0$. Fix $N \in \mathbb{N}$ such that $\varepsilon_N = 1/N \le \varepsilon$. Take $\delta_N = \delta_{\varepsilon_N} > 0$ from (1.6) and the points $t_{k,N}$ from 1). By part 2) there is 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 $k \in \{1, \dots, m_N\}$. Let $t \in K$. There is an index $l \in \{1, \dots, m_N\}$ with $d(t, t_{l,N}) < \delta_N$ due to part 1). We can thus estimate

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

for all $i, j \ge J_\varepsilon$ and all $t \in K$; i.e., $(f_{n_j})_j$ is Cauchy in $C(K)$. This space is complete (which can be shown as in Example 1.4) so that f_{n_j} converges in $C(K)$ to some f in $\overline{F}$ as $j \to \infty$. The compactness of $\overline{F}$ now follows from Corollary 1.38. $\square$

Corollary 1.47. *Let $\alpha \in (0,1]$ and K be a compact metric space. Let (f_n) be bounded and uniformly Hölder (Lipschitz if $\alpha = 1$) continuous, i.e.,*

$$\exists c \geq 0: \quad |f_n(t)| \leq c, \quad |f_n(t) - f_n(s)| \leq c\, d(t,s)^\alpha \quad \text{for all } \; t,s \in K, \; n \in \mathbb{N}. \tag{1.7}$$

Then there is a subsequence f_{n_j} and a fucntion $f \in C(K)$ such that $\|f_{n_j} - f\|_\infty \to 0$ as $j \to \infty$. Moreover, f satisfies (1.7).

Proof. The first part follows from Theorem 1.46 with $F = \{f_n \,|\, n \in \mathbb{N}\}$, noting that (1.6) holds with $\delta_\varepsilon = (\varepsilon/c)^{1/\alpha}$. The last claim can be shown as in Example 1.14. $\square$

Condition (1.7) is true for $K = [0,1]$, $\alpha = 1$ and bounded sequences in $C^1([0,1])$.

Example 1.48. a) The sequence (f_n) in Example 1.40c) is not equicontinuous since $\big|f_n(\frac{3}{2}2^{-n}) - f_n(2^{-n})\big| = 1$, but $\frac{3}{2}2^{-n} - 2^{-n} = 2^{-n-1} \to 0$ as $n \to \infty$.

b) Arzela–Ascoli is wrong for $C_b(\mathbb{R})$ or $C_0(\mathbb{R})$: 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 $|t-n| < 1/2$, which 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 $Tf(t) = \int_0^1 k(t,s)f(s)\,ds$ for $t \in [0,1]$ and $f \in X = C([0,1])$. Then the set $F = \{Tf \,|\, f \in \overline{B}_X(0,1)\}$ is relatively compact.

Proof. Let $f \in \overline{B}_X(0,1)$. By Analysis 1+2 the function $g = Tf$ belongs to X. Clearly, $\|Tf\|_\infty \leq \|k\|_\infty \|f\|_\infty \leq \|k\|_\infty$, so that F is bounded. Let $\varepsilon > 0$. Since k is uniformly continuous, there is a $\delta > 0$ 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 since

$$|Tf(t') - Tf(t)| \leq \int_0^1 |k(t,s) - k(t',s)|\,|f(s)|\,ds \leq \delta$$

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

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

Definition 1.49. *A metric space is called* separable *if it contains a countable dense subset.*

In an exercise it will be shown that separability is preserved under homeomorphisms.

Lemma 1.50. *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

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

is countable, since Y is countable. Let $x \in X$ and $\varepsilon > 0$. By assumption, there exists a $y \in \operatorname{lin} Y$ with $\|x - y\| \leq \varepsilon$. We can also find a $z \in \operatorname{lin}_{\mathbb{Q}} Y$ with $\|y - z\| \leq \varepsilon$. Hence, $\|x - z\| \leq 2\varepsilon$. $\square$

Example 1.51. a) The spaces $\ell^p, 1 \leq p < \infty$, 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' approximation theorem from Analysis 3.

c) The space $L^p(\mathbb{R}^d)$ is separable for $p \in [1,\infty)$. In fact, due to Korollar IV.7.8 in [Wer06] the linear hull of $Y_1 := \{\mathbb{1}_I \,|\, I \text{ is an interval in } \mathbb{R}^d\}$ is dense in $L^p(\mathbb{R}^d)$. Let Y be the subset of Y_1 where the intervals have rational vertices. For all $\varphi \in Y_1$

and $\varepsilon > 0$ there is a $\psi \in Y$ such that $\|\varphi - \psi\|_p \leq \varepsilon$. Hence, the linear hull of the countable set Y is dense in $L^p(\mathbb{R}^d)$.

d) The space ℓ^∞ is not separable. In fact, the set Ω of $\{0,1\}$–valued sequences is uncountable and two different sequences in Ω have the distance 1. Suppose that $\{v_k \,|\, k \in \mathbb{N}\}$ were dense in ℓ^∞. Then, $\Omega \subseteq \bigcup_k B(v_k, 1/4)$ and thus each $\omega \in \Omega$ belongs to exactly one of the balls, say to $B(v_{k(\omega)}, 1/4)$. This means that the map $\Omega \to \mathbb{N};\ \omega \mapsto k(\omega)$ is injective, which contradicts the uncountability of Ω. $\Diamond$

Chapter 2

Continuous linear operators

2.1 Basic properties and examples

Let X and Y be normed vector spaces. We denote their norms by $\|\cdot\|_X$ and $\|\cdot\|_Y$, or just by $\|\cdot\|$. The space of linear mappings $T : X \to Y$ is designated by $L(X,Y)$. We 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. It is known from Linear Algebra that $L(X,Y)$ is a vector space with respect to the operations

$$(T+S)x := Tx + Sx, \quad (\alpha T)x = \alpha Tx \qquad (\forall\, x \in X,\ \alpha \in \mathbb{F},\ T,S \in L(X,Y)).$$

Note that $T0 = 0$. If $\dim X < \infty$ and $\dim Y < \infty$, each operator $T \in L(X,Y)$ can be represented by a matrix, and it is thus continuous. In infinite dimensional spaces there are discontinuous linear maps.

Example 2.1. a) By $T(x_k) = (kx_k)$ we define a linear map $T : c_{00} \to c_{00}$. It holds $T(v_n) = n^{1/2}e_n$ for $v_n = n^{-1/2}e_n$ and $n \in \mathbb{N}$. Since $\|v_n\|_p = n^{-1/2} \to 0$ and $\|T(v_n)\|_p = n^{1/2} \to \infty$ as $n \to \infty$, T is not continuous for any p–norm on c_{00}.

b) The map $Tf = f'$ is linear from $C^1([0,1])$ to $C([0,1])$, but discontinuous if we endow both spaces with the norm $\|f\|_\infty$. (Consider $f_n(t) = n^{-1/2}\sin(nt)$, where $\|f_n\|_\infty \le n^{-1/2}$ and $\|f_n'\|_\infty \ge |f_n'(0)| = n^{1/2}$.) On the other, T is continuous if we equip $C^1([0,1])$ with the norm $\|f\|_{C^1} = \|f\|_\infty + \|f'\|_\infty$ and $C([0,1])$ with $\|f\|_\infty$.

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*: There is a $c > 0$ with $\|Tx\|_Y \le c\,\|x\|_X$ for every $x \in X$.*

Proof. a) The implications a) $\Rightarrow$ b) $\Rightarrow$ c) are clear.
c) $\Rightarrow$ d): By c) and $T0 = 0$ there exists a $\delta > 0$ such that $\|Tz\| \le 1$ for all $z \in X$ with $\|z\| \le \delta$. Let $x \in X \setminus \{0\}$, and set $z = \frac{\delta}{\|x\|}x$. Since T is linear, we deduce $1 \ge \|Tz\| \ge \frac{\delta}{\|x\|}\|Tx\|$ and hence d) with $c = \frac{1}{\delta}$.
d) $\Rightarrow$ a): The linearity of T and d) yield $\|Tx - Tz\| = \|T(x-z)\| \le c\,\|x - z\|$ for all $x, z \in X$. $\square$

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

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

and put $\mathcal{L}(X,X) = \mathcal{L}(X)$. The dual space *$\mathcal{L}(X,\mathbb{F})$ of X is denoted by $X^\star$. An element $x^\star \in X^\star$ is called linear functional on X, and one often writes $\langle x, x^\star\rangle$ instead of $x^\star(x)$. For $T \in \mathcal{L}(X,Y)$ the* operator norm *is given by*

$$\|T\|_{\mathcal{L}(X,Y)} = \|T\| = \inf\{c \geq 0 \,|\, \forall\, x \in X \quad \text{we have} \quad \|Tx\|_Y \leq c\,\|x\|_X\} < \infty.$$

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

a) $\mathcal{L}(X,Y)$ does not change if we replace the norms on X or Y by equivalent ones, though $\|T\|$ may change.

b) $\|I\| = 1$, where $I : X \to X;\ Ix = x$.

c) $\|T\| \stackrel{(1)}{=} \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} \stackrel{(2)}{=} \sup_{\|x\|\leq 1} \|Tx\| \stackrel{(3)}{=} \sup_{\|x\|=1} \|Tx\| =: s$.

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

e) $ST \in \mathcal{L}(X,Z)$ and $\|ST\| \leq \|S\|\,\|T\|$.

Proof. a) and b) are clear.
c) Let $x \neq 0$. For every $\varepsilon > 0$, Definition 2.3 yields $\|Tx\| \leq (\|T\| + \varepsilon)\,\|x\|$. This inequality gives "$\geq$" in (1). We clearly have "$\geq$" in (2) and (3). Further,

$$\frac{\|Tx\|}{\|x\|} = \Big\|T\big(\tfrac{1}{\|x\|}x\big)\Big\| \leq s,$$

so that $\|T\| \leq s$.
d) is obvious for $x = 0$. For $x \neq 0$, from part c) we deduce $\frac{\|Tx\|}{\|x\|} \leq \|T\|$.
e) Part d) implies that $\|STx\| \leq \|S\|\,\|Tx\| \leq \|S\|\,\|T\|\,\|x\|$, whence e) follows. □

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

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

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

Thus, $\mathcal{L}(X,Y)$ is a normed vector space.

Assume that Y is a Banach space. Let (T_n) be a Cauchy sequence in $\mathcal{L}(X,Y)$; i.e., for every $\varepsilon > 0$ there is an $N_\varepsilon \in \mathbb{N}$ such that $\|T_n - T_m\| \leq \varepsilon$ for all $n, m \geq N_\varepsilon$. Let $x \in X$. Since $T_n - T_m \in \mathcal{L}(X,Y)$, Remark 2.4 yields

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

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

for all $n,m \geq N_\varepsilon$, so that $(T_n x)$ is a Cauchy sequence in Y. Hence, there exists exactly one $y \in Y$ with $y = \lim_{n\to\infty} T_n x =: Tx$. Let $\alpha, \beta \in \mathbb{F}$ and $x, z \in X$. We deduce from the linearity of T_n that

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

Since (T_n) is Cauchy, there is a $c > 0$ with $\|T_n\| \leq c$ for all n, whence $\|Tx\| = \lim_{n\to\infty} \|T_n x\| \leq c\,\|x\|$. So far we have shown that $T \in \mathcal{L}(X,Y)$. The above displayed estimate further yields

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

i.e., $\|T - T_n\| \leq \varepsilon$ for all $n \geq N_\varepsilon$. □

Example 2.6 (Multiplication operators)**.** a) Let $X = C([0,1])$, and let $m \in C([0,1])$ be given. Set $Tf = mf$ for every $f \in X$.
1) We then have $Tf \in X$ and $T(\alpha f + \beta g) = \alpha m f + \beta m g = \alpha T f + \beta T g$ for all $f, g \in X$ and $\alpha, \beta \in \mathbb{F}$. Hence, $T : X \to X$ is linear.
2) Moreover, $\|Tf\|_\infty = \sup_t |m(t)|\,|f(t)| \leq \|m\|_\infty \|f\|_\infty$, and thus $T \in \mathcal{L}(X)$ with $\|T\| \leq \|m\|_\infty$.
3) Since $\|\mathbb{1}\|_\infty = 1$, we further obtain $\|T\| \geq \|T\mathbb{1}\|_\infty = \|m\mathbb{1}\|_\infty = \|m\|_\infty$. So we have shown that $\|T\| = \|m\|_\infty$.

b) Let $X = L^p(\mathbb{R}^d)$ and $1 \leq p \leq \infty$, and let $\hat{m} \in L^\infty(\mathbb{R}^d)$ be given. For $\hat{f} \in X$ we define $T\hat{f} = mf + \mathcal{N}$ for any representatives $f \in \hat{f}$ and $m \in \hat{m}$.[2]
1) Take any $f_1 \in \hat{f}$ and $m_1 \in \hat{m}$. Then there are null sets N' and N'' such that $f(t) = f_1(t)$ for all $t \notin N'$ and $m(t) = m_1(t)$ for all $t \notin N''$. Therefore, $m(t)f(t) = m_1(t)f_1(t)$ for all $t \notin N' \cup N''$, which is a null set. As a consequence, $T\hat{f}$ does not depend on the representatives and it is thus well–defined.
2) Clearly, mf is measurable. Let $c > \|m\|_\infty$. Then there is a null set N such that $|m(t)f(t)| \leq c\,|f(t)|$ for all $t \notin N$. It follows that $mf \in \mathcal{L}^p(\mathbb{R}^d)$ and $\|mf\|_p \leq \|m\|_\infty \|f\|_p$. As a result, T maps X into X and $\|T\hat{f}\|_p = \|mf\|_p \leq \|\hat{m}\|_\infty \|\hat{f}\|_p$.
3) For $f, g \in \mathcal{L}(\mathbb{R}^d)$ and $\alpha, \beta \in \mathbb{F}$ we have

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

Consequently, $T \in \mathcal{L}(X)$ and $\|T\| \leq \|\hat{m}\|_\infty$.
4) We now claim that $\|T\| = \|\hat{m}\|_\infty$. This claim holds if $\hat{m} = 0$. So let $\|\hat{m}\|_\infty = \|m\|_\infty > 0$ for any $m \in \hat{m}$. Take $\varepsilon \in (0, \|m\|_\infty)$ and $n \in \mathbb{N}$. We put $A_{\varepsilon,n} = \{t \in B(0,n) \mid |m(t)| \geq \|m\|_\infty - \varepsilon\}$. By the definition of the essential supremum, there is a $k \in \mathbb{N}$ such that $\lambda(A_{\varepsilon,k}) > 0$. We can thus set $f_\varepsilon = \|\mathbb{1}_{A_{\varepsilon,k}}\|_p^{-1} \mathbb{1}_{A_{\varepsilon,k}} \in \mathcal{L}^p(\mathbb{R}^d)$. It holds $\|f_\varepsilon\|_p = 1$ and thus

$$\|T\| \geq \|mf_\varepsilon\|_p = \frac{1}{\|\mathbb{1}_{A_{\varepsilon,k}}\|_p} \left(\int_{A_{\varepsilon,k}} |m(t)|^p \, dt \right)^{\frac{1}{p}} \geq \|m\|_\infty - \varepsilon$$

for all ε, where we assume that $p < \infty$. Taking the supremum over $\varepsilon > 0$, we deduce $\|T\| = \|\hat{m}\|_\infty$. The case $p = \infty$ is treated in the same way.

Example 2.7 (Integral operators)**.** Let $X = C([0,1])$, and let $k \in C([0,1]^2)$ be given. For every $f \in X$ we set

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

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

1) From Analysis 1+2, we know that $Tf \in X$ and that $T : X \to X$ is linear.
2) Set $\kappa := \sup_{t\in[0,1]} \int_0^1 |k(t,s)|\,ds \le \|k\|_\infty$. We then obtain

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

for all $f \in X$, whence $T \in \mathcal{L}(X)$ and $\|T\| \le \kappa$.
3) By continuity there exists a $t_0 \in [0,1]$ such that $\kappa = \int_0^1 |k(t_0,s)|\,ds$. Set

$$f_n(s) = \frac{\bar{k}(t_0,s)}{|k(t_0,s)| + \frac{1}{n}}$$

for all $s \in [0,1]$ and $n \in \mathbb{N}$. Since $f_n \in X$ and $\|f_n\|_\infty \le 1$, we conclude that

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

as $n \to \infty$, using e.g. Lebesgue's theorem. Consequently, $\|T\| = \kappa$. Note that for $k \ge 0$ one obtains this equality more easily since $\|T\| \ge \|T\mathbb{1}\|_\infty = \kappa$ in this case.

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

b) Let $X = L^p(\mu)$ for a measure space and $1 \le p \le \infty$, and let $g \in L^{p'}(\mu)$ be fixed. By Hölder's inequality $\varphi(f) = \int fg\,d\mu \in \mathbb{F}$ exists for all $f \in X$, and $|\varphi(f)| \le \|f\|_p\,\|g\|_{p'}$. Since linearity is clear, we see that $\varphi \in X^\star$ with $\|\varphi\| \le \|g\|_{p'}$. (Equality is shown in Section 4.2.)

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 $0 \le t \le \frac{1}{n}$ 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 $n \in \mathbb{N}$.

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

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

Clearly, $R, L : X \to X$ are linear and $\|Rx\|_p = \|x\|_p$ and $\|Lx\|_p \le \|x\|_p$ for all $x \in X$, as well as $\|Le_2\|_p = 1$. We thus obtain that $R, L \in \mathcal{L}(X)$ with norm 1. Observe that $LR = I$, $RLx = (0, x_2, x_3, \dots)$, 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 for $T \in L(X)$ with $\dim X < \infty$ injectivity and surjectivity are equivalent, and that a right or left inverse is automatically an inverse!

Definition 2.10. *Let X and Y be normed vector spaces and $T : X \to Y$ be linear.*

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

b) An injective operator $T \in \mathcal{L}(X,Y)$ is called an embedding*, and one then writes $X \hookrightarrow Y$.*

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

d) A map $T \in L(X,Y)$ is called 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 $y_n = Jx_n$. Hence, $x_n = J^{-1}y_n$ and so (x_n) converges if and only if (y_n) converges. In particular, X is a Banach space if and only if Y is a Banach space. Let $C \subseteq X$ and $D = JC = (J^{-1})^{-1}(C) \subseteq Y$. Hence, $C = J^{-1}(D)$ and so C is open (closed) if and only if D is open (closed).

b) $N(T)$ is closed for any $T \in \mathcal{L}(X, Y)$. 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 $T \in \mathcal{L}(X, Y)$ be an isometry. Then its inverse $T^{-1} : R(T) \to X$ is linear and isometric. In fact, for $z = Tx \in R(T)$ we compute $\|T^{-1}z\| = \|x\| = \|Tx\| = \|z\|$.

d) Let X be a Banach space and let $T \in \mathcal{L}(X, Y)$ satisfy $\|Tx\| \geq c\,\|x\|$ for some $c > 0$ and all $x \in X$ (e.g., if T is isometric). Then $R(T)$ is closed in Y. In fact, let $y_n = Tx_n \to y$ in Y as $n \to \infty$. The assumption yields $\|x_n - x_m\| \leq c^{-1}\|y_n - y_m\| \to 0$ as $n, m \to \infty$. Since X is a Banach space, there exists $x = \lim_{n\to\infty} x_n$. The continuity of T then implies that $y = Tx \in R(T)$; 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 $\|y\|_X = \|Iy\|_X \leq c\,\|y\|_Y$ for all $y \in Y$ and a constant $c \geq 0$ if and only if $\|\cdot\|_Y$ is finer than $\|\cdot\|_X$. Examples: $\ell^p \hookrightarrow \ell^q$ if $1 \leq p \leq q \leq \infty$; $C^1([0,1]) \hookrightarrow C([0,1])$. ♢

Example 2.12. a) Let $\emptyset \neq B \in \mathcal{B}_d$ 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 $L^p(B)$, $p \in [1, \infty]$, 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 $x_n \to x$. So $f(x) = 0$ 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

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

Clearly, $Jf \in C_0(\mathbb{R})$ and J is linear and isometric. Hence, $J : C([-1,1]) \hookrightarrow C_0(\mathbb{R})$ is an isometric embedding. (One can similarly embed $C(K)$ into $C_0(\mathbb{R}^d)$ for any compact $K \subseteq \mathbb{R}^d$ using Tietze's extension theorem, see Theorem 20.1 in [Rud87].)

c) Let $X = \{f \in C^1([0,1]) \mid f(0) = 0\}$ 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 given by

$$D^{-1}g(t) = \int_0^t g(s)\,ds, \quad t \in [0,1],\ g \in C([0,1]).$$

In the view of the previous Remark, the Banach space structures of X and $C([0,1])$ are the 'same'. However, this isomorphism destroys other properties such as positivity (e.g., $f(t) = t(1-t)$ is positive on $[0,1]$, in contrast to $Df(t) = f'(t) = 1 - 2t$).

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 $T_0 \in \mathcal{L}(D, Y)$. Then there exists exactly one extension $T \in \mathcal{L}(X, Y)$ of T_0 i.e., $T_0x = Tx$ for all $x \in D$. Moreover, $\|T_0\| = \|T\|$ and T is isometric if T_0 is isometric.*

Proof. Let $x \in X$. Choose $x_n \in D$ such that $x_n \to x$ in X as $n \to \infty$. Since $\|T_0x_n - T_0x_m\| \leq \|T_0\|\,\|x_n - x_m\|$, the sequence (T_0x_n) is Cauchy and thus converges to some Tx in Y. Let also $\widetilde{x}_n \in D$ tend to x. Because of $\|T_0x_n - T_0\widetilde{x}_n\| \leq$

$\|T_0\| \, \|x_n - \widetilde{x}_n\| \to 0$ 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 $z_n \to z$. We then 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

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

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

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

so that $\|T_0\| = \|T\|$. Let $S \in \mathcal{L}(X, Y)$ 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

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

Convolution and Young's inequality

In this paragraph we derive the important Young inequality for convolutions. As a preparation we recall Fubini's theorem from Analysis 3, see also Theorem 8.8 in [Rud87]. Let $A \in \mathcal{B}_m$, $B \in \mathcal{B}_n$, and $(x, y) \in A \times B \subseteq \mathbb{R}^{m+n}$. It can be seen that $A \times B \in \mathcal{B}_{m\times n}$. Let $f : A \times B \to \tilde{\mathbb{F}}$ with $\tilde{\mathbb{F}} \in \{[0, \infty], \mathbb{F}\}$ be measurable. We set

$$f^y : A \to \tilde{\mathbb{F}}, \quad f^y(x) = f(x, y) \qquad \text{for each fixed } \; y \in B,$$
$$f_x : B \to \tilde{\mathbb{F}}, \quad f_x(y) = f(x, y) \qquad \text{for each fixed } \; x \in A.$$

These functions are measurable (for each fixed $y \in B$, resp. $x \in A$) since e.g. $f_x = f \circ j_x$ for each $x \in A$ and the continuous function $j_x : B \to A \times B$, $j_x(y) = (x, y)$.

Theorem 2.14 (Fubini). *a) Let $f : A \times B \to [0, \infty]$ be measurable. Then the functions $F : A \to [0, \infty]$ and $G : B \to [0, \infty]$, given by*

$$F(x) := \int_B f_x(y)\, dy \quad \textit{for} \;\; x \in A, \qquad G(y) := \int_A f^y(x)\, dx \quad \textit{for} \;\; y \in B,$$

are measurable, and it holds

$$\int_{A\times B} f(x, y)\, d(x, y) = \int_A \left(\int_B f(x, y)\, dy \right) dx = \int_B \left(\int_A f(x, y)\, dx \right) dy. \tag{2.1}$$

b) Let $f \in L^1(A \times B)$. Then there are nullsets $N_A \subseteq A$ and $N_B \subseteq B$ such that f_x is integrable for all $x \in A \setminus N_A$ and f^y is integrable for all $y \in B \setminus N_B$. We define F and G as above for $x \in A \setminus N_A$ and for $y \in B \setminus N_B$, respectively, and we put $F(x) = 0$ and $G(y) = 0$ for $x \in N_A$ and $y \in N_B$, respectively. Then F and G are integrable and formula (2.1) holds.

We next introduce and discuss the *convolution* $f * g$ of functions in $L^p(\mathbb{R}^d)$ for suitable p. Note that for measurable $f, g : \mathbb{R}^d \to \mathbb{F}$ the map

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

is measurable as a combination of measurable maps.

(1) Let $f, g \in L^1(\mathbb{R}^d)$. By means of Fubini a) and the transformation $z = x - y$, we derive

$$\int_{\mathbb{R}^{2d}} \varphi(x,y) d(x,y) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |f(x-y)|\, dx\, |g(y)|\, dy$$
$$= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |f(z)|\, dz\, |g(y)|\, dy = \|f\|_1 \|g\|_1.$$

Hence, $\varphi \in L^1(\mathbb{R}^{2d})$ and Fubini b) shows that the *convolution*

$$(f * g)(x) := \int_{\mathbb{R}^d} f(x-y) g(y)\, dy \tag{2.2}$$

is defined for a.e. $x \in \mathbb{R}^d$, that it is integrable, and that

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

(2) Let $q \in [1, \infty)$, $f \in L^1(\mathbb{R}^d)$, and $g \in L^q(\mathbb{R}^d)$. Fubini a) again yields the existence of

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

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

$$\psi(x)^q \le \Big(\int_{\mathbb{R}^d} |f(x-y)|\, dy \Big)^{\frac{q}{q'}} \int_{\mathbb{R}^d} |f(x-y)|\, |g(y)|^q\, dy$$
$$= \|f\|_1^{q-1} \int_{\mathbb{R}^d} |f(x-y)|\, |g(y)|^q\, dy,$$

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

$$\|\psi\|_q^q = \int_{\mathbb{R}^d} \psi^q\, dx \le \|f\|_1^{q-1}\, \|\, |f| * |g|^q\, \|_1 \le \|f\|_1^{q-1}\, \|f\|_1 \int_{\mathbb{R}^d} |g|^q\, dx = \|f\|_1^q\, \|g\|_q^q\,.$$

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

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

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 $\mathbb{R}^{2d}$. Fubini b) thus shows that $f * g$ is defined by (2.2) for a.e. $x \in B(0,n)$ and each $n \in \mathbb{N}$ (and hence for a.e. $x \in \mathbb{R}^d$) and that the map $\mathbb{1}_{B(0,n)} f * g$ is measurable on $\mathbb{R}^d$. Letting $n \to \infty$, we see that the pointwise limit $f * g$ is measurable on $\mathbb{R}^d$. The above estimate on $\|\psi\|_q^q$ finally yields

$$\|f * g\|_q^q = \int_{\mathbb{R}^d} \Big| \int_{\mathbb{R}^d} f(x-y) g(y)\, dy \Big|^q dx \le \int_{\mathbb{R}^d} \psi^q\, dx \le \|f\|_1^q\, \|g\|_q^q.$$

We have thus proved a part of the next proposition, see Theorem 4.33 in [Bre11] for the remaining cases. A different proof for the full result is given in Section 2.3.

Proposition 2.15. *Let $1 \le p, q, r \le \infty$ with $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Take $f \in L^p(\mathbb{R}^d)$ and $g \in L^q(\mathbb{R}^d)$. Then the convolution $(f * g)(x)$ in (2.2) is defined for a.e. $x \in \mathbb{R}^d$ and gives a function in $L^r(\mathbb{R}^d)$. We further have* Young's inequality

$$\|f * g\|_r \le \|f\|_p\, \|g\|_q.$$

2.2 Standard constructions

A) Product spaces

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

Definition 2.16. *Let X_1 and X_2 be* **closed** *linear subspaces of a normed vector space X such that $X_1 + X_2 = X$ and $X_1 \cap X_2 = \{0\}$. Then we say that X is the* direct sum *of X_1 and X_2 and that X_2 is the* complement *of X_1, and we write $X = X_1 \oplus X_2$. Let Y be a vector space. A map $P \in L(Y)$ is called* projection *if $P^2 = P$.*

Lemma 2.17. *Let X be a normed vector space and $P \in \mathcal{L}(X)$ be a projection. Then the operator $I - P \in \mathcal{L}(X)$ is also a projection, and we have*

$$R(P) = N(I-P) =: X_1, \qquad N(P) = R(I-P) =: X_2,$$

and $X = X_1 \oplus X_2$. Moreover, $\|P\| \geq 1$ if $P \neq 0$.

Proof. From $P = P^2$ we deduce $(I-P)^2 = I - 2P + P^2 = I - P$. If $y \in R(P)$, then there is an $x \in X$ 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 $\|P\| = \|P^2\| \leq \|P\|^2$. □

Remark 2.18. a) Let $X = X_1 \oplus X_2$. For each $x \in X$ we then have unique $x_1 \in X_1$ 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. The definition easily yields $P^2 = P$, $R(P) = X_1$ and $N(P) = X_2$. Let $x, y \in X$ and $\alpha, \beta \in \mathbb{F}$. Then there are $x_k, y_k \in X_k$ such that $x = x_1 + x_2$ and $y = y_1 + y_2$. We now obtain

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

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

b) Let X be a Banach space. We will see in Proposition 3.25 that the projection P from a) is continuous and that $X_1 \oplus X_2 \cong X_1 \times X_2$. ♢

Example 2.19. a) Let $X = \mathbb{R}^2$, $t \in \mathbb{R}$, and $P = \begin{pmatrix} 1 & t \\ 0 & 0 \end{pmatrix}$. 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 $X = L^1(\mathbb{R})$ and $Pf = \mathbb{1}_{\mathbb{R}_+} f$ for $f \in X$. Clearly, $\|Pf\|_1 \le \|f\|_1$ and $P^2 = P$, so that $P \in \mathcal{L}(X)$ 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 $J : R(P) \to L^1(\mathbb{R}_+)$, $Jf = f|_{\mathbb{R}_+}$ with inverse 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 $L^1(\mathbb{R}_+) \oplus L^1(\mathbb{R}_-)$, considering $L^1(\mathbb{R}_\pm)$ as subspaces of $L^1(\mathbb{R})$ by extending functions by 0.

c) c_0 has no complement in ℓ^∞, see Satz IV.6.5 in [Wer05].

C) Quotient spaces

Let X be a normed vector space, Y a linear subspace and

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

be the quotient space. The quotient map

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

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

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

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

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

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

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

Since $\varepsilon > 0$ is arbitrary, the quotient norm is a seminorm.

Now, let Y be closed. If $\|\hat{x}\| = 0$ for some $\hat{x} \in X/Y$, then there exist $y_n \in Y$ with $\|x - y_n\| \to 0$. From the closedness of Y it follows that $x \in Y$, and hence $\hat{x} = 0$. So far we have established that X/Y is a normed vector space for the quotient norm.

Because of $\|Qx\| = \|\hat{x}\| \le \|x\|$, the quotient map Q is continuous with $\|Q\| \le 1$. Moreover, for each $\delta \in (0, 1)$, Lemma 1.43 gives a $\bar{x} \in X$ with $\|\bar{x}\| = 1$ and

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

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

Proposition 2.20. *Let X be a Banach space and Y be a linear subspace.*

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

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 a Banach space

Proof. 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

$$\|\hat{x}_m - \hat{x}_{n_k}\| \le 2^{-k} \qquad \text{for all} \quad m \ge n_k. \tag{$*$}$$

Consequently, there are $y_{n_k} \in Y$ with $\|x_{n_{k+1}} - x_{n_k} - y_{n_k}\| \le 2 \cdot 2^{-k}$ 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 $x = \lim_{N\to\infty} v_N$ in X. 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 $\hat{v}_N = \hat{x}_{n_{N+1}}$. Since Q is continuous, it follows $\hat{x}_{n_{N+1}} - \hat{x} = Q(v_N - x) \to 0$ in X/Y as $N \to \infty$. Finally, for $\varepsilon > 0$ there is an $N = N_\varepsilon \in \mathbb{N}$ such that $2^{-N} \le \varepsilon$ and $\|\hat{x} - \hat{x}_{n_N}\| \le \varepsilon$. Using ($*$), we deduce

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

for all $m \ge n_N$. □

Example 2.21. 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 $y \in Y$ 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 3.21 below. As a result, $Z \simeq X/Y$ via J. In Example 2.19b) we thus obtain $L^1(\mathbb{R})/L^1(\mathbb{R}_+) \simeq L^1(\mathbb{R}_-)$.

b) Note that the quotient construction is more general than the direct sum. For instance, ℓ^∞/c_0 exists, though c_0 has no complement in ℓ^∞.

D) Completion

Proposition 2.22. *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. Let E by the vector space of all Cauchy sequences $v = (x_n)_{n\in\mathbb{N}} \subseteq X$. Note that for $(x_n) \in E$ the sequence $(\|x_n\|)$ is Cauchy in $\mathbb{R}$ and thus it exists $p(v) := \lim_{n\to\infty} \|x_n\|$ 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 $p(v) \le p(v - w) + p(w) = p(w)$ as well as $p(w) \le p(v)$. 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 $N = N_\varepsilon \in \mathbb{N}$ such that $\|x_k - x_N\| \le \varepsilon$ for all $k \ge N$. Because of $|||\tilde{v} - Jx_N||| = \lim_{k\to\infty} \|x_k - x_N\| \le \varepsilon$, the range of J is dense in $\tilde{X}$.

Now 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

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

We define $w = (y_m)_m = (x_{m,j_m})_m$. From the above inequality we deduce

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

so that $w \in E$. Let $\varepsilon > 0$. Since w is a Cauchy sequence in X, there is an $N_\varepsilon \in \mathbb{N}$ such that $\|y_n - y_m\| \le \varepsilon$ for all $n, m \ge N_\varepsilon$. We can thus estimate

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

for all $m \ge N_\varepsilon$.

It remains to show 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 that it is isometric with 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 2.23. Usually one identifies X with the subspace JX of $\tilde{X}$ (as one does with $\mathbb{Q}$ and $\mathbb{R}$). Every $T \in \mathcal{L}(X,Y)$ can uniquely be extended to an operator $\tilde{T} \in \mathcal{L}(\tilde{X},Y)$ by means of Lemma 2.13. Also, $\tilde{T}$ is isometric if T is isometric. ♢

Example 2.24. 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. We see in Proposition 3.12 below that J has dense range. The above remark yields 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 X and Y be Banach spaces which are linear subspaces of a vector space Z that possesses a metric d for addition and scalar mulitplication 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 was 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 2.25. *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 $\mathbb{R}_+$. 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

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

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 $j = 1, 2$. Since $z_1 + z_2 = (x_1 + x_2) + (y_1 + y_2)$, we conclude

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

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 linear, isometric and surjective and that D is closed in $X \times Y$. In view of Remark 2.11 and Proposition 2.20, 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 $(x, y) \in X \times Y$. The operator J is isometric since

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

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

Proposition 2.26. *Let X_j and Y_j (with $j = 0, 1$) be Banach spaces which are linear subspaces of vector spaces V and W with compatible metrics, respectively. We assume that these inclusion maps are continuous. Let $T_0 \in \mathcal{L}(X_0, Y_0)$ and $T_1 \in \mathcal{L}(X_1, Y_1)$ 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 $\widetilde{T} \in \mathcal{L}(X_0 + X_1, Y_0 + Y_1)$ such that $\widetilde{T} x_j = T_j x_j$ for all $x_j \in X_j$ and $j = 0, 1$.*

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 $x'_j \in X_j$, then 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 $j = 0, 1$. 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,

$$\|\widetilde{T}x\|_{Y_0+Y_1} \le \|T_0x_0\|_{Y_0} + \|T_1x_1\|_{Y_1} \le \max\{\|T_0\|,\ \|T_1\|\}\,(\|x_0\| + \|y_1\|).$$

Taking the infimum over all decompositions $x = x_0 + x_1$ in $X_0 + X_1$, we derive that $\widetilde{T}$ is bounded. Let $S \in \mathcal{L}(X_0 + X_1, Y_0 + Y_1)$ 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 theory is an important branch of functional analysis which treats the following problem. Let X_j and Y_j (with $j = 0, 1$) be Banach spaces being linear subspaces of vector spaces V and W, 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) below, we can extend T to a bounded linear opeartor $\widetilde{T} : X_0 + X_1 \to Y_0 + Y_1$ where the sum space

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

is endowed with the complete norm

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

for Banach spaces X and Y being linear subspaces of vector space Z. 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 [Lun09] 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 $p_0, p_1, q_0, q_1 \in [1, \infty]$. 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 any $p \in [p_0, p_1]$ if $p_0 \le p_1$ and any $p \in [p_1, p_0]$ 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 $\theta \in [0, 1]$. The spaces $L^{p_0}(\mu)$ and $L^{p_1}(\mu)$ need to be 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 an exercise or in Analysis 3, Hölder's inequality shows that $U \subseteq L^p(\mu)$ and $V \subseteq L^q(\nu)$ with

$$\|f\|_p \le \|f\|_{p_0}^{1-\theta}\ \|f\|_{p_1}^{\theta} \quad \text{and} \quad \|g\|_q \le \|g\|_{q_0}^{1-\theta}\ \|g\|_{q_1}^{\theta} \tag{2.3}$$

for all $f \in U$ and $g \in V$. We recall that the space of simple functions

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

with a support of finite measure is dense in $L^r(\mu)$ if $r \in [1, \infty)$ and that the space of simple functions

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

is dense in $L^\infty(\mu)$. The proof given there also shows that for $f \in U$ with $p_0 \ne p_1$ 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 (2.3).

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 $f \in L^p(\mu)$, we set $\widetilde{f} = \|f\|_p^{-1} f$ so that $\|\widetilde{f}\|_p = 1$. Because of $\widetilde{f} \in L^p(\mu)$ the set $\{\widetilde{f} \ge 1\} := \{\omega \in \Omega \,|\, \widetilde{f}(\omega) \ge 1\}$ has finite measure, 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 (2.3). Therefore the functions $\widetilde{f} = f_0 + f_1$ and f belong to $L^{p_0}(\mu) + L^{p_1}(\mu)$. Using $\|\widetilde{f}\|_p = 1$, we further compute

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

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

Theorem 2.27 (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 $T_j \in \mathcal{L}(L^{p_j}(\mu), L^{q_j}(\nu))$ 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 $T_\theta \in \mathcal{L}(L^p(\mu), L^q(\nu))$. Moreover, T_θ is the restriction of $\widetilde{T} \in \mathcal{L}(L^{p_0}(\mu) + L^{p_1}(\mu), L^{q_0}(\nu) + L^{q_1}(\nu))$ which is the unique extension of T_0 and T_1 to this space (see Proposition 2.26), and we have

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

We note that the theorem also holds for $\mathbb{F} = \mathbb{R}$ with an additional multiplicative constant in the estimate, see Satz II.4.2 in [Wer05].

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 (2.3),

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

so that the theorem holds in this case. Next, let $p_0 \ne p_1$. In this case, we have $p < \infty$, and thus $E(\mathcal{A})$ is dense in $L^p(\mu)$. We consider

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

where we may assume that the sets A_j are pairwise disjoint and have finite measure. The assumptions again yield $Tf \in V$. Due to (4.6) (see also Satz II.2.4 in [Wer05]), it holds

$$\|Tf\|_q = \sup_{g\in E(\mathcal{B}),\|g\|_{q'}\le 1} \Big| \int_\Lambda (Tf)\, g \, d\nu \Big| \tag{2.4}$$

if $q' < \infty$. If $q' = \infty$ (i.e., $q = 1$ which holds if and only if $(q_0, q_1) = (1, 1)$), then (2.4) 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

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

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

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

for $y \in \Lambda$ with $g(y) \neq 0$, and put $(G(z))(y) = 0$ if $g(y) = 0$. If $q' = \infty$, we simply set $G(z) = g$. 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

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

for $z \in S$. The right hand side easily yields that $\varphi \in C(S)$ is holomorphic on S°.

We want to apply the Three–Lines-Theorem, see Satz II.4.3 in [Wer05], and thus check the estimates assumed in this result. Writing $z = s + it \in S$ with $s \in [0,1]$ and $t \in \mathbb{R}$, we compute

$$|F(z)(x)| = |f(x)|^{p\left(\frac{1-s}{p_0}+\frac{s}{p_1}\right)}, \qquad |G(z)(y)| = |g(y)|^{q'\left(\frac{1-s}{q_0'}+\frac{s}{q_1'}\right)} \tag{2.5}$$

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

$$\Big|\int_\Lambda (Tf) g\, d\nu\Big| = |\varphi(\theta)| \le \Big(\sup_{t\in\mathbb{R}} |\varphi(it)|\Big)^{1-\theta} \Big(\sup_{t\in\mathbb{R}} |\varphi(1+it)|\Big)^{\theta} \tag{2.6}$$

Hölder's inequality, the assumptions and estimate (2.5) with $s = 0$ further imply

$$\begin{aligned}
|\varphi(it)| &\le \|T_0 F(it)\|_{q_0}\, \|G(it)\|_{q_0'} \le \|T_0\|\, \|F(it)\|_{p_0}\, \|G(it)\|_{q_0'} \\
&= \|T_0\| \Big(\int_\Omega |f|^{\frac{pp_0}{p_0}}\, d\mu\Big)^{\frac{1}{p_0}} \Big(\int_\Lambda |g|^{\frac{q'q_0'}{q_0'}}\, d\nu\Big)^{\frac{1}{q_0'}} \\
&= \|T_0\|\, \|f\|_p^{\frac{p}{p_0}}\, \|g\|_{q'}^{\frac{q'}{q_0'}} \le \|T_0\|\, \|f\|_p^{\frac{p}{p_0}},
\end{aligned}$$

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

$$\begin{aligned}
|\varphi(1+it)| &\le \|T_1 F(1+it)\|_{q_1}\, \|G(1+it)\|_{q_1'} \le \|T_1\|\, \|F(1+it)\|_{p_1}\, \|G(1+it)\|_{q_1'} \\
&= \|T_1\| \Big(\int_\Omega |f|^{\frac{pp_1}{p_1}}\, d\mu\Big)^{\frac{1}{p_1}} \Big(\int_\Lambda |g|^{\frac{q'q_1'}{q_1'}}\, d\nu\Big)^{\frac{1}{q_1'}} \\
&= \|T_1\|\, \|f\|_p^{\frac{p}{p_1}}\, \|g\|_{q'}^{\frac{q'}{q_1'}} \le \|T_1\|\, \|f\|_p^{\frac{p}{p_1}}.
\end{aligned}$$

Formulas (2.4) and (2.6) now lead to

$$\|Tf\|_q \le \|T_0\|^{1-\theta}\, \|f\|_p^{p(1-\theta)/p_0}\, \|T_1\|^{\theta}\, \|f\|_p^{p\theta/p_1} = \|T_0\|^{1-\theta}\, \|T_1\|^{\theta}\, \|f\|_p \tag{2.7}$$

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 (2.3). Inequality (2.7) thus holds for all $f \in U$. Lemma 2.13 then allows to extend T uniquely from its domain U to an operator $T_\theta \in \mathcal{L}(L^p(\mu), L^q(\mu))$ with norm less or equal $\|T_0\|^{1-\theta} \|T_1\|^\theta$. Another observation before the theorem yields that $L^p(\mu) \hookrightarrow L^{p_0}(\mu) + L^{p_1}(\mu)$ 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 Proposition 2.15.

(1a) Recall the definition (2.2) of the convolution $f * g$ for $f, g \in L^1(\mathbb{R}^d)$. There we have also shown that

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

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

$$\begin{aligned}\int_{B(0,n)\times\mathbb{R}^d} |\varphi(x,y)| \, d(x,y) &\le \int_{B(0,n)} \int_{\mathbb{R}^d} |f(x-y)| \, \|g\|_\infty \, dy \, dx \\ &= \|g\|_\infty \int_{B(0,n)} \int_{\mathbb{R}^d} |f(z)| \, dz \, dx = \lambda(B(0,n) \, \|f\|_1 \, \|g\|_\infty.\end{aligned}$$

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 $\mathbb{R}^d$. Replacing in the above estimate the integral over $x \in B(0,n)$ by a supremum in $x \in \mathbb{R}^d$, we further obtain

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

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

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

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

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

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

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

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

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

Riesz–Thorin now allows to restrict S to an operator $S_p \in \mathcal{L}(L^p(\mathbb{R}^d), L^r(\mathbb{R}^d))$ with norm less or equal $\|g\|_q$. We have thus proved Proposition 2.15.

Chapter 3

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.

3.1 The principle of uniform boundedness and strong convergence

Theorem 3.1 (Baire)**.** *Let M be a complete metric space and $O_n \subseteq M$ be open and dense for each $n \in \mathbb{N}$. Then $O := \bigcap_{n\in\mathbb{N}} O_n$ is dense in M.*

Proof. For every $x_0 \in M$ and $\delta > 0$ we must find an $x \in B_0 \cap O$, where $B_0 := B(x_0, \delta)$. So let $x_0 \in M$ and $\delta > 0$. Since O_1 is open and dense, there is an $x_1 \in O_1 \cap B_0$ and a $\delta_1 \in (0, \frac{1}{2}\delta]$ 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

$$\overline{B_n} \subseteq O_n \cap B_{n-1} \subseteq O_n \cap (O_{n-1} \cap B_{n-2}) \subseteq \cdots \subseteq O_n \cap \ldots \cap O_1 \cap B_0. \qquad (*)$$

Since $\delta_m \le 2^{-m}\delta$, we obtain $x_n \in B_m \subseteq B(x_m, 2^{-m}\delta)$ for all $n \ge m$. Hence, $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and there exists $x = \lim_{n\to\infty} x_n$ which belongs to $\overline{B_m}$ for each $m \in \mathbb{N}$. The inclusions $(*)$ now yield $x \in O \cap B_0$. □

The set $\mathbb{R}^2 \setminus \mathbb{R}$ is open and dense in $\mathbb{R}^2$. A property that holds for all x in an open and dense subset is often called *generic*.

Corollary 3.2. *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 $N \in \mathbb{N}$ such that $A_N^\circ \neq \emptyset$.*

Proof. Suppose that $A_n^\circ = \emptyset$ for all $n \in \mathbb{N}$. Then $O_n = M \setminus A_n$ were open and dense. Theorem 3.1 thus implies that $\bigcap_{n\in\mathbb{N}} O_n$ is dense in M. But, this impossible since $\bigcap_n O_n = M \setminus \bigcup_n A_n = \emptyset$ by the assumption. □

Example 3.3. One needs the completeness in the above corollary. Consider for instance, $(c_{00}, \|\cdot\|_p)$ and the sets $A_n = \{x = (x_1, \ldots, x_n, 0, \ldots) \,|\, x_j \in \mathbb{F}\}$. These sets are closed for all $n \in \mathbb{N}$ by Lemma 1.42, 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 $m \to \infty$. ◊

Theorem 3.4 (Principle of uniform boundedness)**.** *Let X be a Banach space, Y be a normed vector space, and $\mathcal{T} \subseteq \mathcal{L}(X,Y)$. 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. Set $A_n = \{x \in X \,|\, \|Tx\| \leq n \text{ for all } T \in \mathcal{T}\}$. By asssumption, $\bigcup_{n\in\mathbb{N}} A_n = X$. 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 every $n \in \mathbb{N}$. Corollary 3.2 now yields $N \in \mathbb{N}$, $x_0 \in A_N$ and $\varepsilon > 0$ with $\overline{B}(x_0,\varepsilon) \subseteq A_N$. Let $z \in \overline{B}(0,\varepsilon)$. We have $x_0 - z, x_0 + z \in \overline{B}(x_0,\varepsilon) \subseteq A_N$, and hence

$$\|Tz\| = \|T(\tfrac{1}{2}(z+x_0)+\tfrac{1}{2}(z-x_0))\| \leq \tfrac{1}{2}\,\|T(z+x_0)\| + \tfrac{1}{2}\,\|T(x_0-z))\| \leq \tfrac{N}{2} + \tfrac{N}{2} = N.$$

Finally, let $x \in X$ with $\|x\| \leq 1$. Set $z = \varepsilon x \in \overline{B}(0,\varepsilon)$. It follows $N \geq \|Tz\| = \varepsilon\,\|Tx\|$; i.e., $\|T\| \leq \frac{N}{\varepsilon}$ for all $T \in \mathcal{T}$. □

Corollary 3.5 (Banach–Steinhaus)**.** *Let X be a Banach space, Y be a normed vector space, and T_n belong to $\mathcal{L}(X,Y)$ for every $n \in \mathbb{N}$. Assume that $(T_n x)_n$ converges in Y for each $x \in X$. Then $\sup_n \|T_n\| < \infty$.*

Proof. This result follows from the principle of uniform boundedness since $c_x := \sup_n \|T_n x\|$ is finite for each $x \in X$, by the assumption. □

Example 3.6. Let $X = c_{00}$ and $Y = c_0$ be endowed with the supremum norm. Set $T_n x = (x_1, 2x_2, \ldots, nx_n, 0, \ldots)$ for $n \in \mathbb{N}$ and $x \in X$. Then T_n belongs to $\mathcal{L}(X,Y)$ 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, but $(T_n x)_n$ 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 $T_n x = (x_1, 2x_2, \ldots, mx_m, 0, \ldots)$ for all $n \geq m$. So the completeness of X is needed in Theorem 3.4. ◊

Definition 3.7. *Let X and Y be normed vector spaces and $T_n, T \in \mathcal{L}(X,Y)$ for $n \in \mathbb{N}$. We say that T_n* converges strongly *to T as $n \to \infty$ if $T_n x \to Tx$ in Y as $n \to \infty$ for each $x \in X$. One then writes $T_n \xrightarrow{s} T$.*

Remark 3.8. Let X and Y be normed vector spaces and $T_n, T, S_n, S \in \mathcal{L}(X,Y)$ for $n \in \mathbb{N}$. Then the following assertions hold.

a) The strong limit is uniquely determined.

b) If T_n tends to T in operator norm as $n \to \infty$, then T_n converges to T strongly since $\|T_n x - Tx\| \leq \|T_n - T\|\,\|x\|$ for all $x \in X$. The converse is wrong in general. Consider e.g. the operators given by $P_n x = (x_1, \ldots, x_n, 0, \ldots)$ on $X = Y = \ell^2$ which converge strongly to I but $\|P_n - I\| = 1$ for all $n \in \mathbb{N}$.

c) If T_n tends strongly to T, S_n tends strongly to S and $\alpha, \beta \in \mathbb{F}$, then $\alpha T_n + \beta S_n$ converges strongly to $\alpha T + \beta S$ as $n \to \infty$. ◊

Lemma 3.9. *Let X be a normed vector space, Y be a Banach space, $S \subseteq X$ such that $D := \operatorname{lin} S$ is dense in X, and $T_n \in \mathcal{L}(X,Y)$ for all $n \in \mathbb{N}$. 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 $T \in \mathcal{L}(X,Y)$ such that T_n converges strongly to T and $\|T\| \leq \liminf_{n\to\infty} \|T_n\| =: M_0$. If $D = X$, these assertions also hold 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 $x \in D$. One checks as in the proof of Proposition 2.5 that T_0 is linear. Choose a subsequence with $M_0 = \lim_{j\to\infty} \|T_{n_j}\|$. We then obtain $\|T_0 x\| = \lim_{j\to\infty} \|T_{n_j} x\| \leq M_0\,\|x\|$ for

[1] Observe that uniform boundedness is equivalent to: $\exists\, c > 0 \;\forall T \in \mathcal{T},\, x \in \overline{B}(0,1) : \|Tx\| \leq c$.

each $x \in D$, hence $T_0 \in \mathcal{L}(D,Y)$ and $\|T_0\| \le M_0$. 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 gives an extension $T \in \mathcal{L}(X,Y)$ of T_0 with $\|T\| = \|T_0\| \le M_0 \le M$, since Y is a Banach space and D is dense. Let $\varepsilon > 0$ and $x \in X$. Choose $z \in D$ with $\|x - z\| \le \varepsilon$. We estimate

$$\begin{aligned}\limsup_{n\to\infty} \|Tx - T_n x\| &\le \limsup_{n\to\infty} (\|T(x-z)\| + \|Tz - T_n z\| + \|T_n(z-x)\|)\\ &\le 2M\varepsilon + \lim_{n\to\infty} \|T_0 z - T_n z\| = 2M\varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, T_n tends strongly to T. □

Example 3.10. 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 3.6 we see that $T_n \in \mathcal{L}(X)$ with $\|T_n\| = n \to \infty$. Due to Corollary 3.5, the operators T_n do not have a strong limit. 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 3.6. The uniform bound on $\|T_n\|$ is thus needed in Lemma 3.9.

b) Let $X = Y = c_0$ and $T_n x = x_n e_n$ for all $x \in c_0$ and $n \in \mathbb{N}$. Then $\|T_n x\|_\infty = |x_n|$ tend to $T := 0$ as $n \to \infty$ for each $x \in X$; i.e., $T_n \to 0$ strongly. However, $\|T_n e_n\|_\infty \ge 1$ and $\|T_n x\|_\infty \le \|x\|_\infty$ 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. ◊

Example 3.11 (Pointwise divergent Fourier series)**.** We endow $X = \{f \in C([-\pi,\pi]) \,|\, f(-\pi) = f(\pi)\}$ with the supremum norm and $f \in X$. By Analysis 3 there are $a_k, b_k \in \mathbb{F}$ 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 $n \to \infty$. Extend f to a 2π–periodic function on $\mathbb{R}$. It is straightforward to check that

$$S_n(f,t) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(s+t) D_n(s)\, ds \quad \text{with} \quad D_n(t) = \begin{cases} \frac{\sin(n+\frac{1}{2})t}{2\sin\frac{t}{2}}, & t \in [-\pi,\pi]\setminus\{0\},\\ n + \frac{1}{2}, & t = 0,\end{cases}$$

see Werner [Wer05, S.146]. Set $\varphi_n(f) = S_n(f,0)$. Then $\varphi_n \in X^\star$ for each $n \in \mathbb{N}$. We suppose that $\varphi_n(f)$ would converge for every $f \in X$. Corollary 3.5 would then imply $\sup_n \|\varphi_n\| < \infty$. However, the proof of Satz IV.2.10 in [Wer05] yields

$$\|\varphi_n\| = \frac{1}{\pi}\int_{-\pi}^{\pi} |D_n(t)|\, dt \;\longrightarrow \infty \qquad \text{as } \; n \to \infty.$$

Consequently, there must exist a function $f \in X$ whose Fourier series diverges at the point $t = 0$. (A more or less concrete example of a pointwise divergent Fourier series is given in Section 18 of [Koe89].) ◊

To apply the above lemma, one often needs dense subspaces of 'good' functions. The next proposition is one of the fundamental results in this direction. It holds in much greater generality, see Theorem 3.14 in [Rud87].

Proposition 3.12. *Let $U \subseteq \mathbb{R}^d$ be open and $p \in [1,\infty)$. Then the set $C_c(U) = \{f \in C(U) \,|\, \overline{\operatorname{supp}(f)} \subseteq U$ is bounded$\}$ is dense in $L^p(U)$. (More precisely, $C_c(U) + \mathcal{N}$ is dense in $L^p(U)$.)*

Proof. Let $f \in L^p(U)$, $p \in [1,\infty)$, and $\varepsilon > 0$. There is a simple function $g \in L^p(U)$ with $\|f-g\|_p \le \varepsilon$ due to Satz IV.7.6 in [Wer06]. We next employ a cutoff argument to obtain functions with a compact support in U. Put $U_n = B(0,n)$ if $U = \mathbb{R}^d$. Otherwise, we set

$$U_n = \{x \in U \cap B(0,n) \,|\, d(x, \partial U) < \tfrac{1}{n}\}$$

for all $n \in \mathbb{N}$. Observe that the functions $\mathbb{1}_{U_n} g$ converge pointwise to g as $n \to \infty$ and $|\mathbb{1}_{U_n} g| \le |g|$ for all $n \in \mathbb{N}$. Lebesgue's theorem then shows that $\mathbb{1}_{U_n} g$ tends to g in $L^p(U)$. We can thus fix an index $N \in \mathbb{N}$ so that the simple function

$$h := \mathbb{1}_{U_N} g = \sum_{j=1}^m a_j \mathbb{1}_{A_j}$$

satisfies $\|g-h\|_p \le \varepsilon$, for some $m \in \mathbb{N}$, $a_j \in \mathbb{F}$ and $A_j \in \mathcal{B}(U)$ with $\lambda(A_j) < \infty$ and $A_j \subseteq U_N$.

Since λ is 'regular' (see Satz IV.3.15 in [Wer06]) there are compact K_j and open O_j such that $K_j \subseteq A_j \subseteq O_j$ and $\lambda(O_j \setminus K_j) \le (\varepsilon |a_j|^{-1} m^{-1})^p$ for each j. Replacing O_j by $O_j \cap U_N$ if necessary, we may assume that $O_j \subseteq U_N$. By an exercise there is a function $\varphi_j \in C(U)$ such that $\varphi_j = 1$ on K_j, $\operatorname{supp} \varphi_j \subseteq \overline{U_N} \subseteq U_{N+1}$ and $0 \le \varphi_j \le 1$ for all $j = 1, \dots, m$. Hence, $\varphi := \sum_{j=1}^m a_j \varphi_j$ belongs to $C_c(U)$ and

$$\|\varphi - h\|_p \le \sum_{j=1}^m |a_j| \, \|\varphi_j - \mathbb{1}_{A_j}\|_p \le \sum_{j=1}^m |a_j| \, \lambda(O_j \setminus K_j)^{\frac{1}{p}} \le \varepsilon.$$

It follows that $\|f - \varphi\|_p \le 3\varepsilon$. □

Example 3.13 (Left translations)**.** Let $X \in L^p(\mathbb{R})$ for some $1 \le p < \infty$. 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,

$$\|T(t)f\|_p^p = \int_{\mathbb{R}} |f(s+t)|^p \, ds = \|f\|_p^p \,,$$

and $T(t)$ has the inverse $T(-t)$. Hence, $T(t) \in \mathcal{L}(X)$ is an isometric isomorphism. (This assertion can similarly be shown for $p = \infty$.) For any fixed $t > 0$ we have

$$T(t)\mathbb{1}_{[0,t]}(s) = \mathbb{1}_{[0,t]}(s+t) = \left\{ \begin{matrix} 1, & -t \le s \le 0 \\ 0, & \text{otherwise} \end{matrix} \right\} = \mathbb{1}_{[-t,0]}(s)$$

for all $s \in \mathbb{R}$. Setting $f = t^{-1/p}\mathbb{1}_{[0,t]}$ we thus obtain $\|f\|_p = 1$ and $\|T(t)f - f\|_p^p = t^{-1} \int_{-t}^t \mathbb{1}^p ds = 2$. Therefore the map $\mathbb{R} \ni t \mapsto T(t) \in \mathcal{L}(X)$ is **not** continuous with respect to the operator norm. However, it is *strongly continuous*, i.e., the maps $\mathbb{R} \ni t \mapsto T(t)f \in X$ are continuous for every $f \in X$. In fact, because $\|T(t)\| = 1$ for all $t \in \mathbb{R}$ and $C_c(\mathbb{R})$ is dense in X by Proposition 3.12, it suffices to consider $f \in C_c(\mathbb{R})$ due to Lemma 3.9. Let $t, t_0 \in \mathbb{R}$. Since $f \in C_c(\mathbb{R})$ is uniformly continuous, we derive

$$\|T(t)f - T(t_0)f\|_\infty = \sup_{s \in \mathbb{R}} |f(s+t) - f(s+t_0)| \longrightarrow 0 \quad \text{as} \;\; t \to t_0.$$

Note that there is a compact interval $J \subseteq \mathbb{R}$ such that $\operatorname{supp}(T(t)f - T(t_0)f) \subseteq J$ for $|t - t_0| \le 1$. We now obtain

$$\|T(t)f - T(t_0)f\|_p \le \lambda(J)^{1/p} \, \|T(t)f - T(t_0)f\|_\infty \longrightarrow 0 \qquad \text{as} \;\; t \to t_0.$$

These results also hold for $X = C_0(\mathbb{R})$ with slightly modified proofs. ◊

Mollifier

Let $U \subset \mathbb{R}^d$ be open. We define the space of *test functions* on U by $C_c^\infty(U) := C^\infty(\mathbb{R}^d) \cap C_c(U)$. Observe that the function

$$\varphi(x) = \begin{cases} e^{-\frac{1}{1-|x|_2^2}}, & |x|_2 < 1, \\ 0, & |x|_2 \ge 1, \end{cases}$$

belongs to $C_c^\infty(\mathbb{R}^d)$. Set $k = \|\varphi\|_1^{-1}\varphi$ and $k_\varepsilon(x) = \varepsilon^{-d}k(\frac{1}{\varepsilon}x)$ for all $x \in \mathbb{R}$ and $\varepsilon > 0$. Then $k_\varepsilon \in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp} k_\varepsilon = \overline{B}(0,\varepsilon)$, $k_\varepsilon \ge 0$ and

$$\|k_\varepsilon\|_1 = \int_{\mathbb{R}^d} \varepsilon^{-d}k(\tfrac{1}{\varepsilon}x)dx = \int_{\mathbb{R}^d} k(y)\,dy = 1$$

for all $\varepsilon > 0$, where we have used the transformation $y = \frac{1}{\varepsilon}x$. We write $f \in L^1_{\mathrm{loc}}(U)$ if $f|B \in L^1(B)$ for any bounded Borel set $B \subseteq U$. The extension of f by 0 is denoted by $\tilde{f}$. We define the *mollifier* G_ε by

$$(G_\varepsilon f)(x) = (k_\varepsilon * \tilde{f})(x) = \int_{\mathbb{R}^d} k_\varepsilon(x-y)\tilde{f}(y)\,dy = \int_{U\cap\bar{B}(x,\varepsilon)} k_\varepsilon(x-y)f(y)\,dy \quad (3.1)$$

for all $x \in U$. (The integral exists for all $x \in U$ since f is integrable on $U \cap \bar{B}(x,\varepsilon)$ and k_ε is bounded.) The next result indicates that the operators G_ε are a powerful tool to approximate functions by smooth ones. As the proof indicates, they are often used in combination with cutoff arguments. We improve the proposition considerably in Section 5.2.

Proposition 3.14. *Let $U \subseteq \mathbb{R}^d$ be open, $f \in L^1_{\mathrm{loc}}(U)$, $\varepsilon > 0$ and $1 \le p \le \infty$. Define G_ε by (3.1). The following assertions hold.*

a) $G_\varepsilon f \in C^\infty(U)$. 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 sufficiently small $\varepsilon > 0$.

b) $G_\varepsilon \in \mathcal{L}(L^p(U))$ with $\|G_\varepsilon\| \le 1$. Let $1 \le p < \infty$ and $f \in L^p(U)$. Then $G_\varepsilon f \to f$ in $L^p(U)$ as $\varepsilon \to 0$.

c) Let $1 \le p < \infty$. 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 $1 \le p,q < \infty$ then there are $f_n \in C_c^\infty(U)$ converging to f in $L^p(U)$ and in $L^q(U)$.

Proof. a) Fix $\varepsilon > 0$ and $f \in L^1_{\mathrm{loc}}(U)$. Let $x_0 \in U$. There is an $r > 0$ such that $B(x_0,r) \subseteq U$. Let $x \in B(x_0,r)$, $j \in \{1,\dots,d\}$, and $l = 0,1$. We then have

$$|\partial^l_{x_j} k_\varepsilon(x-y)f(y)| \le \|\partial^l_{x_j} k_\varepsilon\|_\infty\, \mathbb{1}_{B(x_0,r+\varepsilon)\cap U}(y)\,|f(y)| =: h(y)$$

for all $y \in U$. The function h is integrable. A corollary to the theorem of dominated convergence now yields that $k_\varepsilon * f$ is continuous and partially differentiable at x_0. Iterating this argument, one concludes that $G_\varepsilon f \in C^\infty(U)$.

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, $\varepsilon_0 := \frac{1}{2}\inf\{\operatorname{dist}(x,\partial U)\,|\,x \in K\} > 0$ and hence $S = K + \overline{B}(0,\varepsilon_0) \subseteq U$. So (3.1) yields $\operatorname{supp}(G_\varepsilon f) \subset S$ for $\varepsilon \in (0,\varepsilon_0]$. Note that S is compact. Indeed, let $x_n \in S$. Then $x_n = y_n + z_n$ for some $y_n \in K$ and $z_n \in \overline{B}(0,\varepsilon_0)$. Compactness implies that $y_{n_k} \to y$ in K as $k \to \infty$ and $z_{n_{k_j}} \to z$ in $\overline{B}(0,\varepsilon_0)$ as $j \to \infty$. Therefore, $x_{n_{k_j}}$ tends to $x = y + z \in S$.

b) Let $f \in L^p(U)$. Young's inequality (see Proposition 2.15) 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, $G_\varepsilon \in \mathcal{L}(L^p(U))$ with $\|G_\varepsilon\| \le 1$ for all $\varepsilon > 0$. Let $1 \le p < \infty$. To prove the convergence, we only have to consider $g \in C_c(U)$ due to Lemma 3.9 since $\|G_\varepsilon\| \le 1$ and $C_c(U)$ is dense in $L^p(U)$ by

Proposition 3.12. Let $g \in C_c(U)$ and $K = \operatorname{supp}(g)$. Again, $S := K + \overline{B}(0, \varepsilon_0) \subseteq U$ is compact. We then derive

$$\begin{aligned}
\sup_{x\in U} |G_\varepsilon g(x) - g(x)| &= \sup_{x\in U} \left| \int_{\mathbb{R}^d} k_\varepsilon(x-y)\tilde{g}(y)\,dy - \int_{\mathbb{R}^d} k_\varepsilon(x-y)\,dy\; g(x) \right| \\
&\le \sup_{x\in U} \int_{\bar{B}(x,\varepsilon)} k_\varepsilon(x-y)\,|\tilde{g}(y) - g(x)|\,dy \\
&\le \|k_\varepsilon\|_1 \sup_{|x-y|\le\varepsilon} |\tilde{g}(y) - \tilde{g}(x)| \;\longrightarrow 0 \qquad \text{as } \varepsilon \to 0,
\end{aligned}$$

using that $\tilde{g}$ is uniformly continuous and $\|k_\varepsilon\|_1 = 1$. Because of $\operatorname{supp} G_\varepsilon g \subseteq S$ for $\varepsilon \in (0, \varepsilon_0]$ and $\operatorname{supp} g \subseteq S$, the distance $\|G_\varepsilon g - g\|_p \le \lambda(S)^{1/p}\, \|G_\varepsilon g - g\|_\infty$ tends to 0 as $\varepsilon \to 0$.

c) Let $f \in L^p(U) \cap L^q(U)$ for some $1 \le p, q < \infty$. Put $U_n = B(0,n)$ if $U = \mathbb{R}^d$. Otherwise, we set $U_n = \{x \in U \cap B(0,n) \,|\, d(x, \partial U) < \frac{1}{n}\}$ for all $n \in \mathbb{N}$. Using assertion a), for each $n \in \mathbb{N}$ we can choose an $\varepsilon_n \in (0, 1/n]$ such that $f_n := G_{\varepsilon_n}(\mathbb{1}_{U_n} f) \in C_c^\infty(U)$. Since $\mathbb{1}_{U_n} f \to f$ pointwise and $|\mathbb{1}_{U_n} f| \le |f|$, Lebesgue's theorem yields $\mathbb{1}_{U_n} f \to f$ in $L^p(U)$ and in $L^q(U)$ as $n \to \infty$. 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 $r \in \{p, q\}$. □

3.2 The open mapping theorem and invertibility

The invertibility of $T \in \mathcal{L}(X,Y)$ 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 3.15. *Let X, Y, Z be normed vector spaces, $T \in \mathcal{L}(X,Y)$ and $S \in \mathcal{L}(Y,Z)$ be invertible. Then $ST \in \mathcal{L}(X,Z)$ is invertible with the inverse $T^{-1}S^{-1} \in \mathcal{L}(Z,X)$.*

Proof. The operators ST and $T^{-1}S^{-1}$ are continuous. Moreover, $T^{-1}S^{-1}ST = I_X$ and $STT^{-1}S^{-1} = I_Z$. □

Lemma 3.16. *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, (S_n) has a limit S, and $\|S\| = \lim_{n\to\infty} \|S_n\| \le s$. □

Proposition 3.17 (Neumann series)**.** *Let X be Banach space, Y be a normed vector space and $T, S \in \mathcal{L}(X,Y)$. Assume that T is invertible and that $\|S\| < \|T^{-1}\|^{-1}$. Then $S + T$ is invertible and*

$$(S+T)^{-1} = \sum_{n=0}^\infty (-T^{-1}S)^n T^{-1} =: R \qquad \textit{(convergence in } \mathcal{L}(Y,X)\textit{)},$$

$$\|(S+T)^{-1}\| \le \frac{\|T^{-1}\|}{1 - \|T^{-1}S\|}.$$

In particular, the set of invertible operators is open in $\mathcal{L}(X,Y)$.

Proof. We have $q := \|T^{-1}S\|_{\mathcal{L}(X)} < 1$ by assumption. Then $\sum_{n=0}^{\infty} \|(T^{-1}S)^n\| \le 1/(1-q)$. By Proposition 2.5, $\mathcal{L}(X)$ is a Banach space, and hence Lemma 3.16 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 $\mathcal{L}(Y,X)$ and that $\|R\| \le \|T^{-1}\|/(1-q)$. Moreover,

$$R(S+T) = \sum_{n=0}^{\infty} (-T^{-1}S)^n (T^{-1}S + I) = -\sum_{j=1}^{\infty} (-T^{-1}S)^j + \sum_{n=0}^{\infty} (-T^{-1}S)^n = I,$$

$$(S+T)R = \sum_{n=0}^{\infty} (ST^{-1} + I)(-ST^{-1})^n = I. \qquad \square$$

Example 3.18. Let $X = c_{00}$ with $\|\cdot\|_\infty$ and $Tx = (x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \ldots)$. Then $T \in \mathcal{L}(X)$, T is bijective with $T^{-1}y = (y_1, 2y_2, 3y_3, \ldots)$. But $T^{-1} : c_{00} \to c_{00}$ is not continuous by Example 2.1. ♢

Definition 3.19. *Let M and M' be metric spaces. A map $f : M \to M'$ is called* open *if $f(O)$ is open in M' for each open set $O \subseteq M$.*

Remark 3.20. a) A bijective map $f : M \to M'$ is open if and only if $f(O) = \{y \in M' \mid f^{-1}(y) \in O\} = (f^{-1})^{-1}(O)$ is open in M' for each open $O \subseteq M$ if and only if f^{-1} is continuous.

b) In Example 3.18 the maps $T, T^{-1} : c_{00} \to c_{00}$ are bijective, 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. ♢

Theorem 3.21 (Open mapping theorem)**.** *Let X and Y be Banach spaces and $T \in \mathcal{L}(X,Y)$ be surjective. Then T is open. If T is even bijective, then is T invertible.*

Proof. The second assertion is a consequence of the first one because of Remark 3.20. So let T be surjective. Set $U_r = B_X(0,r)$ and $V_r = B_Y(0,r)$ for every $r > 0$.

Claim A). There is an $\varepsilon > 0$ with $V_\varepsilon \subseteq \overline{TU_2}$.

Assume that Claim A) has been shown. Let $O \subseteq X$ be open, $x \in O$, and $y = Tx$. Then there is an $r > 0$ with $B_X(x,r) \subseteq O$. From A) and the linearity of T we deduce

$$B_Y(y, \tfrac{\varepsilon r}{2}) = y + \tfrac{r}{2} V_\varepsilon \subseteq Tx + \tfrac{r}{2} TU_2 = TB_X(x,r) \subseteq TO.$$

Hence, T is open.

Proof of A). The surjectivity of T yields $Y = \bigcup_{n=1}^{\infty} \overline{TU_n}$. Since Y is complete, Corollary 3.2 gives $N \in \mathbb{N}$, $y_0 \in Y$ and $r > 0$ such that

$$B_Y(y_0, r) \subseteq \overline{TU_N} = \overline{(2N)\,TU_{\frac{1}{2}}} = 2N\,\overline{TU_{\frac{1}{2}}}.$$

Hence, $B_Y(z,\varepsilon) \subseteq \overline{TU_{\frac{1}{2}}}$ where $z = (2N)^{-1}y_0$ and $\varepsilon = r/(2N)$. 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. Therefore

$$V_\varepsilon = B_Y(z,\varepsilon) - z \subseteq \overline{TU_{\frac{1}{2}}} - \overline{TU_{\frac{1}{2}}} = 2(\tfrac{1}{2}\overline{TU_{\frac{1}{2}}} + \tfrac{1}{2}\overline{TU_{\frac{1}{2}}}) \subseteq 2\,\overline{TU_{\frac{1}{2}}} = \overline{TU_1}.$$

For later use, we also note that the above inclusion yields

$$V_{\alpha\varepsilon} = \alpha V_\varepsilon \subseteq \alpha\,\overline{TU_1} = \overline{TU_\alpha} \qquad \text{for all } \alpha > 0. \qquad (*)$$

Now, Claim A) and thus the theorem follow from the next assertion.

Claim B). It holds $\overline{TU_1} \subseteq TU_2$.

Proof of B) Let $y \in \overline{TU_1}$. Then there is an $x_1 \in U_1$ 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 (∗). Similarly, we obtain

$$x_2 \in U_{1/2} \quad \text{with} \quad y - Tx_1 - Tx_2 = y - T(x_1 + x_2) \in V_{\frac{\varepsilon}{4}} \subseteq \overline{TU_{\frac{1}{4}}}\,.$$

Inductively, we find $x_n \in U_{2^{1-n}}$ such that

$$y - T(x_1 + \cdots + x_n) \in V_{\varepsilon 2^{-n}}\,. \tag{+}$$

Since X is a Banach space and $\sum_{n=1}^{\infty} \|x_n\| < 2$, by Lemma 3.16 there exists the vector $x := \sum_{n=1}^{\infty} x_n \in U_2$. Hence, $Tx \in TU_2$. Letting $n \to \infty$ in (+), we thus obtain

$$Tx = \sum_{n=1}^{\infty} Tx_n = y \ \in \ TU_2. \qquad \square$$

Corollary 3.22. *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 3.21, the map $I^{-1} : (X, \|\cdot\|) \to (X, |||\cdot|||)$ is also continuous. The assertion now follows from $I^{-1}x = x$. $\square$

Example 3.23. 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 (e.g., the functions $f_n(t) = \sin(nt)$ satisfy $\|f_n\|_\infty \le 1$, but $|||f_n||| \ge |f_n'(0)| = n$). So in Theorem 3.21 also Y must be complete. ♢

Corollary 3.24. *Let X and Y be Banach spaces and $T \in \mathcal{L}(X,Y)$ be injective. The following assertions are equivalent.*

a) The operator $T^{-1} : R(T) \to X$ is continuous.

b) There is a constant $c > 0$ such that $\|Tx\| \ge c\,\|x\|$ for every $x \in X$.

c) The range $R(T)$ is closed.

Proof. a) ⇒ b): For all $x \in X$ we have $\|x\| = \|T^{-1}Tx\| \le \|T^{-1}\|\,\|Tx\|$. The implication 'b)⇒c)' was shown in Remark 2.11; and 'c)⇒ a)' follows from Theorem 3.21 because $R(T)$ is complete by c). $\square$

Proposition 3.25. *Let X be a Banach space and $X = Y \oplus Z$. Then the following asserions hold.*

a) $X \cong Y \times Z$.

b) The projection P with $R(P) = Y$ and $N(P) = Z$ is continuous (cf. Remark 2.18).

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 $\|T(y,z)\| \le \|y\| + \|z\| = \|(y,z)\|$ and bijective (since $X = Y \oplus Z$). Theorem 3.21 now yields a).

b) Let $x = y + z \in Y \oplus Z$. Then $Px = y$. The operator T from a) satisfies $(y,z) = T^{-1}x$. Hence, part a) yields

$$\|Px\| = \|y\| \le \|y\| + \|z\| = \|T^{-1}x\|_{Y \times Z} \le \|T^{-1}\|\,\|x\|. \qquad \square$$

Chapter 4

Duality

Let X be a normed vector space. We study the 'duality pairing'

$$X \times X^\star \to \mathbb{F}, \quad (x, x^\star) \mapsto x^\star(x) =: \langle x, x^\star \rangle = \langle x, x^\star \rangle_X.$$

This map is linear in both components, and it is continuous since $|\langle x, x^\star \rangle| \leq \|x\|_X \|x^\star\|_{X^\star}$ (see an exercise). We will first determine the dual space $X^\star$ if X is a Hilbert space and if $X = L^p(\mu)$ for $p \in [1, \infty)$.

4.1 Hilbert spaces

Definition 4.1. *A* scalar product *on a vector space X is a map $(\cdot|\cdot) : X^2 \to \mathbb{F}$ with*

a) $(\alpha x_1 + \beta x_2 | y) = \alpha\,(x_1|y) + \beta\,(x_2|y)$,

b) $(x|y) = \overline{(y|x)}$,

c) $(x|x) \geq 0, \quad (x|x) = 0 \iff x = 0$,

for all $x_1, x_2, x, y \in X$ and $\alpha, \beta \in \mathbb{F}$. The map is called sesquilinear form *if a) and b) hold, and* positive definite *if c) holds. The pair $(X, (\cdot|\cdot))$ is called a* pre Hilbert space*. We set $\|x\| = \sqrt{(x|x)}$.*[1]

Remark 4.2. Let $(X, (\cdot|\cdot))$ be a pre Hilbert space.
(i) Properties a) and b) yield $(0|y) = 0$, $(x|0) = 0$, $(x|x) \in \mathbb{R}$, and

$$(x|\alpha y_1 + \beta y_2) = \overline{(\alpha y_1 + \beta y_2|x)} = \bar{\alpha}\,\overline{(y_1|x)} + \bar{\beta}\,\overline{(y_2|x)} = \bar{\alpha}\,(x|y_1) + \bar{\beta}\,(x|y_2)$$

for all $\alpha, \beta \in \mathbb{F}$ and $x, y_1, y_2 \in X$.
(ii) We have the *Cauchy–Schwarz inequality*:

$$|(x|y)| \leq \|x\|\,\|y\| \qquad \text{for all } x, y \in X. \tag{CS}$$

Equality holds if and only if x and y are linearly dependent. See Lineare Algebra or Proposition I.1.4 in [Con90].
(iii) $(X, \|\cdot\|)$ is a normed vector space, where $\|\cdot\|$ is given by Definition 4.1. In fact, let $x, y \in X$ and $\alpha \in \mathbb{F}$. Using (CS), we compute

$$\|x\| = 0 \iff (x|x) = 0 \iff x = 0,$$
$$\|\alpha x\| = \sqrt{\alpha\bar{\alpha}\,(x|x)} = |\alpha|\,\|x\|,$$

[1] Here one also uses the notions 'inner product' and 'inner product space'.

$$\begin{aligned}\|x+y\|^2 &= (x+y|x+y) = \|x\|^2 + (x|y) + (y|x) + \|y\|^2 \\ &= \|x\|^2 + 2\,\mathrm{Re}\,(x|y) + \|y\|^2 \le (\|x\| + \|y\|)^2 . \end{aligned} \tag{4.1}$$

(iv) The scalar product is Lipschitz on each ball of X^2 and thus continuous from X^2 to $\mathbb{F}$. In fact, (CS) yields

$$\begin{aligned}|(x_1|y_1) - (x_2|y_2)| &\le |(x_1 - x_2|y_1)| + |(x_2|y_1 - y_2)| \le r\,\|x_1 - x_2\| + r\,\|y_1 - y_2\| \\ &\le \sqrt{2}r\,\|(x_1, y_1) - (x_2, y_2)\|\end{aligned}$$

for all $x_k, y_k \in X$ with $\|(x_k, y_k)\| := (\|x_k\|^2 + \|y_k\|^2)^{\frac{1}{2}} \le r$.

(v) From (4.1) we deduce the *parallelogramm identity*

$$\begin{aligned}\|x+y\|^2 + \|x-y\|^2 &= \|x\|^2 + 2\,\mathrm{Re}\,(x|y) + \|y\|^2 + \|x\|^2 - 2\,\mathrm{Re}\,(x|y) + \|y\|^2 \\ &= 2\,\|x\|^2 + 2\,\|y\|^2 \qquad \text{for all } \ x, y \in X. \end{aligned} \tag{4.2}$$

Definition 4.3. *Let $(\cdot|\cdot)$ be a scalar product on X. Then the norm $\|\cdot\|$ given in Definition 4.1 is called* Hilbert norm. *If $\|\cdot\|$ is complete, then $(X, (\cdot|\cdot))$ is called* Hilbert space.

Example 4.4. a) $X = \mathbb{F}^d$ is a Hilbert space with

$$(x|y) = \sum_{k=1}^{d} x_k \overline{y}_k \qquad \text{and} \qquad \|x\|_2^2 = \sum_{k=1}^{d} |x_k|^2 .$$

Note that $(X, |\cdot|_p)$ is not induced by a scalar product if $d \ge 2$ and $p \ne 2$, since $|e_1 + e_2|_p^2 + |e_1 - e_2|_p^2 = 2^{2/p} + 2^{2/p} \ne 2(|e_1|_p^2 + |e_2|_p^2) = 4$ contradicting (4.2).

b) $X = \ell^2$ is a Hilbert space with

$$(x|y) = \sum_{k=1}^{\infty} x_k \overline{y}_k \qquad \text{and} \qquad \|x\|_2^2 = \sum_{k=1}^{\infty} |x_k|^2 .$$

The first series converges absolutely due to Hölder's inequality with $p = p' = 2$.

c) Let $(S, \mathcal{A}, \mu)$ be a measure space. Then $X = L^2(\mu)$ is a Hilbert space with

$$(f|g) = \int_S f\overline{g}\,d\mu \qquad \text{and} \qquad \|f\|_2^2 = \int_S |f|^2\,d\mu.$$

The product $f\overline{g}$ is integrable by Hölder's inequality with $p = p' = 2$.

d) A linear subspace of a pre Hilbert space is a pre Hilbert space with the restricted scalar product. ◊

In pre Hilbert spaces one can also define angles, and not only distances as in normed vector spaces. We restrict ourselves to the angle $\pi/2$.

Definition 4.5. *Two elements x and y of a pre Hilbert space X are called* orthogonal *if $(x|y) = 0$. Two subsets $A, B \subseteq X$ are called* orthogonal *if $a \perp b$ for all $a \in A$ and $b \in B$. One then writes $x \perp y$ and $A \perp B$, and also $x \perp A$ instead of $\{x\} \perp A$. The* orthogonal complement *of A is given by*

$$A^\perp = \{x \in X \,|\, x \perp a \ \textit{ for every } \ a \in A\}.$$

A projection $P \in L(X)$ is called orthogonal *if $R(P) \perp N(P)$.*

Example 4.6. a) Functions $f, g \in L^2(\mu)$ with disjoint support are orthogonal since then $(f|g) = \int f\overline{g}\,d\mu = 0$.

b) The functions $f_n(t) = e^{int}$ for $n \neq m$ in $\mathbb{Z}$ are orthogonal in $L^2(0, 2\pi)$ since

$$(f_n|f_m) = \int_0^{2\pi} e^{int} e^{-imt}\, dt = \frac{1}{i(n-m)}\, e^{i(n-m)t}\Big|_0^{2\pi} = 0.$$

c) Let $f \in L^2(\mathbb{R})$ be even and $g \in L^2(\mathbb{R})$ be odd. Then $f \perp g$, since $f\overline{g}$ is odd and hence $\int f\overline{g}\, dt = 0$. ♢

Remark 4.7. Let X be a pre Hilbert space, $A, B \subseteq X$, and $x, y \in X$. It holds:

a) $x \perp x$ holds if and only if $x = 0$; and thus $X^\perp = \{0\}$. $x \perp y$ holds if and only if $y \perp x$. $\{0\}^\perp = X$.

b) If $x \perp y$, then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ due to (4.1). (**Pythagoras**)

c) $A \cap A^\perp \subseteq \{0\}$ and $A \subseteq (A^\perp)^\perp =: A^{\perp\perp}$, because of a).

d) $A^\perp$ is a closed linear subspace. In fact, it is easy to check that it is a linear subspace. Take $x_n \perp A$ with $x_n \to x$. Then $(x|a) = \lim_{n\to\infty}(x_n|a) = 0$ for all $a \in A$, so that $x \perp A$ 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 $(x|y) = (z|y)$ for some $x, z \in X$ and all y from a dense subset $D \subseteq X$, then $x = z$. In fact, by approximation we obtain $x - z \perp X$, and thus $x = z$ by a). ♢

Many of the special properties of Hilbert spaces rely on the next theorem.

Theorem 4.8 (Projection theorem)**.** *Let X be a Hilbert space and $Y \subseteq X$ be a closed linear subspace. It then holds $X = Y \oplus Y^\perp$, and there is a unique orthogonal projection $P \in \mathcal{L}(X)$ with $R(P) = Y$, $N(P) = Y^\perp$, and $\|P\| = 1$ (if $Y \neq \{0\}$). We further have $Y^{\perp\perp} = Y$, $X/Y \simeq Y^\perp$, and*

$$\|x - Px\| = \inf_{y\in Y} \|x - y\| \qquad \textit{for every } \ x \in X.$$

Proof. Due to Remark 4.7, $Y^\perp$ is a closed linear subspace of X and $Y \cap Y^\perp = \{0\}$. Let $x \in X$. To construct P, we look for $y_x \in Y$ with $\|x - y_x\| = \inf_{y\in Y}\|x - y\| =: \delta$. There are $y_n \in Y$ with $\|x - y_n\| \to \delta$ as $n \to \infty$. From (4.2) and $\frac{1}{2}(y_n + y_m) \in Y$ we deduce

$$\begin{aligned} 0 &\le \|\tfrac{1}{2}(y_n - y_m)\|^2 = \|\tfrac{1}{2}(y_n - x) - \tfrac{1}{2}(y_m - x)\|^2 \\ &= \tfrac{1}{2}\|y_n - x\|^2 + \tfrac{1}{2}\|y_m - x\|^2 - \|\tfrac{1}{2}(y_n + y_m) - x\|^2 \\ &\le \tfrac{1}{2}\|y_n - x\|^2 + \tfrac{1}{2}\|y_m - x\|^2 - \delta^2 \longrightarrow 0 \qquad \text{as } n, m \to \infty. \end{aligned}$$

Since X is complete, it exists $y_x = \lim_{n\to\infty} y_n$ in Y and we obtain $\|y_x - x\| = \lim_{n\to\infty}\|y_n - x\| = \delta$.

Set $z_x := x - y_x$. We want to show that $z_x \in Y^\perp$. Take $w \in Y \setminus \{0\}$ and put $\alpha = \|w\|^{-2}(w|z_x)$. We then obtain $\alpha w + y_x \in Y$ and thus

$$\begin{aligned} \delta^2 &\le \|x - (\alpha w + y_x)\|^2 = \|z_x - \alpha w\|^2 \\ &= \|z_x\|^2 - 2\operatorname{Re}\alpha(z_x|w) + |\alpha|^2\|w\|^2 = \delta^2 - \frac{|(z_x|w)|^2}{\|w\|^2}, \end{aligned}$$

using (4.1) and the definition of z_x and α. Therefore, $(z_x|w) = 0$ and thus $z_x \perp Y$. Taking into account $x = y_x + z_x$, we have shown that $Y + Y^\perp = X$, and hence $X = Y \oplus Y^\perp$. Remark 2.18 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., $P \in \mathcal{L}(X)$ and $\|P\| \le 1$. Lemma 2.17 then yields $\|P\| = 1$ if $Y \neq \{0\}$.

It holds $Y \subseteq Y^{\perp\perp}$ due to Remark 4.7. If $x \perp Y^\perp$, then $0 = (x|z_x) = (y_x + z_x|z_x) = \|z_x\|^2$ so that $x = y_x \in Y$; i.e., $Y = Y^{\perp\perp}$. Example 2.21 further implies that $Y^\perp = X/Y$. □

Example 4.9. a) Let $X = \mathbb{F}^2$ and $Y = \mathbb{F} \times \{0\}$. Then $Y^\perp = \{0\} \times \mathbb{F}$ and $P = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$.

b) Let $X = L^2(\mathbb{R})$ and $Y = \{f \in L^2(\mathbb{R}) \,|\, f = 0 \text{ a.e. on } \mathbb{R}_-\} \cong L^2(\mathbb{R}_+)$. Then $Y^\perp = \{f \in L^2(\mathbb{R}) \,|\, f = 0 \text{ a.e. on } \mathbb{R}_+\} \cong L^2(\mathbb{R}_-)$ with $Pf = \mathbb{1}_{\mathbb{R}_+} f$. Compare Example 2.19 for a) and b).

c) Let $X = L^2(\mathbb{R})$ and $Y = \{f \in L^2(\mathbb{R}) \,|\, f(t) = f(-t) \text{ for a.e. } t \in \mathbb{R}\}$ be the subspace of even functions. As in Example 1.33 one sees that the linear subspace Y is closed in X. Set $Pf(t) = \frac{1}{2}(f(t) + f(-t))$ for a.e. $t \in \mathbb{R}$ and $f \in X$. Then $Pf \in Y$, $P^2 = P$ and $\|Pf\|_2 \le \frac{2}{2}\|f\|_2$ for $f \in X$. As a result, $P \in \mathcal{L}(X)$ with $\|P\| = 1$, $R(P) = Y$, and $N(P) = \{f \in L^2(\mathbb{R}) \,|\, f(t) = -f(-t) \text{ for a.e. } t \in \mathbb{R}\}$ is the space of odd functions in X. The projection P is orthogonal by Example 4.6.

d) Let A be a subset of a Hilbert space X. Set $Y = \overline{\operatorname{lin} A}$. Remark 4.7 and the projection theorem then yield that $A^{\perp\perp} = Y^{\perp\perp} = Y$. ◊

Let X be a Hilbert space. For every $y \in X$ we define $\Phi(y) : X \to \mathbb{F}$ by setting

$$[\Phi(y)](x) := (x|y) \qquad \text{for all } x \in X. \tag{4.3}$$

The map $\Phi(y)$ is linear, and it holds $|[\Phi(y)](x)| \le \|x\|\,\|y\|$ by (CS). Hence, $\Phi(y) \in X^\star$ with $\|\Phi(y)\|_{X^\star} \le \|y\|_X$.

Theorem 4.10 (Riesz)**.** *Let X be a Hilbert space and define $\Phi_X = \Phi : X \to X^\star$ by (4.3). Then Φ is bijective, isometric and antilinear.*[2] *With $y := \Phi^{-1}(x^\star)$, we thus obtain for each $x^\star \in X^\star$ exactly one $y \in X$ such that $\langle x, x^\star\rangle = (x|y)$ for all $x \in X$, where $\|y\|_X = \|x^\star\|_{X^\star}$.*

Proof. Equation (4.3) and Remark 4.2 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

$$\|\Phi(y)\|_{X^\star} \ge |[\Phi(y)](x)| = \frac{1}{\|y\|}\,(y|y) = \|y\|$$

so that $\|\Phi(y)\|_{X^\star} = \|y\|_X$. This means that Φ is isometric and hence injective. Let $\varphi \in X^\star \setminus \{0\}$. Then $Z := N(\varphi) \ne X$ is a closed linear subspace of X. Theorem 4.8 yields that $Z^\perp \ne \{0\}$. 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, $x - \varphi(x) y_1 \in Z = Z^{\perp\perp}$ 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 claims are clear. □

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.

4.2 The duals of sequence and Lebesgue spaces

Let $X_p = \ell^p$ for $1 \le p < \infty$ and $X_\infty = c_0$ for $p = \infty$, as well as $p' \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{p'} = 1$. For each $y \in \ell^{p'}$, we define

$$\Phi_p(y) : \; X_p \to \mathbb{F}, \quad (\Phi_p(y))(x) = \sum_{k=1}^{\infty} x_k y_k.$$

[2]This means that $\Phi(\alpha y + \beta z) = \bar{\alpha}\Phi(y) + \bar{\beta}\Phi(z)$ for all $\alpha, \beta \in \mathbb{F}$ and $y, z \in X$.

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, we obtain $\Phi_p(y) \in X_p^\star$ with

$$\|\Phi_p(y)\|_{X_p^\star} \le \|y\|_{p'} \, .$$

As a result, the map

$$\Phi_p : \ell^{p'} \to X_p^\star, \quad \langle x, \Phi_p(y)\rangle = \sum_{k=1}^{\infty} x_k y_k \qquad (\forall\, x \in X_p,\ y \in \ell^{p'}) \tag{4.4}$$

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 4.11. *Equation (4.4) defines an isometrical isomorphism $\Phi_p : \ell^{p'} \to X_p^\star$ for all $p \in [1,\infty]$. We thus obtain for each $x^\star \in X_p^\star$ exactly one $y \in \ell^{p'}$ 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,*

$$c_0^\star \simeq \ell^1, \quad (\ell^1)^\star \simeq \ell^\infty, \quad \textit{and} \quad (\ell^p)^\star = \ell^{p'} \;\; \textit{for} \;\; 1 < p < \infty.$$

Proof. In view of the above considerations it remains to show the surjectivity of Φ_p and that Φ_p^{-1} is contractive since then $\|y\|_{p'} = \|\Phi_p^{-1}\Phi_p y\|_{p'} \le \|\Phi_p y\|_{X_p^\star} \le \|y\|_{p'}$ 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 4.14 and the exercises.)

Let $x^\star \in (\ell^1)^\star$. We define $y_k := x^\star(e_k)$ for every $k \in \mathbb{N}$ and $y = (y_k)$. Since $|y_k| \le \|x^\star\| \, \|e_k\|_1 = \|x^\star\|$ for all $k \in \mathbb{N}$, we obtain $y \in \ell^\infty$ and $\|y\|_\infty \le \|x^\star\|$. Equation (4.4) 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, we conclude that $\Phi_1(y) = x^\star$ and so Φ_1 is bijective. It also follows that Φ_1^{-1} is contractive. □

We will see in Example 4.25 that the dual of $\ell^i nfty$ is not isomorphic to ℓ^1. Next, let $(S, \mathcal{A}, \mu)$ be a σ–finite measure space, $1 \le p < \infty$, and $X = L^p(\mu)$. For each $g \in L^{p'}(\mu)$ we define the map

$$\Phi_p(g) : L^p(\mu) \to \mathbb{F}, \quad [\Phi_p(g)](f) = \int_S fg \, d\mu. \tag{4.5}$$

Hölder's inequality (cf. Example 2.8) yields that

$$\Phi_p(g) \in L^p(\mu)^\star \quad \text{ and } \quad \|\Phi_p(g)\|_{X^\star} \le \|g\|_{p'}.$$

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 some preparations.

A map $\nu : \mathcal{A} \to \mathbb{F}$ is called an $\mathbb{F}$-*valued measure*[3] 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 $A_n \in \mathcal{A}$, $n \in \mathbb{N}$. 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 1.2B) 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 $\nu(A) = 0$ for all $A \in \mathcal{A}$ with $\mu(A) = 0$. One then writes $\nu \ll \mu$.

[3]One usually says *signed measure* instead of '$\mathbb{R}$–valued measure'.

Example 4.12. Let μ be a positive measure on $\mathcal{A}$ and $\rho \in L^1(\mu)$. We set $\nu(A) = \int_S \mathbb{1}_A \rho \, d\mu = \int_A \rho \, d\mu \in \mathbb{F}$ for all $A \in \mathcal{A}$. Then ν is an $\mathbb{F}$-valued measure on $\mathcal{A}$ with $\nu \ll \mu$. (In this case one writes $d\nu = \rho \, d\mu$ and calls ρ the *density* of ν with respect to μ.) If also $\rho \geq 0$, then ν is positive measure. For a measurable f with $f \geq 0$ or $\rho f \in L^1(\mu)$, we have $f \in L^1(\nu)$ and $\int f \, d\nu = \int f \rho \, d\mu$ (provided that $\rho \geq 0$).

Proof. Let $A_j \in \mathcal{A}$, $j \in \mathbb{N}$, be pairwise disjoint. Set $A = \bigcup_{j=1}^\infty A_j \in \mathcal{A}$ and $B_n = \bigcup_{j=1}^n A_j \in \mathcal{A}$ for $n \in \mathbb{N}$. Then $\mathbb{1}_{B_n} \rho$ tends to $\mathbb{1}_A \rho$ pointwise and $|\mathbb{1}_{B_n} \rho| \leq |\rho| \in L^1(\mu)$. Dominated convergence thus yields

$$\nu(A) = \int_S \mathbb{1}_A \rho \, d\mu = \lim_{n\to\infty} \int_S \mathbb{1}_{B_n} \rho \, d\mu = \lim_{n\to\infty} \sum_{j=1}^n \int_S \mathbb{1}_{A_j} \rho \, d\mu = \sum_{j=1}^\infty \nu(A_j).$$

It is clear that $\nu \ll \mu$ and that ν is positive if $\rho \geq 0$. The other assertions can easily be checked for simple functions and then follow by approximation. □

There is a converse to the above example which is one of the fundamental results of measure theory. We give a proof based on Riesz' Theorem 4.10, where we restrict ourselves to the main special case.

Theorem 4.13 (Radon–Nikodym)**.** *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 $w \in L^1(\mu)$ such that $d\nu = w \, d\mu$. If ν is a positive measure, then $w \geq 0$.*

Proof. We show the theorem if μ and ν are positive and bounded. For the general case we refer to Theorem 6.10 in [Rud87] or Appendix C in [Con90].

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 $m \in \mathbb{N}$, $a_j \in \mathbb{F}$ and $A_j \in \mathcal{A}$ (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. Hence, $\|f\|_{L^1(\mu)} \leq \mu(S)^{1/2} \, \|f\|_{L^2(\tau)}$ for all simple functions $f \in L^2(\tau)$. By approximation, each $f \in L^2(\tau)$ belongs to $L^1(\mu)$ and $\|f\|_{L^1(\mu)} \leq \mu(S)^{1/2} \, \|f\|_{L^2(\tau)}$. As a result, by

$$\varphi : L^2(\tau) \to \mathbb{F}, \quad \varphi(f) = \int_S f \, d\mu,$$

we can define a linear map which is continuous since $|\varphi(f)| \leq \|f\|_{L^1(\mu)} \leq \mu(S)^{1/2} \, \|f\|_{L^2(\tau)}$ for all $f \in L^2(\tau)$.

2) Theorem 4.10 now yields a function $\overline{g} \in L^2(\tau) \hookrightarrow L^1(\tau)$ such that

$$0 \leq \mu(A) = \int_S \mathbb{1}_A \, d\mu = \varphi(\mathbb{1}_A) = \int_S \mathbb{1}_A \overline{\overline{g}} \, d\tau = \int_A g \, d\tau$$

for all $A \in \mathcal{A}$; i.e, $d\mu = g \, d\tau$. Set $A_n = \{g \leq -\frac{1}{n}\} := \{s \in S \,|\, g(s) \leq -\frac{1}{n}\} \in \mathcal{A}$ for $n \in \mathbb{N}$. The above inequality implies that

$$0 \leq \int_{A_n} g \, d\tau \leq -\frac{\tau(A_n)}{n},$$

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. Hence, $g \geq 0$ a.e. In the same way, one obtains a function $0 \leq h \in L^1(\tau)$ with $d\nu = h \, d\tau$.

3) Set $N = \{g = 0\}$. Then $\mu(N) = \int_N g\,d\tau = 0$. From the assumption $\nu \ll \mu$ we infer that $\nu(N) = 0$. We now define the measurable function

$$0 \le w(s) = \begin{cases} \frac{h(s)}{g(s)}, & s \in S \setminus N, \\ 0, & s \in N. \end{cases}$$

For every $A \in \mathcal{A}$, we then compute

$$\nu(A) = \nu(A \cap N^c) = \int_{A\cap N^c} h\,d\tau = \int_{A\cap N^c} wg\,d\tau = \int_{A\cap N^c} w\,d\mu = \int_A w\,d\mu,$$

using Example 4.12. This means that $d\nu = w\,d\mu$. Moreover, $\|w\|_1 = \int_S w\,d\mu = \nu(S)$ is finite. To show uniqueness, take $\widetilde{w} \in L^1(\mu)$ with $\nu(A) = \int_A \widetilde{w}\,d\mu$ for all $A \in \mathcal{A}$. As in 2), one sees that $\widetilde{w} \ge 0$. Set $B_n = \{\widetilde{w} \ge w + \frac{1}{n}\}$ for $n \in \mathbb{N}$. Since

$$0 = \nu(B_n) - \nu(B_n) = \int_{B_n} (w - \widetilde{w})\,d\mu \le -\frac{\mu(B_n)}{n},$$

we deduce $\widetilde{w} \le w$ μ–a.e. as in 2). By symmetry, one als has $w \le \widetilde{w}$ μ–a.e., and so $w = \widetilde{w}$ in $L^1(\mu)$. □

Theorem 4.14. *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 (4.5) is an isometric isomorphism, and thus $L^p(\mu)^\star \cong L^{p'}(\mu)$ via*

$$\forall\, \varphi \in L^p(\mu)^\star \quad \exists!\, g \in L^{p'}(\mu) \quad \forall\, f \in L^p(\mu) : \; \langle f, \varphi \rangle = \int_S fg\,d\mu.$$

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 $1 < p < \infty$. Then $p' = \frac{p}{p-1}$. (The general case is treated in Appendix B of [Con90], see also Theorem II.2.4 of [Wer05] for σ–finite measure spaces.)

1) Let $\varphi \in L^p(\mu)^\star$. Since $\mu(S) < \infty$, we have $\mathbb{1}_A \in L^p(\mu)$ for all $A \in \mathcal{A}$. Define $\nu(A) = \varphi(\mathbb{1}_A)$. Take $A_j \in \mathcal{A}$ with $A_j \cap A_k = \emptyset$ for all $j, k \in \mathbb{N}$ with $j \ne k$. Set $A = \bigcup_{j=1}^\infty A_j \in \mathcal{A}$ 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. Moreover, $\mathbb{1}_{B_n} = \mathbb{1}_{A_1} + \cdots + \mathbb{1}_{A_n}$ for all $n \in \mathbb{N}$. Using the continuity and linearity of φ, we then conclude that

$$\nu(A) = \varphi(\lim_{n\to\infty} \mathbb{1}_{B_n}) = \lim_{n\to\infty} \sum_{k=1}^n \varphi(\mathbb{1}_{A_k}) = \sum_{k=1}^\infty \nu(A_k).$$

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 thus gives a function $g \in L^1(\mu)$ with

$$\varphi(\mathbb{1}_A) = \nu(A) = \int_S \mathbb{1}_A\, g\,d\mu \qquad \text{for all} \quad A \in \mathcal{A}.$$

The linearity of φ yields

$$\varphi(f) = \int_S fg\,dx \qquad \text{for every simple function} \quad f : S \to \mathbb{F}. \tag{$*$}$$

Since we only know that $g \in L^1(\mu)$, at first we can extend this equation to $f \in L^\infty(A)$ only. By Analysis 3 (see also Satz IV.4.6 in [Wer06]) there are simple functions $f_n : S \to \mathbb{F}$ converging to f uniformly as $n \to \infty$. Since $\mu(S) < \infty$, Hölder's inequality (Proposition 1.35) implies that $f_n \to f$ in $L^p(\mu)$, and hence $\varphi(f_n) \to \varphi(f)$ as $n \to \infty$. Moreover,

$$\Big| \int_S (f - f_n) g \, d\mu \Big| \le \|f - f_n\|_\infty \, \|g\|_1 \;\longrightarrow\; 0$$

as $n \to \infty$. As a result, $(*)$ holds for all $f \in L^\infty(A)$.

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 $n \in \mathbb{N}$. Then the functions $\mathbb{1}_{A_n} g$ 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 $(*)$, we can now compute

$$\begin{aligned} \int_S \mathbb{1}_{A_n} |g|^{p'} \, d\mu &= \int_S \mathbb{1}_{A_n} hg \, d\mu = \varphi(\mathbb{1}_{A_n} h) \le \|\varphi\|_{X^\star} \, \|\mathbb{1}_{A_n} h\|_p \\ &= \|\varphi\|_{X^\star} \Big(\int_S \mathbb{1}_{A_n} |h|^p \, d\mu \Big)^{\frac{1}{p}} = \|\varphi\|_{X^\star} \Big(\int_S \mathbb{1}_{A_n} |g|^{p'} \, d\mu \Big)^{\frac{1}{p}}, \end{aligned}$$

$$\Big(\int_S \mathbb{1}_{A_n} \, |g|^{p'} \, d\mu \Big)^{\frac{1}{p'}} \le \|\varphi\|_{X^\star}.$$

Letting $n \to \infty$, we arrive at $g \in L^{p'}(\mu)$ and $\|g\|_{p'} \le \|\varphi\|_{X^\star}$.

3) Let $f \in L^p(\mu)$. There are simple functions f_n converging to f in $L^p(\mu)$. Hence, $\varphi(f_n) \to \varphi(f)$ and

$$\Big| \int_S (f - f_n) g \, d\mu \Big| \le \|f - f_n\|_p \, \|g\|_{p'} \;\longrightarrow\; 0$$

as $n \to \infty$. We deduce that $(*)$ holds for all $f \in L^p(\mu)$, i.e., $\Phi_p(g) = \varphi$ and Φ_p is surjective. Finally, $\|g\|_{p'} \le \|\varphi\|_{X^\star} = \|\Phi_p(g)\|_{X^\star}$ by 2). □

Remark. Usually one identifies $L^{p'}(\mu)$ with $L^p(\mu)^\star$ for $1 \le p < \infty$ and writes $\langle f, g \rangle = \int fg \, d\mu$ for the duality, and analogously for the sequence spaces.

4.3 The theorem of Hahn-Banach

Many of the non–trivial properties of duality rely on the Hahn-Banach theorem proved below. We start with a more algebraic version. 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 $\lambda \in \mathbb{R}_+$. A typical example is a seminorm.

Theorem 4.15 (Order theoretic Hahn-Banach)**.** *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 $\varphi : X \to \mathbb{R}$ with $\varphi(y) = \varphi_0(y)$ for all $y \in Y$ and $\varphi(x) \le p(x)$ for all $x \in X$.*

Proof. 1) We define

$$\mathcal{M} = \{(Z,\psi)\,|\, Z \text{ is a linear subspace of } X,\ Y \subseteq Z,\ \psi : Z \to \mathbb{R} \text{ linear with } \psi|_Y = \varphi_0,\ \psi \le p|_Z\},$$

which contains (Y, φ_0). On $\mathcal{M}$ we set $(Z,\psi) \le (Z',\psi')$ if $Z \subseteq Z'$ and $\psi'|_Z = \psi$. Straightforward calculations show that this defines a partial order on $\mathcal{M}$. Let $\mathcal{K}$ be a totally ordered non–empty subset of $\mathcal{M}$; i.e., if $(Z,\psi),(Z',\psi') \in \mathcal{K}$, then $(Z,\psi) \le (Z',\psi')$ or $(Z',\psi') \le (Z,\psi)$. 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 also $x \in \hat{Z}$ for some $(\hat{Z},\hat{\psi}) \in \mathcal{K}$, then e.g. $(\hat{Z},\hat{\psi}) \le (Z,\psi)$ and so $\hat{\psi}(x) = \psi(x)$. Hence, $f : U \to \mathbb{R}$ is well defined. Take further $y \in U$ and $\alpha, \beta \in \mathbb{R}$. Choose a pair $(Z',\psi') \in \mathcal{K}$ with $y \in U'$ and $(Z,\psi) \le (Z',\psi')$, say. The asserted linearity follows from $\alpha x + \beta y \in Z'$ and

$$f(\alpha x + \beta y) = \psi'(\alpha x + \beta y) = \alpha\psi'(x) + \beta\psi'(y) = \alpha f(x) + \beta f(y).$$

Similarly one checks that $f|_Y = \varphi_0$ and $f \le p|_U$ so that $(U, f) \in \mathcal{M}$. By construction, $(Z,\psi) \le (U, f)$ for all $(Z,\psi) \in \mathcal{K}$.

Zorn's Lemma (see Theorem 1.2.7 of [DuS57]) 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) Assume that $V \ne X$ and fix some $x_0 \in X \setminus V$. We define 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 $v \in V$ and $t \in \mathbb{R}$ with $x = v + tx_0$. We will construct a linear map $\tilde{\varphi} : \tilde{V} \to \mathbb{R}$ such that $(V,\varphi) \lneqq (\tilde{V},\tilde{\varphi}) \in \mathcal{M}$. This fact contradicts the maximality of (V,φ), so that $V = X$ and the theorem will be established. Let $v, w \in V$. We then obtain

$$\varphi(v) + \varphi(w) = \varphi(v + w) \le p(v + w) \le p(v + x_0) + p(w - x_0),$$

and thus $\varphi(w) - p(w - x_0) \le p(v + x_0) - \varphi(v)$ for all $v, w \in V$. Consequently, there exists a number

$$\alpha \in \left[\sup_{w\in V}(\varphi(w) - p(w - x_0)),\ \inf_{v\in V}(p(v + x_0) - \varphi(v))\right].$$

We now define $\tilde{\varphi}(x) = \varphi(v) + \alpha t$ for every $x = v + tx_0 \in \tilde{V}$. Then $\tilde{\varphi} : \tilde{V} \to \mathbb{R}$ is linear. For $y \in Y \subseteq V$, we have $\tilde{\varphi}(y) = \varphi(y) = \varphi_0(y)$. For $x = v \in V$ (i.e., $t = 0$), it holds $\tilde{\varphi}(x) = \varphi(v) \le p(v) = p(x)$. Using the definition of α and the sublinearity of p, we obtain for $t > 0$ that

$$\tilde{\varphi}(x) \le \varphi(v) + t(p(\tfrac{1}{t}v + x_0) - \varphi(\tfrac{1}{t}v)) = p(v + tx_0) = p(x).$$

and for $t < 0$ that

$$\tilde{\varphi}(x) \le \varphi(v) + t(\varphi(-\tfrac{1}{t}v) - p(-\tfrac{1}{t}v - x_0)) = p(v + tx_0) = p(x).$$

Hence, $\tilde{\varphi} \le p|_{\tilde{V}}$. Summing up, we have shown that $(V,\varphi) \lneqq (\tilde{V},\tilde{\varphi}) \in \mathcal{M}$. □

Lemma 4.16. *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 $\xi^\star : X \to \mathbb{R}$ 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 $x^\star(x) = \xi^\star(x) - i\xi^\star(ix)$ for all $x \in X$. Then $x^\star \in X^\star$ with $\|x^\star\| = \|\xi^\star\|$ and $\operatorname{Re} x^\star = \xi^\star$.

Proof. The $\mathbb{R}$–linearity of $\xi^\star$ in a) and the $\mathbb{C}$–linearity of $x^\star$ in b) can be checked in a straightforward way. Clearly, $\xi^\star = \operatorname{Re} x^\star$ in b), and hence $\|\xi^\star\| \le \|x^\star\|$ in a) and b). Take $x \in X$ with $\|x\| = 1$. Set $\alpha = 1$ if $x^\star(x) = 0$ and $\alpha = x^\star(x)/|x^\star(x)|$ otherwise. Then $\|\alpha^{-1}x\| \le 1$ and so 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 $\|x\| \le 1$, we obtain $\|x^\star\| \le \|\xi^\star\|$ in a) and b). □

Due to their importance, the Hahn–Banach theorem, the principle of uniform boundedness and the open mapping theorem are called the three basic principles of functional analysis.

Theorem 4.17 (Main version of Hahn-Banach)**.** *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 an $x^\star \in X^\star$ 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 4.15 yields a linear $x^\star : X \to \mathbb{R}$ 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, $\|x^\star\| \le \|y^\star\|$ and $x^\star \in X^\star$. 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 4.16a) 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 $\eta^\star$ has an extension $\xi^\star \in (X_\mathbb{R})^\star$ with $\|\xi^\star\| = \|\eta^\star\| = \|y^\star\|$. Lemma 4.16b) then provides an $x^\star \in X^\star$ with $\|x^\star\| = \|\xi^\star\| = \|y^\star\|$ and

$$x^\star(y) = \xi^\star(y) - i\xi^\star(iy) = \operatorname{Re} y^\star(y) - i \operatorname{Re} y^\star(iy) = \operatorname{Re} y^\star(y) - i \operatorname{Re}(iy^\star(y)) = y^\star(y)$$

for all $y \in Y$, where we used the $\mathbb{C}$–linearity of $y^\star$. □

Example 4.18. 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 (4.4) since otherwise it would follow both $z \neq 0$ and

$$0 = \langle e_n, x^\star\rangle_{\ell^\infty} = \sum_{j=1}^\infty (e_n)_j z_j = z_n \qquad \text{for all } \; n \in \mathbb{N}. \qquad \Diamond$$

The next result uses a basic construction of functionals $y^\star$ and allows to distinguish between a linear subspace Y and a vector $x_0 \notin \overline{Y}$. This fact leads to several fundamental consquences of the Hahn–Banach theorem.

Proposition 4.19. *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 an $x^\star \in X^\star$ such that $x^\star(y) = 0$ for all $y \in Y$, $x^\star(x_0) = d(x_0, Y) := \inf_{y\in Y} \|x_0 - y\| > 0$, and $\|x^\star\| = 1$.*

Proof. We define the linear subspace $Z = Y + \mathrm{lin}\{x_0\}$ of X and the linear map $z^\star : Z \to \mathbb{F}$ by $z^\star(y + tx_0) = t\,d(x_0, Y)$ for all $y \in Y$ and $t \in \mathbb{F}$, using that $Y \cap \mathrm{lin}\{x_0\} = \{0\}$. Clearly, $z^\star|_Y = 0$ and $z^\star(x_0) = d(x_0, Y)$. 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. Take $y_n \in Y$ with $\|x_0 - y_n\| \to d(x_0, Y)$ (> 0, since otherwise $x_0 \in \bar{Y} = Y$). The properties of $z^\star$ now yield

$$\|z^\star\| \ge \left|\left\langle \frac{1}{\|x_0 - y_n\|}(x_0 - y_n), z^\star \right\rangle\right| = \frac{d(x_0, Y) - 0}{\|x_0 - y_n\|} \longrightarrow 1 \qquad \text{as} \quad n \to \infty,$$

so that $\|z^\star\| = 1$. A Hahn–Banach extension of $z^\star$ yields the required $x^\star \in X^\star$. □

Corollary 4.20. *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) If $x \neq 0$, then there exists an $x^\star \in X^\star$ with $\langle x, x^\star\rangle = \|x\|$ and $\|x^\star\| = 1$.

b) If $x_1 \neq x_2$, then there exists an $x^\star \in X^\star$ 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 4.19 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. □

Example 4.21. 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 4.20 and Theorem 4.14 further yield

$$\|f\|_p = \sup_{g\in D, \|g\|_{p'}\le 1} \left|\int_S f g \, d\mu\right|. \tag{4.6}$$

Corollary 4.22. *Let X be a normed vector space and $M \subseteq X$. The 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 $x^\star \in X^\star$. The assumption yields $T_x \in \mathcal{L}(X^\star, \mathbb{F})$ with $|T_x(x^\star)| \le c(x^\star) := \sup_{x\in M} |x^\star(x)| < \infty$ for all $x^\star \in X^\star$ and $x \in M$. Since $X^\star$ is a Banach space, the principle of uniform boundedness gives 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 4.20c). □

Corollary 4.23. *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 an $x^\star \in X^\star \setminus \{0\}$ with $\langle y, x^\star\rangle = 0$ for all $y \in Y$.*

Proof. The implication "⇒" follows from Proposition 4.19. If Y is dense and $x^\star \in X^\star$ vanishes on Y, then $x^\star$ must be 0 by continuity. □

Corollary 4.24. *Let X be a normed vector space and $X^\star$ be separable. Then X is separable.*

Proof. By an exercise, we have a dense subset $\{x_n^\star \,|\, n \in \mathbb{N}\}$ in $\partial B_{X^\star}(0,1)$. There are $y_n \in X$ 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 4.23 yields an $x^\star \in X^\star$ such that $\|x^\star\| = 1$ and $\langle y, x^\star\rangle = 0$ for all $y \in Y$. There exists a $j \in \mathbb{N}$ with $\|x^\star - x_j^\star\| \leq \frac{1}{4}$. We then deduce

$$\frac{1}{2} \leq \left|\langle y_j, x_j^\star\rangle\right| = \left|\langle y_j, x_j^\star - x^\star\rangle\right| \leq \frac{1}{4}$$

which is a contradiction. □

Example 4.25. The spaces c_0 and $\ell^1 = c_0^\star$ are separable, whereas $\ell^\infty = (\ell^1)^\star$ is not separable, see Example 1.51. (Since separability is preserved under isomorphisms by an exercise, we can omit here the isomorphisms from Proposition 4.11.) 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 4.11. ♢

Example 4.26 (Operators with finite rank)**.** Let X and Y be normed vector spaces with $\dim X \geq n$, $x_1, \dots, x_n \in X$ 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 4.19 provides us with an $x_k^\star \in X^\star$ 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, $T \in \mathcal{L}(X,Y)$ 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

$$T^2 x = \sum_{j=1}^{n} \sum_{k=1}^{n} \langle x, x_k^\star\rangle \langle x_k, x_j^\star\rangle x_j = Tx. \qquad \diamondsuit$$

Proposition 4.27. *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 4.26 with $x_k = y_k$. Then $T \in \mathcal{L}(X)$ is a projection with range U and thus Lemma 2.17 implies assertion a).

b) Let $Q : X \to X/Y$, $Qx = x + Y$, be the quotient map (see Proposition 2.20) and let $\{b_1, \dots, b_n\}$ a basis of X/Y. Since Q is surjective, there are $x_k \in X$ 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 $k = 1, \dots, n$. Therefore, $\{x_1, \dots, x_n\}$ is 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 $x \in V$. On the other hand, $Qx = 0$ which yields $0 = \sum_{k=1}^n \beta_k b_k$ and thus $\beta_k = 0$ for all $k = 1, \dots, n$. As a result, $x = 0$. Take $x \in X$. There are $\alpha_k \in \mathbb{F}$ such that $Qx = \alpha_1 b_1 + \ldots + \alpha_n b_n$. Set $v = \alpha_1 x_1 + \ldots + \alpha_n x_n$. We then obtain $Q(x - v) = 0$, and thus $x - v \in Y$ and $x = x - v + v \in Y + V$. Consequently, $X = Y \oplus V$. □

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* 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 (see exercises). 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 different halfspaces $H_\pm := \{x \in X \mid x^\star(x) \gtrless \gamma\}$ 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 $A \subseteq X$. We define the *Minkowski functional* $p_A : X \to [0, \infty]$ by setting

$$p_A(x) = \inf\{\lambda > 0 \mid \tfrac{1}{\lambda} x \in A\}$$

for $x \in X$, where $\inf \emptyset = \infty$. Note that $p_{B(0,1)}(x) = \|x\|$.

Remark 4.28. 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 $a_i \in A$, $b_i \in B$, and $t \in [0, 1]$ for $i = 1, 2$. 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 $A + B$. The first assertion is shown similarly. ♢

Lemma 4.29. *Let X be a normed vector space and $A \subseteq X$ convex with $0 \in A^\circ$. (Thus there exists a $\delta > 0$ with $\overline{B}(0, \delta) \subseteq A$.) Then the following assertions hold.*

a) We have $p_A(x) \le \frac{1}{\delta}\|x\|$ 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 $\frac{\delta}{\|x\|} x \in A$ for all $x \in X \setminus \{0\}$, so that $p_A(x) \le \frac{1}{\delta}\|x\|$.

b) Let $t > 0$, $x, y \in X$, and $\varepsilon > 0$. Clearly, $p_A(0x) = 0p_A(x)$. Moreover,

$$p_A(tx) = \inf\{\lambda > 0 : \ \tfrac{t}{\lambda} x \in A\} = \inf\{t\mu > 0 : \ \tfrac{1}{\mu} x \in A\} = t\, p_A(x).$$

Further, take $0 < \lambda \le p_A(x) + \varepsilon$ and $0 < \mu \le p_A(y) + \varepsilon$ with $\frac{1}{\lambda}x, \frac{1}{\mu}y \in A$. Since A is convex, the vector

$$\frac{1}{\lambda + \mu}(x + y) = \frac{\lambda}{\lambda + \mu}\frac{1}{\lambda}x + \frac{\mu}{\lambda + \mu}\frac{1}{\mu}y$$

belongs to A, so that $p_A(x + y) \le \lambda + \mu \le p_A(x) + p_A(y) + 2\varepsilon$. Letting $\varepsilon \to 0$, we deduce assertion b).

c) Now, let A be open. First take x with $p_A(x) < 1$. Then there exists a $\lambda \in (0, 1)$ 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, take x with $p_A(x) \ge 1$. Then $\frac{1}{\lambda}x$ is contained in $X \setminus A$ for all $\lambda < 1$. Since $X \setminus A$ is closed, we obtain $x \in X \setminus A$ letting $\lambda \to 1$. □

Theorem 4.30 (Separation theorems)**.** *Let X be a normed vector space, $A, B \subseteq X$ be convex and non–empty, and $A \cap B = \emptyset$. Then the following assertions hold.*

a) If A and B are open, there exists an $x^\star \in X^\star$ separating A and B.

b) *If A is closed and B is compact, there exists an $x^\star \in X^\star$ separating A and B strictly. (Special case: $B = \{x\}$.)*

Proof. We only treat $\mathbb{F} = \mathbb{R}$. The case $\mathbb{F} = \mathbb{C}$ can be deduced from the case $\mathbb{F} = \mathbb{R}$ by means of Lemma 4.16.

a) Let A and B be open. Fix some $x_0 \in A - B$ and set $C = A - B - x_0 = \bigcup_{b\in B} A - b - x_0$. Then C is open because of Proposition 1.15, C is convex by Remark 4.28, and we have $0 \in C$ as well as $y_0 := -x_0 \notin C$ (since $0 \notin A - B$). Thanks to Lemma 4.29, the Minkowski functional p_C is sublinear and $p_C(y_0) \geq 1$. We define $y^\star(ty_0) = tp_C(y_0)$ for all $t \in \mathbb{R}$. The map $y^\star : \mathrm{lin}\{y_0\} \to \mathbb{R}$ 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 4.15 now gives a linear $x^\star : X \to \mathbb{R}$ such that $x^\star(x) \leq p_C(x)$ for all $x \in X$ and $x^\star(y_0) = y^\star(y_0) = p_C(y_0) \geq 1$. Let $x \in X$. Lemma 4.29 a) implies that

$$|x^\star(x)| = \max\{x^\star(x), x^\star(-x)\} \leq \max\{p_C(x), p_C(-x)\} \leq \frac{1}{\delta}\,\|x\|$$

for some $\delta > 0$, hence $x^\star \in X^\star$. Let $x = a - b - x_0 \in C$ for any $a \in A$ and $b \in B$. Since $y_0 = -x_0$, Lemma 4.29 c) 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 assertion a) holds.

b) Let A be closed and B be compact. We can thus find an $\varepsilon > 0$ such that $\|a - b\| \geq 3\varepsilon$ for all $a \in A$ and $b \in B$. (See the exercises.) Therefore, the sets $A_\varepsilon = A + B(0,\varepsilon) = \bigcup_{a\in A} B(a,\varepsilon)$ and $B_\varepsilon = B + B(0,\varepsilon)$ are disjoint and, as above, also open and convex. Part a) gives an $x^\star \in X^\star$ with $\langle a+x, x^\star\rangle < \langle b+y, x^\star\rangle$ for all $a \in A$, $b \in B$ and $x, y \in B(0,\varepsilon)$. Choosing $y = 0$ and $x = \pm\varepsilon z$ with $z \in B(0,1)$, we conclude that $\varepsilon\,|\langle z, x^\star\rangle| < \langle b - a, x^\star\rangle$ for all $a \in A$, $b \in B$, and $z \in B(0,1)$. Taking the supremum over $z \in B(0,1)$, we obtain $0 < \varepsilon\,\|x^\star\| \leq \langle b - a, x^\star\rangle$ for all $a \in A$ and $b \in B$ which implies assertion b). □

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 given by

$$A^\perp = \{x^\star \in X^\star \,|\, \forall\, a \in A \ \text{ we have } \ \langle a, x^\star\rangle = 0\} \subseteq X^\star,$$
$${}^\perp B_\star = \{x \in X \,|\, \forall\, b^\star \in B_\star \ \text{ we have } \ \langle x, b^\star\rangle = 0\} \subseteq X.$$

In view of Riesz' Theorem 4.10, in a Hilbert space $X \cong X^\star$ these two sets are isomorphic to the othorgonal complement.

Remark 4.31. Let X be a normed space, $A \subseteq X$ and $B_\star \subseteq X^\star$ be non–empty.

a) As in Remark 4.7 one verifies that $A^\perp$ and ${}^\perp B_\star$ are closed linear subspaces of $X^\star$ and X, respectively. Also, $A \subseteq {}^\perp(A^\perp)$, $(\overline{\mathrm{lin}\, A})^\perp = A^\perp$ and ${}^\perp(\overline{\mathrm{lin}\, B_\star}) = {}^\perp B_\star$.

b) Corollary 4.23 shows that $A^\perp = \{0\}$ if and only if $\overline{\mathrm{lin}\, A} = X$. From Corollary 4.20 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 4.20, $\overline{\mathrm{lin}\, B_\star} = X^\star$ implies that ${}^\perp B_\star = \{0\}$.

d) Let $X = \ell^1$ and $B_\star = \{e_n \,|\, n \in \mathbb{N}\} \subseteq X^\star = \ell^\infty$. We then have $\overline{\mathrm{lin}\, B_\star} = c_0$, but ${}^\perp B_\star = \{0\}$ since $\ell^1 = c_0^\star$. ◊

Proposition 4.32. *For a normed vector space X and a non–empty subset $A \subseteq X$, we have $\overline{\mathrm{lin}\, A} = {}^\perp(A^\perp)$.*

Proof. Remark 4.31 yields $\overline{\mathrm{lin}\, A} \subseteq {}^\perp(A^\perp)$. Suppose there were an $x_0 \in {}^\perp(A^\perp)\backslash\overline{\mathrm{lin}\, A}$. Theorem 4.30 b) then gives an $x^\star \in X^\star$ with

$$s := \sup_{x\in\overline{\mathrm{lin}\, A}} \mathrm{Re}\langle x, x^\star\rangle < \mathrm{Re}\langle x_0, x^\star\rangle.$$

If $r := \mathrm{Re}\langle x, x^\star\rangle \neq 0$ for some $x \in \overline{\mathrm{lin}\, A}$, then we deduce for $t \in \mathbb{R}$ that

$$\mathrm{Re}\langle tx, x^\star\rangle = tr \to \infty \begin{cases} \text{as} \quad t \to \infty, & \text{if} \quad r > 0, \\ \text{as} \quad t \to -\infty, & \text{if} \quad r < 0. \end{cases}$$

This contradicts the above estimate for s, and thus $\mathrm{Re}\langle x, x^\star\rangle = 0$ for all $x \in \overline{\mathrm{lin}\, A}$. Hence, $s = 0$. Using it instead of t, we similarly obtain $\mathrm{Im}\langle x, x^\star\rangle = 0$ for all $x \in \overline{\mathrm{lin}\, A}$ so that $x^\star \in A^\perp$, which yields the contradiction $0 = s < \mathrm{Re}\langle x_0, x^\star\rangle = 0$. □

Let $X = \ell^1$ and $B_\star = \{e_n \,|\, n \in \mathbb{N}\} \subseteq X^\star = \ell^\infty$. In Remark 4.31d), we have seen that $\overline{\mathrm{lin}\, B_\star} = c_0$ and that ${}^\perp B_\star = \{0\}$, hence $({}^\perp B_\star)^\perp = X^\star = \ell^\infty \neq \overline{\mathrm{lin}\, B_\star}$. See Theorem 4.7 in [Rud91] for more information.

Proposition 4.33. *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 $Q : X \to X/Y$ is the quotient map $Qx = x + Y$.

This result is shown in the exercises.

4.4 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 $J_X(x) : X^\star \to \mathbb{F}$ by

$$\langle x^\star, J_X(x)\rangle_{X^\star} = \langle x, x^\star\rangle_X \qquad \text{for all} \quad x^\star \in X^\star. \tag{4.7}$$

Clearly, $J_X(x)$ is linear in $x^\star$ and we have $|[J_X(x)](x^\star)| \le \|x\|\,\|x^\star\|$, so that $J_X(x) \in X^{\star\star}$. Moreover, the map J_X is linear in x, and by means of Corollary 4.20 and equation (4.7) we obtain

$$\|x\|_X = \sup_{\|x^\star\| \le 1} |\langle x, x^\star\rangle| = \|J_X(x)\|_{X^{\star\star}}.$$

Proposition 4.34. *Let X be a normed vector space. Then equation (4.7) defines a linear isometry $J_X : X \to X^{\star\star}$.* □

Definition 4.35. *X is called* reflexive *if J_X from (4.7) is surjective.*

Remark 4.36. a) If X is reflexive, then $X \cong (X^\star)^\star$ and thus is X a Banach space. However, there are non–reflexive Banach spaces which are isomorphic to its bidual (with an isomorphism different from J_X), see Example 1.d.2 in [LiT96].

b) If X reflexive and $B_\star \subseteq X^\star$, then $B_\star^\perp = J_X({}^\perp B_\star)$ by (4.7). Here (and in some other points below) one sees that reflexive Banach spaces share some properties of Hilbert spaces which are not true in a general Banach space.

c) In Corollary 4.61 we will show that reflexivity is preserved under isomorphisms.

d) Usually one identifies X with $R(J_X)$ and reflexive spaces X with $X^{\star\star}$. ◊

Example 4.37. a) Hilbert spaces X are reflexive.
Proof. Recall from Theorem 4.10 that a Hilbert space Y is isomorphic to its dual space via the isomorphism $\Phi_Y : Y \to Y^\star$ given by $\langle z, \Phi_Y(y)\rangle = (z|y)$ for all $z \in Y$. The dual space $X^\star$ of X is also a Hilbert space equipped with the scalar product

$(x^\star|y^\star)_{X^\star} := (\Phi_X^{-1}y^\star|\Phi_X^{-1}x^\star)_X$ (check!). Now take any $x^{\star\star} \in X^{\star\star}$. Set $x^\star = \Phi_{X^\star}^{-1}x^{\star\star} \in X^\star$ and $x = \Phi_X^{-1}x^\star \in X$. We then obtain

$$\langle z^\star, x^{\star\star}\rangle_{X^\star} = (z^\star|x^\star)_{X^\star} = (\Phi_X^{-1}x^\star|\Phi_X^{-1}z^\star)_X = \langle x, z^\star\rangle_X$$

for every $z^\star \in X^\star$; 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 (e.g., $X = \ell^p$).

Proof. Theorem 4.14 says that the map $\Phi_r : L^{r'}(\mu) \to X^\star$ defined by $\langle \varphi, \Phi_r(\psi)\rangle = \int \varphi\psi \, d\mu$ for all $\varphi \in L^r(\mu)$ and $\psi \in L^{r'}(\mu)$ is an isometric isomorphism, where $r \in (1,\infty)$. Take $\phi \in X^{\star\star}$. Then $\phi \circ \Phi_p$ belongs $L^{p'}(\mu)^\star$. Since $p'' = p$, Theorem 4.14 yields an $f \in L^p(\mu)$ with $\Phi_{p'}(f) = \phi \circ \Phi_p$. Let $\Phi_p(g)$ with $g \in L^{p'}(\mu)$ be an arbitrary element of $X^\star$. We then compute

$$\langle \Phi_p(g), \varphi\rangle_{X^\star} = \phi(\Phi_p(g)) = \langle g, \Phi_{p'}(f)\rangle_{L^{p'}} = \int_S gf \, d\mu = \langle f, \Phi_p(g)\rangle_{L^p}.$$

Hence, $\phi = J_X(f)$ as asserted. □

c) Proposition 4.11 shows that $c_0^{\star\star} \cong \ell^\infty$. Moreover, $\ell^\infty \ncong c_0$ since c_0 is separable and ℓ^∞ is not separable (see Example 1.51). Hence, c_0 is not reflexive. ◊

In the above example and also below, we use that separability is preserved under isomorphisms (see exercises).

Proposition 4.38. *For a normed vector space X, the following assertions hold.*

a) If X is reflexive, then each closed linear subspace Y of X is reflexive (where Y is endowed with the norm of X).

b) Let X be a Banach space. Then 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 $y^{\star\star} \in Y^{\star\star}$. Each $x^\star \in X^\star$ satisfies $x^\star|_Y \in Y^\star$ with $\|x^\star|_Y\|_{Y^\star} \le \|x^\star\|_{X^\star}$. We define the linear map $x^{\star\star}$ 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 $x^{\star\star}$ belongs to $X^{\star\star}$. By the reflexivity of X, there is a vector $y \in X$ with

$$\langle x^\star|_Y, y^{\star\star}\rangle_{Y^\star} = \langle x^\star, x^{\star\star}\rangle_{X^\star} = \langle y, x^\star\rangle_X \qquad \text{for all } \ x^\star \in X^\star.$$

Suppose that $y \notin Y$. Since Y is closed, Proposition 4.19 yields an $\tilde{x}^\star \in X^\star$ with $\tilde{x}^\star|_Y = 0$ and $\langle y, \tilde{x}^\star\rangle_X \neq 0$. This fact contradicts the above equation in display and we thus obtain $y \in Y$. Take any $y^\star \in Y^\star$. Let $x^\star \in X^\star$ be a Hahn–Banach extension of $y^\star$. Then, $\langle y, y^\star\rangle_Y = \langle y, x^\star\rangle_X = \langle y^\star, y^{\star\star}\rangle_{Y^\star}$ by the above considerations, and thus $J_Y(y) = y^{\star\star}$.

b) Let X be reflexive. Take $x^{\star\star\star} \in X^{\star\star\star}$. 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 $X^\star$. Every $x^{\star\star} \in X^{\star\star}$ can be written as $x^{\star\star} = J_X(x)$ for some $x \in X$, so that $\langle x^\star, x^{\star\star}\rangle_{X^\star} = \langle x^{\star\star}, x^{\star\star\star}\rangle_{X^{\star\star}}$ for all $x^{\star\star} \in X^{\star\star}$. Thus, $X^\star$ is reflexive.

Conversely, assume that X is not reflexive. By Proposition 4.19 there exists an $x^{\star\star\star} \in X^{\star\star\star} \setminus \{0\}$ with $\langle J_X(x), x^{\star\star\star}\rangle_{X^{\star\star}} = 0$ for all $x \in X$. Suppose that $X^\star$ was reflexive. Then there must exist an $x^\star \in X^\star$ with $x^{\star\star\star} = J_{X^\star}(x^\star)$. Hence,

$$0 = \langle J_X(x), J_{X^\star}(x^\star)\rangle_{X^{\star\star}} = \langle x^\star, J_X(x)\rangle_{X^\star} = \langle x, x^\star\rangle_X \qquad \text{for all } \ x \in X,$$

which means that $x^\star = 0$, contradicting $x^{\star\star\star} \neq 0$.

c) The implication "$\Leftarrow$" was shown in Corollary 4.24. If X is separable, then $X^{\star\star} \cong X$ is also separable. Hence, $X^\star$ is separable by Corollary 4.24. □

Example 4.39. a) The space $X = \ell^1$ is not reflexive, because it is separable and its dual $X^\star \cong \ell^\infty$ is non-separable. Since c_0 is not reflexive by Example 4.37 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 (see exercises). ♢

In order to obtain fundamental compactness results, one introduces new convergence concepts.

Definition 4.40. *Let X be a normed vector space.*

a) A sequence (x_n) in X converges weakly *to $x \in X$ if*

$$\forall\, x^\star \in X^\star: \quad \langle x_n, x^\star\rangle \to \langle x, x^\star\rangle \quad \textit{as} \quad n \to \infty.$$

b) A sequence $(x_n^\star)$ in $X^\star$ converges weakly* *to $x^\star \in X^\star$ if*

$$\forall\, x \in X: \quad \langle x, x_n^\star\rangle \to \langle x, x^\star\rangle \quad \textit{as} \quad n \to \infty.$$

We then write $x_n \rightharpoonup x$ or $x_n \xrightarrow{\sigma} x$ or σ–$\lim_{n\to\infty} x_n = x$, respectively, $x_n \overset{\star}{\rightharpoonup} x$ or $x_n^\star \xrightarrow{\sigma^\star} x^\star$ or $\sigma^\star$–$\lim_{n\to\infty} x_n^\star = x^\star$. One often replaces here the letter 'σ' by 'w'.

Remark 4.41. (Positive results.) 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 $X^\star = \mathcal{L}(X,\mathbb{F})$ is just the strong convergence of a sequence of operators in $\mathcal{L}(X,\mathbb{F})$.

b) Weak and weak* convergence is linear.

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 $X = \mathbb{F}^d$ 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) Weak and weak* limits are unique.

Proof. Let $x_n \xrightarrow{\sigma} x$ and $x_n \xrightarrow{\sigma} y$. If $x \neq y$, then Corollary 4.20 yields an $x^\star \in X^\star$ with $\langle x, x^\star\rangle \neq \langle y, x^\star\rangle$, but $\langle x_n, x^\star\rangle$ converges to both $\langle x, x^\star\rangle$ and $\langle y, x^\star\rangle$, which is impossible. The second part follows from a). □

e) For $x \in X$ we have $\langle x, x^\star\rangle_X = \langle x^\star, J_X(x)\langle_{X^\star}$ 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.

f) If $(\langle x, x_n^\star\rangle)_n$ is Cauchy in $\mathbb{F}$ for each $x \in X$ and a sequence $(x_n^\star)$ in $X^\star$, then there exists an $x^\star \in X^\star$ such that $x_n^\star \xrightarrow{\sigma^\star} x^\star$ as $n \to \infty$, due to Lemma 3.9. In this sense, $X^\star$ is 'weakly* sequentially complete'.

g) Reflexive spaces are 'weakly sequentially complete' due to e) and f). ♢

Remark 4.42. (Negative results.) Let $n \to \infty$.

a) Weakly or weakly* convergent sequences may diverge in norm. Weak or weak* limits may have a strictly smaller norm.

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 $\|e_n - e_m\|_2 = 2^{1/2}$ for all $n \neq m$ so that (e_n) diverges in ℓ^2. Moreover, $\|e_n\| = 1$ and the weak limit 0 has a strictly smaller norm. □

b) A weakly* convergent sequence may have no 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 divergent subsequence $n_j \to \infty$ and choose $y \in \ell^\infty = (\ell^1)^\star$ such that $(y_{n_j})_j$ diverges. Then $\langle e_{n_j}, y\rangle_{\ell^1} = y_{n_j}$ does not converge as $j \to \infty$. □

c) There are (nonreflexive) 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 $y \in \ell^1$. But, if (v_n) had a weak limit x in c_0, then x would also be a weak* limit of (v_n) in ℓ^∞ and thus $x = \mathbb{1}$ by the uniqueness of weak* limits, which is impossible. □

Proposition 4.43. *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 $\overline{\operatorname{lin} D} = X$. We then have:*

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) holds, then $\|x\| \le \underline{\lim}_{n\to\infty}\|x_n\|$; and if b) holds, then $\|x^\star\| \le \underline{\lim}_{n\to\infty}\|x_n^\star\|$. Moreover, in b) the implication '$\Leftarrow$' holds for all normed vector spaces X.

Proof. The proposition follows from Corollary 3.5 and Lemma 3.9 using in a) that $J_X(x_n) \in X^{\star\star} = \mathcal{L}(X^\star, \mathbb{F})$. □

Example 4.44. a) Let $X = c_0$ oder $X = \ell^p$ for $1 < p < \infty$, $(v_n) \subseteq X$ be bounded, and $x \in X$. Then Proposition 4.43 with $D_\star = \{e_n \mid n \in \mathbb{N}\}$ yields

$$v_n \xrightarrow{\sigma} x \iff [v_n]_k = \langle v_n, e_k\rangle \to \langle x, e_k\rangle = x_k \quad \text{for all } k \in \mathbb{N} \qquad (n \to \infty).$$

(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 $D = \{e_n \mid n \in \mathbb{N}\} \subseteq c_0$ or $\subseteq \ell^1$

b) The implication '$\Leftarrow$' fails in a) for $p = 1$. In fact, the sequence (e_n) converges componentwise to 0, but it is boundad and diverges weakly in ℓ^1 by the proof of Remark 4.42b).

c) The example $v_n = ne_n$ shows that the assumption of boundedness cannot be omitted in Proposition 4.43.

d) Let $(S, \mathcal{A}, \mu)$ be a measure space, $X = L^p(\mu)$, $1 < p < \infty$, and (f_n) be bounded in X. Then

$$f_n \xrightarrow{\sigma} f \iff \int_A f_n \, d\mu \to \int_A f \, d\mu \quad \text{for all } A \in \mathcal{A} \quad \text{with } \mu(A) < \infty \qquad (n \to \infty)$$

due to Proposition 4.43 since the space $D_\star$ of simple functions is dense in $L^{p'}(\mu)$. An analogous result holds for the weak* convergence in $L^\infty(\mu)$, if the measure space is σ–finite. ◊

Remark 4.45. If $X^\star$ is separable, then the weak sequential convergence in bounded subsets of X is given by the metric

$$d(x,y) = \sum_{j=1}^{\infty} 2^{-j} \frac{\left|\langle x - y, x_j^\star\rangle\right|}{1 + \left|\langle x - y, x_j^\star\rangle\right|}$$

where $D_\star = \{x_j^\star \mid j \in \mathbb{N}\}$ is dense in $X^\star$. This result follows Propositions 4.43 and 1.8 using the seminorms $p_j(x) = \left|\langle x, x_j^\star\rangle\right|$. 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 4.20 and the density of $D_\star$. Similarly, if X is separable then the weak* sequential convergence in bounded subsets of $X^\star$ is given by an analogous metric. ◊

The next theorem says that closed convex sets are 'weakly sequentially closed'. It is a consequence of the separation theorem.

Theorem 4.46 (Mazur)**.** *Let X be a normed vector space, K be a closed and convex subset of X, and let $x_n \in K$ converge weakly to some $x \in X$ as $n \to \infty$. Then $x \in K$, and there is a sequence (y_N) of convex combinations of $\{x_n \,|\, n \geq N\}$ converging to x in norm as $N \to \infty$.*

Proof. 1) Suppose that x were not contained in K. Theorem 4.30b) (with $A = K$ and $B = \{x\}$) then gives an $x^\star \in X^\star$ with $\sup_n \operatorname{Re}\langle x_n, x^\star\rangle < \operatorname{Re}\langle x, x^\star\rangle$. But this inequality cannot hold since $\langle x_n, x^\star\rangle$ converges to $\langle x, x^\star\rangle$ as $n \to \infty$. Hence, $x \in K$.

2) Let $N \in \mathbb{N}$ and define the (convex!) set

$$C_N = \Big\{ y = \sum_{j=N}^{m} t_j x_j \,\Big|\, m \in \mathbb{N},\ m \geq N,\ t_j \geq 0,\ t_N + \cdots + t_m = 1 \Big\}.$$

Then $\overline{C_N}$ is closed (in norm) and convex (see Corollary 1.18), and $x_n \in C_N$ for all $n \geq N$. By part 1), x belongs to each $\overline{C_N}$. So we can choose $y_N \in C_N$ such that $\|x - y_N\| \leq 1/N$ for every $N \in \mathbb{N}$. □

Remark 4.47. a) One needs convexity in Mazur's theorem: For instance, the vectors e_n in the (closed) unit sphere S of ℓ^2 converge weakly in ℓ^2 to $0 \notin S$.

b) Mazur's theorem can fail for the weak* convergence: For instance, the vectors e_n converge weakly* to 0 in $X^\star = \ell^1$, and they belong to the closed and affine (thus convex) subspace $A = \{y \in \ell^1 \,|\, \langle y, \mathbb{1}\rangle = 1\}$, but $0 \notin A$. ◇

We now prove an important extension of the Bolzano–Weierstraß theorem to infinite dimensional Banach spaces: In dual spaces the balls $\overline{B}(0,r)$ are 'weakly* sequentially compact' (instead of being compact as in finite dimensions).

Theorem 4.48 (Banach–Alaoglu, simplified version)**.** *Let X be a separable normed vector space. Let $(x_n^\star)$ be a bounded sequence in $X^\star$. Then there is an $x^\star \in X^\star$ and a subsequence $x_{n_j}^\star$ converging weakly* to $x^\star$ as $j \to \infty$, where $\|x^\star\| \leq \underline{\lim}_{n\to\infty} \|x_n^\star\|$.*

Proof. Let $\{x_k \,|\, k \in \mathbb{N}\}$ 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 $(\langle x_1, x_{\nu_1(j)}^\star\rangle)_j$ converging in $\mathbb{F}$ as $j \to \infty$. Since also $(\langle x_2, x_{\nu_1(j)}^\star\rangle)_j$ is bounded, there is a converging subsequence $(\langle x_2, x_{\nu_2(j)}^\star\rangle)_j$. Iteratively, we obtain subsequences $\nu_k \subseteq \nu_{k-1}$ such that $(\langle x_k, x_{\nu_k(j)}^\star\rangle)_j$ converges for each $k \in \mathbb{N}$. We define $y_m^\star = x_{\nu_m(m)}^\star$ for each $m \in \mathbb{N}$. Then $(\langle x_k, y_m^\star\rangle)_m$ converges as $m \to \infty$ for all $k \in \mathbb{N}$ because of $(\nu_m(m))_{m\geq k} \subseteq (\nu_k(j))_{j\in\mathbb{N}}$. Since $\{x_k \,|\, k \in \mathbb{N}\}$ is dense and the subsequence $(y_m^\star)$ of $(x_n^\star)$ is also bounded, Lemma 3.9 implies the assertion. □

Example 4.49. a) The above version of the theorem of Banach–Alaoglu can fail if X is not separable. As an example consider $X = \ell^\infty$ and the map $\varphi_n(x) = x_n$ for $x \in X$. Then $\varphi_n \in X^\star$ for all $n \in \mathbb{N}$ and $\|\varphi_n\| = 1$. Take any subsequence $(\varphi_{n_k})_k$. For $j \in \mathbb{N}$, we set $x_j = (-1)^k$ if $j = n_k$ and $x_j = 0$ otherwise, and put $x = (x_j) \in \ell^\infty$. Then $\langle x, \varphi_{n_k}\rangle = (-1)^k$ diverges, and thus (φ_n) has no weakly* convergent subsequence. We refer to Theorem V.3.1 in [Con90] for the full version of the theorem of Banach–Alaoglu which holds without the separability assumption.

b) The theorem of Banach–Alaoglu can fail for the weak convergence. As an example, consider $e_n \in \ell^1 = X$ for all $n \in \mathbb{N}$. Take any subsequence e_{n_k} and choose $y \in \ell^\infty = (\ell^1)^\star$ such that y_{n_k} diverges. Then $\langle e_{n_k}, y\rangle = y_{n_k}$ does not converge. ◇

Theorem 4.50. *Let X be a reflexive Banach space. Then every bounded sequence (x_n) in X has a weakly convergent subsequence.*

Proof. Let $Y = \overline{\mathrm{lin}}\{x_n \,|\, n \in \mathbb{N}\}$. By Proposition 4.38, the space Y is reflexive, and it is separable by Lemma 1.50. Proposition 4.38 now shows that $Y^\star$ is separable. Applying Theorem 4.48 to $Y^\star$, we obtain a weakly* convergent subsequence $(J_Y(x_{n_k}))_k$ of $(J_Y(x_n))_n$ in $Y^{\star\star}$. Since Y is reflexive, there exists an $x \in Y$ with

$$\langle x_{n_k}, y^\star\rangle_Y = \langle y^\star, J_Y(x_{n_k})\rangle_{Y^\star} \longrightarrow \langle y^\star, J_Y(x)\rangle_{Y^\star} = \langle x, y^\star\rangle_Y$$

for all $y^\star \in Y^\star$, as $k \to \infty$. Let $x^\star \in X^\star$. Then the restriction $x^\star|_Y$ belongs to $Y^\star$. As a result,

$$\langle x_{n_k}, x^\star\rangle_X = \langle x_{n_k}, x^\star|_Y\rangle_Y \longrightarrow \langle x, x^\star|_Y\rangle_Y = \langle x, x^\star\rangle_X,$$

which means that $x_{n_k} \xrightarrow{\sigma} x$ as $k \to \infty$. □

Example 4.51. Let X be reflexive and $C \subseteq X$ be closed and convex. Then there exists a vector $\overline{x} \in C$ with minimal norm.

Proof. If $0 \in C$, then we can take $\overline{x} = 0$. Otherwise, $d := \inf\{\|x\| \,|\, x \in C\} > 0$, and there are $x_n \in C$ such that $\|x_n\| \to d$ as $n \to \infty$. Since the sequence (x_n) is bounded, it has a subsequence (x_{n_j}) weakly converging to some $\overline{x} \in X$, due to Theorem 4.50. Moreover, $\|\overline{x}\| \le \liminf_{n\to\infty} \|x_n\| = d$. Mazur's Theorem 4.46 further yields that $\overline{x} \in C$, and thus $\|\overline{x}\| \ge d$. □

4.5 Adjoint operators

Let X and Y be normed vector spaces and $T \in \mathcal{L}(X, Y)$. 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 $\|\varphi_{y^\star}\| \le \|T\|\,\|y^\star\|$. We now define $T^\star y^\star := \varphi_{y^\star} \in X^\star$, which means that

$$\langle Tx, y^\star\rangle_Y = \langle x, T^\star y^\star\rangle_X \qquad (\forall\, x \in X,\ y^\star \in Y^\star). \tag{4.8}$$

Equation (4.8) determines $T^\star y^\star$ uniquely as an element of $X^\star$. Therefore (4.8) defines a map $T^\star : Y^\star \to X^\star$ which is called the *adjoint* of T.

If X and Y are pre Hilbert spaces, analogously we introduce the *Hilbert space adjoint* $T' : Y \to X$ of T by

$$(Tx|y)_Y = (x|T'y)_X \qquad (\forall\, x \in X,\ y \in Y). \tag{4.9}$$

Let X and Y be Hilbert spaces. Then $T' = \Phi_X^{-1} T^\star \Phi_Y$ for the Riesz isomorphisms from Theorem 4.10.

Proposition 4.52. *Let X, Y, Z be normed vector spaces, $S, T \in \mathcal{L}(X, Y)$, $R \in \mathcal{L}(Y, Z)$, and $\alpha \in \mathbb{F}$. The following assertions hold.*

a) $T^\star \in \mathcal{L}(Y^\star, X^\star)$ *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 $T = (T')' =: T''$ hold 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 4.20 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) holds. We further calculate

$$\langle RTx, z^\star\rangle = \langle Tx, R^\star z^\star\rangle = \langle x, T^\star R^\star z^\star\rangle$$

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 $T \in \mathcal{L}(X, Y)$ possesses its *bi-adjoint* $T^{\star\star} := (T^\star)^\star \in \mathcal{L}(X^{\star\star}, Y^{\star\star})$ with $\|T\| = \|T^{\star\star}\|$.

Definition 4.53. *Let X and Y be Hilbert spaces and $T \in \mathcal{L}(X, Y)$. 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., $T = T'$).*

Example 4.54. a) For $X = \mathbb{F}^d$ and a matrix $T = [a_{kl}]$, the adjoint $T^\star$ is given by $[a_{lk}]$ and T' by $[\overline{a_{lk}}]$.

b) On $X = c_0$ or $X = \ell^p$ with $1 \le p < \infty$ consider the shift operators $Lx = (x_{n+1})_n$ and $Rx = (0, x_1, x_2, \dots)$. Take $x \in X$ and $y \in X^\star$. We calculate

$$\langle Lx, y\rangle = \sum_{k=1}^\infty x_{k+1} y_k = \sum_{n=2}^\infty x_n y_{n-1} = \langle x, Ry\rangle$$

so that $L^\star = R$. Similarly one derives $R^\star = L$ 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 3.13. For $f, g \in X$ we compute

$$(T(t)f|g) = \int_{\mathbb{R}} f(s+t)\overline{g(s)}\, ds = \int_{\mathbb{R}} f(\tau)\overline{g(\tau - t)}\, d\tau = (f|T(-t)g),$$

i.e., $T(t)' = T(-t) = T(t)^{-1}$. 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 $t \in \mathbb{R}$.

d) Let $A \in \mathcal{B}_d$, $1 \le p < \infty$, and $k : A \times A \to \mathbb{F}$ be measurable with

$$\kappa_p = \left(\int_A \left(\int_A |k(x,y)|^{p'}\, dy \right)^{p/p'} dx \right)^{1/p} < \infty, \qquad \text{if } \ p > 1,$$

$$\kappa_1 = \operatorname*{ess\,sup}_{y \in A} \int_A |k(x,y)|\, dx < \infty, \qquad \text{if } \ p = 1.$$

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 a) and Hölder (first in the y and then in the x integral), we deduce

$$\int_{A \times A} |\varphi|\, d(x,y) = \int_A \int_A |k(x,y)|\, |f(y)|\, dy\, |g(x)|\, dx$$

$$\leq \int_A \Big(\int_A |k(x,y)|^{p'} \, dy \Big)^{\frac{1}{p'}} \Big(\int_A |f(y)|^p \, dy \Big)^{\frac{1}{p}} |g(x)| \, dx$$
$$\leq \kappa_p \, \|f\|_p \Big(\int_A |g(x)|^{p'} \, dx \Big)^{\frac{1}{p'}} = \kappa_p \, \|f\|_p \, \|g\|_{p'} < \infty.$$

This means that φ is integrable on $A \times A$, and so Fubini b) yields that

$$h_g(x) := \int_A k(x,y) \, f(y) \, g(x) \, dy = g(x) \int_A k(x,y) \, f(y) \, dy$$

exists for $x \notin N_{fg}$ and a null set N_{fg}, and that 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

$$Tf(x) := \begin{cases} \int_A k(x,y) \, f(y) \, dy, & x \in A \setminus N_f, \\ 0, & x \in N_f, \end{cases}$$

exists and that it is measurable. Note that this definition does not depend on the representative of f. Using again Hölder in the y–integral, we further estimate

$$\|Tf\|_p^p = \int_A \Big| \int_A k(x,y) \, f(y) \, dy \Big|^p dx \leq \int_A \Big(\int_A |k(x,y)|^{p'} \, dy \Big)^{\frac{p}{p'}} \Big(\int_A |f(y)|^p \, dy \Big)^{\frac{p}{p}} dx$$
$$= \kappa_p^p \, \|f\|_p^p \, .$$

As a consequence, $Tf \in L^p(A)$. The linearity of T is clear, and so $T \in \mathcal{L}(L^p(A))$ with norm less or equal κ_p. Since φ is integrable, Fubini b) finally yields

$$\langle Tf, g \rangle = \int_A \int_A k(x,y) f(y) \, dy \, g(x) \, dx = \int_A \int_A f(y) k(x,y) g(x) \, dx \, dy$$
$$= \int_A f(y) \int_A k(x,y) g(x) \, dx \, dy = \langle f, T^\star g \rangle.$$

This means that

$$T^\star g(y) = \int_A k(x,y) g(x) \, dx \qquad \text{for a.e. } y \in A$$

and for all $g \in L^{p'}(A)$. For $p = 2$ one derives analogously $T'g(y) = \int_A \overline{k(x,y)} g(x) \, dx$ for a.e. $y \in A$ and all $g \in L^2(A)$. Hence, T is self adjoint if $k(x,y) = \overline{k(y,x)}$ for all $x, y \in A$. For $p = 1$ one can further show that $T \in \mathcal{L}(L^1(A))$ with $\|T\| \leq \kappa_1$. ◊

Proposition 4.55. *For normed spaces X and Y, the following assertions hold.*

a) For $T \in \mathcal{L}(X,Y)$ 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 $y^\star \in Y^\star$. Employing (4.8) and (4.7), 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). □

We mostly identify T and $T^{\star\star}$ if X and Y are reflexive.

Range and kernel of bounded linear operators

Let $T \in \mathcal{L}(X,Y)$. Then the kernel $N(T)$ is closed, but $R(T)$ need not to be closed. An example is the Volterra operator given by

$$Tf(t) = \int_0^t f(s)ds \qquad (t \in [0,1])$$

for $f \in X = C([0,1])$. Here, $T \in \mathcal{L}(X)$ and $R(T) = \{g \in C^1([0,1]) \,|\, g(0) = 0\}$ is different from $\overline{R(T)} = \{g \in X \,|\, g(0) = 0\}$.

Proposition 4.56. *Let X and Y be normed vector spaces and $T \in \mathcal{L}(X,Y)$. Then the following assertions hold.*

a) $R(T)^\perp = N(T^\star)$.

b) $\overline{R(T)} = {}^\perp N(T^\star)$. *Hence, $R(T)$ 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 vector $y^\star$ 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 4.32 and part a) show that $\overline{R(T)} = {}^\perp(R(T)^\perp) = {}^\perp N(T^\star)$. The second assertion now follows from Remark 4.31.

c) Let $x \in X$. Due Corollary 4.20, the vector x belongs to $N(T)$ 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 4.31 implies the second assertion.

d) Proposition 4.32 and part a) imply that $\overline{R(T^\star)} = {}^\perp(R(T^\star)^\perp) = {}^\perp N(T^{\star\star})$. We thus have to prove that ${}^\perp N(T^{\star\star}) \subseteq N(T)^\perp$. So let $y^\star \in {}^\perp N(T^{\star\star})$ and take any $x \in N(T)$. Then Proposition 4.55 yields $T^{\star\star} J_X(x) = J_Y Tx = 0$ so that $J_X(x) \in N(T^{\star\star})$ and thus $\langle x, y^\star\rangle = \langle y^\star, J_X(x)\rangle = 0$. This means that $y^\star \in N(T)^\perp$, as required.

e) Let X be reflexive. It remains to show that $N(T)^\perp \subseteq {}^\perp N(T^{\star\star})$ in view of the proof of part 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 an $x \in X$ such that $J_X(x) = x^{\star\star}$. Proposition 4.55 yields that $J_Y Tx = T^{\star\star} J_X x = 0$. Since J_Y is injective by Proposition 4.34, the vector x belongs to $N(T)$, and hence $\langle y^\star, x^{\star\star}\rangle = \langle x, y^\star\rangle = 0$, as required. □

Remark 4.57. In 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 $N(T) = \{0\}$. In Example 1.25 of the lecture notes [ST] it is shown that the range of $T^\star = I - R$ in $X^\star = \ell^1$ is not dense, hence $\overline{R(T^\star)} \neq X^\star = N(T)^\perp$. ♢

The next result is an easy consequence of Proposition 4.56 b).

Corollary 4.58. *Let X and Y be normed vector spaces, $y \in Y$, and let $T \in \mathcal{L}(X,Y)$ have* closed range. *Then the equation $Tx = y$ has a solution $x_0 \in X$ if and only if $\langle y, y^\star\rangle = 0$ for all $y^\star \in N(T^\star)$, and 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.*

Corollary 4.59. *Let X and Y be Banach spaces and $T \in \mathcal{L}(X,Y)$. 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. Note that statement b) implies that T is injective. From Corollary 3.24 we deduce that b) holds if and only if T is injective and $R(T)$ is closed. Hence, Corollary 4.58 implies that the validity of a) and b) is equivalent to the bijectivity of T, and thus to its invertibility. □

Theorem 4.60. *Let X and Y be Banach spaces and $T \in \mathcal{L}(X,Y)$. The operator T is invertible if and only if $T^\star \in \mathcal{L}(Y^\star, X^\star)$ is invertible. In this case we have $(T^{-1})^\star = (T^\star)^{-1}$.*

Proof. Let T be invertible. We then have $I_X = T^{-1}T$, and thus $I_{X^\star} = (I_X)^\star = T^\star(T^{-1})^\star$. Similarly, $I_{Y^\star} = (T^{-1})^\star T^\star$. Thus $T^\star$ has the inverse $(T^{-1})^\star \in \mathcal{L}(X^\star, Y^\star)$. Let $T^\star$ be invertible, and thus injective. Then $T^{\star\star}$ is invertible. Let $x \in X$. Proposition 4.55 now implies that

$$\|x\| = \|J_X(x)\| = \|(T^{\star\star})^{-1}T^{\star\star}J_X(x)\| \leq \|(T^{\star\star})^{-1}\|\,\|T^{\star\star}J_X(x)\| = \|(T^{\star\star})^{-1}\|\,\|Tx\|.$$

From Corollary 4.59 we now deduce the invertibility of T. □

The above results will be used and extended in the lecture Spectral Theory.

Corollary 4.61. *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 (4.7) and (4.8), we compute

$$\begin{aligned}\langle y, y^\star\rangle_Y &= \langle \Phi J_X^{-1}(\Phi^{\star\star})^{-1}y^{\star\star}, y^\star\rangle_Y = \langle J_X^{-1}(\Phi^{\star\star})^{-1}y^{\star\star}, \Phi^\star y^\star\rangle_X \\ &= \langle \Phi^\star y^\star, (\Phi^{\star\star})^{-1}y^{\star\star}\rangle_{X^\star} = \langle y^\star, \Phi^{\star\star}(\Phi^{\star\star})^{-1}y^{\star\star}\rangle_{Y^\star} = \langle y^\star, y^{\star\star}\rangle_{Y^\star}.\end{aligned}$$

Hence, $J_Y(y) = y^{\star\star}$. □

Proposition 4.62. *Let X and Y be Hilbert spaces and $T \in \mathcal{L}(X,Y)$. It then holds:*
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 $\alpha \in \mathbb{F}$ and $x, z \in X$. Using (4.1) and that T is isometric, we calculate

$$\begin{aligned}(T(x+\alpha z)|T(x+\alpha z)) &= \|Tx\|^2 + 2\,\mathrm{Re}\,(Tx|\alpha Tz) + \|\alpha Tz\|^2 \\ &= \|x\|^2 + 2\,\mathrm{Re}\,\bar\alpha\,(Tx|Tz) + |\alpha|^2\,\|z\|^2, \\ \|T(x+\alpha z)\|^2 = (x+\alpha z|x+\alpha z) &= \|x\|^2 + 2\,\mathrm{Re}\,\bar\alpha\,(x|z) + |\alpha|^2\,\|z\|^2.\end{aligned}$$

Since $(T(x+\alpha z)|T(x+\alpha z)) = \|T(x+\alpha z)\|^2$, it follows $\mathrm{Re}\,\bar\alpha\,(Tx|Tz) = \mathrm{Re}\,\bar\alpha\,(x|z)$. Choosing $\alpha = 1$ and $\alpha = i$ (if $\mathbb{F} = \mathbb{C}$), we deduce assertion a).

b) The second equivalence is a consequence of part a). To show the first equivalence, take $x, z \in X$. If T is unitary we obtain $(Tx|Tz) = (x|T^{-1}Tz) = (x|z)$, so that T is isometric by a). If T is isometric, 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 $T' = T^{-1}$. □

This proposition is needed in the next section.

Chapter 5

The Fourier transform and generalized derivatives

The Fourier transform is a fundamental tool in many branches of mathematics and its applications. In the first section we study its basic properties in an L^2 context. If one wants to treat partial differential equations in L^2 (or L^p) spaces, one needs weak derivatives and the Sobolev spaces $W^{k,p}$. The second section gives a brief introduction in these topics and establishes important links between the Fourier transform and the spaces $W^{k,2}(\mathbb{R}^d)$. In the last section we take a glimpse into the theory of (tempered) distributions.

5.1 The Fourier transform

Definition 5.1. *Let $f \in L^1(\mathbb{R}^d)$ and $\xi \in \mathbb{R}^d$. Then*

$$\hat{f}(\xi) = (\mathcal{F}f)(\xi) := \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-i\,\xi\cdot x} f(x)\,dx \tag{5.1}$$

is the Fourier transform *of f, where $\xi \cdot x := \sum_{k=1}^d \xi_k x_k$ for $\xi = (\xi_k)_k \in \mathbb{C}^d$ and $x = (x_k)_k \in \mathbb{C}^d$. Hence, $|x|_2^2 = x \cdot x$ for $x \in \mathbb{R}^d$.*

We set $\varphi(x,\xi) = e^{-i\,\xi\cdot x} f(x)$. Observe that $|\varphi(x,\xi)| = |f(x)|$ is integrable in $x \in \mathbb{R}^d$ for every $\xi \in \mathbb{R}^d$ and that $\mathbb{R}^d \ni \xi \mapsto \varphi(\xi, x)$ is continuous for a.e. $x \in \mathbb{R}^d$. Using a corollary to the theorem of dominated convergence, we thus conclude

$$\hat{f} \text{ is continuous on } \mathbb{R}^d \quad \text{and} \quad \|\hat{f}\|_\infty \le (2\pi)^{-\frac{d}{2}}\,\|f\|_1 \quad \text{for each } f \in L^1(\mathbb{R}^d). \tag{5.2}$$

Example 5.2. a) Let $d = 1$ and $f = \mathbb{1}_{[a,b]}$. We then have $\hat{f}(0) = (b-a)/\sqrt{2\pi}$ and

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-i\xi x}\,dx = \frac{i(e^{-ib\xi} - e^{-ia\xi})}{\sqrt{2\pi}\,\xi}\,, \qquad \xi \neq 0.$$

b) Let $d = 1$ and $f(x) = (1 + x^2)^{-1}$. Using complex curve integrals, it can be shown that $\hat{f}(\xi) = \sqrt{\pi/2}\,e^{-i\,|\xi|}$ for $\xi \in \mathbb{R}$, see also §9.7 in [Rud87].

In the above two examples (non-)rapid decay and (non-)smoothness on f correspond to (non-)smoothness and (non-)rapid decay of $\hat{f}$, respectively, cf. Lemma 5.7. In the next example, both f and $\hat{f}$ are smooth and decay rapidly.

c) Let $\gamma(x) = \exp(-\frac{1}{2}\,|x|_2^2)$ for $x \in \mathbb{R}^d$ be the standard Gaussian. We show that γ is a fix vector of the Fourier transform, i.e., $\widehat{\gamma} = \gamma$. Let $\xi \in \mathbb{R}^d$. Observe that

$\frac{1}{2}(x+i\xi)\cdot(x+i\xi) = \frac{1}{2}|x|_2^2 + i\xi\cdot x - \frac{1}{2}|\xi|_2^2$. We then obtain

$$\begin{aligned}\hat{\gamma}(\xi) &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-(i\xi\cdot x + \frac{1}{2}|x|_2^2)}\,dx = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{1}{2}|\xi|_2^2}\, e^{-\frac{1}{2}(x+i\xi)\cdot(x+i\xi)}\,dx \\ &= e^{-\frac{1}{2}|\xi|_2^2} \prod_{k=1}^d \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{1}{2}(x_k+i\xi_k)^2} dx_k = e^{-\frac{1}{2}|\xi|_2^2} \prod_{k=1}^d \frac{1}{\sqrt{2\pi}} \int_{i\xi_k+\mathbb{R}} e^{-\frac{1}{2}z^2} dz \\ &= e^{-\frac{1}{2}|\xi|_2^2} \prod_{k=1}^d \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{1}{2}t^2} dt = \gamma(\xi),\end{aligned}$$

using the formula $\int_{\mathbb{R}} e^{-\frac{1}{2}t^2}\,dt = \sqrt{2\pi}$ from Analysis 3. In the penultimate equality we shifted the path of integration in the complex plane. To justify this shift, we fix $\xi \in \mathbb{R}\setminus\{0\}$ and use the rectangular path Γ_n with vertices $-n$, n, $n+i\xi$ and $-n+i\xi$. Cauchy's theorem yields $\int_{\Gamma_n} e^{-\frac{1}{2}z^2}\,dz = 0$. The two vertical lines $S_n^{\pm}$ in Γ_n have length $|\xi|$, and on $S_n^{\pm}$ it holds

$$\left|e^{-\frac{1}{2}z^2}\right| = e^{-\frac{1}{2}\operatorname{Re}(\pm n+i\tau)^2} = e^{-\frac{1}{2}n^2}\, e^{\frac{1}{2}|\xi|^2},$$

where $0 \le |\tau| \le |\xi|$. Hence $\int_{S_n^{\pm}} e^{-\frac{1}{2}z^2}\,dz$ tends to 0 as $n\to\infty$, and above used equality is true. ♢

Let $f \in L^p(\mathbb{R}^d)$, $t, x \in \mathbb{R}^d$, and $1 \le p \le \infty$. To describe important mapping properties of $\mathcal{F}$, we set

$$e_t(x) = e^{i\,t\cdot x},$$

and introduce the *translation* operator T_t by

$$(T_t f)(x) = f(x+t).$$

As in Example 3.13 one sees that $T_t : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$ is an isometric isomorphism with inverse T_{-t}. For $a > 0$ we further define the *dilation* operator D_a by

$$(D_a f)(x) = f(ax).$$

Observe that $D_{1/a}D_a = D_a D_{1/a} = I$, and that the substitution $y = ax$ yields

$$\|D_a f\|_p^p = \int_{\mathbb{R}^d} |f(ax)|^p\,dx = \int_{\mathbb{R}^d} a^{-d}\,|f(y)|^p\,dy = a^{-d}\,\|f\|_p^p$$

for $p < \infty$ (and analogously for $p = \infty$). As a result, $a^{-d/p}D_a : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$ is an isometric isomorphism. Finally, also the *reflection* operator

$$R : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d); \quad Rf(x) = f(-x),$$

is an isometric isomorphism with $R^2 = I$, i.e., $R^{-1} = R$.

Proposition 5.3. *Let $f, g \in L^1(\mathbb{R}^d)$, $t \in \mathbb{R}^d$ and $a > 0$. The following formulas hold.*

a) $\mathcal{F}(T_t f) = e_t \hat{f}$.
b) $\mathcal{F}(e_t f) = T_{-t}\hat{f}$.
c) $\mathcal{F}(D_a f) = a^{-d}\, D_{\frac{1}{a}} \hat{f}$.
d) $\mathcal{F}(f * g) = (2\pi)^{\frac{d}{2}}\, \hat{f}\hat{g}$.

Proof. Let $f, g \in L^1(\mathbb{R}^d)$, $t \in \mathbb{R}^d$, $a > 0$, and $\xi \in \mathbb{R}^d$. Using the substitutions $y = x+t$ and $z = ax$, we check the assertions a), b) and c) by the calculations

$$\mathcal{F}(T_t f)(\xi) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-i\xi\cdot x} f(x+t)\,dx = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-i\xi\cdot(y-t)} f(y)\,dy = e^{i\xi\cdot t}\hat{f}(\xi),$$

$$\mathcal{F}(e_t f)(\xi) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-i\xi\cdot x} e^{it\cdot x} f(x)\,dx = \hat{f}(\xi - t),$$
$$\mathcal{F}(D_a f)(\xi) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-i\xi\cdot x} f(ax)\,dx = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} a^{-d}\, e^{-i\,\frac{1}{a}\xi\cdot z} f(z)\,dz = a^{-d} \hat{f}(\tfrac{1}{a}\xi).$$

To prove d), we first recall from (the proof of) Proposition 2.15 that $f * g \in L^1(\mathbb{R}^d)$ and that the map $\mathbb{R}^{2d} \ni (x,y) \mapsto f(y-x)g(x)$ is integrable. Hence, also the function $(x,y) \mapsto e^{-i\,\xi\cdot y} f(y-x)g(x)$ belongs to $L^1(\mathbb{R}^{2d})$. Fubini's theorem thus yields

$$\begin{aligned}
\mathcal{F}(f*g)(\xi) &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{-i\,\xi\cdot y} f(y-x)g(x)\,dx\,dy \\
&= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{-i\,\xi\cdot(y-x)} f(y-x) e^{-i\,\xi\cdot x} g(x)\,dy\,dx \\
&= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{-i\,\xi\cdot z} f(z)\,dz\; e^{-i\,\xi\cdot x} g(x)\,dx = (2\pi)^{\frac{d}{2}} \hat{f}(\xi)\hat{g}(\xi),
\end{aligned}$$

where we also employed the subtitution $z = y - x$ in one of y–integrals. □

Example 5.4. We set $f(x) = \exp(-\frac{a}{2}|x-v|^2)$ for all $x \in \mathbb{R}^d$ and some $a > 0$ and $v \in \mathbb{R}^d$. The Fourier transform of this Gaussian function is given by $\hat{f}(\xi) = a^{-d/2}\exp(-i\,v\cdot\xi)\exp(-\frac{1}{2a}|\xi|_2^2)$ for all $\xi \in \mathbb{R}^d$. In fact, we have $f = T_{-v} D_{\sqrt{a}}\gamma$. Proposition 5.3 and Example 5.2 thus yield

$$\hat{f} = e_{-v}\mathcal{F}(D_{\sqrt{a}}\gamma) = e_{-v} a^{-\frac{d}{2}} D_{\frac{1}{\sqrt{a}}}\hat{\gamma} = a^{-\frac{d}{2}} e_{-v} D_{\frac{1}{\sqrt{a}}}\gamma,$$

as asserted. ♢

As one of its main properties, the Fourier transform maps derivatives into multiplication by polynomials, and vice versa. To state this fact concisely, we use the *multi-index notation*: For $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ and $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, we set

$$|\alpha| := \alpha_1 + \cdots + \alpha_n, \qquad x^\alpha := x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}, \qquad \partial^\alpha := \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.$$

We further denote the function $\mathbb{R}^d \ni x \to x^\alpha f(x)$ by $x^\alpha f$. Observe that

$$|x^\alpha| = |x_1|^{\alpha_1} \cdot \ldots \cdot |x_n|^{\alpha_n} \le |x|_2^{|\alpha|} \le 1 + |x|_2^m$$

for $x \in \mathbb{R}^n$ and $|\alpha| \le m$.

To relate the Fourier transform with derivatives we need a space of smooth functions. Unfortunately, the space $C_c^\infty(\mathbb{R}^d)$ is not invariant under the Fourier transform. Instead one uses the (somewhat less convenient) 'Schwartz space' on which $\mathcal{F}$ becomes a bijection, as seen below.

Definition 5.5. *For $f \in C^\infty(\mathbb{R}^d)$, $m \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^d$, we set*

$$p_{m,\alpha}(f) = \sup_{x\in\mathbb{R}^d} |x|_2^m\,|\partial^\alpha f(x)|.$$

We define the Schwartz space $\mathcal{S}_d$ *by*

$$\mathcal{S}_d = \{f \in C^\infty(\mathbb{R}^d) \,|\, p_{m,\alpha}(f) < \infty \;\; \textit{for all} \;\; m \in \mathbb{N}_0,\ \alpha \in \mathbb{N}_0^d\}.$$

Notice that $\mathcal{S}_d$ is a vector space and that all derivatives of $f \in \mathcal{S}_d$ decay faster than $|x|_2^{-m}$ for any $m \in \mathbb{N}$, as $|x|_2 \to \infty$. One thus calls $f \in \mathcal{S}_d$ *rapidly decreasing.* It is straightforward to check that the function $\gamma(x) = e^{-|x|_2^2/2}$ belongs to $\mathcal{S}_d$. Moreover, one can replace $p_{m,\alpha}(f)$ by $p_{2m,\alpha}(f)$ in the definition of $\mathcal{S}_d$ without changing $\mathcal{S}_d$.

Remark 5.6. a) Let $f \in \mathcal{S}_d$, $m \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^d$. We estimate

$$\begin{aligned}|x|_2^m\,|\partial^\alpha f(x)| &= (1+|x|_2^{d+1})^{-1}\,(|x|_2^m + |x|_2^{m+d+1)})\,|\partial^\alpha f(x)|\\ &\le (1+|x|_2^{d+1})^{-1}\,(p_{m,\alpha}(f) + p_{d+m+1,\alpha}(f))\end{aligned}$$

for all $x \in \mathbb{R}^d$. Since the function $x \mapsto (1+|x|_2^{d+1})^{-1}$ is integrable on $\mathbb{R}^d$ (see Analysis 3), we deduce $|x|_2^m\partial^\alpha f \in L^1(\mathbb{R}^d)\cap C_0(\mathbb{R}^d)$, and hence $|x|_2^m\partial^\alpha f$ belongs to $L^p(\mathbb{R}^d)$ for all $p \in [1,\infty]$ by (2.3).

b) Because $C_c^\infty(\mathbb{R}^d) \subset \mathcal{S}_d \subset L^p(\mathbb{R}^d)$, Proposition 3.14 yields that $\mathcal{S}_d$ is dense in $L^p(\mathbb{R}^d)$ for every $p \in [1,\infty)$.

c) Observe that $p_{m,\alpha}$ is a seminorm on $\mathcal{S}_d$ for all $m \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^d$, where p_{00} is the supnorm. We order these seminorms as a sequence $(p_j)_{j\in\mathbb{N}}$. Due to Proposition 1.8, the Schwartz space $\mathcal{S}_d$ has the metric

$$d(f,g) = \sum_{j=1}^\infty 2^{-j}\,\frac{p_j(f-g)}{1+p_j(f-g)},$$

and $d(f_n,f) \to 0$ as $n\to\infty$ if and only if $p_{m,\alpha}(f-f_n) \to 0$ as $n\to\infty$ for all $m \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^d$. One can verify that $\mathcal{S}_d$ is complete for this metric, cf. Example 1.9. $\Diamond$

The next lemma deals with the announced relation between Fourier transform and derivatives. We use the *Laplace operator* given by $\Delta = \partial_1^2 + \cdots + \partial_d^2$ and the space of smooth, polynomially bounded functions

$$\mathcal{E}_d = \{f \in C^\infty(\mathbb{R}^d)\,|\,\forall\,\alpha \in \mathbb{N}_0^d\ \ \exists\, n_\alpha \in \mathbb{N}_0 : \sup_{|x|_2\ge 1} |x|_2^{-n_\alpha}\,|\partial^\alpha f(x)| < \infty\}.$$

Note that Schwartz functions and polynomials belong to $\mathcal{E}_d$.

Lemma 5.7. *Let $f \in \mathcal{S}_d$, $g \in \mathcal{E}_d$,and $\alpha \in \mathbb{N}_0^d$. Then the following assertions hold.*

a) $\widehat{f} \in C^\infty(\mathbb{R}^d)$, $\partial^\alpha \widehat{f} = (-i)^{|\alpha|}\,\mathcal{F}(x^\alpha f)$, $\mathcal{F}(\partial^\alpha f) = i^{|\alpha|}\,\xi^\alpha\,\widehat{f}$.

b) $\mathcal{F}\Delta f = \mathcal{F}\partial_1^2 f + \cdots + \mathcal{F}\partial_d^2 f = i^2(\xi_1^2 + \cdots \xi_d^2)\mathcal{F}f = -|\xi|_2^2\mathcal{F}f$.

c) The mappings $f \mapsto gf$ and $f \mapsto \partial^\alpha f$ are continuous from $\mathcal{S}_d$ to $\mathcal{S}_d$.

d) The Fourier transform is continuous from $\mathcal{S}_d$ to $\mathcal{S}_d$.

Proof. Let $\xi, x \in \mathbb{R}^d$, $f, f_n \in \mathcal{S}_d$, $g \in \mathcal{E}_d$, $\alpha \in \mathbb{N}_0^d$, and $m \in \mathbb{N}_0$.

We show a) for $\alpha = e_k$, the assertion then follows by induction. There exists

$$\frac{\partial}{\partial \xi_k}\,e^{-i\,\xi\cdot x} f(x) = -ix_k e^{-i\,\xi\cdot x} f(x) =: \varphi_k(\xi,x),$$

and $\mathbb{R}^d \ni x \mapsto |\varphi_k(\xi,x)| = |x_k f(x)|$ is integrable by Remark 5.6. A corollary of the theorem of dominated convergence thus shows that

$$\exists\ \frac{\partial}{\partial \xi_k}\,\hat{f}(\xi) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} -ix_k e^{-i\,\xi\cdot x} f(x)\,dx = -i\mathcal{F}(x_k f)(\xi).$$

For the second part we write $[-n,n]^k = C_n^k$ and $x = (x', x_k)$ with $x' = (x_1,\dots,x_{k-1},x_{k+1},\dots,x_d) \in \mathbb{R}^{d-1}$. Using that $\partial_k f \in L^1(\mathbb{R}^d)$ and integrating by parts in x_k, we compute

$$\begin{aligned}\mathcal{F}(\partial_k f)(\xi) &= (2\pi)^{-\frac{d}{2}} \lim_{n\to\infty} \int_{C_n^{d-1}} \int_{-n}^{n} e^{-i\,\xi\cdot(x',x_k)} \partial_k f(x',x_k)\,dx_k\,dx'\\ &= (2\pi)^{-\frac{d}{2}} \lim_{n\to\infty} \Big[\int_{C_n^d} i\xi_k e^{-i\,\xi\cdot x} f(x)\,dx + \int_{C_n^{d-1}} e^{-i\,\xi\cdot x} f(x',x_k)\Big|_{x_k=-n}^{n} dx'\Big]\\ &= i\xi_k\,\widehat{f}(\xi).\end{aligned}$$

Here the second integral J_n in the second line tends to 0 as $n \to \infty$ since

$$|J_n| \le \sum_{N=\pm n} \int_{C_n^{d-1}} |(x',N)|_2^{-d}\, |(x',N)|_2^d\, |f(x',N)|\, dx' \le 2^d n^{d-1} n^{-d} p_{d,0}(f).$$

Assertion b) is a consequence of a). Thanks to Leibniz' rule, the function $|x|_2^m \partial^\alpha(fg)$ is a linear combination of terms $|x|_2^{m+n_\gamma}(\partial^\beta f)\, |x|_2^{-n_\gamma} \partial^\gamma g$ for all $\beta, \gamma \in \mathbb{N}_0^d$ with $\beta + \gamma = \alpha$. Since $g \in \mathcal{E}_d$, we obtain $p_{m,\alpha}(fg) \le c\,(p_{m+k,\alpha}(f) + p_{m,\alpha}(f))$, where $k = \max_{|\gamma| \le |\alpha|} n_\gamma$ and c only depends on m, α, d and g. Hence, fg belongs to $\mathcal{S}_d$. The asserted continuity follows by replacing f by $f - f_n$. Similarly, one checks the second part of c). Using a) and b), we further compute

$$|\xi|_2^{2m} \partial^\alpha \widehat{f} = (-i)^{|\alpha|}(\xi_1^2 + \cdots + \xi_d^2)^m \mathcal{F}(x^\alpha f) = (-i)^{|\alpha|}(-i)^{2m} \mathcal{F}(\Delta^m(x^\alpha f)).$$

Due to c) and Remark 5.6, the function $\Delta^m(x^\alpha f)$ belongs to $\mathcal{S}_d \subseteq L^1(\mathbb{R}^d)$ so that its Fourier transform is bounded by (5.2). This means that $\widehat{f} \in \mathcal{S}_d$ and

$$p_{2m,\alpha}(\mathcal{F}(f - f_n)) \le c\, \|\Delta^m(x^\alpha(f - f_n))\|_1.$$

Using again c) and Remark 5.6, the term on the right hand side can be bounded by a linear combination of certain seminorms $p_{k,\beta}(f - f_n)$; i.e., $\mathcal{F}_d : \mathcal{S}_d \to \mathcal{S}_d$ is continuous. □

Corollary 5.8 (Riemann-Lebesgue lemma)**.** *If $f \in L^1(\mathbb{R}^d)$, then $\hat{f} \in C_0(\mathbb{R}^d)$. Hence, $\mathcal{F} \in \mathcal{L}(L^1(\mathbb{R}^d), C_0(\mathbb{R}^d))$.*

Proof. Let $f \in L^1(\mathbb{R}^d)$. Due to Remark 5.6, there are $f_n \in \mathcal{S}_d$ converging to f in $L^1(\mathbb{R}^d)$. Lemma 5.7 yields $\hat{f}_n \in \mathcal{S}_d \subset C_0(\mathbb{R}^d)$. By (5.2), the functions $\hat{f}_n$ converge to $\hat{f}$ in supnorm so that $\hat{f} \in C_0(\mathbb{R}^d)$. The second assertion then follows from (5.2). □

The next lemma is the crucial step towards the main results of this section. Observe that in its second part a double integral disappears due to cancellations of the highly oscillating integrands.

Lemma 5.9. *The following assertions hold.*

a) $\int_{\mathbb{R}^d} \hat{f} g\, dx = \int_{\mathbb{R}^d} f \hat{g}\, dx$ *for all* $f, g \in \mathcal{S}_d$.

b) $\mathcal{F}^2 = R$, *i.e.,* $(\mathcal{F}\mathcal{F}f)(x) = f(-x)$ *for all* $f \in \mathcal{S}_d$ *and* $x \in \mathbb{R}^d$.

Proof. Let $f, g \in \mathcal{S}_d$. Since $(x,y) \mapsto e^{-i\, y \cdot x} f(x) g(y)$ is integrable on $\mathbb{R}^{2d}$, Fubini's theorem yields

$$\begin{aligned}\int_{\mathbb{R}^d} \hat{f}(y) g(y)\, dy &= \int_{\mathbb{R}^d} (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-i\, y\cdot x} f(x) g(y)\, dx\, dy \\ &= \int_{\mathbb{R}^d} f(x)\, (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-i\, y\cdot x} g(y)\, dy\, dx = \int_{\mathbb{R}^d} f(x) \hat{g}(x)\, dx.\end{aligned}$$

In the second assertion one is led to the integrand $e^{-i\,\xi\cdot y} e^{-i\, y\cdot x} f(x)$ which is not integrable for $(y,x) \in \mathbb{R}^{2d}$. So Fubini's theorem does not apply directly, and one has to use a regularization. To that purpose, fix $\xi \in \mathbb{R}^d$ and $a > 0$. Set $h_a = e_{-\xi} D_a \gamma \in \mathcal{S}_d$, i.e., $h_a(y) = e^{-i\,\xi\cdot y} \exp(-\frac{a^2}{2}|y|_2^2)$ for $y \in \mathbb{R}^d$. Due to the theorem of dominated convergence with the majorant $|\hat{f}|$, the integral

$$J_a := \int_{\mathbb{R}^d} \hat{f}(x) h_a(x)\, dx = \int_{\mathbb{R}^d} \hat{f}(y) e^{-i\,\xi\cdot y} \gamma(ay)\, dy$$

converges to $(2\pi)^{\frac{d}{2}}(\mathcal{F}\hat{f})(\xi)$ as $a \to 0$. On the other hand, part a), Proposition 5.3 and Example 5.2 imply that

$$\begin{aligned}J_a &= \int_{\mathbb{R}^d} f(x)\,\mathcal{F}(e_{-\xi}D_a\gamma)(x)\,dx = \int_{\mathbb{R}^d} f(x)\,a^{-d}(T_\xi D_{\frac{1}{a}}\gamma)(x)\,dx\\ &= \int_{\mathbb{R}^d} f(x)\,a^{-d}\gamma(\tfrac{1}{a}(x+\xi))\,dx = \int_{\mathbb{R}^d} f(az-\xi)\gamma(z)\,dz,\end{aligned}$$

where we also use the substitution $z = \frac{1}{a}(x+\xi)$. We can now apply the theorem of dominated convergence with the majorant $\|f\|_\infty\gamma$ to conclude that $J_a \to f(-\xi)\|\gamma\|_1 = (2\pi)^{d/2}f(-\xi)$ as $a \to 0$, which shows assertion b). □

Proposition 5.10. *The Fourier transform $\mathcal{F} : \mathcal{S}_d \to \mathcal{S}_d$ is bijective with $\mathcal{F}^4 = I$. For all $f, g \in \mathcal{S}_d$ and $x \in \mathbb{R}^d$, we have*

$$\mathcal{F}^{-1}g(x) = \frac{1}{(2\pi)^{\frac{d}{2}}}\int_{\mathbb{R}^d} e^{i\,x\cdot\xi}g(\xi)\,d\xi, \tag{5.3}$$

$$(\mathcal{F}f|\mathcal{F}g)_{L^2} = (f|g)_{L^2}, \tag{5.4}$$

$$f * g \in \mathcal{S}_d, \tag{5.5}$$

$$\mathcal{F}(fg) = (2\pi)^{-\frac{d}{2}}\hat{f} * \hat{g}. \tag{5.6}$$

Proof. Lemma 5.9 shows that $I = R^2 = \mathcal{F}^4 = \mathcal{F}\mathcal{F}^3 = \mathcal{F}^3\mathcal{F}$ on $\mathcal{S}_d$ so that $\mathcal{F} : \mathcal{S}_d \to \mathcal{S}_d$ has the inverse $\mathcal{F}^3 = R\mathcal{F}$. This fact already gives (5.3). Let $f, g \in \mathcal{S}_d$ and $x \in \mathbb{R}^d$. Equation (5.3) further yields

$$\begin{aligned}\mathcal{F}(\overline{\mathcal{F}g})(\xi) &= (2\pi)^{-\frac{d}{2}}\int_{\mathbb{R}^d} e^{-i\,x\cdot\xi}\overline{\hat{g}(\xi)}\,d\xi = (2\pi)^{-\frac{d}{2}}\overline{\int_{\mathbb{R}^d} e^{i\,x\cdot\xi}\hat{g}(\xi)\,d\xi}\\ &= \overline{(\mathcal{F}^{-1}\mathcal{F}g)(x)} = \overline{g(x)}.\end{aligned}$$

So we can deduce from Lemma 5.9a) that

$$(\hat{f}|\hat{g})_{L^2} = \int_{\mathbb{R}^d} \hat{f}\,\overline{\hat{g}}\,d\xi = \int_{\mathbb{R}^d} f\,\mathcal{F}(\overline{\mathcal{F}g})\,dx = \int_{\mathbb{R}^d} f\overline{g}\,dx = (f|g)_{L^2}.$$

For the final two assertions, Proposition 5.3 and Lemma 5.7 imply that $\mathcal{F}(f*g) = (2\pi)^{d/2}\hat{f}\hat{g} =: \varphi$ belongs to $\mathcal{S}_d$. This fact yields $f * g = \mathcal{F}^{-1}\varphi \in \mathcal{S}_d$ and further

$$\mathcal{F}(\hat{f} * \hat{g}) = (2\pi)^{d/2}\mathcal{F}^2(f)\mathcal{F}^2(g) = (2\pi)^{d/2}R(fg) = (2\pi)^{d/2}\mathcal{F}^2(fg)$$

using that $R = \mathcal{F}^2$. If we apply $\mathcal{F}^{-1}$, we derive (5.6). □

The equality (5.4) shows that $\|\mathcal{F}f\|_2 = \|f\|_2$ for all $f \in \mathcal{S}_d$. Since $\mathcal{S}_d$ is dense in $L^2(\mathbb{R}^d)$ by Remark 5.6, we can extend $\mathcal{F}$ to an isometry $\mathcal{F}_2 : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ which is also called *Fourier transform* (use Lemma 2.13). Let $f \in L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$. In Proposition 3.14 we have constructed functions $f_n \in C_c^\infty(\mathbb{R}^d) \subset \mathcal{S}_d$ which converge to f in $L^2(\mathbb{R}^d)$ and in $L^1(\mathbb{R}^d)$. Since $\mathcal{F}f_n \to \mathcal{F}_2 f$ in $L^2(\mathbb{R}^d)$, there is a subsequence $\mathcal{F}f_{n_j}$ converging to $\mathcal{F}_2 f$ a.e. as $j \to \infty$ due to Riesz–Fischer. On the other hand, $\mathcal{F}f_{n_j}$ converges uniformly to $\mathcal{F}f$ by (5.2). Thus, $\mathcal{F}_2 f = \mathcal{F}f$ a.e.. We now write $\mathcal{F} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ instead of $\mathcal{F}_2$, and also $\mathcal{F}_2 f = \hat{f}$.

Warning: $\mathcal{F}f$ is **not** given by the formula (5.1) if $f \in L^2(\mathbb{R}^d) \setminus L^1(\mathbb{R}^d)$.

In the next theorem we collect the main properties of $\mathcal{F}$ on $L^2(\mathbb{R}^d)$, except for its behavior under derivatives which will be dealt with in the following two sections.

Theorem 5.11. *The Fourier transform on $\mathcal{S}_d$ extends to a unitary operator $\mathcal{F} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ which is given by (5.1) on $L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$.*

Let $f, g \in L^2(\mathbb{R}^d)$, $h \in L^1(\mathbb{R}^d)$, $\varphi = \mathcal{F}^{-1}\psi$ for some $\psi \in L^1(\mathbb{R}^d)$, $t \in \mathbb{R}^d$, and $a > 0$. Then the following assertions hold.

a) $\mathcal{F}^2 = R$, $\mathcal{F}^4 = I$, $\mathcal{F}^{-1} = \mathcal{F}^3 = R\mathcal{F}$.

b) $\mathcal{F}^{-1}h(x) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{i\,x\cdot\xi} h(\xi)\, d\xi$ *for* $x \in \mathbb{R}^d$ *(inversion formula).*

c) $(\mathcal{F}f|\mathcal{F}g)_{L^2} = (f|g)_{L^2}$ *(Plancherel identity).*

d) $\int_{\mathbb{R}^d} \hat{f} g\, dx = \int_{\mathbb{R}^d} f\hat{g}\, dx$ *(jumping hat lemma).*

e) $\mathcal{F}(T_t f) = e_t \hat{f}$, $\mathcal{F}(e_t f) = T_{-t}\hat{f}$, $\mathcal{F}(D_a f) = a^{-d} D_{\frac{1}{a}} \hat{f}$.

f) $\mathcal{F}(h * f) = (2\pi)^{\frac{d}{2}}\, \hat{h}\hat{f}$, $\mathcal{F}(\varphi f) = (2\pi)^{-\frac{d}{2}}\, \hat{\varphi} * \hat{f}$ *(convolution theorem).*

g) $\mathcal{F} \in \mathcal{L}(L^p(\mathbb{R}^d), L^{p'}(R^d))$ *for* $p \in [1, 2]$ *(Hausdorff–Young inequality).*

Proof. Recall that $\mathcal{F}$ is an isometry on $L^2(\mathbb{R}^d)$. The equations $\mathcal{F}^2 = R$, $\mathcal{F}^4 = I$, and those in c)–e) hold on the dense subspace $\mathcal{S}_d$ as shown in Proposition 5.3, Lemma 5.9 and Proposition 5.10. Since the maps $\mathcal{F}$, R, T_t, D_a, $f \mapsto e_t f$ and the scalar product are continuous from $L^2(\mathbb{R}^d)$, resp. from $L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$, to $L^2(\mathbb{R}^d)$, these identities can be extended to $L^2(\mathbb{R}^d)$ by approximation.

The equation $\mathcal{F}^4 = I$ implies $I = \mathcal{F}\mathcal{F}^3 = \mathcal{F}^3\mathcal{F}$ so that $\mathcal{F} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ has the inverse $\mathcal{F}^{-1} = \mathcal{F}^3 = R\mathcal{F}$ which also yields b). As a bijective isometry on a Hilbert space, $\mathcal{F}$ is unitary by Proposition 4.62.

For f), note that $f \mapsto h * f$ is continuous on $L^2(\mathbb{R}^d)$ by Proposition 2.15 (Young's inequality) and $h \in L^1(\mathbb{R}^d)$. Hence the first part of f) can be deduced from Proposition 5.3 by approximation. For the second part, take $f_n, \psi_n \in \mathcal{S}_d$ so that $f_n \to f$ in $L^2(\mathbb{R}^d)$ and $\psi_n \to \psi$ in $L^1(\mathbb{R}^d)$ as $n \to \infty$. Then $\varphi_n := \mathcal{F}^{-1}\psi_n$ tends to $\varphi = \mathcal{F}^{-1}\psi$ in $L^\infty(\mathbb{R}^d)$ by (5.2). Hence, $\varphi_n f_n$ converges to φf and $\psi_n * \hat{f}_n$ to $\psi * \hat{f} = \hat{\varphi} * \hat{f}$ in $L^2(\mathbb{R}^d)$. Equation (5.6) now implies the second part of f).

Assertion g) is a consequence of the Riesz–Thorin Theorem 2.27 with $\theta = 2/p' \in [0, 1]$, since $\mathcal{F} \in \mathcal{L}(L^1(\mathbb{R}^d), L^\infty(R^d))$ and $\mathcal{F} \in \mathcal{L}(L^2(\mathbb{R}^d))$. □

Let $f \in L^2(\mathbb{R}^d)$. Set $f_n = \mathbb{1}_{B(0,n)} f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ for $n \in \mathbb{N}$. Since $|f_n| \le |f| \in L^2(\mathbb{R}^d)$, we have $f_n \to f$ in $L^2(\mathbb{R}^d)$ as $n \to \infty$ by the theorem of dominated convergence. Therefore, $\hat{f}_n \to \hat{f}$ in $L^2(\mathbb{R}^d)$ as $n \to \infty$, where

$$\hat{f}_n(\xi) = (2\pi)^{-\frac{d}{2}} \int_{B(0,n)} e^{-i\xi\cdot x} f(x)\, dx, \qquad \xi \in \mathbb{R}^d,$$

is the *truncated Fourier transform.*

Let $J \subseteq \mathbb{R}$ be an interval and X be a Banach space. A function $u : J \to X$ is differentiable at $t \in J$ if the limit

$$u'(t) := \lim_{h\to 0, t+h\in J} \tfrac{1}{h}(u(t+h) - u(t))$$

exists in X. Then u is continuous at t. We write $u \in C^1(J, X)$ if $u'(t)$ exists for all $t \in J$ and $u' : J \to X$ is continuous.

Example 5.12. We consider the diffusion equation

$$\begin{aligned} \partial_t u(t,x) &= \Delta u(t,x), \qquad t \ge 0,\ x \in \mathbb{R}^d, \\ u(0,x) &= u_0(x), \qquad x \in \mathbb{R}^d, \end{aligned} \tag{5.7}$$

for a given initial value $u_0 \in \mathcal{S}_d$. Let us assume for moment that we have a function $u : \mathbb{R}_+ \times \mathbb{R}^d \mapsto \mathbb{C}$ such that the function $u(t) : x \mapsto u(t,x)$ belongs to $\mathcal{S}_d$ for all $t \ge 0$, $u \in C^1(\mathbb{R}_+, L^2(\mathbb{R}^d))$ and u satisfies (5.7). We set $\hat{u}(t,\xi) = (\mathcal{F}u(t))(\xi)$ for all $t \ge 0$ and $\xi \in \mathbb{R}^d$. Since $\mathcal{F}$ is continuous on $L^2(\mathbb{R}^d)$, we have

$$\mathcal{F}u'(t) = \mathcal{F} \lim_{h\to 0} \tfrac{1}{h}(u(t+h) - u(t)) = \lim_{h\to 0} \tfrac{1}{h}(\widehat{u}(t+h) - \widehat{u}(t))$$

for all $t \geq 0$, so that $\widehat{u} \in C^1(\mathbb{R}_+, L^2(\mathbb{R}^d))$ and $\partial_t \widehat{u} = \mathcal{F}\partial_t u$. Applying $\mathcal{F}$ to (5.7), we then deduce from Lemma 5.7 that

$$\partial_t \hat{u}(t) = \mathcal{F}\partial_t u(t) = \mathcal{F}\Delta u(t) = -|\xi|_2^2 \, \hat{u}(t).$$

For each fixed $\xi \in \mathbb{R}^d$ we thus arrive at the ordinary differential equation

$$\begin{aligned} \partial_t \hat{u}(t,\xi) &= -|\xi|_2^2 \, \hat{u}(t,\xi), \qquad t \geq 0, \ \xi \in \mathbb{R}^d, \\ \hat{u}(0,\xi) &= \widehat{u_0}(\xi), \qquad \xi \in \mathbb{R}^d, \end{aligned}$$

which has the solution

$$\hat{u}(t,\xi) = e^{-t\,|\xi|_2^2} \widehat{u_0}(\xi) = (D_{\sqrt{2t}}\gamma)(\xi)\, \widehat{u_0}(\xi).$$

Using the properties of $\mathcal{F}$, we compute

$$u(t) = \mathcal{F}^{-1}[(\mathcal{F}\mathcal{F}^{-1}D_{\sqrt{2t}}\gamma)\widehat{u_0}] = (2\pi)^{-\frac{d}{2}}\,(\mathcal{F}^{-1}D_{\sqrt{2t}}\gamma) * u_0 = (4\pi t)^{-\frac{d}{2}}\,(D_{1/\sqrt{2t}}\gamma) * u_0$$

for all $t > 0$. This gives

$$u(t,x) = \int_{\mathbb{R}^d} (4\pi t)^{-\frac{d}{2}} \, e^{-\frac{|x-y|_2^2}{4t}} u_0(x)\, dx =: (\gamma_t * u_0)(x)$$

for all $t > 0$ and $x \in \mathbb{R}^d$.

Define u by this equation. Since $\gamma_t \in \mathcal{S}_d$, we can reverse the above arguments (or check directly) that this u solves (5.7) and has the asserted regularity. As a result, we have found a unique solution of (5.7) in the class indicated above.

Moreover, $\|u(t)\| = \|\hat{u}(t)\|_2 \leq \|\widehat{u_0}\|_2 = \|u_0\|_2$ since the Fourier transform is an isometry and $|D_{\sqrt{2t}}\gamma| \leq 1$. ◊

The above result is restricted to initial values u_0 in $\mathcal{S}_d$. At the end of the next section we solve the diffusion equation on $\mathbb{R}^d$ for $u_0 \in L^2(\mathbb{R}^d)$ again using the Fourier transform. To that purpose we have to extend the differentiation formulas in Lemma 5.7a) to a much larger class of functions, the Sobolev spaces.

5.2 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 some basic results on the spatial domain $U = \mathbb{R}^d$. For an introduction to this area we refer to the books [Bre11] and [Dob06], and also to [ST].

To motivate the definition below, we consider $u \in C^1(\mathbb{R}^d)$. Take any $\varphi \in C_c^\infty(\mathbb{R}^d)$. There is an $a > 0$ such that $\operatorname{supp}\varphi \subseteq (-a,a)^d$. Integrating by parts, we obtain

$$\begin{aligned} \int_{\mathbb{R}^d} \partial_1 u \, \varphi \, dx &= \int_{C_a} \int_{-a}^{a} \partial_1 u(x_1,x')\, \varphi(x_1,x')\, dx_1\, dx' \\ &= \int_{C_a} \Big[-\int_{-a}^{a} u(x_1,x')\, \partial_1\varphi(x_1,x')\, dx_1 + \Big[u(x_1,x')\varphi(x_1,x') \Big]_{x_1=-a}^{a} \Big] dx' \\ &= -\int_{\mathbb{R}^d} u\, \partial_1\varphi\, dx, \end{aligned} \tag{5.8}$$

where $C_a = [-a,a]^{d-1}$. Other partial derivatives can be treated in the same way. The following definition now relies on the observation that the right hand side of (5.8) is defined for all u in the vector space

$$L^1_{\text{loc}}(\mathbb{R}^d) = \{ f : \mathbb{R}^d \to \mathbb{C} \mid f \text{ is measurable, } f_{|K} \in L^1(K) \text{ for all compact } K \subseteq \mathbb{R}^d \}.$$

It thus can serve as definition of $\partial_1 u$ on the left hand side of (5.8). Observe that $L^p(\mathbb{R}^d) \subseteq L^1_{\text{loc}}(\mathbb{R}^d)$ for all $1 \leq p \leq \infty$.

Definition 5.13. *Let $u \in L^1_{loc}(\mathbb{R}^d)$, $\alpha \in \mathbb{N}_0^d$, $k \in \mathbb{N}^d$ and $1 \le p \le \infty$. If there is a function $f \in L^1_{loc}(\mathbb{R}^d)$ such that*

$$\int_{\mathbb{R}^d} f\varphi\,dx = (-1)^{|\alpha|} \int_{\mathbb{R}^d} u\,\partial^\alpha \varphi\,dx,$$

for all $\varphi \in C_c^\infty(\mathbb{R}^d)$, then $f =: \partial^\alpha u$ is called α-th weak derivative *of u. We write D_α for the space of $u \in L^1_{loc}(\mathbb{R}^d)$ such that a weak derivative $\partial^\alpha u \in L^1_{loc}(\mathbb{R}^d)$ exists. One further defines the* Sobolev spaces *by*

$$W^{k,p}(\mathbb{R}^d) = \{u \in L^p(\mathbb{R}^d) \,|\, u \in D_\alpha \text{ and } \partial^\alpha u \in L^p(\mathbb{R}^d) \text{ for all } |\alpha| \le k\}$$

and endows them with

$$\|u\|_{k,p} = \begin{cases} \Big(\sum\limits_{0 \le |\alpha| \le k} \|\partial^\alpha u\|_p^p \Big)^{\frac{1}{p}}, & 1 \le p < \infty, \\ \max\limits_{0 \le |\alpha| \le k} \|\partial^\alpha u\|_\infty, & p = \infty, \end{cases}$$

where $\partial^0 f := f$. We write ∂_j instead of ∂^{e_j}.

As usually, the spaces $L^1_{\text{loc}}(\mathbb{R}^d)$ and $W^{k,p}(\mathbb{R}^d)$ are spaces of equivalence classes modulo the space of null functions $\mathcal{N}$. In the above concepts one can replace $\mathbb{R}^d$ by any nonempty open subset U.

The above definition first raises the question whether the weak derivative is uniquely determined. Already to settle this basic question one needs the mollifier G_ε from Proposition 3.14 whose properties are used below intensively.

Lemma 5.14. *Let $f \in L^1_{loc}(\mathbb{R}^d)$ satisfy*

$$\int_{\mathbb{R}^d} f\varphi\,dx = 0 \qquad \text{for all } \varphi \in C_c^\infty(U).$$

Then $f = 0$ a.e.. Hence, a function $u \in D_\alpha$ has exactly one α-th weak derivative.

Proof. Assume that $f \neq 0$ on a Borel set B with $\lambda(B) > 0$. By Satz IV.3.15 of [Wer06], there is a compact set $K \subseteq B$ with $\lambda(K) > 0$. The set $K_1 = K + \overline{B}(0,1)$ is also compact. Proposition 3.14 then shows that $\varphi = G_1 \mathbb{1}_{K_1}$ belongs to $C_c^\infty(\mathbb{R}^d)$. Moreover, the definition (3.1) of G_ε yields

$$\varphi(x) = \int_{B(x,1)} k_1(x-y)\,\mathbb{1}_{K_1}(y)\,dy = \int_{B(x,1)} k_1(x-y)\,dy = 1$$

for all $x \in K$. Since $\varphi f \in L^1(\mathbb{R}^d)$, the functions $G_\varepsilon(\varphi f)$ converge to φf in $L^1(\mathbb{R}^d)$ as $\varepsilon \to 0$, due to Proposition 3.14. There thus exist a nullset N and a subsequence $\varepsilon_j \to 0$ such that $(G_{\varepsilon_j}(\varphi f))(x) \to \varphi(x)f(x) = f(x) \neq 0$ 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

$$(G_{\varepsilon_j}(\varphi f))(x) = \int_{\mathbb{R}^d} k_{\varepsilon_j}(x-y)\varphi(y)f(y)\,dy = 0$$

from the assumption, since the function $y \mapsto k_{\varepsilon_j}(x-y)\varphi(y)$ belongs to $C_c^\infty(\mathbb{R}^d)$. This contradiction implies the assertions. □

Remark 5.15. a) In view of formula (5.8), we have $C^k(\mathbb{R}^d) + \mathcal{N} \subseteq D_\alpha$ for all $\alpha \in \mathbb{N}_0^d$ mit $|\alpha| \le k$, and weak and classical derivatives coincide for $u \in C^k(\mathbb{R}^d)$. This fact justifies to use the same notation for both them.

b) The set D_α is a vector space and $\partial^\alpha : D_\alpha \to L^1_{\mathrm{loc}}(\mathbb{R}^d)$ is linear for all $\alpha \in \mathbb{N}_0^d$.

c) Let $1 \le p \le \infty$ and $k \in \mathbb{N}$. It is straightforward to check that $(W^{k,p}(\mathbb{R}^d), \|\cdot\|_{k,p})$ is a normed vector space. Moreover, a sequence (u_n) converges in $W^{k,p}(\mathbb{R}^d)$ if and only if every sequence $(\partial^\alpha u_n)$ converges in $L^p(\mathbb{R}^d)$ for $|\alpha| \le k$.

d) Let $1 \le p \le \infty$ and $k \in \mathbb{N}$. The map

$$J : W^{k,p}(\mathbb{R}^d) \to L^p(\mathbb{R}^d)^m; \quad u \mapsto (\partial^\alpha u)_{|\alpha| \le m},$$

is a linear isometry, where $m = 1 + d + \ldots + d^k$ and $L^p(\mathbb{R}^d)^m$ is endowed with the norm $\|(f_j)_{j=1}^m\| := |\,(\|f_j\|_p)_j\,|_p$. (We see in the proof of the next proposition that $W^{k,p}(\mathbb{R}^d)$ is isometrically isomorphic to a closed subspace of $L^p(\mathbb{R}^d)^m$.) Since the p-norm and the 1-norm on $\mathbb{R}^m$ are equivalent, there are constants $C_k, c_k > 0$ with

$$c_k \sum_{0 \le |\alpha| \le k} \|\partial^\alpha u\|_p \le \|u\|_{k,p} \le C_k \sum_{0 \le |\alpha| \le k} \|\partial^\alpha u\|_p$$

for all $u \in W^{k,p}(\mathbb{R}^d)$. ◊

The following lemma first gives a convergence result in $L^1_{\mathrm{loc}}(\mathbb{R}^d)$ for weak derivatives which is analogous to a well-known result for uniform limits and the classical derivative. Second, it makes clear that the mollifiers fit perfectly well to weak derivatives. Here and below we use the fact that the map

$$L^p(\mathbb{R}^d) \times L^{p'}(\mathbb{R}^d) \to \mathbb{F}; \quad (f,g) \mapsto \int_{\mathbb{R}^d} fg\,dx,$$

is continuous for all $p \in [1,\infty]$, due to Hölder's inequality.

Lemma 5.16. *a) Let $u_n \in D_\alpha$ and $u, f \in L^1_{loc}(\mathbb{R}^d)$ such that $u_n \to u$ and $\partial^\alpha u_n \to f$ in $L^1_{loc}(\mathbb{R}^d)$ as $n \to \infty$. Then $u \in D_\alpha$ and $\partial^\alpha u = f$.*

b) Let $u \in D_\alpha$ for some $\alpha \in \mathbb{N}_0^d$ and $\varepsilon > 0$. Then $\partial^\alpha G_\varepsilon u = G_\varepsilon \partial^\alpha u$.

Proof. a) Let $\varphi \in C_c^\infty(\mathbb{R}^d)$. In a) we have convergence on $L^1(\operatorname{supp}\varphi)$. Using also Definition 5.13, we compute

$$\begin{aligned}\int_{\mathbb{R}^d} u\,\partial^\alpha\varphi\,dx &= \lim_{n\to\infty}\int_{\mathbb{R}^d} u_n\,\partial^\alpha\varphi\,dx = \lim_{n\to\infty}(-1)^{|\alpha|}\int_{\mathbb{R}^d}(\partial^\alpha u_n)\,\varphi\,dx\\ &= (-1)^{|\alpha|}\int_{\mathbb{R}^d} f\,\varphi\,dx,\end{aligned}$$

so that a) holds.

b) Fix $\varepsilon > 0$, $u \in D_\alpha$ and $x \in \mathbb{R}^d$. The function $y \mapsto \varphi(y) := k_\varepsilon(x-y)$ belongs to $C_c^\infty(\mathbb{R}^d)$. From a corollary to Lebesgue's theorem and Definition 5.13 it follows

$$\begin{aligned}(\partial^\alpha G_\varepsilon u)(x) &= \int_{\mathbb{R}^d} \partial_x^\alpha k_\varepsilon(x-y)u(y)\,dy = (-1)^{|\alpha|}\int_{\mathbb{R}^d}(\partial^\alpha\varphi)(y)u(y)\,dy\\ &= \int_{\mathbb{R}^d}\varphi(y)\,\partial^\alpha u(y)\,dy = (G_\varepsilon\partial^\alpha u)(x).\end{aligned}$$ □

The above convergence result is a crucial tool when extending properties from classical to weak derivatives, see Proposition 5.19 below. We first use it to show the completeness of Sobolev spaces.

Proposition 5.17. *Let $1 \le p \le \infty$ and $k \in \mathbb{N}$. Then $W^{k,p}(\mathbb{R}^d)$ is a Banach space. It is separable if $1 \le p < \infty$ and reflexive if $1 < p < \infty$. Moreover, $W^{k,2}(\mathbb{R}^d)$ is a Hilbert space endowed with the scalar product*

$$(f|g)_{k,2} = \sum_{|\alpha|\le k}\int_{\mathbb{R}^d}(\partial^\alpha f)\,\overline{\partial^\alpha g}\,dx.$$

Proof. Let $(u_n)_n$ be a Cauchy sequence in $W^{k,p}(\mathbb{R}^d)$. Then $(\partial^\alpha u_n)_n$ is a Cauchy sequence in $L^p(\mathbb{R}^d)$ for every $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \le k$, and thus $\partial^\alpha u_n \to f_\alpha$ in $L^p(\mathbb{R}^d)$ for some $f_\alpha \in L^p(\mathbb{R}^d)$ as $n \to \infty$. Setting $u := f_0$, we see that $u_n \to u$ in $L^p(\mathbb{R}^d)$. Lemma 5.16 now implies that $u \in D_\alpha$ and $f_\alpha = \partial^\alpha u$. Hence, $W^{k,p}(\mathbb{R}^d)$ is a Banach space. Using Remark 2.11, we then deduce from Remark 5.15d) that $W^{k,p}(\mathbb{R}^d)$ is isometrically isomorphic to a closed subspace of $L^p(\mathbb{R}^d)^m$. The remaining assertions now follow by isomorphy from known results. □

Example 5.18. a) Let $u \in C(\mathbb{R})$ be such that $u_\pm := u_{|\mathbb{R}_\pm}$ belong to $C^1(\mathbb{R}_\pm)$. We then have $u \in D_1$ with

$$\partial_1 u = \left\{ \begin{array}{ll} u_+' & \text{on } (0,\infty) \\ u_-' & \text{on } (-\infty,0) \end{array} \right\} =: f.$$

For $u(x) = |x|$, we thus obtain $\partial_1 f = \mathbb{1}_{(0,\infty)} - \mathbb{1}_{(-\infty,0)}$.

Proof. For every $\varphi \in C_c^\infty(\mathbb{R})$, integration by parts yields

$$\begin{aligned}\int_{\mathbb{R}} u\varphi'\,dt &= \int_{-\infty}^0 u_-\varphi'\,dt + \int_0^\infty u_+\varphi'\,dt \\ &= -\int_{-\infty}^0 u_-'\varphi\,dt + u_-\varphi|_{-\infty}^0 - \int_0^\infty u_+'\varphi dt + u_+\varphi'|_0^\infty \\ &= -\int_{\mathbb{R}} f\,\varphi\,dt,\end{aligned}$$

since $u_+(0) = u_-(0)$ by the continuity of u. □

b) The function $u = \mathbb{1}_{\mathbb{R}_+}$ does not belong to D_1.

Proof. Assume there would exist $f = \partial_1 u \in L^1_{\mathrm{loc}}(\mathbb{R})$. For every $\varphi \in C_c^\infty(\mathbb{R})$ we then obtain

$$\int_{\mathbb{R}} f\varphi\,dt = -\int_{\mathbb{R}} \mathbb{1}_{\mathbb{R}_+}\varphi'\,dt = -\int_0^\infty \varphi'(t)\,dt = \varphi(0).$$

Taking φ with $\operatorname{supp}\varphi \subseteq (0,\infty)$, we deduce from a variant of Lemma 5.14 that $f = 0$ on $(0,\infty)$. Similarly, it follows that $f = 0$ on $(-\infty, 0)$. Hence, $f = 0$ and so $\varphi(0) = 0$ for all $\varphi \in C_c^\infty(\mathbb{R})$, which is wrong. □

c) Let $p \in [1,\infty)$ and $\beta \in (1 - \frac{d}{p}, 1]$. Set $u(x) = |x|_2^\beta$ for $0 < |x|_2 < 1$ and $f_j(x) = \beta x_j |x|_2^{\beta-2}$ for $0 < |x|_2 < 1$ and $j \in \{1,\dots,d\}$, as well as and $u(0) = f_j(0) = 0$. Then $u \in W^{1,p}(B(0,1))$ and $\partial_j u = f_j$. Observe that u is unbounded and has no continuous extension at $x = 0$ if $\beta < 0$ (which is possible if $d \ge 2$).

Proof. The functions $r \mapsto r^{\beta p} r^{d-1}$ and $r \mapsto r^{(\beta-1)p} r^{d-1}$ are integrable on $(0,1)$ since $\beta > 1 - \frac{d}{p}$. Using polar coordinates, we infer that $u, f_j \in L^p(B(0,1))$ for $j \in \{1,\dots,d\}$. We introduce the regularized functions $u_n(x) = (n^{-2} + |x|_2^2)^{\beta/2}$ for $n \in \mathbb{N}$ and $x \in B(0,1)$. Then $u_n \in C_b^1(\mathbb{R}^d) \subset W^{1,p}(B(0,1))$ and $\partial_j u_n(x) = \beta x_j\,(n^{-2} + |x|_2^2)^{\frac{\beta}{2}-1}$. Observe that $u_n(x)$ and $\partial_j u_n(x)$ tend to $u(x)$ and $f_j(x)$ for $x \ne 0$ as $n \to \infty$, respectively. Moreover, $|u_n| \le \max\{|u|, 2^{\beta/2}\}$ and $|\partial_j u_n| \le |f_j|$ a.e. on $B(0,1)$ for all n and j. Hence, $u_n \to u$ and $\partial_j u_n \to f_j$ in $L^p(B(0,1))$ as $n \to \infty$ due to dominated convergence. The assertion thus follows from Lemma 5.16. □

Under suitable regularity and integrability assumptions, the product and substitution rules also hold for weak derivatives. Here we only present the basic version of the product rule that is needed below.

Proposition 5.19. *Let $p \in [1,\infty)$, $u \in W^{1,p}(\mathbb{R}^d)$ and $v \in W^{1,p'}(\mathbb{R}^d)$. Then, $uv \in W^{1,1}(\mathbb{R}^d)$ and $\partial_j(uv) = (\partial_j u)v + u\partial_j v$ for every $j \in \{1,\dots,d\}$.*

Proof. Hölder's inequality implies that uv, $(\partial_j u)v$ and $u\partial_j v$ belong to $L^1(\mathbb{R}^d)$. Set $u_n = G_{1/n}u \in C^\infty(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$ and $v_n = G_{1/n}v \in C^\infty(\mathbb{R}^d) \cap L^{p'}(\mathbb{R}^d)$ for $n \in \mathbb{N}$.

First, let $p > 1$ so that $p' < \infty$. Due to Proposition 3.14, the functions u_n converge to u in $L^p(\mathbb{R}^d)$ and v_n converge to v in $L^{p'}(\mathbb{R}^d)$ as $n \to \infty$. Lemma 5.16 further yields $\partial_j u_n = G_{1/n}\partial_j u$ and $\partial_j v_n = G_{1/n}\partial_j v$. Since $\partial_j u \in L^p(\mathbb{R}^d)$ and $\partial_j v \in L^{p'}(\mathbb{R}^d)$, Proposition 3.14 also implies that $\partial_j u_n \to \partial_j u$ in $L^p(\mathbb{R}^d)$ and $\partial_j v_n \to \partial_j v$ in $L^{p'}(\mathbb{R}^d)$ as $n \to \infty$, for each $j \in \{1,\dots,d\}$. In view of (5.8) and using the product rule for C^1–functions, for each $\varphi \in C_c^\infty(\mathbb{R}^d)$ we can now conclude that

$$\begin{aligned}\int_{\mathbb{R}^d} uv\,\partial_j\varphi\,dx &= \lim_{n\to\infty}\int_{\mathbb{R}^d} u_n v_n\,\partial_j\varphi\,dx = -\lim_{n\to\infty}\int_{\mathbb{R}^d}[(\partial_j u_n)v_n + u_n(\partial_j v_n)]\,\varphi\,dx\\ &= -\int_{\mathbb{R}^d}[(\partial_j u)v + u(\partial_j v)]\,\varphi\,dx. \end{aligned} \tag{5.9}$$

Hence, the weak derivative $\partial_j(uv) = (\partial_j u)v + u(\partial_j v) \in L^1(\mathbb{R}^d)$ exists.

Let $p = 1$. The above used convergence in $L^{p'}(\mathbb{R}^d)$ may now fail since $p' = \infty$. The functions u_n and $\partial_j u_n$ still converge to u and $\partial_j u$ in $L^1(\mathbb{R}^d)$. Passing to a subsequence, we can thus assume that they converge pointwise a.e. and that $|u_n|, |\partial_j u_n| \le g$ a.e. for some $g \in L^1(\mathbb{R}^d)$ and all n and j. We further use that $\|v_n\|_\infty \le \|v\|_\infty$ and $\|\partial_j v_n\|_\infty \le \|\partial_j v\|_\infty$ by Proposition 3.14 and that the restrictions of v and $\partial_j v$ belong $L^1(B(0,k))$ for each $k > 0$. If we apply Proposition 3.14 for each $L^1(B(0,k))$ and iteratively choose subsequences, we obtain (diagonal) subsequences $(v_{n_m})_m$ and $(\partial_j v_{n_m})_m$ which converge to v and $\partial_j v$ pointwise a.e. as $m \to \infty$. By means of the dominated convergence theorem, we can now conclude the assertion as in (5.9). □

We come now to an important density result. In the proof we first use a cutoff argument to obtain a compact support and then perform a mollification. The cutoff must be chosen so that the extra terms caused by the $W^{k,p}$-norm vanish in the limit.

Theorem 5.20. *The spaces $C_c^\infty(\mathbb{R}^d)$ and $\mathcal{S}_d$ are dense in $W^{k,p}(\mathbb{R}^d)$ for all $p \in [1,\infty)$ and $k \in \mathbb{N}$.*

Proof. Remark 5.6a) yields $\mathcal{S}_d \subseteq W^{k,p}(\mathbb{R}^d)$ so that we only have to consider $C_c^\infty(\mathbb{R}^d)$. We restrict ourselves to $k = 1$, the general case is treated similarly.

1) Let $u \in W^{1,p}(\mathbb{R}^d)$. Take any $\phi \in C^\infty(\mathbb{R})$ with $0 \le \phi \le 1$, $\phi = 1$ on $[0,1]$ and $\phi = 0$ on $[2,\infty)$. Set

$$\varphi_n(x) = \phi\left(\tfrac{1}{n}\,|x|_2\right) \qquad (\text{“cut-off function”})$$

for $n \in \mathbb{N}$ and $x \in \mathbb{R}^d$. We then have $\varphi_n \in C_c^\infty(\mathbb{R}^d)$, $0 \le \varphi_n \le 1$ 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 $x \in \mathbb{R}^d$ as $n \to \infty$. Thus $\|\varphi_n u - u\|_p \to 0$ as $n \to \infty$ by Lebesgue's convergence theorem with majorant $|u|$. Let $j \in \{1,\dots,d\}$. Proposition 5.19 further implies that

$$\|\partial_j(\varphi_n u - u)\|_p = \|(\varphi_n\partial_j u - \partial_j u) + (\partial_j\varphi_n)u\|_p \le \|\varphi_n\partial_j u - \partial_j u\|_p + \tfrac{1}{n}\|\phi'\|_\infty\,\|u\|_p$$

Again by Lebesgue, the right hand side tends to 0 as $n \to \infty$. Given $\varepsilon > 0$, we can thus fix an $m \in \mathbb{N}$ such that $\|\varphi_m u - u\|_{1,p} \le \varepsilon$.

2) Proposition 3.14 implies that the functions $G_{\frac{1}{n}}(\varphi_m u)$ belong to $C_c^\infty(\mathbb{R}^d)$ for all $n \in \mathbb{N}$ and that $G_{\frac{1}{n}}(\varphi_m u) \to \varphi_m u$ in $L^p(\mathbb{R}^d)$ as $n \to \infty$. Moreover, Lemma 5.16 yields $\partial_j G_{\frac{1}{n}}(\varphi_m u) = G_{\frac{1}{n}}\partial_j(\varphi_m u)$, so that $\partial_j G_{\frac{1}{n}}(\varphi_m u)$ converges to $\partial_j(\varphi_m u)$ in $L^p(\mathbb{R}^d)$ as $n \to \infty$. As a result, $\|G_{\frac{1}{n}}(\varphi_m u) - u\|_{1,p} \le \|G_{\frac{1}{n}}(\varphi_m u) - \varphi_m u\|_{1,p} + \varepsilon \le 2\varepsilon$ for all sufficiently large n. □

We now come to main result of this section which describes the space $W^{k,2}(\mathbb{R}^d)$ via the Fourier transform in a very convenient way. Recall that $\|u\|_2 = \|\widehat{u}\|_2$ for $u \in L^2(\mathbb{R}^d)$ by Theorem 5.11.

Theorem 5.21. *Let $k \in \mathbb{N}$ and $\alpha \in N_0^d$ with $|\alpha| \le k$. We then have*

$$W^{k,2}(\mathbb{R}^d) = \{u \in L^2(\mathbb{R}^d) \mid \ |\xi|_2^k \,\widehat{u} \in L^2(\mathbb{R}^d)\} =: H^k$$

and the norm of $W^{k,2}(\mathbb{R}^d)$ is equivalent to $(\|u\|_2^2 + \|\, |\xi|_2^k \,\widehat{u}\|_2^2)^{\frac{1}{2}}$. For $u \in W^{k,2}(\mathbb{R}^d)$ it further holds

$$\mathcal{F}(\partial^\alpha u) = i^{|\alpha|}\, \xi^\alpha \widehat{u}. \tag{5.10}$$

Proof. Due to Lemma 5.7, Schwartz functions $u \in \mathcal{S}_d$ satisfy (5.10). We thus obtain

$$\|u\|_{k,2}^2 = \sum_{|\alpha|\le k} \|\mathcal{F}\partial^\alpha u\|_2^2 = \sum_{|\alpha|\le k} \|\xi^\alpha \widehat{u}\|_2^2 = \int_{\mathbb{R}^d} \sum_{|\alpha|\le k} |\xi^\alpha|^2\, |\widehat{u}|_2^2 \,d\xi$$
$$\begin{cases} \le c_1\,(\|u\|_2^2 + \|\, |\xi|_2^k \,\widehat{u}\|_2^2) \\ \ge c_2\,(\|u\|_2^2 + \|\, |\xi|_2^k \,\widehat{u}\|_2^2) \end{cases} \tag{5.11}$$

for $u \in \mathcal{S}_d$ and constants $c_j > 0$.

Let $u \in W^{k,2}(\mathbb{R}^d)$. Theorem 5.20 gives $u_n \in \mathcal{S}_d$ which converge to u in $W^{k,2}(\mathbb{R}^d)$ as $n \to \infty$. Since $\mathcal{F}$ is continuous on $L^2(\mathbb{R}^d)$, the functions $\widehat{u_n}$ tend to $\widehat{u}$ in $L^2(\mathbb{R}^d)$ and (possibly after passing to a subsequence) pointwise a.e., as $n \to \infty$. Hence, the functions $\xi^\alpha \widehat{u_n}$ converge pointwise a.e. to $\xi^\alpha \widehat{u}$. On the other hand, equation (5.11) yields that $(\xi^\alpha \widehat{u_n})_n$ is Cauchy in $L^2(\mathbb{R}^d)$ for each $|\alpha| \le k$, and thus $\xi^\alpha \widehat{u_n}$ converge to $\xi^\alpha \widehat{u}$ in $L^2(\mathbb{R}^d)$ as $n \to \infty$. We conclude that every $u \in W^{k,2}(\mathbb{R}^d)$ fulfills (5.11) and that $W^{k,2}(\mathbb{R}^d) \subseteq H^k$.

Conversely, take $u \in H^k$. Let $\varphi \in C_c^\infty(\mathbb{R}^d)$ and $|\alpha| \le k$. From Theorem 5.11 and Lemma 5.7 we deduce

$$\int_{\mathbb{R}^d} u\, \partial^\alpha \varphi \,dx = (u|\partial^\alpha \overline{\varphi}) = (\mathcal{F}u|\mathcal{F}\partial^\alpha \overline{\varphi}) = (\widehat{u}|i^{|\alpha|}\, \xi^\alpha\, \widehat{\overline{\varphi}}) = ((-i)^{|\alpha|}\, \xi^\alpha\, \widehat{u}|\mathcal{F}\overline{\varphi})$$
$$= (\mathcal{F}'((-i)^{|\alpha|}\, \xi^\alpha\, \widehat{u})|\overline{\varphi}) = (-1)^{|\alpha|} \int_{\mathbb{R}^d} \varphi\, \mathcal{F}^{-1}(i^{|\alpha|}\, \xi^\alpha\, \widehat{u}) \,dx.$$

Therefore $u \in D_\alpha$ and $\partial^\alpha u = \mathcal{F}^{-1}(i^{|\alpha|}\, \xi^\alpha\, \widehat{u}) \in L^2(\mathbb{R}^d)$, and hence $u \in W^{k,2}(\mathbb{R}^d)$; i.e., $W^{k,2}(\mathbb{R}^d) = H^k$. Applying $\mathcal{F}$ to the equation in the previous sentence, we also derive (5.10) for all $u \in W^{k,2}(\mathbb{R}^d)$. □

There are $u \in W_2^k(\mathbb{R}^d)$ such that $x^\alpha u \notin L^2(\mathbb{R})$ for $|\alpha| = k$, e.g., the function

$$u(t) = \begin{cases} |t|^{-3/2}, & |t| \ge 1, \\ 1, & |t| < 1, \end{cases}$$

for $k = 1$, cf. Example 5.18. Moreover, one can show directly that here $\widehat{u}'$ is even not integrable near 0. Therefore the other equation in Lemma 5.7a), namely $\partial^\alpha \widehat{u} = (-i)^{|\alpha|}\, \mathcal{F}(x^\alpha u)$, cannot be extended from $u \in \mathcal{S}_d$ to every $u \in W_2^k(\mathbb{R}^d)$ in the present framework, but we obtain it in a more general setting in Theorem 5.27.

Example 5.22. We consider the diffusion equation

$$\begin{aligned} \partial_t u(t,x) &= \Delta u(t,x), \qquad t \ge 0,\ x \in \mathbb{R}^d, \\ u(0,x) &= u_0(x), \qquad x \in \mathbb{R}^d, \end{aligned} \tag{5.12}$$

for a given initial value $u_0 \in L^2(\mathbb{R}^d)$. We show that there is a unique *solution* of (5.12); i.e., a function $u \in C([0,\infty), L^2(\mathbb{R}^d)) \cap C^1((0,\infty), L^2(\mathbb{R}^d))$ such that $u(t)$

belongs to $W_2^2(\mathbb{R}^d)$ for all $t > 0$ and u satisfies (5.12) as equations in $L^2(\mathbb{R}^d)$. To that purpose, we define

$$u(t) = \mathcal{F}^{-1}(m_t \widehat{u_0}) \quad \text{with} \quad m_t(\xi) = e^{-t\,|\xi|_2^2}$$

for $t > 0$ and $\xi \in \mathbb{R}^d$ as in Example 5.12. Since $|\xi|_2^k\,\widehat{u}(t) = |\xi|_2^k\, m_t\widehat{u_0}$ belongs to $L^2(\mathbb{R}^d)$, Theorem 5.21 implies that $u(t) \in W^{k,2}(\mathbb{R}^d)$ for all $k \in \mathbb{N}$ and $t > 0$. From (5.10) we then infer

$$\mathcal{F}\Delta u(t) = -|\xi|_2^2\,\widehat{u}(t) = -|\xi|_2^2\, m_t\widehat{u_0}, \qquad \Delta u(t) = -\mathcal{F}^{-1}(|\xi|_2^2\, m_t\widehat{u_0}).$$

Let $v(t) = \mathcal{F}u(t) = m_t\widehat{u_0}$ for $t > 0$. Clearly, $\frac{1}{h}(v(t+h)-v(t))$ converges pointwise to $-|\xi|_2^2\, m_t\widehat{u_0}$ as $h \to 0$ and $|\frac{1}{h}(v(t+h) - v(t))| \le |\xi|_2^2\, m_t \in L^1(\mathbb{R}^d)$. Dominated convergence then implies that v has the (continuous) derivative $v'(t) = -|\xi|_2^2\, m_t\widehat{u_0}$ in $L^2(\mathbb{R}^d)$ for $t > 0$. The continuity of $\mathcal{F}^{-1}$ on $L^2(\mathbb{R}^d)$ thus yields that $u \in C^1((0,\infty), L^2(\mathbb{R}^d))$ and

$$u'(t) = -\mathcal{F}^{-1}(|\xi|_2^2\, m_t\widehat{u_0}) = \Delta u(t)$$

for $t > 0$. Finally, $m_t\widehat{u_0}$ tends to $\widehat{u_0}$ in $L^2(\mathbb{R}^d)$ as $t \to 0$ by Lebesgue's theorem with majorant $|\widehat{u_0}|$. Hence, $u \in C([0,\infty), L^2(\mathbb{R}^d))$ with $u(0) = u_0$.

As in Example 5.12 one sees that every solution is given by the above formula so that solutions are unique. Moreover, $\|u(t)\|_2 \le \|u_0\|_2$ for all $t \ge 0$. ◊

5.3 Tempered distributions

The theory of distributions allows to define derivatives of any order for rather general objects such as locally integrable functions or measures on open subsets of $\mathbb{R}^d$, see e.g. [Rud91]. Here we only discuss the subclass of tempered distributions on $\mathbb{R}^d$ to which one can extend the Fourier transform in a very convenient way.

Definition 5.23. Tempered distributions *are continuous linear functionals on $\mathcal{S}_d$ One writes $\mathcal{S}_d^* := \{u : \mathcal{S}_d \to \mathbb{F} \,|\, u$ is linear and continuous$\}$ for the space of tempered distributions and $\langle \varphi, u\rangle_{\mathcal{S}_d} = u(\varphi)$ for $u \in \mathcal{S}_d^*$ and $\varphi \in \mathcal{S}_d$.*

Recall that $\varphi_n \to \varphi$ in $\mathcal{S}_d$ means that $p_{m,\alpha}(\varphi_n - \varphi) = \||x|^m \partial^\alpha(\varphi_n - \varphi)\|_\infty \to 0$ for all $m \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^d$, as $n \to \infty$. It is possible to define weak* convergence in $\mathcal{S}_d^*$, but we will not deal with such questions.

Example 5.24. a) (regular tempered distributions) Let $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ satisfy

$$a_k(f) := \int_{k \le |x|_2 < k+1} |f(x)|\, dx \le c k^\kappa$$

for all $k \in \mathbb{N}$ and some $\kappa, c \ge 0$. This condition is satisfied by polynomially bounded f and by $f \in L^p(\mathbb{R}^d)$ with $p \in [1,\infty]$ (because then $a_k(f) \le c\,\|f\|_p\, k^{(d-1)/p'}$ by Hölder's inequality). For $\varphi \in \mathcal{S}_d$, define

$$u_f(\varphi) = \int_{\mathbb{R}^d} \varphi f\, dx.$$

Let φ_n tend to φ in $\mathcal{S}_d$. Take $m \in \mathbb{N}$ with $m \ge \kappa + 2$. Inserting $|x|^{-m}|x|^m$ in the integrands for $k \ge 1$, we estimate

$$|u_f(\varphi - \varphi_n)| \le \sum_{k=0}^{\infty} \int_{k \le |x|_2 < k+1} |\varphi - \varphi_n|\, |f(x)|\, dx$$

$$\leq \|f\|_{L^1(B(0,1))} p_{0,0}(\varphi - \varphi_n) + p_{m,0}(\varphi - \varphi_n) \sum_{k=1}^{\infty} c k^{\kappa - m}$$
$$\leq c' \left(p_{0,0}(\varphi - \varphi_n) + p_{m,0}(\varphi - \varphi_n) \right)$$

for a constant $c' > 0$. Hence, $u_f : \mathcal{S}_d \to \mathbb{F}$ is continuous. The linearity of u_f is clear, and so it belongs to $\mathcal{S}_d^*$. One often writes f instead of u_f.

b) (measures) Let μ be a measure on $\mathcal{B}_d$ with $\mu(B(0,1)) < \infty$ and $\mu(B(0,k+1) \setminus B(0,k)) \leq ck^\kappa$ for all $k \in \mathbb{N}$ and some $\kappa, c \geq 0$. Then one sees as in a) that

$$u_\mu(\varphi) = \int_{\mathbb{R}^d} \varphi \, d\mu, \qquad \varphi \in \mathcal{S}_d,$$

defines a tempered distribution u_μ, which is often simply denoted by μ.

c) (Dirac distributions) Let $y \in \mathbb{R}^d$ and $\alpha \in \mathbb{N}_0^d$. We set $\delta_y^\alpha(\varphi) = \partial^\alpha \varphi(y)$ for $\varphi \in \mathcal{S}_d$. Let φ_n converge to φ in $\mathcal{S}_d$. Then $|\delta_y^\alpha(\varphi) - \delta_y^\alpha(\varphi_n)| \leq p_{0,\alpha}(\varphi - \varphi_n)$ tends to 0, so that $\delta_y^\alpha \in \mathcal{S}_d^*$. ◊

Definition 5.25. *Let $u \in \mathcal{S}_d^*$, $g \in \mathcal{E}_d$, and $\alpha \in \mathbb{N}_0^d$. For all $\varphi \in \mathcal{S}_d$ we define*

a) $(gu)(\varphi) = \langle \varphi, gu \rangle_{\mathcal{S}_d} := \langle g\varphi, u \rangle_{\mathcal{S}_d} = u(g\varphi)$,

b) $(\partial^\alpha u)(\varphi) = \langle \varphi, \partial^\alpha u \rangle_{\mathcal{S}_d} := (-1)^{|\alpha|} \langle \partial^\alpha \varphi, u \rangle_{\mathcal{S}_d} = u((-1)^{|\alpha|} \partial^\alpha \varphi)$,

c) $\widehat{u}(\varphi) := (\mathcal{F}u)(\varphi) = \langle \varphi, \mathcal{F}u \rangle_{\mathcal{S}_d} := \langle \mathcal{F}\varphi, u \rangle_{\mathcal{S}_d} = u(\mathcal{F}\varphi)$,

d) $(Ru)(\varphi) = \langle \varphi, Ru \rangle_{\mathcal{S}_d} := \langle R\varphi, u \rangle_{\mathcal{S}_d} = u(R\varphi)$,

e) $(\varphi * u)(x) := \langle T_{-x} R\varphi, u \rangle_{\mathcal{S}_d} = u(T_{-x} R\varphi)$ for every $x \in \mathbb{R}^d$.

In view of Lemma 5.7, the maps gu, $\partial^\alpha u$, $\mathcal{F}u$, and Ru are continuous and linear from $\mathcal{S}_d$ to $\mathbb{F}$, and hence they belong to $\mathcal{S}_d^*$. Moreover, $T_{-x} R\varphi \in \mathcal{S}_d$.

Observe that we multiply and convolve tempered distributions only with the (very regular) functions in $\mathcal{E}_d$ and $\mathcal{S}_d$, respectively. The following examples and the theorem below indicate that the above definitions extend the known concepts in a natural way and that they allow to generalize several main properties of the Fourier transform to the space $\mathcal{S}_d^*$.

Example 5.26. Let $\varphi \in \mathcal{S}_d$, $g \in \mathcal{E}_d$, $\alpha \in \mathbb{N}_0^d$, and $x, y \in \mathbb{R}^d$.

a) Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ be as in Example 5.24a). Then $gu_f = u_{gf}$ since

$$(gu_f)(\varphi) = \int_{\mathbb{R}^d} \varphi g f \, dx = u_{gf}(\varphi).$$

b) Let $f \in W^{k,p}(\mathbb{R}^d)$ for some $p \in [1, \infty]$ and $|\alpha| \leq k \in \mathbb{N}$. Then $\partial^\alpha u_f = u_{\partial^\alpha f}$ since the definitions yield

$$\langle \varphi, \partial^\alpha u_f \rangle_{\mathcal{S}_d} = (-1)^{|\alpha|} \langle \partial^\alpha \varphi, u_f \rangle_{\mathcal{S}_d} = (-1)^{|\alpha|} \int_{\mathbb{R}^d} (\partial^\alpha \varphi) f \, dx = \int_{\mathbb{R}^d} \varphi \, \partial^\alpha f \, dx$$
$$= \langle \varphi, u_{\partial^\alpha f} \rangle_{\mathcal{S}_d}.$$

c) Let $f \in L^2(\mathbb{R}^d)$. Then $\mathcal{F}u_f = u_{\mathcal{F}f}$ since Theorem 5.11 implies

$$\langle \varphi, \mathcal{F}u_f \rangle_{\mathcal{S}_d} = \langle \mathcal{F}\varphi, u_f \rangle_{\mathcal{S}_d} = \int_{\mathbb{R}^d} \widehat{\varphi} f \, dx = \int_{\mathbb{R}^d} \varphi \widehat{f} \, dx = \langle \varphi, u_{\mathcal{F}f} \rangle_{\mathcal{S}_d}.$$

d) We have $\partial^\alpha \delta_y = (-1)^{|\alpha|} \delta_y^\alpha$ since

$$\langle \varphi, \partial^\alpha \delta_y \rangle_{\mathcal{S}_d} = (-1)^{|\alpha|} \langle \partial^\alpha \varphi, \delta_y \rangle_{\mathcal{S}_d} = (-1)^{|\alpha|} \partial^\alpha \varphi(y) = (-1)^{|\alpha|} \delta_y^\alpha(\varphi).$$

e) We have $\mathcal{F}\delta_y = (2\pi)^{-d/2} e_{-y}$ since

$$\langle \varphi, \mathcal{F}\delta_y \rangle_{\mathcal{S}_d} = \langle \mathcal{F}\varphi, \delta_y \rangle_{\mathcal{S}_d} = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-iy \cdot x} \varphi(x) \, dx = \langle \varphi, (2\pi)^{-d/2} e_{-y} \rangle_{\mathcal{S}_d}.$$

f) We have $\mathcal{F}e_y = (2\pi)^{d/2}\delta_y$ since Proposition 5.10 implies

$$\langle \varphi, \mathcal{F}e_y\rangle_{\mathcal{S}_d} = \langle \mathcal{F}\varphi, e_y\rangle_{\mathcal{S}_d} = \int_{\mathbb{R}^d} \widehat{\varphi}(\xi)e^{iy\cdot\xi}\,d\xi = (2\pi)^{d/2}(\mathcal{F}^{-1}\widehat{\varphi})(y) = (2\pi)^{d/2}\varphi(y).$$

Assertion f) can also be deduced from e) since $\mathcal{F}^2$ is equal to R in $\mathcal{S}_d^*$, too, as shown in the next theorem (with a similar proof as above).

g) Let $f \in L^1(\mathbb{R}^d)$. Then $\varphi * u_f = \varphi * f$, since

$$\varphi * u_f(x) = \langle T_{-x}R\varphi, u_f\rangle_{\mathcal{S}_d} = \int_{\mathbb{R}^d} \varphi(-(z-x))f(z)\,dz = \varphi * f(x). \qquad \lozenge$$

Theorem 5.27. *Let $u \in \mathcal{S}_d^*$, $\varphi, \psi \in \mathcal{S}_d$, and $\alpha \in \mathbb{N}_0^d$. The following assertions hold.*
a) $\mathcal{F} : \mathcal{S}_d^* \to \mathcal{S}_d^*$ is bijective with $\mathcal{F}^4 = I$ and $\mathcal{F}^{-1} = \mathcal{F}^3 = R\mathcal{F}$.
b) $\mathcal{F}(\partial^\alpha u) = i^{|\alpha|}\xi^\alpha \mathcal{F}u$ and $\partial^\alpha(\mathcal{F}u) = (-i)^{|\alpha|}\mathcal{F}(x^\alpha u)$.
c) $\varphi * u \in \mathcal{E}_d$, and hence $\varphi * u$ induces a tempered distribution.
d) $\partial^\alpha(\varphi * u) = (\partial^\alpha\varphi) * u = \varphi * \partial^\alpha u$.
e) $\mathcal{F}(\varphi * u) = (2\pi)^{d/2}\widehat{\varphi}\widehat{u}$ and $\mathcal{F}(\varphi u) = (2\pi)^{-d/2}\widehat{\varphi} * \widehat{u}$.

Proof. Let $u \in \mathcal{S}_d^*$, $\varphi \in \mathcal{S}_d$, and $\alpha \in \mathbb{N}_0^d$. In a), Proposition 5.10 yields

$$\langle \varphi, \mathcal{F}^4 u\rangle_{\mathcal{S}_d} = \langle \mathcal{F}\varphi, \mathcal{F}^3 u\rangle_{\mathcal{S}_d} = \cdots = \langle \mathcal{F}^4\varphi, u\rangle_{\mathcal{S}_d} = \langle \varphi, u\rangle_{\mathcal{S}_d},$$

so that $\mathcal{F}^4 = I$ on $\mathcal{S}_d^*$ and $\mathcal{F} : \mathcal{S}_d^* \to \mathcal{S}_d^*$ is bijective with inverse $\mathcal{F}^{-1} = \mathcal{F}^3$. Similarly, we show the remaining equality $\mathcal{F}^2 = R$ by computing

$$\langle \varphi, \mathcal{F}^2 u\rangle_{\mathcal{S}_d} = \langle \mathcal{F}^2\varphi, u\rangle_{\mathcal{S}_d} = \langle R\varphi, u\rangle_{\mathcal{S}_d}.$$

The first equality in b) follows from Lemma 5.7 and

$$\begin{aligned}\langle \varphi, \mathcal{F}\partial^\alpha u\rangle_{\mathcal{S}_d} &= \langle \mathcal{F}\varphi, \partial^\alpha u\rangle_{\mathcal{S}_d} = (-1)^{|\alpha|}\langle \partial^\alpha\mathcal{F}\varphi, u\rangle_{\mathcal{S}_d} = i^{|\alpha|}\langle \mathcal{F}(x^\alpha\varphi), u\rangle_{\mathcal{S}_d}\\ &= \langle \varphi, i^{|\alpha|}\partial^\alpha\mathcal{F}u\rangle_{\mathcal{S}_d}.\end{aligned}$$

The second part of b) established in the same way.

For the proof of assertions c) and d) we refer to Theorem 7.19 in [Rud91]. To show e), first take $\psi \in C_c^\infty(\mathbb{R}^d)$. There is a closed interval $I \subset \mathbb{R}^d$ such that $\operatorname{supp}\psi \subset I$. Using the definitions, part a) and Proposition 5.3 in the last step, we compute

$$\begin{aligned}\langle \widehat{\psi}, \mathcal{F}(\varphi * u)\rangle_{\mathcal{S}_d} &= \langle R\psi, \varphi * u\rangle_{\mathcal{S}_d} = \int_{\mathbb{R}^d}\psi(-x)(\varphi * u)(x)\,dx = \int_{-I}\psi(-x)u(T_{-x}R\varphi)\,dx\\ &= \int_I u(\psi(z)T_zR\varphi)\,dz = u\Big(\int_I \psi(z)T_zR\varphi\,dz\Big) \qquad (5.13)\\ &= \langle R(\psi * \varphi), u\rangle_{\mathcal{S}_d} = \langle \mathcal{F}(\psi * \varphi), \mathcal{F}u\rangle_{\mathcal{S}_d} = (2\pi)^{d/2}\langle \widehat{\psi}\widehat{\varphi}, \mathcal{F}u\rangle_{\mathcal{S}_d}.\end{aligned}$$

Here the second integral in (5.13) is understood as an $\mathcal{S}_d$–valued Riemann integral on I; i.e., as the limit in $\mathcal{S}_d$ of Riemann sums such as

$$S_n(y) = \sum_{j=1}^{m_n}\psi(z_{j,n})(R\varphi)(y + z_{j,n})\operatorname{vol}(Q_{j,n}), \qquad y \in \mathbb{R}^d,$$

where $z_{j,n} \in Q_{j,n}$, the rectangles $Q_{j,n}$, $j \in \{1,\ldots,m_n\}$, subdivide I and $\max_j \operatorname{vol}(Q_{j,n})$ tends to 0 as $n \to \infty$. Clearly, S_n belongs to $\mathcal{S}_d$. We omit the somewhat tedious, but elementary proof that S_n indeed converges in $\mathcal{S}_d$. Hence, u can be taken out of the approximating Riemann sums by its linearity and out of the limit by its continuity. This fact justifies that we have interchanged u and the integral in (5.13). So far we have shown

$$\langle \widehat{\psi}, \mathcal{F}(\varphi * u)\rangle_{\mathcal{S}_d} = (2\pi)^{d/2}\langle \widehat{\psi}, \widehat{\varphi}\widehat{u}\rangle_{\mathcal{S}_d} \tag{5.14}$$

for all $\psi \in C_c^\infty(\mathbb{R}^d)$. Arguing as in the first part of the proof of Theorem 5.20, one can show that $C_c^\infty(\mathbb{R}^d)$ is dense in $\mathcal{S}_d$, see Theorem 7.10 in [Rud91]. Since the Fourier transform is continuous on $\mathcal{S}_d$ by Lemma 5.7, the identity (5.14) is thus valid for all $\widehat{\psi}$ with $\psi \in \mathcal{S}_d$ due to an approximation argument. We can now replace here $\widehat{\psi}$ by $\psi \in \mathcal{S}_d$ using that $\mathcal{F}$ is bijective on $\mathcal{S}_d$ thanks to Proposition 5.10. So the first part of e) is shown. For the second part, observe that

$$\langle \psi, R(\varphi)R(u)\rangle_{\mathcal{S}_d} = \langle \psi R\varphi, Ru\rangle_{\mathcal{S}_d} = \langle R(\psi)R^2\varphi, u\rangle_{\mathcal{S}_d} = \langle R\psi, \varphi u\rangle_{\mathcal{S}_d} = \langle \psi, R(\varphi u)\rangle_{\mathcal{S}_d},$$

for all $\psi \in \mathcal{S}_d$; i.e., $R(\varphi)R(u) = R(\varphi u)$. Employing also a), we then calculate

$$\mathcal{F}(\widehat{\varphi} * \widehat{u}) = (2\pi)^{d/2}\mathcal{F}^2(\varphi)\mathcal{F}^2(u) = (2\pi)^{d/2}R(\varphi u) = (2\pi)^{d/2}\mathcal{F}^2(\varphi u),$$

which yields the second part of e) if we apply $\mathcal{F}^{-1}$. □

Bibliography

[Bre11] H. Brézis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Springer, 2011.

[Con90] J.B. Conway, *A Course in Functional Analysis.* 2nd edition. Springer–Verlag, 1990.

[Dob06] M. Dobrowolski, *Angewandte Funktionalanalysis.* Springer–Verlag, 2005.

[DuS57] N. Dunford and J.T. Schwartz, *Linear Operators Part I: General Theory.* John Wiley & Sons, 1957.

[Koe89] T.W. Körner, *Fourier Analysis.* Cambridge University Press, 1989.

[LiT96] J. Lindenstraus and L. Tzafriri, *Classical Banach Spaces 1: Sequence spaces.* Springer–Verlag, 1996.

[Lun09] A. Lunardi, *Interpolation Theory.* 2nd Edition. Edizioni della Normale, Pisa, 2009.

[Rud87] W. Rudin, *Real and Complex Analysis.* McGraw–Hill, 1987.

[Rud91] W. Rudin, *Functional Analysis.* 2nd edition. McGraw–Hill, 1991.

[ST15] R. Schnaubelt, *Spectral Theory.* Class Notes. Karlsruhe, 2015.

[Wer05] D. Werner, *Funktionalanalysis.* 5te erweiterte Auflage. Springer–Verlag, 2005.

[Wer06] D. Werner, *Einführung in die höhere Analysis.* Springer–Verlag, 2006.