Skript zur Vorlesung
"Funktionalanalysis"

Bachelor Mathematik
Sommersemester 2020

Univ.-Prof. Dr. Radu Ioan Boţ

Preface

“Mathematics is the most beautiful and most powerful creation of the human spirit.” (Stefan Banach)

Contents

Chapter I

Normed spaces

1 Definitions and examples

In this lecture we consider vector spaces over the fields $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. We use the syntagm vector space over $\mathbb{K}$ in order to indicate that both a real and a complex vector space can be meant. If not otherwise stated, we will tacitly assume that the vector space is not the trivial one $\{0\}$.

Definition 1.1 (Seminorm and norm) Let X be a vector space over $\mathbb{K}$. A function $p : X \to [0, +\infty)$ is called seminorm if

(a) $p(\lambda x) = |\lambda| p(x)\ \forall \lambda \in \mathbb{K}\ \forall x \in X$;

(b) (triangle inequality) $p(x + y) \leq p(x) + p(y)\ \forall x, y \in X$.

In this case (X, p) is called a seminormed space.

The function p is called norm if, in addition,

(c) $p(x) = 0 \Leftrightarrow x = 0$.

Norms are usually denoted with the symbol $\|\cdot\|$ and $(X, \|\cdot\|)$ is called a normed space.

One can easily see that (a) (for $\lambda = 0$ and $x = 0$) yields $p(0) = 0$.

On a normed space $(X, \|\cdot\|)$ one can induce in a natural way the metric

$$d : X \times X \to [0, +\infty), \quad d(x, y) = \|x - y\|.$$

It is easy to see that the metric axioms are, for arbitrary $x, y, z, \in X$, fulfilled:

(a) $d(x, y) \geq 0$;

(b) $d(x, y) = d(y, x)$;

(c) $d(x, y) \leq d(x, z) + d(y, z)$;

(d) $d(x, y) = 0 \Leftrightarrow x = y$.

Definition 1.2 (Cauchy sequence and convergent sequence) Let $(X, \|\cdot\|)$ be a normed space and $(x_n)_{n\in\mathbb{N}}$ a sequence in X. We say that

(a) $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence if

$$\forall\varepsilon > 0 \text{ there exists } N \in \mathbb{N} \text{ such that } \forall m, n \geq N \text{ it holds } \|x_m - x_n\| < \varepsilon.$$

(b) $(x_n)_{n\in\mathbb{N}}$ converges to $x \in X$ (and write $\lim_{n\to+\infty} x_n = x$) if

$$\forall\varepsilon > 0 \text{ there exists } N \in \mathbb{N} \text{ such that } \forall n \geq N \text{ it holds } \|x_n - x\| < \varepsilon.$$

In a normed space every convergent sequence is a Cauchy sequence, however, as we will see later, the opposite statement is in general not true.

Definition 1.3 (Banach space) A metric space (X, d) is said to be complete if every Cauchy sequence in X converges to an element in X. A complete normed space $(X, \|\cdot\|)$ is called Banach space.

In Chapter III we will show that every normed space which is not complete can be "embedded" into a Banach space.

Example 1.4 The vector space $\mathbb{K}^n$ endowed with each of the following norms (defined for $x = (x_1, ..., x_n) \in \mathbb{K}^n$)

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$
$$\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}$$
$$\|x\|_\infty = \max_{i=1,\dots,n} |x_i|$$

is a Banach space. Every sequence in $\mathbb{K}^n$ which is Cauchy/convergent with respect to one of the norms is Cauchy/convergent with respect to each of the other two norms.

Example 1.5 (The space $\ell^\infty(T)$) Let T be a nonempty set and

$$\ell^\infty(T) := \{x : T \to \mathbb{K} \mid x \text{ is a bounded function}\}.$$

It is easy to see that, for $x \in \ell^\infty(T)$,

$$\|x\|_\infty = \sup_{t\in T} |x(t)|$$

defines a norm on on $\ell^\infty(T)$, which is the so-called supremum norm.

$(\ell^\infty(T), \|\cdot\|_\infty)$ is a Banach space.

We will show that this normed space is complete. Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\ell^\infty(T)$. Let $t \in T$ be fixed. Since

$$|x_m(t) - x_n(t)| \leq \|x_m - x_n\|_\infty \ \forall m, n \in \mathbb{N},$$

it follows that $(x_n(t))_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{K}$. Since $\mathbb{K}$ is complete, there exists $x(t) := \lim_{n\to+\infty} x_n(t) \in \mathbb{K}$. This defines a function $x : T \to \mathbb{K}$. We show that $(x_n)_{n\in\mathbb{N}}$ converges to x in the supremum norm (which actually means that the sequence converges uniformly to x on $\mathbb{K}$).

Let $\varepsilon > 0$ be fixed. Then there exists $N \in \mathbb{N}$ such that

$$\|x_m - x_n\|_\infty < \frac{\varepsilon}{2} \ \forall m, n \geq N.$$

Let $t \in T$. Then there exists $M \geq N$ such that

$$|x_M(t) - x(t)| < \frac{\varepsilon}{2}.$$

For every $n \geq N$ it holds

$$\begin{aligned} |x_n(t) - x(t)| \leq & |x_n(t) - x_M(t)| + |x_M(t) - x(t)| \\ \leq & \|x_n - x_M\|_\infty + \frac{\varepsilon}{2} < \varepsilon. \end{aligned}$$

On the one hand, for every $t \in T$ it holds

$$|x(t)| \leq |x_N(t)| + |x(t) - x_N(t)| \leq \|x_N\|_\infty + \varepsilon,$$

which shows that x is bounded. On the other hand, we have

$$\|x_n - x\|_\infty < \varepsilon \ \forall n \geq N,$$

which proves that $x_n \to x$ as $n \to +\infty$.

Lemma 1.6 *(a) If X is a Banach space and M is a closed linear subspace of X, then M is complete.*

(b) If X is a normed space and M a complete linear subspace of X, then M is closed.

Proof. (a) Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in M. Since X is a Banach space, there exists $x := \lim_{n\to+\infty} x_n \in X$. Taking into account that M is closed, $x \in M$.

(b) Let $(x_n)_{n\in\mathbb{N}} \subseteq M$ be a given sequence and $x := \lim_{n\to+\infty} x_n \in X$. Since the sequence $(x_n)_{n\in\mathbb{N}}$ is convergent, it is also a Cauchy sequence. From the fact that M is complete it yields that $(x_n)_{n\in\mathbb{N}}$ must have a limit in M. Using that the limit of the sequence is unique, it follows that $x \in M$. This proves that M is closed. ■

Lemma 1.6 can be used to prove the completeness of linear subspaces of Banach spaces. To this end one just has to prove that the respective linear subspaces are closed.

Example 1.7 (The space of continuous functions) Let T be a topological (metric) space and

$$C^b(T) := \{x : T \to \mathbb{K} \mid x \text{ is a bounded and continuous function}\}.$$

It is obvious that $C^b(T)$ is a linear subspace of $\ell^\infty(T)$.

$C^b(T)$ endowed with the supremum norm is a Banach space.

To prove this, according to Lemma 1.6, it is enough to show that if $(x_n)_{n\in\mathbb{N}} \subseteq C^b(T)$ is a sequence, which converges uniformly to a bounded function x on T, then x is continuous on T.

To this end, let $t_0 \in T$ be fixed and $\varepsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $\|x_N - x\|_\infty < \frac{\varepsilon}{3}$. Since x_N is continuous, there exists a neighbourhood $U(t_0)$ of t_0 in T such that for all $t \in U(t_0)$ it holds $|x_N(t) - x_N(t_0)| < \frac{\varepsilon}{3}$. From here we have for all $t \in U(t_0)$ that

$$\begin{aligned}|x(t) - x(t_0)| \leq & |x(t) - x_N(t)| + |x_N(t) - x_N(t_0)| + |x_N(t_0) - x(t_0)| \\ \leq & 2\|x - x_N\|_\infty + |x_N(t) - x_N(t_0)| < \varepsilon.\end{aligned}$$

This shows that x is continuous at t_0. Since t_0 was arbitrarily chosen, it follows that x is continuous.

If T is a compact topological (metric) space, then every continuous function $x : T \to \mathbb{K}$ is bounded. In such a case, instead of $C^b(T)$, one writes $C(T)$.

Example 1.8 (Spaces of differentiable functions) Let

$$C^1[a,b] := \{x : [a,b] \to \mathbb{K} \mid x \text{ is continuously differentiable on } [a,b]\}.$$

$C^1[a,b]$ is a linear subspace of $C[a,b]$.

$C^1[a,b]$ endowed with the supremum norm is not a Banach space.

This is a consequence of Lemma 1.6(b) and of the fact that $C^1[a,b]$ is not closed with respect to the supremum norm. Indeed, define for every $n \geq 1$

$$x_n : [-1,1] \to \mathbb{R}, \quad x_n(t) = \left(t^2 + \frac{1}{n}\right)^{\frac{1}{2}}.$$

We obtain in this way a sequence of continuously differentiable functions, which converges uniformly to the absolute value function $|\cdot|$, which is obviously not a continuously differentiable function.

We consider on $C^1[a,b]$ the following norms (which is easy to see)

$$\begin{aligned}\|x\| =& \max\{\|x\|_\infty, \|x'\|_\infty\} = \sup_{t\in[a,b]} \max\{|x(t)|, |x'(t)|\}\\ |||x||| =& \|x\|_\infty + \|x'\|_\infty.\end{aligned}$$

$C^1[a,b]$ endowed with each of the norms $\|\cdot\|$ and $|||\cdot|||$ is a Banach space.
It holds

$$\|x\| \leq |||x||| \leq 2\|x\| \ \forall x \in C^1[a,b].$$

Thus, $(C^1[a,b], \|\cdot\|)$ is complete if and only if $(C^1[a,b], |||\cdot|||)$ is complete. Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence with respect to one of the two norms. Then $(x_n)_{n\in\mathbb{N}}$ and $(x'_n)_{n\in\mathbb{N}}$ are Cauchy sequences with respect to $\|\cdot\|_\infty$. Since $C[a,b]$ is complete with respect to the supremum norm, the two sequences have both limits in $C[a,b]$; let these be x and y, respectively. Since $(x_n)_{n\in\mathbb{N}}$ converges uniformly to x and $(x'_n)_{n\in\mathbb{N}}$ converges uniformly to y, it follows that x is differentiable and $x' = y$. In conclusion, $x \in C^1[a,b]$ and $\lim_{n\to+\infty} x_n = x$.

Analogously, for $r \in \mathbb{N}, r \geq 1$, let be

$$C^r[a,b] := \{x : [a,b] \to \mathbb{K} \mid x \text{ is } r-\text{times continuously differentiable on } [a,b]\}.$$

$C^r[a,b]$ endowed with the norm $|||x||| = \sum_{i=0}^r \|x^{(i)}\|_\infty$ is a Banach space.

Similarly, one can endow spaces of r-times continuously differentiable functions of several real variables with corresponding norms in order to obtain Banach spaces.

Example 1.9 (The sequence spaces $c_{00}, c_0, c, \ell^\infty$) Consider the vector spaces

$$\begin{aligned} c_{00} :=& \{(t_n)_{n\in\mathbb{N}} \subseteq \mathbb{K} \mid t_n \neq 0 \text{ for at most finitely many } n\}\\ c_0 :=& \{(t_n)_{n\in\mathbb{N}} \subseteq \mathbb{K} \mid \lim_{n\to+\infty} t_n = 0\}\\ c :=& \{(t_n)_{n\in\mathbb{N}} \subseteq \mathbb{K} \mid (t_n)_{n\in\mathbb{N}} \text{ is convergent}\}\\ \ell^\infty := \ell^\infty(\mathbb{N}) =& \{(t_n)_{n\in\mathbb{N}} \subseteq \mathbb{K} \mid (t_n)_{n\in\mathbb{N}} \text{ is bounded}\}\end{aligned}$$

endowed with the supremum norm

$$\|x\|_\infty = \sup_{n\in\mathbb{N}} |t_n|, \quad \text{for } x = (t_n)_{n\in\mathbb{N}}.$$

It holds

$$c_{00} \subset c_0 \subset c \subset \ell^\infty$$

and, as seen in Example 1.5, ℓ^∞ is a Banach space.

c_0, c and ℓ^∞ are Banach spaces with respect to the supremum norm, c_{00} is not a Banach space.

c_{00} is not closed (and the conclusion follows from Lemm 1.6 (b)). Let $x = (\frac{1}{n})_{n\in\mathbb{N}}$ and, for every $n \in \mathbb{N}$,

$$x_n := \left(1, \frac{1}{2}, ..., \frac{1}{n}, 0, 0, ...\right).$$

Then $(x_n)_{n\in\mathbb{N}} \subseteq c_{00}$ and $\|x_n - x\|_\infty = \frac{1}{n+1} \to 0$ as $n \to +\infty$, however, $x \notin c_{00}$.

c is closed. Let $(x_n)_{n\in\mathbb{N}} \subseteq c$ and $x_n \to x \in \ell^\infty$ as $n \to +\infty$. Let $x_n := (t_{n,k})_{k\in\mathbb{N}}$ and $\lim_{k\to+\infty} t_{n,k} = t_{n,\infty}$ for every $n \in \mathbb{N}$, and $x := (t_k)_{k\in\mathbb{N}}$. Since $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, $(t_{n,\infty})_{n\in\mathbb{N}}$ is a Cauchy sequence, too, thus convergent in $\mathbb{K}$. Let $t_\infty := \lim_{n\to+\infty} t_{n,\infty}$. We show that $\lim_{k\to+\infty} t_k = t_\infty$, which will prove that $x \in c$.

To this end, let be $\varepsilon > 0$ and $N \in \mathbb{N}$ such that

$$\|x_N - x\|_\infty < \frac{\varepsilon}{3} \text{ and } |t_{N,\infty} - t_\infty| < \frac{\varepsilon}{3}.$$

Let $K \in \mathbb{N}$ be such that for all $k \geq K$

$$|t_{N,k} - t_{N,\infty}| < \frac{\varepsilon}{3}.$$

Therefore,

$$\begin{aligned} |t_k - t_\infty| \leq& |t_k - t_{N,k}| + |t_{N,k} - t_{N,\infty}| + |t_{N,\infty} - t_\infty| \\ \leq& \|x_N - x\|_\infty + \frac{2\varepsilon}{3} < \varepsilon \ \forall k \geq K. \end{aligned}$$

c_0 is closed. Let $(x_n)_{n\in\mathbb{N}} \subseteq c_0$ and $x_n \to x \in \ell^\infty$ as $n \to +\infty$. Using the notations above, we have that $x \in c$ and $\lim_{k\to+\infty} t_k = t_\infty = \lim_{n\to+\infty} t_{n,\infty} = 0$. In conclusion, $x \in c_0$.

Example 1.10 (The sequence spaces ℓ^p $(1 \leq p < \infty)$) Let

$$\ell^p := \left\{(t_n)_{n\in\mathbb{N}} \subseteq \mathbb{K} \mid \sum_{n=1}^{+\infty} |t_n|^p < +\infty\right\}$$

and, for $x = (t_n)_{n\in\mathbb{N}}$,

$$\|x\|_p := \left(\sum_{n=1}^{+\infty} |t_n|^p\right)^{\frac{1}{p}}.$$

$(\ell^p, \|\cdot\|_p)$ is a Banach space.

For $\lambda \in \mathbb{K}$ and $x \in \ell^p$ we obviously have that $\lambda x \in \ell^p$.

Let now $x = (t_n)_{n\in\mathbb{N}}, y = (s_n)_{n\in\mathbb{N}} \in \ell^p$ and $n \geq 1$ be fixed. By the Minkowski inequality (see Übungsbeispiel 2) it holds for all $m \geq n$

$$\left(\sum_{k=1}^{n} |t_k + s_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{n} |t_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |s_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{m} |t_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{m} |s_k|^p\right)^{\frac{1}{p}}.$$

By letting $m \to +\infty$ we see that

$$\left(\sum_{k=1}^{n} |t_k + s_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{+\infty} |t_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{+\infty} |s_k|^p\right)^{\frac{1}{p}} = \|x\|_p + \|y\|_p < +\infty.$$

By letting now $n \to +\infty$ we obtain the Minkowski inequality for sequences

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p, \tag{1.1}$$

which yields $x + y \in \ell^p$.

Therefore, ℓ^p is a vector space. It is easy to see that (a), (b) (by taking into account (1.1)) and (c) in Definition 1.1 are verified, thus $\|\cdot\|_p$ is a norm on ℓ^p.

It remains to show that $(\ell^p, \|\cdot\|_p)$ is complete. To this end we consider a Cauchy sequence $(x_n)_{n\in\mathbb{N}}$ in ℓ^p, where $x_n := (t_{n,k})_{k\in\mathbb{N}}$ for all $n \in \mathbb{N}$. Let $k \in \mathbb{N}$ be fixed. Since $|t_{m,k} - t_{n,k}| \leq \|x_m - x_n\|_p$ for all $m, n \in \mathbb{N}$, it follows that $(t_{n,k})_{k\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{K}$. Thus there exists $t_k := \lim_{n\to+\infty} t_{n,k} \in \mathbb{K}$. Let $x := (t_k)_{k\in\mathbb{N}}$. We will prove that $x \in \ell^p$ and $\|x_n - x\|_p \to 0$ as $n \to +\infty$.

Let $\varepsilon > 0$ be fixed and $N \in \mathbb{N}$ such that $\|x_m - x_n\|_p < \frac{\varepsilon}{2}$ for all $m, n \geq N$. Let $K \in \mathbb{N}$ be fixed. It holds

$$\left(\sum_{k=1}^{K} |t_{m,k} - t_{n,k}|^p\right)^{\frac{1}{p}} \leq \|x_m - x_n\|_p < \frac{\varepsilon}{2} \ \forall m, n \geq N.$$

For all $n \geq N$ and all $K \in \mathbb{N}$, by letting $m \to +\infty$, we get

$$\left(\sum_{k=1}^{K} |t_k - t_{n,k}|^p\right)^{\frac{1}{p}} \leq \frac{\varepsilon}{2}$$

and further, by letting $K \to +\infty$,

$$\left(\sum_{k=1}^{+\infty} |t_k - t_{n,k}|^p\right)^{\frac{1}{p}} = \|x - x_n\|_p < \varepsilon \ \forall n \geq N.$$

Thus $\|x_n - x\|_p \to 0$ as $n \to +\infty$ and $x_N - x \in \ell^p$, which leads to $x = x - x_N + x_N \in \ell^p$.

Remark 1.11 (Hölder inequality for sequences) For $x \in \ell^1$ and $y \in \ell^\infty$ it holds $xy \in \ell^1$ and

$$\|xy\|_1 \leq \|x\|_1 \|y\|_\infty.$$

By using the Hölder inequality (see Übungsbeispiel 1 (b)) and the same technique as in Example 1.10 one can prove that, for $p, q > 1$ fulfilling $\frac{1}{p} + \frac{1}{q} = 1$ and $x \in \ell^p$ and $y \in \ell^q$, it holds $xy \in \ell^1$ and

$$\|xy\|_1 \leq \|x\|_p \|y\|_q.$$

Before introducing a new class of Banach function spaces we will prove the following lemma, which characterizes complete seminormed spaces.

Lemma 1.12 *Let $(X, \|\cdot\|)$ be a seminormed space. The following statements are equivalent:*

(i) X is complete, namely, every Cauchy sequence in X (in the sense of Definiton 1.2(a)) converges to an element in X (in the sense of Definiton 1.2(b)).

(ii) Every absolutely convergent series in X is convergent, in other words, for every sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ fulfilling $\sum_{n=1}^{+\infty} \|x_n\| < +\infty$ there exists $x \in X$ such that $\lim_{N\to+\infty} \left\|\sum_{n=1}^{N} x_n - x\right\| = 0$.

Proof. $(i) \Rightarrow (ii)$ For every $n \geq 1$ let be $s_n := \sum_{k=1}^{n} x_n$. For every $n > m \geq 1$ we have

$$\|s_n - s_m\| = \left\|\sum_{k=m+1}^{n} x_k\right\| \leq \sum_{k=m+1}^{n} \|x_k\|.$$

Thus $\lim_{m,n\to+\infty} \|s_n - s_m\| = 0$, which means that $(s_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. From here and the fact that X is complete the conclusion follows.

$(ii) \Rightarrow (i)$ Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in X. For every $k \in \mathbb{N}$, let be $N_k \in \mathbb{N}$ such that

$$\|x_m - x_n\| < \frac{1}{2^k} \quad \forall m, n \geq N_k.$$

This leads to a sequence $(x_{n_k})_{k\in\mathbb{N}}$ fulfilling

$$\|x_{n_{k+1}} - x_{n_k}\| < \frac{1}{2^k} \quad \forall k \in \mathbb{N}.$$

We notice that $\sum_{k=1}^{+\infty} \|x_{n_{k+1}} - x_{n_k}\| < +\infty$, which means that there exists $y \in X$ such that

$$0 = \lim_{K\to+\infty} \left\|\sum_{k=1}^{K} (x_{n_{k+1}} - x_{n_k}) - y\right\| = \lim_{K\to+\infty} \left\|x_{n_{K+1}} - x_{n_1} - y\right\|.$$

We will show that the whole sequence $(x_n)_{n\in\mathbb{N}}$ converges to $x_{n_1} + y$ (in the sense of Definiton 1.2(b)). Let $\varepsilon > 0$ and $N \in \mathbb{N}$ be such that

$$\|x_m - x_n\| < \frac{\varepsilon}{2} \quad \forall m, n \geq N.$$

Let $k \in \mathbb{N}$ be such that $n_k \geq N$ and

$$\|x_{n_k} - (x_{n_1} + y)\| < \frac{\varepsilon}{2}.$$

Then for all $n \geq N$ it holds

$$\|x_n - (x_{n_1} + y)\| \leq \|x_n - x_{n_k}\| + \|x_{n_k} - (x_{n_1} + y)\| < \varepsilon,$$

and the statement is proved. ■

Example 1.13 (The spaces L^p $(1 \leq p < \infty)$) Let $I \subseteq \mathbb{R}$ be an interval, Σ the Borel σ-algebra on I (the smallest σ-algebra containing all open subsets of I) and λ the Lebesgue measure on I. Let (if f is measurable, then $|f|^p$ is also measurable)

$$\mathcal{L}^p(I) := \left\{ f : I \to \mathbb{K} \mid f \text{ is measurable and } \int_I |f|^p d\lambda < +\infty \right\}$$

and, for $f \in \mathcal{L}^p(I)$,

$$\|f\|_p^* := \left(\int_I |f|^p d\lambda \right)^{\frac{1}{p}}.$$

For $\alpha \in \mathbb{K}$ and $f \in \mathcal{L}^p(I)$ we obviously have that $\alpha f \in \mathcal{L}^p(I)$. Choose now $f, g \in \mathcal{L}^p(I)$. Then $f + g$ is measurable and it holds

$$\begin{aligned}
\int_I |f(t) + g(t)|^p d\lambda(t) &\leq \int_I (|f(t)| + |g(t)|)^p d\lambda(t) \leq \int_I (2 \max\{|f(t)|, |g(t)|\})^p d\lambda(t) \\
&= 2^p \int_I \max\{|f(t)|^p, |g(t)|^p\} d\lambda(t) \\
&\leq 2^p \left(\int_I |f(t)|^p d\lambda(t) + \int_I |g(t)|^p d\lambda(t) \right) < +\infty.
\end{aligned}$$

This proves that $\mathcal{L}^p(I)$ is a vector space.

$(\mathcal{L}^p(I), \|\cdot\|_p^*)$ is a complete seminormed space.

It holds

$$\|\alpha f\|_p^* = \left(\int_I |\alpha f(t)|^p d\lambda(t) \right)^{\frac{1}{p}} = |\alpha| \left(\int_I |f|^p d\lambda \right)^{\frac{1}{p}} = |\alpha| \|f\|_p^* \ \forall \alpha \in \mathbb{K}\ \forall f \in \mathcal{L}^p(I).$$

In order to prove the triangle inequality we will derive first variants of the Hölder and the Minkowski inequality in the context of function spaces.

A variant of the Hölder inequality for $\mathcal{L}^p(I)$: Let $p, q > 1$ be such that $\frac{1}{p} + \frac{1}{q} = 1$, $f \in \mathcal{L}^p(I)$ and $g \in \mathcal{L}^q(I)$. Then $fg \in \mathcal{L}^1(I)$ and

$$\|fg\|_1^* \leq \|f\|_p^* \|g\|_q^*. \tag{1.2}$$

We denote $A := (\|f\|_p^*)^p$ and $B := (\|g\|_q^*)^q$. If $A = 0$, then $f = 0$ almost everywhere (a.e.), thus $fg = 0$ a.e. and the conclusion follows. We come to the same conclusion if $B = 0$, thus we can assume that $A, B > 0$. The Young inequality (see Übungsbeispiel 1(a)) yields for every $t \in I$

$$\frac{|f(t)|}{A^{\frac{1}{p}}} \frac{|g(t)|}{B^{\frac{1}{q}}} = \left(\frac{|f(t)|^p}{A} \right)^{\frac{1}{p}} \left(\frac{|g(t)|^q}{B} \right)^{\frac{1}{q}} \leq \frac{1}{p} \frac{|f(t)|^p}{A} + \frac{1}{q} \frac{|g(t)|^q}{B}.$$

The conclusion follows by integration.

A variant of the Minkowski inequality for $\mathcal{L}^p(I)$: Let $p \geq 1$ and $f, g \in \mathcal{L}^p(I)$. Then

$$\|f+g\|_p^* \leq \|f\|_p^* + \|g\|_p^*. \tag{1.3}$$

We assume that $p > 1$ (the case $p = 1$ is trivial) and define $q := \frac{p}{p-1}$. Then $\frac{1}{p} + \frac{1}{q} = 1$. We apply twice (1.2):

$$\begin{aligned}
(\|f+g\|_p^*)^p &= \int_I |f(t)+g(t)|^p d\lambda(t) \\
&\leq \int_I |f(t)||f(t)+g(t)|^{p-1} d\lambda(t) + \int_I |g(t)||f(t)+g(t)|^{p-1} d\lambda(t) \\
&\leq \|f\|_p^* \||f+g|^{p-1}\|_q^* + \|g\|_p^* \||f+g|^{p-1}\|_q^* \\
&= (\|f\|_p^* + \|g\|_p^*)(\|f+g\|_p^*)^{p-1}.
\end{aligned}$$

From here, the conclusion follows.

The triangle inequality for $\|\cdot\|_p^*$ is nothing else than (1.3). Thus $(\mathcal{L}^p(I), \|\cdot\|_p^*)$ is a seminormed space. One can easily see that $\|\cdot\|_p^*$ is “only” a seminorm, since $\|f\|_p^* = 0$ if and only if $f = 0$ almost everywhere. However, as we will see in the following, $(\mathcal{L}^p(I), \|\cdot\|_p^*)$ is complete.

We will use Lemma 1.12 $(ii) \Rightarrow (i)$. Let $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{L}^p(I)$ be such that $M := \sum_{n=1}^{+\infty} \|f_n\|_p^* < +\infty$. We define for every $n \in \mathbb{N}$ the functions $g_n(t) = \sum_{k=1}^{n} |f_k(t)|$ for all $t \in I$. Since $\mathcal{L}^p(I)$ is a vector space, $g_n \in \mathcal{L}^p(I)$ for every $n \in \mathbb{N}$. In addition, according to the Minkowski inequality, we have

$$\|g_n\|_p^* \leq \sum_{k=1}^{n} \|f_k\|_p^* \leq M < +\infty \quad \forall n \in \mathbb{N}.$$

Since for every $t \in I$, $0 \leq g_n(t) \leq g_{n+1}(t)$ for every $n \in \mathbb{N}$, there exists $g(t) := \lim_{n\to+\infty} g_n(t) = \sum_{k=1}^{+\infty} |f_k(t)| \in [0, +\infty]$. The function g is measurable. Making use of the Theorem of Beppo Levi we can conclude that

$$\int_I g(t)^p d\lambda(t) = \lim_{n\to+\infty} \int_I g_n(t)^p d\lambda(t) \leq M^p.$$

This mean that g^p and, consequently, g is finite almost everywhere. In other words, there exists a set $N \in \Sigma$ of measure zero such that

$$g(t) = \sum_{k=1}^{+\infty} |f_k(t)| \in \mathbb{R} \quad \forall t \in I \setminus N.$$

Using the Cauchy criterion for series we can conclude that

$$f(t) := \sum_{k=1}^{+\infty} f_k(t) \in \mathbb{K}$$

exists for every $t \in I \setminus N$. We define $f(t) := 0$ for all $t \in N$ and get in this way a measurable function $f : I \to \mathbb{K}$.

We define for every $n \in \mathbb{N}$ the functions $s_n(t) = \sum_{k=1}^{n} f_k(t)$ for all $t \in I$. Then

$$|s_n(t)| \le g_n(t) \le g(t) \quad \forall n \in \mathbb{N}\ \forall t \in I \setminus N.$$

Therefore $|f(t)| \le g(t)$ for all $t \in I \setminus N$, which implies that $\int_I |f(t)|^p d\lambda(t) \le \int_I g(t)^p d\lambda(t) \le M^p$, thus $f \in \mathcal{L}^p(I)$.

Finally, we have $\lim_{n\to+\infty} |s_n(t) - f(t)| = 0$, thus $\lim_{n\to+\infty} |s_n(t) - f(t)|^p = 0$ for every $t \in I \setminus N$. On the other hand,

$$0 \le |s_n(t) - f(t)|^p = \lim_{m\to+\infty} \left| \sum_{k=n}^{m} f_k(t) \right|^p \le \left(\lim_{m\to+\infty} \sum_{k=n}^{m} |f_k(t)| \right)^p = g(t)^p$$

for every $t \in I \setminus N$ and g is Lebesgue integrable. By the Lebesgue dominated convergence Theorem we conclude that

$$\lim_{n\to+\infty} \left\| \sum_{k=1}^{n} f_k - f \right\|_p^* = \lim_{n\to+\infty} \|s_n - f\|_p^* = \lim_{n\to+\infty} \left(\int_I |s_n(t) - f(t)|^p d\lambda(t) \right)^{\frac{1}{p}} = 0.$$

In order to obtain a complete normed space we have to identify functions which coincide almost everywhere. This approach uses the notion of quotient (semi) normed space.

Definition 1.14 Let $(X, \|\cdot\|)$ be a seminormed space and $A \subseteq X$. The distance function to the set A is defined as

$$\operatorname{dist}(\cdot, A) : X \to \mathbb{R}, \quad \operatorname{dist}(x, A) := \inf\{\|x - a\| : a \in A\}.$$

It holds $\operatorname{dist}(x, A) = 0 \Leftrightarrow x \in \overline{A}$.

Theorem 1.15 *Let X be a seminormed space over $\mathbb{K}$ and $M \subseteq X$ a nonempty linear subspace. Let*

$$X/M := \{[x] = x + M \mid x \in X\}$$

be the quotient space of X by M. Recall that X/M with the operations

$$[x] + [y] := [x + y] \quad \forall x, y \in X$$

and

$$\lambda[x] := [\lambda x] \quad \forall \lambda \in \mathbb{K}\ \forall x \in X$$

is a vector space over $\mathbb{K}$. The following statements are true:

(a) $\|[x]\| := \operatorname{dist}(x, M)$ *defines a seminorm on X/M.*

(b) If M is closed, then $\|\cdot\|$ is a norm on X/M.

(c) If X is complete and M is closed, then X/M is complete.

Proof. (a) We prove first that $\|\cdot\|$ is well-defined. If $x+M=y+M$, then there exists $m\in M$ such that $x=y+m$. Thus $\operatorname{dist}(x,M)=\inf\{\|y+m-m'\|: m'\in M\}=\inf\{\|y-m'\|: m'\in M\}=\operatorname{dist}(y,M)$.

It is clear that $\operatorname{dist}(x,M)\geq 0$ for all $x\in X$.

For $\lambda\in\mathbb{K}$ and $x\in X$ it holds

$$\begin{aligned}\|\lambda[x]\| &= \operatorname{dist}(\lambda x,M)=\inf\{\|\lambda x-m\|: m\in M\}=\inf\{\|\lambda x-\lambda m\|: m\in M\}\\ &= |\lambda|\inf\{\|x-m\|: m\in M\}=|\lambda|\operatorname{dist}(x,M)=|\lambda|\|[x]\|.\end{aligned}$$

Let $x,y\in X$ and $\varepsilon>0$. Then there exist $m,n\in M$ such that $\|x-m\|\leq\|[x]\|+\varepsilon$ and $\|y-n\|\leq\|[y]\|+\varepsilon$. It holds

$$\begin{aligned}\|[x]+[y]\| &= \|[x+y]\|\leq\|(x+y)-(m+n)\|\leq\|x-m\|+\|y-n\|\\ &\leq \|[x]\|+\|[y]\|+2\varepsilon.\end{aligned}$$

Letting $\varepsilon\to 0$, it yields $\|[x]+[y]\|\leq\|[x]\|+\|[y]\|$.

(b) We have that $\|[x]\|=0\Leftrightarrow\operatorname{dist}(x,M)=0\Leftrightarrow x\in\overline{M}=M\Leftrightarrow[x]=[0]$.

(c) Let $(x_n)_{n\in\mathbb{N}}\subseteq X$ be such that $\sum_{n=1}^{+\infty}\|[x_n]\|<+\infty$. For every $n\in\mathbb{N}$ there exists $m_n\in M$ such that $\|x_n-m_n\|\leq\|[x_n]\|+2^{-n}$. Then $\sum_{n=1}^{+\infty}\|x_n-m_n\|<+\infty$ and, according to Lemma 1.12 $(i)\Rightarrow(ii)$, there exists $x:=\sum_{n=1}^{+\infty}(x_n-m_n)$. This implies for every $N\in\mathbb{N}$

$$\begin{aligned}\left\|[x]-\sum_{n=1}^{N}[x_n]\right\| &= \left\|\left[x-\sum_{n=1}^{N}x_n\right]\right\| = \operatorname{dist}\left(x-\sum_{n=1}^{N}x_n, M\right)\\ &\leq \left\|x-\sum_{n=1}^{N}(x_n-m_n)\right\|.\end{aligned}$$

Letting $N\to+\infty$, it yields $\sum_{n=1}^{+\infty}[x_n]=[x]$. In the light of Lemma 1.12 $(ii)\Rightarrow(i)$ we obtain that X/M is complete. ■

Remark 1.16 If $(X,\|\cdot\|)$ is a seminormed space, then $N:=\{x\in X:\|x\|=0\}$ is a closed linear subspace of X. For every $x\in X$ it holds $\operatorname{dist}(x,N)=\|x\|$. Thus, according to Theorem 1.15, $\|[x]\|:=\|x\|$ for every $x\in X$, defines a norm on X/N and, if X is complete, then X/N is a Banach space.

Coming back to the L^p spaces discussed in Example 1.13, we consider the kernel of the seminorm $\|\cdot\|_p^*$ (which actually does not depend on p)

$$N_p:=\{f \text{ is measurable} \mid f=0 \text{ a.e.}\}=\{f \text{ is measurable} \mid \|f\|_p^*=0\}$$

and the quotient vector space

$$L^p(I) := \mathcal{L}^p(I)/N_p.$$

We denote the quotient norm on $L^p(I)$ by $\|\cdot\|_p$ or by $\|\cdot\|_{L^p}$. In view of Remark 1.16 we have that:

$(L^p(I), \|\cdot\|_p)$ is a Banach space for $p \geq 1$.

The elements of $L^p(I)$ are equivalence classes of functions, however, they are handled as functions. This means that instead of $[f] \in L^p(I)$ we write $f \in L^p(I)$.

Having a measure space (Ω, Σ, μ) one can similarly introduce for $p \geq 1$ the space $L^p(\Omega, \Sigma, \mu)$, which is also shortly denoted by $L^p(\mu)$, and prove that:

$(L^p(\mu), \|\cdot\|_p)$ is a Banach space for $p \geq 1$.

For $\Omega = \mathbb{N}$, $\Sigma =$ the power set of the set of natural numbers, and $\mu =$ the counting measure, it holds $L^p(\Omega, \Sigma, \mu) = \ell^p$.

Example 1.17 (The spaces L^∞) Let $I \subseteq \mathbb{R}$ be an interval, Σ the Borel σ-algebra on I and λ the Lebesgue measure on I. Further, let be

$$\begin{aligned}\mathcal{L}^\infty(I) := \{f : I \to \mathbb{K} \mid f \text{ is measurable and} \\ \exists N \in \Sigma, \lambda(N) = 0 \text{ such that } f|_{I\setminus N} \text{ is bounded}\}\end{aligned}$$

and, for $f \in \mathcal{L}^\infty(I)$,

$$\|f\|^*_{L^\infty} := \inf_{N\in\Sigma, \lambda(N)=0} \sup_{t\in I\setminus N} |f(t)| = \inf_{N\in\Sigma, \lambda(N)=0} \|f|_{I\setminus N}\|_\infty \ (< +\infty).$$

It is easy to see that $\mathcal{L}^\infty(I)$ is a vector space.

We notice that for every $f \in \mathcal{L}^\infty(I)$ there exists a set $N = N(f) \in \Sigma$ of measure zero such that $\|f\|^*_{L^\infty} = \|f|_{I\setminus N}\|_\infty$. Indeed, for every $k \in \mathbb{N}$, let be $N_k \in \Sigma$ with $\lambda(N_k) = 0$ such that $\|f|_{I\setminus N_k}\|_\infty \leq \|f\|^*_{L^\infty} + \frac{1}{k}$. Define $N := \cup_{k\in\mathbb{N}} N_k \in \Sigma$. Then $\lambda(N) = 0$ and

$$\|f\|^*_{L^\infty} \leq \|f|_{I\setminus N}\|_\infty \leq \|f|_{I\setminus N_k}\|_\infty \leq \|f\|^*_{L^\infty} + \frac{1}{k} \quad \forall k \in \mathbb{N}.$$

Letting $k \to +\infty$, the conclusion follows.

$(\mathcal{L}^\infty(I), \|\cdot\|^*_\infty)$ is a complete seminormed space.

For every $\alpha \in \mathbb{K}$ and every $f \in \mathcal{L}^\infty(I)$ it holds

$$\|\alpha f\|^*_\infty = \inf_{N\in\Sigma, \lambda(N)=0} \|\alpha f|_{I\setminus N}\|_\infty = |\alpha| \inf_{N\in\Sigma, \lambda(N)=0} \|f|_{I\setminus N}\|_\infty = |\alpha| \|f\|^*_\infty.$$

Let $f, g \in \mathcal{L}^\infty(I)$ and $N(f), N(g) \in \Sigma$ such that $\lambda(N(f)) = \lambda(N(g)) = 0$, $\|f\|^*_{L^\infty} = \|f|_{I\setminus N(f)}\|_\infty$ and $\|g\|^*_{L^\infty} = \|g|_{I\setminus N(g)}\|_\infty$. We have

$$\begin{aligned}\|f+g\|^*_{L^\infty} &\leq \|(f+g)|_{I\setminus(N(f)\cup N(g))}\|_\infty \leq \|f|_{I\setminus(N(f)\cup N(g))}\|_\infty + \|g|_{I\setminus(N(f)\cup N(g))}\|_\infty \\ &\leq \|f|_{I\setminus N(f)}\|_\infty + \|g|_{I\setminus N(g)}\|_\infty = \|f\|^*_{L^\infty} + \|g\|^*_{L^\infty}.\end{aligned}$$

Next we will prove that $\mathcal{L}^\infty(I)$ is complete. Let $(f_n)_{n\in\mathbb{N}}$ a Cauchy sequence in $\mathcal{L}^\infty(I)$ and, for every pair $(n,m)\in\mathbb{N}\times\mathbb{N}$, let $N_{n,m}\in\Sigma$ such that $\lambda(N_{n,m})=0$ and $\|f_n-f_m\|^*_{L^\infty}=\|(f_n-f_m)|_{I\setminus N_{n,m}}\|_\infty$. Define $N:=\cup_{n,m\in\mathbb{N}}N_{n,m}\in\Sigma$. It holds $\lambda(N)=0$ and for all $n,m\in\mathbb{N}$

$$\|(f_n-f_m)|_{I\setminus N}\|_\infty\le\|(f_n-f_m)|_{I\setminus N_{n,m}}\|_\infty=\|f_n-f_m\|^*_{L^\infty}.$$

This implies that $(f_n|_{I\setminus N})_{n\in\mathbb{N}}$ is a Cauchy sequence in $(\ell^\infty(I\setminus N),\|\cdot\|_\infty)$ and, consequently, convergent to a function $f:I\setminus N\to\mathbb{K}$. We define $f(t)=0$ for all $t\in N$ and notice that f is the uniform limit of the sequence of measurable and bounded functions $(f_n\chi_{I\setminus N})_{n\in\mathbb{N}}$, where, for a set $A\subseteq I$, $\chi_A(t)=1$ for $t\in A$ and $\chi_A(t)=0$ for $t\notin A$. This implies that

$$\lim_{n\to+\infty}\|f_n-f\|^*_{L^\infty}\le\lim_{n\to+\infty}\|(f_n-f)|_{I\setminus N}\|_\infty=0.$$

The kernel of the seminorm $\|\cdot\|^*_\infty$ is

$$N_\infty:=\{f\text{ is measurable}\mid f=0\text{ a.e.}\}$$

and the corresponding quotient vector space is denoted by

$$L^\infty(I):=\mathcal{L}^\infty(I)/N_\infty.$$

The quotient norm on $L^\infty(I)$ is denoted by $\|\cdot\|_{L^\infty}$ and it is called the essential supremum norm. In view of Remark 1.16 we have that

$(L^\infty(I),\|\cdot\|_{L^\infty})$ is a Banach space.

The elements of $L^\infty(I)$ are handled as functions, which means that instead of $[f]\in L^\infty(I)$ we write $f\in L^\infty(I)$.

For a given measure space (Ω,Σ,μ) one can introduce in a similar way the space $L^\infty(\Omega,\Sigma,\mu)$, which is also shortly denoted by $L^\infty(\mu)$, and prove that

$(L^\infty(\mu),\|\cdot\|_{L^\infty})$ is a Banach space.

2 Properties of normed spaces

The following theorem shows that addition, multiplication with scalar and $\|\cdot\|$ are continuous mappings on normed spaces.

Theorem 2.1 *Let X a normed space. The following statements are true:*

(a) If $x_n\to x$ and $y_n\to y$ as $n\to+\infty$, then $x_n+y_n\to x+y$ as $n\to+\infty$.

(b) If $\lambda_n\to\lambda$ in $\mathbb{K}$ and $x_n\to x$ as $n\to+\infty$, then $\lambda_nx_n\to\lambda x$ as $n\to+\infty$.

(c) If $x_n\to x$ as $n\to+\infty$, then $\|x_n\|\to\|x\|$ as $n\to+\infty$.

Proof. (a) We have

$$\|(x_n + y_n) - (x + y)\| \leq \|x_n - x\| + \|y_n - y\| \to 0 \quad (n \to +\infty).$$

(b) We have

$$\|\lambda_n x_n - \lambda x\| \leq |\lambda_n - \lambda|\|x_n\| + |\lambda|\|x_n - x\| \to 0 \quad (n \to +\infty).$$

(c) We have

$$|\|x_n\| - \|x\|| \leq \|x_n - x\| \to 0 \quad (n \to +\infty).$$

■

As a direct consequence of Theorem 2.1, we obtain that the closure of a linear subspace of a normed space is also a linear subspace.

Definition 2.2 (Equivalent norms) Two norms $\|\cdot\|$ and $|||\cdot|||$ defined on a vector space X are said to be equivalent if there exist two positive numbers $0 < m \leq M$ such that

$$m\|x\| \leq |||x||| \leq M\|x\| \quad \forall x \in X.$$

Theorem 2.3 *Let $\|\cdot\|$ and $|||\cdot|||$ be two norms on X. The following statements are equivalent:*

(i) $\|\cdot\|$ and $|||\cdot|||$ are equivalent.

(ii) A sequence in X is convergent with respect to $\|\cdot\|$ if and only if it is convergent with respect to $|||\cdot|||$.

(c) A sequence in X converges to zero with respect to $\|\cdot\|$ if and only if it converges to zero with respect to $|||\cdot|||$.

Proof. The implications $(i) \Rightarrow (ii) \Rightarrow (iii)$ are trivial.

$(iii) \Rightarrow (i)$. We assume that there is no $M > 0$ such that $|||x||| \leq M\|x\|$ for every $x \in X$. This means that for every $n \in \mathbb{N}$ there exists $x_n \in X$ such that $|||x_n||| > n\|x_n\|$. We define $y_n := \frac{x_n}{n\|x_n\|}$ for every $n \in \mathbb{N}$ and notice that $\|y_n\| = \frac{1}{n} \to 0$ $(n \to +\infty)$. Since $|||y_n||| > 1$ for every $n \in \mathbb{N}$, $(y_n)_{n\in\mathbb{N}}$ does not converge to zero with respect to $|||\cdot|||$. Contradiction! The existence of $m > 0$ can be shown analogously. ■

If the norms $\|\cdot\|$ and $|||\cdot|||$ defined on a vector space X are equivalent, then

$$\begin{aligned} &(x_n)_{n\in\mathbb{N}} \text{ is a Cauchy sequence in } (X, \|\cdot\|) \\ \Leftrightarrow\ &(x_n)_{n\in\mathbb{N}} \text{ is a Cauchy sequence in } (X, |||\cdot|||), \end{aligned}$$

which means that

$$(X, \|\cdot\|) \text{ is complete} \ \Leftrightarrow\ (X, |||\cdot|||) \text{ is complete}.$$

Example 2.4 (a) The norms $\|\cdot\|$ and $|||\cdot|||$ on $C^1[a,b]$ introduced in Example 1.8 are equivalent. They are not equivalent to $\|\cdot\|_\infty$.

(b) The norms $\|\cdot\|_1$ (see Übungsbeispiel 7) and $\|\cdot\|_\infty$ on $C[0,1]$ are not equivalent. It holds $\|x\|_1 \le \|x\|_\infty$ for every $x \in C[0,1]$. However, $(C[0,1], \|\cdot\|_\infty)$ is complete (see Example 1.7), while $(C[0,1], \|\cdot\|_1)$ is not complete (see Übungsbeispiel 7).

(c) For $\alpha > 0$, we define on $C[0,1]$ the norm

$$|||x|||_\alpha := \sup_{0\le t\le 1} |x(t)|e^{-\alpha t}.$$

Such constructions are useful in the theory of ordinary differential equations. It holds

$$|||x|||_\alpha \le \|x\|_\infty \le e^\alpha |||x|||_\alpha \quad \forall x \in C[0,1],$$

which means that $|||\cdot|||_\alpha$ and $\|\cdot\|_\infty$ are equivalent.

Theorem 2.5 *Any two norms defined on a finite-dimensional space are equivalent.*

Proof. Let X be a vector space of dimension n, $\{e_1, ..., e_n\}$ a basis of X and $\|\cdot\|$ a norm on X. We will show that $\|\cdot\|$ is equivalent to the Euclidean norm $\|\cdot\|_2$. The triangle inequality and the Cauchy-Schwarz (Hölder) inequality yield

$$\begin{aligned}\left\|\sum_{i=1}^n \alpha_i e_i\right\| &\le \sum_{i=1}^n |\alpha_i|\|e_i\| \le \left(\sum_{i=1}^n |\alpha_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^n \|e_i\|^2\right)^{\frac{1}{2}} \\ &= \left(\sum_{i=1}^n \|e_i\|^2\right)^{\frac{1}{2}} \left\|\sum_{i=1}^n \alpha_i e_i\right\|_2 \qquad \forall(\alpha_1, ..., \alpha_n) \in \mathbb{K}^n,\end{aligned}$$

therefore

$$\|x\| \le \left(\sum_{i=1}^n \|e_i\|^2\right)^{\frac{1}{2}} \|x\|_2 \quad \forall x \in X.$$

This implies that $\|\cdot\|$ is continuous in $(X, \|\cdot\|_2)$, since, for $x_n \xrightarrow{\|\cdot\|_2} x$ $(n \to +\infty)$ we have (see Theorem 2.1(c)) $\|x_n\| \to \|x\|$ $(n \to +\infty)$.

Using that $S = \{x \in X : \|x\|_2 = 1\}$ is compact in $(X, \|\cdot\|_2)$, there exists $m > 0$ such that

$$m := \min_{x\in S} \|x\|.$$

Since $\frac{1}{\|x\|_2} x \in S$ for every $x \neq 0$, we obtain that

$$m\|x\|_2 \le \|x\| \quad \forall x \in X.$$

This shows that every norm on X is equivalent to $\|\cdot\|_2$, and from here the conclusion follows. ■

The following lemma will play an important role in the characterization of finite-dimensional normed spaces.

Lemma 2.6 (Riesz Lemma) *Let X be a normed space and $M \subseteq X$ a closed linear subspace of X such that $M \neq X$ and $0 < \delta < 1$. Then there exists $x_\delta \in X$ such that $\|x_\delta\| = 1$ and*

$$\|x_\delta - m\| \geq 1 - \delta \quad \forall m \in M.$$

Proof. Let $x \in X \setminus M$. Since M is closed, $\operatorname{dist}(x, M) = \inf\{\|x - m\| : m \in M\} > 0$ (see Definition 1.14). Since $\operatorname{dist}(x, M) < \frac{1}{1-\delta} \operatorname{dist}(x, M)$, there exists $m_\delta \in M$ such that $\|x - m_\delta\| < \frac{1}{1-\delta} d(x, M)$. Define

$$x_\delta := \frac{x - m_\delta}{\|x - m_\delta\|}.$$

Then $\|x_\delta\| = 1$ and for every $m \in M$ it holds (notice that $m_\delta + \|x - m_\delta\| m \in M$)

$$\|x_\delta - m\| = \frac{1}{\|x - m_\delta\|} \|x - (m_\delta + \|x - m_\delta\| m)\| \geq \frac{1-\delta}{\operatorname{dist}(x, M)} \operatorname{dist}(x, M) = 1 - \delta.$$

■

Remark 2.7 Riesz Lemma does not hold in general for $\delta = 0$. In order to see this, consider $X := \{x \in C[0, 1] \mid x(0) = 0\}$ endowed with the supremum norm and $M := \{x \in X \mid \int_0^1 x(t)dt = 0\}$ (see Übungsbeispiel 13).

Theorem 2.8 *Let X be a normed space. The following statements are equivalent:*

(i) $\dim X < \infty$.

(ii) $B_X := \{x \in X \mid \|x\| \leq 1\}$ *is compact.*

(iii) Every bounded sequence in X has a convergent subsequence.

Proof. $(i) \Rightarrow (ii)$. Follows from the equivalence of norms in finite-dimensional normed spaces and the Theorem of Heine-Borel, which says that a subset of $\mathbb{R}^n$ is compact if and only if it is bounded and closed.

$(ii) \Rightarrow (iii)$. Let $(x_n)_{n\in\mathbb{N}}$ and $M \geq 0$ be such that $\|x_n\| \leq M$ for all $n \in \mathbb{N}$. Then $(x_n)_{n\in\mathbb{N}} \subseteq MB_X$, which, according to (ii), is a compact set. The conclusion follows from the Theorem of Bolzano-Weierstraß.

$(iii) \Rightarrow (i)$. Assume that $\dim X = \infty$. Let $x_1 \in X$ with $\|x_1\| = 1$ and define $M_1 := \operatorname{lin}\{x_1\}$, where $\operatorname{lin} A$ denotes the linear subspace generated by the set A. Then M_1 is finite-dimensional, thus $M_1 \neq X$ and, according to Lemma 1.6 (b), it is a closed linear subspace of X. According to Riesz Lemma, there exists $x_2 \in X$

such that $\|x_2\| = 1$ and $\|x_2 - x_1\| \geq \frac{1}{2}$. We define $M_2 := \text{lin}\{x_1, x_2\}$ and make use of the same arguments to obtain an element $x_3 \in X$ such that $\|x_3\| = 1$ and $\|x_3 - x_1\| \geq \frac{1}{2}$, $\|x_3 - x_2\| \geq \frac{1}{2}$. In this way we construct inductively a sequence $(x_n)_{n\in\mathbb{N}}$ with the property that $\|x_n\| = 1$ for every $n \in \mathbb{N}$ and $\|x_n - x_m\| \geq \frac{1}{2}$ for every $n \neq m$. The sequence $(x_n)_{n\in\mathbb{N}}$ is bounded, however it cannot have a convergent subsequence, since it has no Cauchy subsequence. ■

In the final part of this chapter we will introduce and discuss the notion of separability of normed spaces.

Definition 2.9 (Separable space) A topological (metric) space T is called separable if it has a countable dense subset, in other words, if there exists a countable set $D \subseteq T$ such that $\overline{D} = T$.

$\mathbb{R}^n$ is separable, since $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$. Since every subset of a separable metric space is separable, it follows that every $T \subseteq \mathbb{R}^n$ is a separable space.

Lemma 2.10 *Let X be a normed space. The following statements are equivalent:*

(i) X is separable.

(ii) There exists a countable set A such that $X := \overline{\text{lin}\, A}$.

Proof. *(i)* $\Rightarrow$ *(ii)*. Since X is separable, $X = \overline{A}$, for a countable set $A \subseteq X$. Then $X = \overline{A} \subseteq \overline{\text{lin}\, A} \subseteq X$ and the conclusion follows.

(ii) $\Rightarrow$ *(i)*. We consider first the case when $\mathbb{K} = \mathbb{R}$ and denote

$$B := \left\{ \sum_{k=1}^{n} \lambda_k x_k : n \in \mathbb{N}, \lambda_k \in \mathbb{Q}, x_k \in A, k = 1, ..., n \right\}.$$

It is obvious that B is countable. We will prove that $\overline{B} = X$. Let $x \in X$ and $\varepsilon > 0$. Choose $a_0 \in \text{lin}\, A, a_0 = \sum_{k=1}^{n} \lambda_k x_k$, where $n \in \mathbb{N}, \lambda_k \in \mathbb{R}$ and $x_k \in A, k = 1, ..., n$, such that $\|x - a_0\| < \frac{\varepsilon}{2}$. For every $k = 1, ..., n$, choose $\mu_k \in \mathbb{Q}$ such that $\left(\sum_{k=1}^{n} \|x_k\|\right) |\lambda_k - \mu_k| < \frac{\varepsilon}{2}$. Then, for $a := \sum_{k=1}^{n} \mu_k x_k \in B$, it holds

$$\|x - a\| \leq \|x - a_0\| + \|a_0 - a\| \leq \frac{\varepsilon}{2} + \max_{k=1,\dots,n} |\lambda_k - \mu_k| \left(\sum_{k=1}^{n} \|x_k\| \right) < \varepsilon.$$

In the case when $\mathbb{K} = \mathbb{C}$ one just needs to replace in the above proof $\mathbb{Q}$ by $\mathbb{Q} + i\mathbb{Q}$. ■

Example 2.11

(a) The space ℓ^p is separable for all $1 \leq p < \infty$. For $n \in \mathbb{N}$, let

$$e_n = (0, ..., 0, 1, 0, ...),$$

where the n-th element is 1 and all other elements are 0. For $A := \{e_n \mid n \in \mathbb{N}\}$ it holds $\ell^p = \overline{\operatorname{lin} A}$. Indeed, for $x = (t_n)_{n\in\mathbb{N}}$ we have

$$\left\| x - \sum_{k=1}^{n} t_k e_k \right\|_p = \left(\sum_{k=n+1}^{+\infty} |t_k|^p \right)^{\frac{1}{p}} \to 0 \quad (n \to +\infty).$$

(b) The spaces c_{00}, c_0 and c are separable.

(c) The space ℓ^∞ is not separable. Assume that ℓ^∞ is separable and let D be a countable set in ℓ^∞ such that $\overline{D} = \ell^\infty$. Let $\mathcal{F}$ be the family of all subsets of $\mathbb{N}$. Then $\mathcal{F}$ is uncountable. For every $F \in \mathcal{F}$, let $\chi_F \in \ell^\infty$ be the sequence defined as $\chi_F(n) = 1$, for $n \in F$, and $\chi_F(n) = 0$, otherwise. If $F, G \in \mathcal{F}$ with $F \neq G$, then $\|\chi_F - \chi_G\|_\infty = 1$. For every $F \in \mathcal{F}$, let $d_F \in D$ be such that $\|\chi_F - d_F\|_\infty < \frac{1}{4}$. We notice that, if $F, G \in \mathcal{F}$ with $F \neq G$, then $\|d_F - d_G\|_\infty > \frac{1}{4}$. Indeed, assuming that $\|d_F - d_G\|_\infty \leq \frac{1}{4}$, it holds

$$\|\chi_F - \chi_G\|_\infty \leq \|\chi_F - d_F\|_\infty + \|d_F - d_G\|_\infty + \|d_G - \chi_G\|_\infty \leq \frac{3}{4}.$$

Contradiction! This proves that $F \mapsto d_F$ is injective and, since it maps an uncountable set into a countable one, we obtain the desired contradiction.

Example 2.12

(a) The space $C[a, b]$ is separable. According to the Stone-Weierstrass Theorem (see [4, Theorem 7.26]), the linear subspace $P[a, b]$ of polynomial functions defined on $[a, b]$ is dense in $(C[a, b], \|\cdot\|_\infty)$. Thus $C[a, b] = \overline{\operatorname{lin}\{1, t, t^2, ...\}}$ for $\mathbb{K} = \mathbb{R}$, and $C[a, b] = \overline{\operatorname{lin}\{1, i, t, it, t^2, it^2, ...\}}$ for $\mathbb{K} = \mathbb{C}$.

(b) The space $L^p[a, b]$ is separable for all $1 \leq p < \infty$. We will show that $P[a, b]$ is dense in $(L^p[a, b], \|\cdot\|_{L^p})$. To this end we will use that $C[a, b]$ is dense in $(L^p[a, b], \|\cdot\|_{L^p})$ (see [5, Theorem 3.14]). Let $f \in L^p[a, b]$. Then there exists $(f_n)_{n\in\mathbb{N}} \subseteq C[a, b]$ such that $\|f_n - f\|_{L^p} \to 0$ $(n \to +\infty)$. For all $n \in \mathbb{N}$ there exists $g_n \in P[a, b]$ such that $\|f_n - g_n\|_\infty \leq \frac{1}{n}$. Since $\|f_n - g_n\|_{L^p} \leq (b-a)^{\frac{1}{p}} \|f_n - g_n\|_\infty \leq (b-a)^{\frac{1}{p}} \frac{1}{n}$ for all $n \in \mathbb{N}$, it holds $\|f_n - g_n\|_{L^p} \to 0$ $(n \to +\infty)$. The conclusion follows from the fact that

$$\|g_n - f\|_{L^p} \leq \|g_n - f_n\|_{L^p} + \|f_n - f\|_{L^p} \quad \forall n \in \mathbb{N}.$$

(c) The space $L^\infty[a, b]$ is not separable.

Chapter II

Continuous linear operators

3 Properties and examples of continuous linear operators

Definition 3.1 A continuous linear mapping between two normed spaces is called continuous linear operator. A continuous linear mapping from a normed space into its field of scalars is called continuous linear functional.

In the following, for a continuous linear operator $T : X \to Y$, we will write Tx instead of $T(x)$.

Theorem 3.2 *Let X and Y be two normed spaces and $T : X \to Y$ a linear operator. The following statements are equivalent:*

(i) T is continuous.

(ii) T is continuous at 0.

(iii) There exists $M \geq 0$ such that

$$\|Tx\| \leq M\|x\| \quad \forall x \in X.$$

(iv) T is uniformly continuous.

Proof. $(iii) \Rightarrow (iv)$. Let $\varepsilon > 0$ and $\delta > 0$ be such that $M\delta < \varepsilon$. Then for every $x, y \in X$ with $\|x-y\| \leq \delta$ it holds $\|Tx-Ty\| = \|T(x-y)\| \leq M\|x-y\| \leq M\delta < \varepsilon$.

$(iv) \Rightarrow (i) \Rightarrow (ii)$. Everything is clear.

$(ii) \Rightarrow (iii)$. Assuming that (iii) does not hold, for all $n \in \mathbb{N}$ there exists $x_n \in X$ such that $\|Tx_n\| > n\|x_n\|$. It is obvious that $x_n \neq 0$ for all $n \in \mathbb{N}$. Define $y_n := \frac{x_n}{n\|x_n\|}$ and notice that $\|y_n\| = \frac{1}{n}$ for all $n \in \mathbb{N}$. Thus $y_n \to 0$ as $n \to +\infty$, however, $\|Ty_n\| = \frac{\|Tx_n\|}{n\|x_n\|} > 1$ for all $n \in \mathbb{N}$, thus $(Ty_n)_{n\in\mathbb{N}}$ does not converge to zero. Contradiction! ■

Definition 3.3 We denote the smallest constant M in Theorem 3.2 (iii) by $\|T\|$, namely,

$$\|T\| := \inf\{M \geq 0 \mid \|Tx\| \leq M\|x\| \ \forall x \in X\}.$$

It is easy to see that

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\| = \sup_{\|x\| \leq 1} \|Tx\|$$

and

$$\|Tx\| \leq \|T\|\|x\| \quad \forall x \in X.$$

Continuous linear operators map the unit ball $B_X = \{x \in X \mid \|x\| \leq 1\}$ into a bounded set. This is the reason why they are also called bounded linear operators.

It is easy to see that

$$L(X,Y) := \{T : X \to Y \mid T \text{ is linear and continuous}\},$$

endowed with the operations

$$\begin{aligned}(S+T)(x) &:= Sx + Tx \\ (\lambda T)(x) &:= \lambda Tx,\end{aligned}$$

is a vector space. The zero element is the operator $x \mapsto 0$ for all $x \in X$. We also denote $L(X) := L(X,X)$.

Theorem 3.4 *(a) $\|T\| = \sup_{\|x\| \leq 1} \|Tx\|$ defines a norm on $L(X,Y)$, which is the so-called operator norm.*

(b) If Y is complete, then $L(X,Y)$ is also complete.

Proof. (a) It holds $\|T\| \geq 0$ and $\|T\| = 0$ if and only if $T = 0$. In addition,

$$\|\lambda T\| = \sup_{\|x\| \leq 1} \|\lambda Tx\| = |\lambda| \sup_{\|x\| \leq 1} \|Tx\| = |\lambda|\|T\| \quad \forall \lambda \in \mathbb{K} \ \forall T \in L(X,Y).$$

Furthermore, for $S, T \in L(X,Y)$ it holds

$$\begin{aligned}\|S+T\| &= \sup_{\|x\| \leq 1} \|(S+T)x\| = \sup_{\|x\| \leq 1} \|Sx + Tx\| \leq \sup_{\|x\| \leq 1} \|Sx\| + \sup_{\|x\| \leq 1} \|Tx\| \\ &\leq \|S\| + \|T\|.\end{aligned}$$

(b) Let $(T_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $L(X,Y)$ and $x \in X$. Then $(T_n x)_{n\in\mathbb{N}}$ is a Cauchy sequence in Y. Let $Tx := \lim_{n\to+\infty} T_n x$. For all $\lambda, \mu \in \mathbb{K}$ and all $u, v \in X$ it holds

$$T(\lambda u + \mu v) = \lim_{n\to+\infty} T_n(\lambda u + \mu v) = \lim_{n\to+\infty} \lambda T_n u + \mu \lim_{n\to+\infty} T_n v = \lambda Tu + \mu Tv,$$

which shows that T is linear.

Let $\varepsilon > 0$. Then there exists $N \in \mathbb{N}$ such that

$$\|T_n - T_m\| < \frac{\varepsilon}{2} \quad \forall m, n \geq N.$$

Let $x \in X$ with $\|x\| \leq 1$. Choose $M \geq N$ such that

$$\|T_M x - Tx\| < \frac{\varepsilon}{2}.$$

Then, for every $n \geq N$, it holds

$$\|T_n x - Tx\| \leq \|T_n x - T_M x\| + \|T_M x - Tx\| \leq \|T_n - T_M\| + \frac{\varepsilon}{2} < \varepsilon.$$

Taking the supremum over all $\{x \in X \mid \|x\| \leq 1\}$, it yields $\|T_n - T\| < \varepsilon$ for every $n \geq N$, thus T_n converges to T in the operator norm as $n \to +\infty$. In addition, $\|T\| \leq \|T - T_N\| + \|T_N\| < +\infty$, which shows that $T \in L(X, Y)$. ■

Lemma 3.5 *If $S \in L(X, Y)$ and $T \in L(Y, Z)$, then $TS \in L(X, Z)$ and*

$$\|TS\| \leq \|T\| \|S\|.$$

Proof. TS is a linear operator and for every $x \in X$ it holds

$$\|TSx\| \leq \|T\| \|Sx\| \leq \|T\| \|S\| \|x\|,$$

thus TS is continuous and $\|TS\| \leq \|T\| \|S\|$. ■

Please notice that for the operators $S : \mathbb{R}^2 \to \mathbb{R}^2, S(x, y) = (x, 0)$, and $T : \mathbb{R}^2 \to \mathbb{R}^2, T(x, y) = (0, y)$, it holds $0 = \|TS\| < \|T\| \|S\| = 1$.

The following result shows that continuous linear operators defined on a dense linear subspace of a normed space can be extended to the whole space.

Theorem 3.6 *Let X be a normed space, M a dense linear subspace of X, Y a Banach space, and $T \in L(M, Y)$. Then there exists a unique continuous linear extension $\widehat{T} \in L(X, Y)$ of T. In addition, $\|\widehat{T}\| = \|T\|$.*

Proof. Let $x \in X$ and $(x_n)_{n \in \mathbb{N}} \subseteq M$ be such that $x_n \to x$ $(n \to +\infty)$. This means that $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence, which yields that $(Tx_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in Y. Thus $\lim_{n \to +\infty} Tx_n := \widehat{T}x \in Y$ exists. Then $\widehat{T}$ is well-defined, linear and continuous on X. It is easy to see that $\|\widehat{T}\| = \|T\|$.

Assume that $S \in L(X, Y)$ is another continuous linear extension of T and let $x \in X$. Then there exists $(x_n)_{n \in \mathbb{N}} \subseteq M$ such that $x_n \to x$ $(n \to +\infty)$ and we have $Sx = \lim_{n \to +\infty} Sx_n = \lim_{n \to +\infty} Tx_n = \widehat{T}x$. ■

Example 3.7 (a) The identity operator $\mathrm{Id} : X \to X, x \mapsto x$, fulfills $\mathrm{Id} \in L(X)$ and $\|\mathrm{Id}\| = 1$.

(b) Let X be a vector space and $\|\cdot\|$ and $|||\cdot|||$ two norms on X. The two norms are equivalent if and only if both operators

$$\mathrm{Id} : (X, \|\cdot\|) \to (X, |||\cdot|||) \text{ and } \mathrm{Id} : (X, |||\cdot|||) \to (X, ||\cdot||)$$

are continuous.

(c) Let X be a finite-dimensional normed space and Y an arbitrary normed space. Then every linear operator $T : X \to Y$ is continuous. Let $\{e_1, ..., e_n\}$ be a basis of X. According to Theorem 2.5, there exists $M > 0$ such that for every $x \in X$ with $x = \sum_{i=1}^n \alpha_i e_i$ it holds $\sum_{i=1}^n |\alpha_i| \leq M\|x\|$. The conclusion follows from the fact that for every $x = \sum_{i=1}^n \alpha_i e_i \in X$ it holds

$$\|Tx\| = \left\| T\left(\sum_{i=1}^n \alpha_i e_i\right)\right\| \leq \sum_{i=1}^n |\alpha_i|\|Te_i\| \leq \max_{i=1,\dots,m} \|Te_i\| M\|x\|.$$

Example 3.8 (a) Let $T : (C[0,1], \|\cdot\|_\infty) \to \mathbb{K}, Tx = x(0)$. Then T is linear and

$$|Tx| = |x(0)| \leq \|x\|_\infty \quad \forall x \in C[0,1].$$

Thus T is continuous and $\|T\| \leq 1$. For the function $\mathbf{1}(t) = 1$ for every $t \in [0,1]$, it holds $T\mathbf{1} = 1$, thus $\|T\| = 1$.

(b) Let $T : C^1[0,1] \to \mathbb{K}, Tx = x(0) + x'(1)$. T is linear. If we endow $C^1[0,1]$ with the norm $\|x\|_{C^1} = \|x\|_\infty + \|x'\|_\infty$, then we get

$$|Tx| \leq \|x\|_\infty + \|x'\|_\infty = \|x\|_{C^1} \quad \forall x \in C^1[0,1],$$

thus T is continuous and $\|T\| \leq 1$. Since $T\mathbf{1} = 1$, it holds $\|T\| = 1$.

If we endow $C^1[0,1]$ with the norm $|||x||| = \max\{\|x\|_\infty, \|x'\|_\infty\}$, then we get

$$|Tx| \leq \|x\|_\infty + \|x'\|_\infty \leq 2|||x||| \quad \forall x \in C^1[0,1],$$

thus T is continuous and $\|T\| \leq 2$. For $x(t) = \left(t - \frac{1}{2}\right)^2 + \frac{3}{4}$ for every $t \in [0,1]$, it holds $|Tx| = 2$ and $|||x||| = 1$, thus $\|T\| = 2$.

(c) The differential operator $D : C^1[0,1] \to C[0,1]$, $Dx = x'$, is well-defined and linear.

If we endow $C^1[0,1]$ and $C[0,1]$ with the supremum norm, then D is not continuous. Indeed, for $x_n(t) := t^n$, for all $n \geq 1$, we have $\|x_n\|_\infty = 1$ and $\|Dx_n\|_\infty = \sup_{t\in[0,1]} |nt^{n-1}| = n$. Thus, there is no $M \geq 0$ such that $\|Dx_n\|_\infty \leq M\|x_n\|_\infty$ for all $n \geq 1$.

If we endow $C^1[0,1]$ with $\|x\|_{C^1} = \|x\|_\infty + \|x'\|_\infty$ and $C[0,1]$ with the supremum norm, then $\|Dx\|_\infty \leq \|x\|_{C^1}$ for all $x \in C^1[0,1]$, thus D is continuous.

Example 3.9 (a) Let $T : c \to \mathbb{K}, Tx = \lim_{k\to+\infty} t_k$, for $x = (t_k)_{k\in\mathbb{N}}$. Then T is linear and continuous, and $\|T\| = 1$. Notice that $c_0 = T^{-1}(\{0\})$, which is another proof for the fact that c_0 is closed in c.

(b) Let $T : c_{00} \to \mathbb{K}, Tx = \sum_{k=1}^{+\infty} kt_k$, for $x = (t_k)_{k\in\mathbb{N}}$. T is well-defined and linear. If we endow c_{00} with the supremum norm, then T is not continuous. Indeed, for $e_n = (0, ..., 0, 1, 0, ...)$, where the n-th element is 1 and all others are 0, we have $\|e_n\|_\infty = 1$ and $|Te_n| = n$ for all $n \in \mathbb{N}$.

Example 3.10 (integral operators)

(a) Let $g \in C[0, 1]$ and

$$T : C[0, 1] \to \mathbb{K}, \quad Tx = \int_0^1 x(t)g(t)dt.$$

T is linear. For every $x \in C[0, 1]$ it holds

$$|Tx| = \left|\int_0^1 x(t)g(t)dt\right| \le \int_0^1 |x(t)||g(t)|dt \le \left(\int_0^1 |g(t)|dt\right) \|x\|_\infty.$$

This proves that T is continuous and $\|T\| \le \int_0^1 |g(t)|dt$. We will show that actually $\|T\| = \int_0^1 |g(t)|dt$.

Let $\varepsilon > 0$. Define

$$x_\varepsilon(t) = \frac{\overline{g(t)}}{|g(t)| + \varepsilon} \quad \forall t \in [0, 1].$$

Then $x_\varepsilon \in C[0, 1]$, $\|x_\varepsilon\|_\infty \le 1$, and

$$|Tx_\varepsilon| = \int_0^1 \frac{|g(t)|^2}{|g(t)| + \varepsilon} dt \ge \int_0^1 |g(t)|dt - \varepsilon.$$

This gives

$$\|T\| = \sup_{\|x\|_\infty \le 1} |Tx| \ge \sup_{\varepsilon>0} |Tx_\varepsilon| \ge \int_0^1 |g(t)|dt,$$

and proves the statement.

(b) Let $T : C[0, 1] \to \mathbb{K}, Tx = \int_0^1 x(t)dt$. As seen in (a), T is a continuous linear operator with $\|T\| = 1$. Let $X = \{x \in C[0, 1] \mid x(0) = 0\}$ be endowed with the supremum norm and $S := T|_X$. Then

$$\|S\| = \sup_{x\in X, x\neq0} \frac{|Sx|}{\|x\|_\infty} \le \sup_{x\in C[0,1], x\neq0} \frac{|Tx|}{\|x\|_\infty} = \|T\|.$$

On the other hand, for all $n \in \mathbb{N}$ and $x_n(t) = \frac{n+1}{n} t^{\frac{1}{n}}$ we have that $x_n \in X, \|x_n\|_\infty = \frac{n+1}{n}$ and $|Sx_n| = 1$, thus

$$\|S\| \geq \sup_{n\in\mathbb{N}} \frac{|Sx_n|}{\|x_n\|_\infty} = \sup_{n\in\mathbb{N}} \frac{n}{n+1} = 1.$$

This shows that $\|S\| = 1$. One can notice that there exists no $x \in X, \|x\|_\infty \leq 1$, such that $|Sx| = \|S\| = 1$.

(c) Let (Ω, Σ, μ) be a measure space, $1 \leq p \leq \infty$, and $\frac{1}{p} + \frac{1}{q} = 1$ (with $\frac{1}{\infty} = 0$). Then, for every $g \in L^q(\mu)$, the operator

$$Tg : L^p(\mu) \to \mathbb{K}, \quad (Tg)(f) = \int fg d\mu,$$

is linear and continuous with $\|Tg\| \leq \|g\|_{L^q}$. This is a direct consequence of the Hölder inequality in integral form (see Übungsbeispiel 14).

We actually have

$$\|Tg\| = \|g\|_{L^q} \tag{3.1}$$

if

(i) $1 < p \leq \infty$;

(ii) $p = 1$ and μ is semi-finite, namely, for every $M \in \Sigma$ with $\mu(M) = +\infty$ there exists $N \subset M, N \in \Sigma$, with $0 < \mu(N) < +\infty$.

(i) This is obvious if $g = 0$ almost everywhere. Otherwise, let $A := \{x \in \Omega \mid |g(x)| > 0\}$ and choose:

(i1) if $1 < p < \infty$, $f := \frac{\bar{g}}{|g|} \left(\frac{|g|}{\|g\|_{L^q}}\right)^{\frac{q}{p}} \chi_A$. Then $\|f\|_{L^p} = 1$ and $\int fg d\mu = \|g\|_{L^q}$;

(i2) if $p = \infty$, $f := \frac{\bar{g}}{|g|} \chi_A$. Then $\|f\|_{L^\infty} = 1$ and $\int fg d\mu = \|g\|_{L^1}$.

(ii) For the proof of (3.1) in case $p = 1$ (see Übungsbeispiel 21, Gruppe 1).

Example 3.11 (Fredholm integral operator) Let $k : [0,1] \times [0,1] \to \mathbb{K}$ be continuous and

$$T : C[0,1] \to C[0,1], \quad (Tx)(s) := \int_0^1 k(s,t)x(t)dt,$$

the so-called Fredholm integral operator with kernel k.

We show that for all $x \in C[0,1]$ the function $Tx : [0,1] \to \mathbb{K}$ is continuous. Let $\varepsilon > 0$. Since k is uniformly continuous (being continuous on a compact set), there exists $\delta > 0$ such that for all $(s,t), (s',t') \in [0,1] \times [0,1]$ with $\|(s,t) - (s',t')\|_2 < \delta$

it holds $|k(s,t) - k(s',t')| < \varepsilon$. This means that for every $s, s' \in [0,1]$ with $|s - s'| < \delta$ it holds

$$|(Tx)(s) - (Tx)(s')| \leq \int_0^1 |k(s,t) - k(s',t)||x(t)|dt \leq \varepsilon \|x\|_\infty,$$

which proves that Tx is continuous.

T is linear. If $C[0,1]$ is endowed with the supremum norm, then T is continuous. Indeed, we have (according to Example 3.10 (a))

$$\|T\| = \sup_{\|x\|_\infty \leq 1} \|Tx\|_\infty = \sup_{\|x\|_\infty \leq 1} \sup_{s \in [0,1]} |(Tx)(s)| = \sup_{s \in [0,1]} \sup_{\|x\|_\infty \leq 1} \left| \int_0^1 k(s,t)x(t)dt \right|$$
$$= \sup_{s \in [0,1]} \int_0^1 |k(s,t)|dt \leq \|k\|_\infty.$$

In the following we will investigate the invertibility of continuous linear operators in normed spaces. If $T : X \to Y$ is a bijective continuous linear continuous operator between two normed spaces X and Y, then its inverse $T^{-1} : Y \to X$ is linear, but, in general, not continuous. The continuity of the inverse operator plays an important role, for instance, in the characterization of the dependence of a solution of a linear equation from the data. The following result provides a first characterization of the continuity of the inverse operator.

Theorem 3.12 *Let $T : X \to Y$ be a linear operator between two normed spaces X and Y. Then the inverse $T^{-1} : \operatorname{ran} T := T(X) \to X$ exists and is a continuous operator if and only if there exists $m > 0$ such that*

$$\|Tx\| \geq m\|x\| \quad \forall x \in X. \tag{3.2}$$

Proof. "$\Leftarrow$" Notice that $\operatorname{ran} T$ is a linear subspace of Y. If (3.2) holds, then T is injective and its inverse $T^{-1} : \operatorname{ran} T \to X$ exists. For $\lambda, \mu \in \mathbb{K}$ and $Tx, Ty \in \operatorname{ran} T$, it holds

$$T^{-1}(\lambda Tx + \mu Ty) = T^{-1}(T(\lambda x + \mu y)) = \lambda x + \mu y = \lambda T^{-1}(Tx) + \mu T^{-1}(Ty),$$

which shows that the inverse operator is linear. Since, according to (3.2),

$$\|T^{-1}y\| \leq \frac{1}{m}\|y\| \quad \forall y \in \operatorname{ran} T,$$

we obtain that T^{-1} is continuous.

"$\Rightarrow$" Using that $T^{-1} : \operatorname{ran} T \to X$ is continuous, it holds

$$\|T^{-1}y\| \leq \|T^{-1}\|\|y\| \quad \forall y \in \operatorname{ran} T,$$

and, from here,

$$\|x\| = \|T^{-1}(Tx)\| \leq \|T^{-1}\|\|Tx\| \quad \forall x \in X,$$

which shows that (3.2) is fulfilled for $m := \frac{1}{\|T^{-1}\|}$. ■

Example 3.13 The inverse of the identity operator $\mathrm{Id} : (C[0,1], \|\cdot\|_\infty) \to (C[0,1], \|\cdot\|_1)$ is not continuous. Otherwise, there would exists $m > 0$ such that $\|x\|_1 \geq m\|x\|_\infty$ for all $x \in C[0,1]$; in particular, for $x_n(t) = t^n$ it holds $\frac{1}{n+1} = \|x_n\|_1 \geq m\|x_n\|_\infty = m$ for all $n \in \mathbb{N}$. Contradiction!

Definition 3.14 (isomorphism) A continuous linear operator $T : X \to Y$ between two normed spaces X and Y is called isomorphism if T is bijective and T^{-1} is continuous. If for a linear operator $T : X \to Y$ it holds $\|Tx\| = \|x\|$ for all $x \in X$, then T is said to be isometric. Two normed spaces X and Y are said to be (isometrically) isomorphic if the there exists an (isometric) isomorphism between them; we denote this by $(X \cong Y)$ $X \backsimeq Y$.

If $T : X \to Y$ is an isomorphism, then there exist $M \geq m > 0$ such that

$$m\|x\| \leq \|Tx\| \leq M\|x\| \quad \forall x \in X.$$

Example 3.15 ($c \backsimeq c_0$) Let $x = (s_k)_{k\in\mathbb{N}} \in c$ and $l(x) := \lim_{k\to+\infty} s_k \in \mathbb{K}$. We define $y = (t_k)_{k\in\mathbb{N}}$ by $t_1 := l(x)$ and $t_k := s_{k-1} - l(x)$ for all $k \geq 2$. Then $y \in c_0$ and $T : x \mapsto y$ is a linear operator from c to c_0. For all $x \in c$ it holds

$$\|Tx\|_\infty = \sup_{k\in\mathbb{N}} |t_k| \leq \sup_{k\in\mathbb{N}} |s_k| + |l(x)| \leq 2\|x\|_\infty.$$

This proves that T is continuous.

Let $y = (t_k)_{k\in\mathbb{N}} \in c_0$ and define $x = (s_k)_{k\in\mathbb{N}}$ by $s_k := t_{k+1} + t_1$ for all $k \geq 1$. Then $x \in c$ and $S : y \mapsto x$ is a linear operator from c_0 to c. For all $y \in c_0$ it holds

$$\|Sy\|_\infty = \sup_{k\in\mathbb{N}} |s_k| \leq \sup_{k\in\mathbb{N}} |t_{k+1}| + |t_1| \leq 2\|y\|_\infty.$$

This proves that S is continuous.

It is easy to see that $ST = \mathrm{Id}_c$ and $TS = \mathrm{Id}_{c_0}$, thus $S^{-1} = T$ and T is an isomorphism.

Example 3.16 (quotient operator) A linear operator $T : X \to Y$ between two normed spaces X and Y is called quotient operator if T maps the open unit ball in X onto the open unit ball in Y, in other words, if $T(\operatorname{int} B_X) = \operatorname{int} B_Y$. Then T is surjective and continuous with $\|T\| = 1$. It is clear that T is surjective. Let $x \in X$ with $\|x\| \leq 1$. Then $\left\|T\left(\frac{n}{n+1}x\right)\right\| < 1$ or, equivalently, $\|Tx\| < \frac{n+1}{n}$ for all $n \geq 1$, thus $\|Tx\| \leq 1$. This proves that $\|T\| \leq 1$. On the other hand, let $y \in Y$ with $\|y\| = 1$. Then for all $n \in \mathbb{N}$ there exists $x_n \in X, \|x_n\| < 1$, such that $Tx_n = \frac{n}{n+1}y$. For all $n \in \mathbb{N}$ we have $\|T\| \geq \frac{\|Tx_n\|}{\|x_n\|} > \frac{n}{n+1}\|y\| = \frac{n}{n+1}$, thus $\|T\| \geq 1$.

If $M \subseteq X$ is a closed linear subspace of the normed space X, then $T : X \to X/M, x \mapsto [x]$, is a quotient operator. Indeed, if $\|x\| < 1$, then $\|[x]\| = \operatorname{dist}(x, M) \leq \|x\| < 1$. On the other hand, if $[x] \in X/M$ fulfils $\|[x]\| < 1$, then there exists $m \in M$ such that $\|x - m\| < 1$ and it holds $T(x-m) = [x-m] = [x]$.

If $T : X \to Y$ is a quotient operator, then $X/\ker T \cong Y = \operatorname{ran} T$ (first isomorphism theorem), where $\ker T = \{x \in X \mid Tx = 0\}$ (see Übungsbeispiel 23, Gruppe 1).

The following result introduces a method which can be sometimes used to calculate the inverse of an operator. For $T \in L(X)$, we denote $T^0 := \mathrm{Id}$ and $T^n := T \circ ... \circ T$ (n times) for all $n \in \mathbb{N}$.

Theorem 3.17 *(Neumann series) Let X be a normed space and $T \in L(X)$. If $\sum_{n=0}^{+\infty} T^n$ is convergent in $L(X)$, then $\mathrm{Id} - T$ is invertible and*

$$(\mathrm{Id} - T)^{-1} = \sum_{n=0}^{+\infty} T^n.$$

The series $\sum_{n=0}^{+\infty} T^n$ is convergent in $L(X)$ if X is a Banach space and $\|T\| < 1$. In this situation, $\|(\mathrm{Id} - T)^{-1}\| \leq \frac{1}{1-\|T\|}$.

Proof. Let be $m \in \mathbb{N}, m \geq 0$, and $S_m := \sum_{n=0}^{m} T^n$. Then

$$(\mathrm{Id} - T)S_m = S_m(\mathrm{Id} - T) = \mathrm{Id} - T^{m+1} \quad \forall m \geq 0.$$

The Cauchy criterion guarantees that $T^n \to 0$ as $n \to +\infty$. Using that the operators $S \mapsto TS$ and $S \mapsto ST$ are continuous (see Lemma 3.5), we obtain

$$\mathrm{Id} = \lim_{m \to +\infty} (\mathrm{Id} - T^{m+1}) = \lim_{m \to +\infty} (\mathrm{Id} - T)S_m = (\mathrm{Id} - T) \lim_{m \to +\infty} S_m$$

and, similarly,

$$\mathrm{Id} = \lim_{m \to +\infty} S_m(\mathrm{Id} - T),$$

which proves that

$$(\mathrm{Id} - T)^{-1} = \lim_{m \to +\infty} S_m = \sum_{n=0}^{+\infty} T^n.$$

Assume now that $\|T\| < 1$. It holds $\sum_{n=0}^{+\infty} \|T^n\| \leq \sum_{n=0}^{+\infty} \|T\|^n < +\infty$, thus $\sum_{n=0}^{+\infty} T^n$ is absolutely convergent. Since X is complete, $L(X)$ is also complete, thus, according to Lemma 1.12, $\sum_{n=0}^{+\infty} T^n$ is convergent. In addition,

$$\|(\mathrm{Id} - T)^{-1}\| = \left\| \sum_{n=0}^{+\infty} T^n \right\| \leq \sum_{n=0}^{+\infty} \|T\|^n = \frac{1}{1 - \|T\|}.$$

■

4 Dual spaces

Definition 4.1 (dual space) Let X be a normed space over $\mathbb{K}$. The space $L(X, \mathbb{K})$ of continuous linear functionals on X is called dual space of X and is denoted by X^*.

The dual space X^* of a normed space X endowed with the norm

$$\|x^*\| := \sup_{\|x\|\le 1} |x^*(x)|$$

is a Banach space.

Example 4.2 $((\ell^p)^* \cong \ell^q, 1 \le p < \infty; (c_0)^* \cong \ell^1)$

(a) Let $1 \le p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ (with $\frac{1}{\infty} = 0$). The operator

$$T : \ell^q \to (\ell^p)^*, \quad (Tx)(y) = \sum_{n=1}^{+\infty} s_n t_n,$$

where $x = (s_n)_{n\in\mathbb{N}} \in \ell^q$ and $y = (t_n)_{n\in\mathbb{N}} \in \ell^p$, is an isometric isomorphism.

We will prove the statement for $1 < p < \infty$. The proof in case $p = 1$ follows similarly.

Let $x = (s_n)_{n\in\mathbb{N}} \in \ell^q$. According to the Hölder inequality we have

$$|(Tx)(y)| \le \|x\|_q \|y\|_p < +\infty \quad \forall y = (t_n)_{n\in\mathbb{N}} \in \ell^p,$$

which proves that $\sum_{n=1}^{+\infty} s_n t_n$ is convergent. In addition, Tx is linear and continuous, thus, $Tx \in (\ell^p)^*$ and $\|Tx\| \le \|x\|_q$.

It is easy to see that T is linear. Assuming that $Tx = 0$, it follows that $s_n = (Tx)(e_n) = 0$ for all $n \in \mathbb{N}$, thus $x = 0$, which proves that T is injective.

We prove that T is surjective and isometric. To this end, let $y^* \in (\ell^p)^*$ and define

$$s_n := y^*(e_n) \text{ and } x := (s_n)_{n\in\mathbb{N}} \quad \forall n \in \mathbb{N}.$$

We will prove that

$$x \in \ell^q, \quad Tx = y^*, \quad \|x\|_q \le \|y^*\| = \|Tx\|.$$

For all $n \in \mathbb{N}$ we define

$$t_n := \begin{cases} \frac{|s_n|^q}{s_n} & \text{for } s_n \ne 0, \\ 0 & \text{for } s_n = 0. \end{cases}$$

For all $N \in \mathbb{N}$ it holds

$$\sum_{n=1}^{N} |t_k|^p = \sum_{n=1}^{N} |s_n|^{p(q-1)} = \sum_{n=1}^{N} |s_n|^q$$

and

$$\sum_{n=1}^{N}|s_n|^q = \sum_{n=1}^{N} s_n t_n = \sum_{n=1}^{N} t_n y^*(e_n) = y^*\left(\sum_{n=1}^{N} t_n e_n\right)$$
$$\leq \|y^*\|\left(\sum_{n=1}^{N}|t_n|^p\right)^{\frac{1}{p}} = \|y^*\|\left(\sum_{n=1}^{N}|s_n|^q\right)^{\frac{1}{p}},$$

thus $\left(\sum_{n=1}^{N}|s_n|^q\right)^{\frac{1}{q}} \leq \|y^*\|$. Letting $N \to +\infty$ we get $x \in \ell^q$ and $\|x\|_q \leq \|y^*\|$.

By construction, $(Tx)(e_n) = y^*(e_n)$ for all $n \in \mathbb{N}$, thus $(Tx)(y) = y^*(y)$ for all $y \in c_{00} = \operatorname{lin}\{e_n : n \in \mathbb{N}\}$. The continuity of Tx and y^* gives $(Tx)(y) = y^*(y)$ for all $y \in \overline{c_{00}} = \overline{\operatorname{lin}\{e_n : n \in \mathbb{N}\}} = \ell^p$ (see Example 2.11), thus $Tx = y^*$.

(b) In a similar way, one can prove that the operator

$$T : \ell^1 \to (c_0)^*, \quad (Tx)(y) = \sum_{n=1}^{+\infty} s_n t_n,$$

where $x = (s_n)_{n\in\mathbb{N}} \in \ell^1$ and $y = (t_n)_{n\in\mathbb{N}} \in c_0$, is an isometric isomorphism.

(c) We will see in Example 6.13 that $\ell^1 \subsetneq (\ell^\infty)^*$.

Example 4.3 Let $1 \leq p \leq \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ (with $\frac{1}{\infty} = 0$). By using the operator

$$(Tx)(y) = \sum_{i=1}^{n} x_i y_i,$$

for $x = (x_1, ..., x_n) \in \mathbb{K}^n$ and $y = (y_1, ..., y_n) \in \mathbb{K}^n$, it is easy to see that

$$(\mathbb{K}^n, \|\cdot\|_p)^* \cong (\mathbb{K}^n, \|\cdot\|_q).$$

Example 4.4 (a) Let (Ω, Σ, μ) be a finite measure space, $1 \leq p < \infty$, and $\frac{1}{p} + \frac{1}{q} = 1$ (with $\frac{1}{\infty} = 0$). The operator

$$T : L^q(\mu) \to (L^p(\mu))^*, \quad (Tg)(f) = \int_\Omega f g d\mu,$$

is an isometric isomorphism, thus $(L^p(\mu))^* \cong L^q(\mu)$. The spaces $(L^\infty(\mu))^*$ and $L^1(\mu)$ are in general not isomorph.

(b) Let K be a compact topological Hausdorff (metric) space, Σ a σ-algebra on K and $M(K, \Sigma)$ the Banach space of measures on Σ endowed with the total variation norm (see Übungsbeispiel 12, Gruppe 1). The operator

$$T : M(K) \to (C(K))^*, \quad (T\mu)(x) = \int_K x d\mu,$$

is an isometric isomorphism, thus $(C(K))^* \cong M(K)$.

5 Compact operators

Compact operators are meant to compensate in infinite-dimensional normed spaces the missing property that "every bounded sequence has a convergent subsequence".

Definition 5.1 (compact operator) A linear operator T between two normed spaces X and Y is called compact if $T(B_X)$ is relatively compact, which means that $\overline{T(B_X)}$ is compact. Let $K(X,Y) := \{T : X \to Y \mid T \text{ is compact}\}$ and $K(X) := K(X,X)$.

Given a linear operator $T : X \to Y$, the following statements are equivalent:

(i) T is compact;

(ii) For every bounded set $B \subseteq X$ the set $T(B)$ is relatively compact in Y;

(iii) For every bounded sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ the sequence $(Tx_n)_{n\in\mathbb{N}} \subseteq Y$ has a convergent subsequence.

It is easy to see that compact operators are continuous, thus $K(X,Y) \subseteq L(X,Y)$. In the following we will treat compact operators between a normed space and a Banach space, in order to guarantee that the closure of $T(B_X)$ lies in the right space.

Theorem 5.2 *Let X be a normed space and Y a Banach space.*

(a) $K(X,Y)$ is a closed linear subspace of $L(X,Y)$, which means that $K(X,Y)$ is a Banach space.

(b) If Z is a further Banach space, $T \in L(X,Y)$ and $S \in L(Y,Z)$ such that either T or S is compact, then ST is compact.

Proof. (a) If $T \in K(X,Y)$ and $\lambda \in \mathbb{K}$, then, obviously, $\lambda T \in K(X,Y)$. Take now $S,T \in K(X,Y)$ and $(x_n)_{n\in\mathbb{N}} \subseteq X$ a bounded sequence. Then there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that $(Sx_{n_k})_{k\in\mathbb{N}}$ and $(Tx_{n_k})_{k\in\mathbb{N}}$ are convergent, thus $((S+T)x_{n_k})_{k\in\mathbb{N}}$ is convergent. This shows that $K(X,Y)$ is a linear subspace.

We will prove the closedness of this space by using a diagonalization argument. Let $(T_n)_{n\in\mathbb{N}} \subseteq K(X,Y)$ and $T \in L(X,Y)$ with $\|T_n - T\| \to 0$ as $n \to +\infty$. Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ a bounded sequence. Since T_1 is compact, there exists a convergent subsequence $(T_1x_{k_n})_{n\in\mathbb{N}}$. Denote $x_n^1 := x_{k_n}$ for all $n \in \mathbb{N}$. Since T_2 is compact, there exists a convergent subsequence $(T_2x^1_{k_n})_{n\in\mathbb{N}}$. Denote $x_n^2 := x^1_{k_n}$ for all $n \in \mathbb{N}$. Since T_3 is compact, there exists a convergent subsequence $(T_3x^2_{k_n})_{n\in\mathbb{N}}$. Denote $x_n^3 := x^2_{k_n}$ for all $n \in \mathbb{N}$. We continue this process and define in this way a sequence $\xi_1 := x_{k_1}, \xi_2 := x^1_{k_2}, \xi_3 := x^2_{k_3}, ..., \xi_{n+1} := x^n_{k_{n+1}}, ...$ Then, for every $m \in \mathbb{N}$, the sequence $(T_m \xi_n)_{n \in \mathbb{N}}$ is convergent.

We will prove that $(T\xi_n)_{n\in\mathbb{N}}$ is also convergent by proving that it is a Cauchy sequence. Let $L > 0$ be such that $\|x_n\| \le L$ for all $n \in \mathbb{N}$. This means that $\|\xi_n\| \le L$ for all $n \in \mathbb{N}$. Let $\varepsilon > 0$, $M \in \mathbb{N}$ such that $\|T_M - T\| < \frac{\varepsilon}{3L}$ and $N \in \mathbb{N}$ such that

$$\|T_M \xi_m - T_M \xi_n\| < \frac{\varepsilon}{3} \quad \forall m, n \ge N.$$

Then for all $m, n \ge N$ it holds

$$\begin{aligned}\|T\xi_m - T\xi_n\| &\le \|T\xi_m - T_M\xi_m\| + \|T_M\xi_m - T_M\xi_n\| + \|T_M\xi_n - T\xi_n\| \\ &\le 2L\|T - T_M\| + \frac{\varepsilon}{3} < \varepsilon,\end{aligned}$$

which leads to the desired conclusion.

(b) Let $(x_n)_{n\in\mathbb{N}} \subseteq X$ be a bounded sequence and assume that S is compact. Then $(Tx_n)_{n\in\mathbb{N}} \subseteq Y$ is bounded, too, and $(STx_n)_{n\in\mathbb{N}} \subseteq Z$ has a convergent subsequence.

Assume that T is compact. Then there exists a convergent subsequence $(Tx_{n_k})_{k\in\mathbb{N}} \subseteq Y$. Consequently, $(STx_{n_k})_{k\in\mathbb{N}} \subseteq Z$ is convergent.

■

Example 5.3 (a) If X is a finite-dimensional normed space and $T : X \to Y$ is linear, then T is compact. According to Example 3.7 (a), T is continuous and, therefore, since B_X is compact in X, $T(B_X)$ is compact in Y.

(b) If X is a normed space, Y a Banach space, $T \in L(X,Y)$ and $\operatorname{ran} T$ is finite-dimensional, then T is compact, since $T(B_X) \subseteq \operatorname{ran} T$ is bounded and, therefore, relatively compact.

(c) For X a normed space and Y a Banach space, let

$$F(X,Y) := \{T \in L(X,Y) \mid \operatorname{ran} T \text{ is finite-dimensional}\}$$

be the so-called space of finite-rank operators. We denote $F(X) := F(X,X)$. By using (b) and and Theorem 5.2(a), it follows that $\overline{F(X,Y)} \subseteq K(X,Y)$.

(d) If X is a normed space and Y a separable Banach space such that there exists a bounded sequence $(S_n)_{n\in\mathbb{N}}$ in $F(Y)$ with the property

$$\lim_{n\to+\infty} S_n y = y \quad \forall y \in Y, \tag{5.1}$$

then it holds $\overline{F(X,Y)} = K(X,Y)$ (see Übungsbeispiel 26(a)).

Relation (5.1) is fulfilled when $Y = c_0$ or $Y = \ell^p, 1 \leq p < \infty$ by

$$S_n\left((t_k)_{k\in\mathbb{N}}\right) = (t_1, ..., t_n, 0, ...) \quad \forall n \in \mathbb{N}.$$

It is also fulfilled in case $Y = C[0,1]$ and $Y = L^p[0,1], 1 \leq p < \infty$.

The following theorem provides an useful compactness criterion.

Theorem 5.4 (Arzelá-Ascoli Theorem) *Let (T,d) be a compact metric space and $M \subseteq (C(T), \|\cdot\|_\infty))$ with the properties:*

(i) M is bounded;

(ii) M is closed;

(iii) M is equicontinuous, namely,

$$\forall \varepsilon > 0\ \exists \delta > 0 \text{ such that } \forall x \in M\ d(s,t) < \delta \Rightarrow |x(s) - x(t)| < \varepsilon.$$

Then M is compact.

Proof. We will prove that M is sequentially compact, in other words, that every sequence in M has a convergent subsequence in M. The compactness of M will follow from the Bolzano-Weierstraß Theorem (see Übungsbeispiel 26, Gruppe 1).

First we will show that T is separable. Let $n \in \mathbb{N}$. Since $T = \bigcup_{t\in T}\{s \in T \mid d(t,s) < \frac{1}{n}\}$, there exists $t_1^n, ..., t_{m_n}^n \in T$ such that $T = \bigcup_{k=1}^{m_n}\{s \in T \mid d(t_k^n, s) < \frac{1}{n}\}$. This shows that the countable set $\{t_k^n : 1 \leq k \leq m_n, n \in \mathbb{N}\}$ is dense in T.

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in M and $\{t_1, t_2, ...\}$ a countable dense set in T. Since M is bounded, the sequence $(x_n(t_1))_{n\in\mathbb{N}}$ is bounded in $\mathbb{K}$, thus, it has a convergent subsequence $(x_{k_n}(t_1))_{n\in\mathbb{N}}$. Denote $x_n^1 := x_{k_n}$ for all $n \in \mathbb{N}$. The sequence $(x_n^1(t_2))_{n\in\mathbb{N}}$ is bounded in $\mathbb{K}$, thus, it has a convergent subsequence $(x_{k_n}^1(t_2))_{n\in\mathbb{N}}$. Denote $x_n^2 := x_{k_n}^1$ for all $n \in \mathbb{N}$. The sequence $(x_n^2(t_3))_{n\in\mathbb{N}}$ is bounded in $\mathbb{K}$, thus, it has a convergent subsequence $(x_{k_n}^2(t_3))_{n\in\mathbb{N}}$. Denote $x_n^3 := x_{k_n}^2$ for all $n \in \mathbb{N}$. We continue this process and define in this way the diagonal sequence $\xi_1 := x_{k_1}, \xi_2 := x_{k_2}^1, \xi_3 := x_{k_3}^2, ..., \xi_{n+1} := x_{k_{n+1}}^n, ...$. Then, for every $m \in \mathbb{N}$, the sequence $(\xi_n(t_m))_{n\in\mathbb{N}}$ is convergent.

We will prove that $(\xi_n)_{n\in\mathbb{N}}$ is convergent by proving that it is a Cauchy sequence with respect to the supremum norm. Let $\varepsilon > 0$ and $\delta > 0$ such that the property in (iii) holds. Let $B_1, ..., B_p$ be open balls of radius $\frac{\delta}{2}$ such that $T = \cup_{l=1}^p B_p$. For all $l = 1, ..., p$ in every B_l there exists an element t^l of the countable dense set in T. Let $N \in \mathbb{N}$ be such that

$$|\xi_n(t^l) - \xi_m(t^l)| < \frac{\varepsilon}{3} \quad \forall m, n \geq N\ \forall l = 1, ..., p.$$

We choose an arbitrary $t \in T$ and $l \in \{1, ..., p\}$ such that $t \in B_l$. It holds $d(t, t^l) < \delta$ and the equicontinuity of M gives

$$|\xi_n(t) - \xi_n(t^l)| < \frac{\varepsilon}{3} \quad \forall n \in \mathbb{N}.$$

Combining the last two inequalities it yields for all $m, n \geq N$

$$|\xi_m(t) - \xi_n(t)| \leq |\xi_m(t) - \xi_m(t^l)| + |\xi_m(t^l) - \xi_n(t^l)| + |\xi_n(t^l) - \xi_n(t)| < \varepsilon.$$

Since $t \in T$ was arbitrarily chosen, it yields $\|\xi_m - \xi_n\|_\infty \leq \varepsilon$ for all $m, n \geq N$, which provides the desired conclusion. ■

Example 5.5 The Fredholm integral operator

$$T : C[0,1] \to C[0,1], \quad (Tx)(s) = \int_0^1 k(s,t)x(t)dt,$$

with kernel $k \in C([0,1]^2)$ is compact. Let $M := \overline{T(B_X)}$, which is a closed set. Since T is continuous (see Example 3.11), M is bounded. We proved in Example 3.11 that $T(B_X)$ is equicontinuous, thus M is equicontinuous, too. The Arzelá-Ascoli Theorem guarantees that M is compact, which means that T is a compact operator.

Chapter III

The Hahn-Banach Theorem

6 Extension of functionals

In this chapter we will provide results which guarantee the existence of nonzero continuous linear functionals with certain properties.

Definition 6.1 Let X be a vector space. A mapping $p : X \to \mathbb{R}$ is called sublinear if

(a) $p(\lambda x) = \lambda p(x) \quad \forall \lambda \geq 0 \; \forall x \in X$;

(b) $p(x + y) \leq p(x) + p(y) \quad \forall x, y \in X$.

Example 6.2 Every seminorm is sublinear. Every linear functional defined on a real vector space is sublinear.

Theorem 6.3 (Hahn-Banach Theorem for Real Vector Spaces) *Let X be a real vector space and M a linear subspace of X. Further, let $p : X \to \mathbb{R}$ be sublinear and $l : M \to \mathbb{R}$ be linear such that*

$$l(x) \leq p(x) \quad \forall x \in M.$$

Then there exists a linear extension $L : X \to \mathbb{R}$ such that

$$L|_M = l \quad \textit{and} \quad L(x) \leq p(x) \quad \forall x \in X.$$

Proof. Step 1. We will prove the statement in case $\dim X/M = \dim\{[x] = x + M \mid x \in X\} = 1$. This means that there exists $x_0 \in X \setminus M$ such that every $x \in X$ can be represented as

$$x = m + \lambda x_0 \quad \text{for } m \in M, \lambda \in \mathbb{R}.$$

For $r \in \mathbb{R}$ fixed, we introduce the linear functional on X

$$L_r(x) = L_r(m + \lambda x_0) := l(m) + \lambda r.$$

It is clear that $L_r(m) = l(m)$ for all $m \in M$, thus L_r is a linear extension of l. It remains to show that there exists $r \in \mathbb{R}$ such that $L_r \leq p$.

The inequality $L_r \leq p$ holds if and only if

$$l(m) + \lambda r \leq p(m + \lambda x_0) \quad \forall m \in M \ \forall \lambda \in \mathbb{R}. \tag{6.1}$$

Relation (6.1) is obviously fulfilled for $\lambda = 0$. In case $\lambda > 0$, (6.1) is satisfied if and only if

$$r \leq p\left(\frac{m}{\lambda} + x_0\right) - l\left(\frac{m}{\lambda}\right) \ \forall m \in M \Leftrightarrow r \leq \inf_{s \in M} \left(p(s + x_0) - l(s)\right).$$

In case $\lambda < 0$, (6.1) is satisfied if and only if

$$-r \leq p\left(\frac{m}{-\lambda} - x_0\right) - l\left(\frac{m}{-\lambda}\right) \ \forall m \in M \Leftrightarrow r \geq \sup_{t \in M} \left(l(t) - p(t - x_0)\right).$$

This shows that there exists a real number r such that $L_r \leq p$ if and only if

$$l(t) - p(t - x_0) \leq p(s + x_0) - l(s) \quad \forall s, t \in M.$$

Since

$$l(s) + l(t) = l(s + t) \leq p(s + t) \leq p(s + x_0) + p(t - x_0) \quad \forall s, t \in M,$$

such a real number r always exists.

One can notice that r is not uniquely determined, which means that the extension L_r is also not uniquely determined.

Step 2. We prove the statement of the theorem in the general case and use to this end the Zorn Lemma.

Zorn Lemma. If $(A, \leq)$ is a nonempty partially ordered set in which every chain (this is a totally ordered set, namely, a set for the elements of which either $a \leq b$ or $b \leq a$ holds) has an upper bound, then A has a maximal element (this means that there exists $\max \in A$ such that if $\max \leq a$ for $a \in A$, then $\max = a$).

Let

$$A := \left\{ (N, L_N) \mid \begin{array}{l} N \text{ is a linear subspace of } X \text{ with } M \subseteq N, \\ L_N : N \to \mathbb{R} \text{ is linear with } L_N \leq p|_N \text{ and } L_N|_M = l \end{array} \right\}$$

and the partial order on A

$$(N_1, L_{N_1}) \leq (N_2, L_{N_2}) \Leftrightarrow N_1 \subseteq N_2, L_{N_2}|_{N_1} = L_{N_1}.$$

The set A is nonempty since $(M, l) \in A$. Let $((N_i, L_{N_i})_{i \in I})$ be a totally ordered subset of A. Then (N, L_N), where

$$N := \cup_{i \in I} N_i \text{ and } L_N(x) := L_{N_i}(x) \ \forall x \in N_i,$$

is an upper bound of this chain. Notice that, since $((N_i, L_{N_i})_{i\in I})$ is totally ordered, N is a linear subspace of X and L_N is well-defined and linear.

Let (X_0, L_{X_0}) be a maximal element of A. Assuming that $X_0 \neq X$, according to **Step 1**, there would exist a majorant of (X_0, L_{X_0}) in A, which would contradict its maximality. Thus, $X_0 = X$ and $L := L_{X_0}$ solves the extension problem. ∎

The following lemma will allow us to extend Theorem 6.3 to complex vector spaces.

Lemma 6.4 *Let X be a complex vector space.*

(a) If $l : X \to \mathbb{R}$ is a $\mathbb{R}$-linear functional, namely,

$$l(\lambda x + \mu y) = \lambda l(x) + \mu l(y) \quad \forall \lambda, \mu \in \mathbb{R}\ \forall x, y \in X,$$

then

$$\tilde{l} : X \to \mathbb{C}, \quad \tilde{l}(x) := l(x) - il(ix),$$

is a $\mathbb{C}$-linear functional and $\operatorname{Re}\tilde{l} = l$.

(b) If $h : X \to \mathbb{C}$ is a $\mathbb{C}$-linear functional, then $l := \operatorname{Re} h$ is $\mathbb{R}$-linear and, for $\tilde{l}$ defined as above, it holds $\tilde{l} = h$.

(c) If $p : X \to \mathbb{R}$ is a seminorm and $l : X \to \mathbb{C}$ is a $\mathbb{C}$-linear functional, then

$$|l(x)| \le p(x) \quad \forall x \in X \Leftrightarrow |\operatorname{Re} l(x)| \le p(x) \quad \forall x \in X.$$

(d) If X is a normed space and $l : X \to \mathbb{C}$ is $\mathbb{C}$-linear and continuous, then $\|l\| = \|\operatorname{Re} l\|$.

Proof. (a) It is clear that $\tilde{l}$ is $\mathbb{R}$-linear and $\operatorname{Re}\tilde{l} = l$. Since $\tilde{l}(ix) = i\tilde{l}(x)$, we obtain that $\tilde{l}$ is $\mathbb{C}$-linear.

(b) It is clear that $l = \operatorname{Re} h$ is $\mathbb{R}$-linear. In addition, for all $x \in X$ we have

$$\begin{aligned} h(x) =& \operatorname{Re} h(x) + i \operatorname{Im} h(x) = \operatorname{Re} h(x) - i \operatorname{Re} ih(x) = \operatorname{Re} h(x) - i \operatorname{Re} h(ix) \\ =& l(x) - il(ix) = \tilde{l}(x). \end{aligned}$$

(c) The implication "⇒" is obvious. Let now $x \in X$ and $\lambda \in \mathbb{C}, |\lambda| = 1$, such that $l(x) = \lambda |l(x)|$. We have

$$|l(x)| = \lambda^{-1} l(x) = l(\lambda^{-1}x) = |\operatorname{Re} l(\lambda^{-1}x)| \le p(\lambda^{-1}x) = p(x).$$

(d) Is a direct consequence of (c).

∎

Theorem 6.5 (Hahn-Banach Theorem for Complex Vector Spaces) *Let X be a complex vector space and M a linear subspace of X. Further, let $p : X \to \mathbb{R}$ be sublinear and $l : M \to \mathbb{C}$ be linear such that*

$$\operatorname{Re} l(x) \leq p(x) \quad \forall x \in M.$$

Then there exists a linear extension $L : X \to \mathbb{C}$ such that

$$L|_M = l \quad \text{and} \quad \operatorname{Re} L(x) \leq p(x) \quad \forall x \in X.$$

Proof. We apply Theorem 6.3 for $\operatorname{Re} l$ and get a $\mathbb{R}$-linear functional $F : X \to \mathbb{R}$ with $F|_M = \operatorname{Re} l$ and $F(x) \leq p(x)$ for all $x \in X$. According to Lemma 6.4, there exists a $\mathbb{C}$-linear function $L : X \to \mathbb{C}$ with $F = \operatorname{Re} L$. One can easily see that $L|_M = l$. ∎

In the following we apply the algebraic versions of the Hahn-Banach Theorem to normed spaces.

Theorem 6.6 (Hahn-Banach Theorem for Normed Spaces) *Let X be a normed space over $\mathbb{K}$ and M a linear subspace of X. For every continuous linear functional $m^* : M \to \mathbb{K}$ there exists a continuous linear functional $x^* : X \to \mathbb{K}$ such that*

$$x^*|_M = m^* \quad \text{and} \quad \|x^*\| = \|m^*\|.$$

Proof. Assume first that X is a real normed space and define $p(x) = \|m^*\|\|x\|$ for all $x \in X$. Theorem 6.3 provides a linear functional $x^* : X \to \mathbb{R}$ fulfilling $x^*|_M = m^*$ and $x^*(x) \leq p(x)$ for all $x \in X$. Since $-x^*(x) = x^*(-x) \leq p(-x) = p(x)$ for all $x \in X$, it yields $|x^*(x)| \leq \|m^*\|\|x\|$ for all $x \in X$, thus x^* is continuous and $\|x^*\| \leq \|m^*\|$. In addition,

$$\|m^*\| = \sup_{\|m\| \leq 1, m \in M} |m^*(m)| = \sup_{\|m\| \leq 1, m \in M} |x^*(m)| \leq \sup_{\|x\| \leq 1, x \in X} |x^*(x)| = \|x^*\|.$$

For X a complex normed space we can use Theorem 6.5 and the same arguments to obtain a linear functional $x^* : X \to \mathbb{C}$ fulfilling $x^*|_M = m^*$ and $\|\operatorname{Re} x^*\| = \|m^*\|$. To conclude we use Lemma 6.4 which gives $\|\operatorname{Re} x^*\| = \|x^*\|$. ∎

Corollary 6.7 *In every normed space X there exists for every $x \in X, x \neq 0$, a continuous linear functional $x^* \in X^*$ such that*

$$\|x^*\| = 1 \quad \text{and} \quad x^*(x) = \|x\|.$$

In particular, the dual space X^* separates points in X, which means that for all $x, y \in X, x \neq y$, there exists $x^* \in X^*$ such that $x^*(x) \neq x^*(y)$.

Proof. The first statement follows from Theorem 6.6 for $M := \mathrm{lin}\{x\}$ and $m^* : M \to \mathbb{K}$, $m^*(\lambda x) = \lambda\|x\|$. The second statement follows by applying the first statement to $x - y \neq 0$. ■

Corollary 6.7 leads to the following result, which is symmetric to the definition of the norm of a continuous linear functional.

Corollary 6.8 *In every normed space X it holds*

$$\|x\| = \max_{x^* \in B_{X^*}} |x^*(x)| \quad \forall x \in X,$$

*where "*max*" indicates that the supremum is attained.*

Corollary 6.9 *Let X be a normed space, $M \subseteq X$ a closed linear subspace of X and $x \in X \setminus M$. Then there exists $x^* \in X^*$ such that*

$$x^*|_M = 0 \quad \textit{and} \quad x^*(x) \neq 0.$$

Proof. Consider the quotient operator $[\cdot] : X \to X/M$. Then $[m] = 0$ for all $m \in M$ and $[x] \neq 0$. According to Corollary 6.7, there exists $l \in (X/M)^*$ such that $l([x]) \neq 0$. To conclude, we take $x^* := l \circ [\cdot] \in X^*$. ■

The following result is an immediate consequence of Corollary 6.9.

Corollary 6.10 *Let X be a normed space and M a linear subspace of X. The following statements are equivalent:*

(i) M is dense in X.

(ii) If $x^* \in X^*$ fulfills $x^*|_M = 0$, then $x^* = 0$.

We will close this section by providing some direct consequences of the Hahn-Banach Theorem.

Let X be a normed space, $M \subseteq X$ and $N \subseteq X^*$. The closed linear subspace

$$M^\perp := \{x^* \in X^* \mid x^*(x) = 0 \ \forall x \in M\}$$

is called anihilator of M in X^*. The closed linear subspace

$$N_\perp := \{x \in X \mid x^*(x) = 0 \ \forall x^* \in N\}$$

is called anihilator of N in X.

Theorem 6.11 *Let X be a normed space and M a closed linear subspace of X. Then*

$$(X/M)^* \cong M^\perp \tag{6.2}$$

and

$$M^* \cong X^*/M^\perp. \tag{6.3}$$

Proof. The isometric isomorphism proving (6.2) is

$$l \in (X/M)^* \mapsto x^* = l \circ [\cdot] \in M^\perp.$$

The isometric isomorphism proving (6.3) is

$$x^* + M^\perp \in X^*/M^\perp \mapsto x^*|_M \in M^*.$$

■

Theorem 6.12 *Every normed space with a separable dual space is separable.*

Proof. Since X^* is separable we have that $S_{X^*} = \{x^* \in X^* \mid \|x^*\| = 1\}$ is also separable (see Übungsbeispiel 17(a)). Let $\{x_1^*, x_2^*, ...\}$ be a countable dense subset of S_{X^*}. For every $k \in \mathbb{N}$ we can choose (see Übungsbeispiel 18) $x_k \in S_X = \{x \in X \mid \|x\| = 1\}$ such that $|x_k^*(x_k)| \geq \frac{1}{2}$. Define $M := \operatorname{lin}\{x_1, x_2, ...\}$. We will prove that M is dense in X, which, by taking into account Lemma 2.10, will allow us to conclude that X is separable.

Let $x^* \in X^*$ with $x^*|_M = 0$. We will prove that $x^* = 0$ and the conclusion will follow via Corollary 6.10. Assume without loss of generality that $\|x^*\| = 1$. Then there exists $x_K^* \in S_{X^*}$ such that $\|x^* - x_K^*\| \leq \frac{1}{4}$. This yields

$$\frac{1}{2} \leq |x_K^*(x_K)| = |x_K^*(x_K) - x^*(x_K)| \leq \|x_K^* - x^*\| \|x_K\| \leq \frac{1}{4}.$$

Contradiction! ■

Example 6.13 ($\ell^1 \subsetneq (\ell^\infty)^*$) The operator

$$T : \ell^1 \to (\ell^\infty)^*, \quad (Tx)(y) = \sum_{n=1}^{+\infty} s_n t_n,$$

where $x = (s_n)_{n\in\mathbb{N}} \in \ell^1$ and $y = (t_n)_{n\in\mathbb{N}} \in \ell^\infty$, is isometric, but not surjective. It is easy to see that T is isometric.

We will show that T is not surjective. Consider the continuous linear functional $\lim : c \to \mathbb{K}, \lim z := \lim_{k\to+\infty} r_n$, for $z = (r_n)_{n\in\mathbb{N}} \in c$ (see Example 3.9). In the light of Theorem 6.6 we can extend it to a continuous linear functional $x^* : \ell^\infty \to \mathbb{K}$ with $\|x^*\| = \|\lim\| = 1$. Assume that that there exists $x = (s_n)_{n\in\mathbb{N}} \in \ell^1$ such that $x^*(y) = (Tx)(y) = \sum_{n=1}^{+\infty} s_n t_n$ for all $y = (t_n)_{n\in\mathbb{N}} \in \ell^\infty$. For all $n \in \mathbb{N}$ it holds

$$s_n = x^*(e_n) = \lim e_n = 0,$$

thus $x^* = 0$. Contradiction!

One should notice that there exists no isomorphism between ℓ^1 and $(\ell^\infty)^*$, as ℓ^1 is separable (see Example 2.11 (a)) and $(\ell^\infty)^*$ is not separable (otherwise, according to Theorem 6.12, ℓ^∞ would be also separable, which, as seen in Example 2.11 (c), is not the case).

Corollary 6.14 *Let X be a normed space, I an index set, and $x_i \in X$ and $c_i \in \mathbb{K}$ for all $i \in I$. The following statements are equivalent:*

(i) *There exists $x^* \in X^*$ such that $x^*(x_i) = c_i$ for all $i \in I$.*

(ii) *There exists $M \geq 0$ such that for every finite set $F \subseteq I$ it holds*

$$\left| \sum_{i \in F} \lambda_i c_i \right| \leq M \left\| \sum_{i \in F} \lambda_i x_i \right\|.$$

Proof. $(i) \Rightarrow (ii)$. Take $M := \|x^*\|$.

$(ii) \Rightarrow (i)$. Follows from Theorem 6.6 for $M := \operatorname{lin}\{x_i \mid i \in I\}$ and $m^* : M \to \mathbb{K}, m^* \left(\sum_{i \in F} \lambda_i x_i\right) = \sum_{i \in F} \lambda_i c_i$ for every finite set $F \subseteq I$, which is a continuous linear functional on M. ■

7 Separation theorems for convex sets

In this section we will deal with geometric versions of the Hahn-Banach Theorem. Given two disjoint convex sets $U, V \subseteq X$, the aim is to separate them through continuous linear functionals, more precisely, to formulate conditions which guarantee the existence of $x^* \in X^*, x^* \neq 0$, such that

$$\sup_{x \in U} x^*(x) \leq \inf_{x \in V} x^*(x) \quad (\text{in case } \mathbb{K} = \mathbb{R}),$$

respectively,

$$\sup_{x \in U} \operatorname{Re} x^*(x) \leq \inf_{x \in V} \operatorname{Re} x^*(x) \quad (\text{in case } \mathbb{K} = \mathbb{C}).$$

Definition 7.1 Let X be a vector space and $A \subseteq X$.

(i) The set A is called convex if $\lambda x + (1 - \lambda)y \in A$ for all $x, y \in A$ and all $0 \leq \lambda \leq 1$.

(ii) The set $\operatorname{core} A := \{a \in A \mid \forall x \in X\ \exists \delta > 0 \text{ such that } a + \lambda x \in A\ \forall \lambda \in [0, \delta]\}$ is called the algebraic interior of the set A. If A is convex, then $\operatorname{core} A = \{a \in A \mid \cup_{\lambda > 0} \lambda(A - a) = X\}$.

(iii) The set A is called absorbing if $0 \in \operatorname{core} A$.

Definition 7.2 (Minkowski functional) Let X be a vector space and $A \subseteq X$ an absorbing set. The Minkowski functional of A is defined by

$$p_A : X \to \mathbb{R}, \quad p_A(x) := \inf\{\lambda \geq 0 \mid x \in \lambda A\}.$$

In a normed space X we have $p_{B_X} = p_{\operatorname{int} B_X} = \| \cdot \|$. The following proposition collects some properties of the Minkowski functional (see Übungsbeispiel 33).

Proposition 7.3 *Let X be a vector space and $A \subseteq X$ a convex and absorbing set.*

(a) Then p_A is finite, sublinear and $\operatorname{core} A = \{x \in X \mid p_A(x) < 1\}$.

(b) If, in addition, X is a normed space and A is a neighbourhood of 0*, then p_A is continuous,*

$$\operatorname{int} A = \operatorname{core} A = \{x \in X \mid p_A(x) < 1\} \text{ and } \overline{A} = \{x \in X \mid p_A(x) \leq 1\}.$$

Lemma 7.4 *If X is a normed space and $U \subseteq X$ a convex and open set with $0 \notin U$, then there exists $x^* \in X^*$ such that*

$$\operatorname{Re} x^*(x) < 0 \quad \forall x \in U.$$

Proof. If $U = \emptyset$, then the statement follows from Corollary 6.7.

We consider first the case $\mathbb{K} = \mathbb{R}$. Let $x_0 \in U$ and $A := U - x_0$. Then A is an open set with $0 \in A$ and $-x_0 \notin A$. The Minkowski functional p_A is finite, sublinear and it fulfils $p_A(-x_0) \geq 1$.

On $Y := \operatorname{lin}\{x_0\}$ we define the linear functional

$$y^*(tx_0) = (-t)p_A(-x_0) \quad \forall t \in \mathbb{R}.$$

For $t < 0$ we have $y^*(tx_0) = p_A(tx_0)$, while for $t \geq 0$ we have $y^*(tx_0) = (-t)p_A(-x_0) \leq 0 \leq p_A(tx_0)$, thus

$$y^*(y) \leq p_A(y) \quad \forall y \in Y.$$

Theorem 6.3 provides a linear extension x^* of y^* on X such that $x^*(x) \leq p_A(x)$ for all $x \in X$. Let $\varepsilon > 0$ be such that $\{x \in X \mid \|x\| < \varepsilon\} \subseteq A$. Then for every $x \in X, x \neq 0$, it holds $\frac{\varepsilon x}{2\|x\|} \in A = \operatorname{int} A$, thus $p_A(x) < \frac{2}{\varepsilon}\|x\|$. This means that for every $x \in X$ we have

$$|x^*(x)| = \max\{x^*(x), x^*(-x)\} \leq \max\{p_A(x), p_A(-x)\} \leq \frac{2}{\varepsilon}\|x\|,$$

which proves that x^* is continuous.

In addition, for every $x \in U$ it holds

$$x^*(x) = x^*(x - x_0) + x^*(x_0) \leq p_A(x - x_0) - p_A(-x_0) \leq p_A(x - x_0) - 1 < 0,$$

since $x - x_0 \in A$.

In case $\mathbb{K} = \mathbb{C}$ the statement follows by combining Theorem 6.5 with Lemma 6.4. ■

Example 7.5 Consider the normed space $(c_{00}, \|\cdot\|_\infty)$ over $\mathbb{R}$ and

$$U = \{(t_n)_{n\in\mathbb{N}} \in c_{00} \setminus \{0\} \mid t_N > 0 \text{ for } N := \max\{n \in \mathbb{N} \mid t_n \neq 0\}\}.$$

The set U is convex and $0 \notin U$. It is easy to notice that U is not open (for instance, $e_1 \in U \setminus \operatorname{int} U$). However, there exists no $x^* \in (c_{00})^*$ such that $x^*|_U < 0$. Assume that there exists such a continuous linear functional on c_{00}. Then it has a unique continuous linear extension on c_0 (see Theorem 3.6). According to Example 4.2(b) this can be identified with a sequence $(s_n)_{n\in\mathbb{N}} \in \ell^1$. Since $e_n \in U$, $x^*(e_n) = s_n < 0$ for all $n \in \mathbb{N}$. Consider $x := -\frac{s_2}{s_1}e_1 + e_2 \in U$. Then $x^*(x) = 0$, which leads to a contradiction.

Theorem 7.6 (Hahn-Banach Separation Theorem) *Let X be a normed space and $U, V \subseteq X$ two convex sets such that U is open and $U \cap V = \emptyset$. Then there exists $x^* \in X^*$ such that*

$$\operatorname{Re} x^*(u) < \operatorname{Re} x^*(v) \quad \forall u \in U \ \forall v \in V.$$

Proof. The set $U - V = \{u - v \mid u \in U, v \in V\}$ is convex and open (since $U - V = \cup_{v\in V}(U - v)$), and $0 \notin U - V$. Lemma 7.4 provides an element $x^* \in X^*$ such that $\operatorname{Re} x^*(u - v) < 0$ for all $u \in U$ and all $v \in V$, which leads to the desired conclusion. ∎

Theorem 7.7 (Hahn-Banach Strong Separation Theorem) *Let X be a normed space and $U, V \subseteq X$ two convex sets such that U is compact, V is closed and $U \cap V = \emptyset$. Then there exists $x^* \in X^*$ such that*

$$\sup_{u\in U} \operatorname{Re} x^*(u) < \inf_{v\in V} \operatorname{Re} x^*(v).$$

Proof. The set $U - V$ is closed and $0 \notin U - V$. This means that there exists $r > 0$ such that $\{x \in X \mid \|x\| < r\} \cap (U - V) = \emptyset$. According to Theorem 7.6, there exists $x^* \in X^*, x^* \neq 0$, such that

$$\operatorname{Re} x^*(x) < \operatorname{Re} x^*(u - v) \quad \forall x \in X \text{ with } \|x\| < r \ \forall u \in U \ \forall v \in V,$$

thus

$$\sup_{x\in X, \|x\|<r} |\operatorname{Re} x^*(x)| \leq \inf_{u\in U, v\in V} \operatorname{Re} x^*(u - v)$$

or, equivalently,

$$0 < r\|\operatorname{Re} x^*\| \leq \inf_{u\in U, v\in V} \operatorname{Re} x^*(u - v),$$

which finishes the proof. ∎

8 Weak convergence and reflexivity

Let X be a normed space, X^* its dual space, and X^{**} the dual space of X^*. X^{**} is called the bidual space of X.

Theorem 8.1 (the canonical embedding of a normed space into its bidual) *The mapping $i_X : X \to X^{**}$ defined for every $x \in X$ as*

$$i_X(x) : X^* \to \mathbb{K}, \quad (i_X(x))\,(x^*) = x^*(x),$$

is a linear isometry (which is in general not surjective). It is called the canonical embedding of the normed space X into its bidual.

Proof. Let $x \in X$. It is easy to see that $i_X(x)$ is linear. Since $|\,(i_X(x))\,(x^*)| = |x^*(x)| \le \|x^*\|\|x\|$ for all $x \in X$ and all $x^* \in X^*$, $i_X(x)$ is continuous and $\|i_X(x)\| \le \|x\|$. On the other hand, according to Corollary 6.8, there exists $x^* \in X^*, \|x^*\| \le 1$, such hat $|\,(i_X(x))\,(x^*)| = |x^*(x)| = \|x\|$, thus $\|i_X(x)\| = \|x\|$.

Obviously, i_X is linear. ∎

The canonical embedding allows to identify a normed space X with a linear subspace of its bidual. If X is complete, then $i_X(X)$ is also complete, thus it is a closed linear subspace of X^{**} (see Lemma 1.6(b)).

For a normed space X, since X^{**} is a Banach space, we have that $\overline{i_X(X)}$ is a closed linear subspace, thus it is complete (see Lemma 1.6(a)). This means that every normed space is isometrically isomorph to a dense linear subspace of a Banach space.

Definition 8.2 (reflexive space) A Banach space X is said to be reflexive if $i_X : X \to X^{**}$ is surjective.

Since X^{**} is complete, a non complete normed space cannot be reflexive.

Example 8.3 (a) c_0 and ℓ^1 are not reflexive. Let $X = c_0$. We have seen in Example 4.2 that $X^* \cong \ell^1$ and $X^{**} \cong \ell^\infty$. Let $x = (s_n)_{n\in\mathbb{N}} \in c_0$. We associate $x^* \in (c_0)^* = X^*$ with an element $(t_n)_{n\in\mathbb{N}} \in \ell^1$ and get

$$(i_{c_0}(x))\,(x^*) = x^*(x) = \sum_{n=1}^{+\infty} s_n t_n = L_x(x^*),$$

where $L_x \in (\ell^1)^* = X^{**}$ maps $x^* = (t_n)_{n\in\mathbb{N}} \in X^*$ to $\sum_{n=1}^{+\infty} s_n t_n$. Thus $i_{c_0}(x) = L_x \in (c_0)^{**}$. By identifying $(c_0)^{**}$ with ℓ^∞ and L_x with x, we get $i_{c_0} : c_0 \to \ell^\infty, i_{c_0}(x) = x$. Obviously, i_{c_0} is not surjective.

Similarly, for $x = (s_n)_{n\in\mathbb{N}} \in \ell_1$ and $x^* \in (\ell^1)^*$, which we can associate with an element $(t_n)_{n\in\mathbb{N}} \in \ell^\infty$, we have

$$(i_{\ell^1}(x))\,(x^*) = x^*(x) = \sum_{n=1}^{+\infty} s_n t_n.$$

We have seen in Example 6.13 that i_{ℓ^1} is not surjective.

(b) For $1 < p < \infty$, the spaces ℓ^p and $L^p(\mu)$ are reflexive. As above, the canonical embedding i_{ℓ^p} can be identified with the identical operator $\mathrm{Id} : \ell^p \to \ell^p$, therefore it is surjective. Similar arguments prove that the spaces $L^p(\mu), 1 < p < \infty$, are reflexive.

(c) Finite-dimensional normed spaces X are reflexive, since $\dim X = \dim X^* = \dim X^{**}$.

(d) If X is reflexive, then $X \cong X^{**}$. However, the opposite statement is not true in general. For the space J in (Übungsbeispiel 9, Gruppe 1) it holds $J \cong J^{**}$, however, i_J is not surjective.

Theorem 8.4 *(a) Closed linear subspaces of reflexive spaces are reflexive.*

(b) A Banach space X is reflexive if and only if X^* is reflexive.

Proof. (a) Let X be reflexive and $M \subseteq X$ a closed linear subspace. Let $m^{**} \in M^{**}$. The functional $x^* \mapsto m^{**}(x^*|_M)$ on X^* is linear and continuous, since

$$|m^{**}(x^*|_M)| \le \|m^{**}\|\|x^*|_M\| \le \|m^{**}\|\|x^*\|,$$

which means that it is an element of X^{**}. Since X is reflexive, there exists $x \in X$ such that

$$m^{**}(x^*|_M) = x^*(x) \quad \forall x^* \in X^*.$$

We have that $x \in M$. Otherwise, according to Corollary 6.9, there would exists $x^* \in X^*$ such that $x^*|_M = 0$ and $x^*(x) \neq 0$. Contradiction! From now on, we denote x by m.

It remains to show that $m^{**}(m^*) = m^*(m)$ for all $m^* \in M^*$. Indeed, let $m^* \in M^*$ and $x^* \in X^*$ its extension (the existence of which is guaranteed by Theorem 6.6). It holds $m^{**}(m^*) = m^{**}(x^*|_M) = x^*(m) = m^*(m)$. We proved that $i_M(m) = m^{**}$, which shows that M is reflexive.

(b) "$\Rightarrow$" Assume that X is reflexive. Let $x^{***} \in X^{***}$. The functional $x^* : X \to \mathbb{K}, x^*(x) = x^{***}(i_X(x))$, is linear and continuous, thus $x^* \in X^*$. For every $x^{**} \in X^{**}$ there exists $x \in X$ such that $i_X(x) = x^{**}$, thus

$$x^{***}(x^{**}) = x^{***}(i_X(x)) = x^*(x) = (i_X(x))(x^*) = x^{**}(x^*),$$

which shows that $x^{***} = i_{X^*}(x^*)$. In conclusion, i_{X^*} is surjective.

"$\Leftarrow$" If X^* is reflexive, then X^{**} is also reflexive and, according to (a), the closed linear subspace $i_X(X)$ is reflexive, too. Since $X \cong i_X(X)$, it follows that X is also reflexive. ∎

Theorem 8.4 and Theorem 6.12 give the following corollary.

Corollary 8.5 *A reflexive normed space is separable if and only if its dual space is separable.*

Example 8.6 The spaces ℓ^∞, $L^1[0,1]$, $L^\infty[0,1]$ and $C[0,1]$ are not reflexive.

Definition 8.7 (weak convergence) Let X be a normed space. A sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ is said be weakly convergent to $x \in X$ if

$$\lim_{n\to+\infty} x^*(x_n) = x^*(x) \quad \forall x^* \in X^*.$$

We write $x_n \xrightarrow{w} x$ $(n \to +\infty)$.

Since X^* separates points in X (see Corollary 6.7), the limit is uniquely determined. Every convergent sequence is weakly convergent. The sequence $(e_n)_{n\in\mathbb{N}}$ in $\ell^p, 1 < p < \infty$, or in c_0 converges weakly to 0, while $\|e_n\| = 1$ for all $n \in \mathbb{N}$, which shows that weakly convergent sequences are in general not convergent. However, in finite-dimensional spaces, a sequence is convergent if and only if it is weakly convergent.

Theorem 8.8 *In a reflexive space X every bounded sequence has a weakly convergent subsequence.*

Proof. First we assume that X is separable. According to Corollary 8.5, X^* is also separable; assume that $X^* = \overline{\{x_1^*, x_2^*, ...\}}$. Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in X. By using a similar diagonalization technique as in the proof of the Arzelá-Ascoli Theorem, we construct a subsequence $(\xi_n)_{n\in\mathbb{N}}$ such that the sequence $(x_m^*(\xi_n))_{n\in\mathbb{N}}$ is convergent for all $m \in \mathbb{N}$.

Let $x^* \in X^*$ and $\varepsilon > 0$. Let $M > 0$ such that $\|x_n\| \leq M$ for all $n \in \mathbb{N}$ and $K \in \mathbb{N}$ such that $\|x_K^* - x^*\| < \frac{\varepsilon}{3M}$. Then there exists $N \in \mathbb{N}$ such that for all $m, n \geq \mathbb{N}$ it holds

$$|x^*(\xi_n) - x^*(\xi_m)| \leq 2M\|x_K^* - x^*\| + |x_K^*(\xi_n) - x_K^*(\xi_m)| < \varepsilon.$$

This shows that $(x^*(\xi_n))_{n\in\mathbb{N}}$ is a Cauchy sequence, therefore, a convergent sequence.

Consider the functional on X^*

$$\lim : X^* \to \mathbb{K}, \quad \lim(x^*) = \lim_{n\to+\infty} x^*(\xi_n).$$

It is well-defined and linear. Since

$$|\lim(x^*)| = \left| \lim_{n\to+\infty} x^*(\xi_n) \right| = \lim_{n\to+\infty} |x^*(\xi_n)| \le \|x^*\| M \ \forall x^* \in X^*,$$

$\lim \in X^{**}$. Since X is reflexive, there exists $\xi \in X$ such that $i_X(\xi) = \lim$, which means that

$$x^*(\xi) = \lim(x^*) = \lim_{n\to+\infty} x^*(\xi_n) \quad \forall x^* \in X^*$$

and therefore shows that $\xi_n \overset{w}{\to} \xi$ $(n \to +\infty)$.

Consider now the general case of a reflexive normed space X. Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in X and the closed linear subspace $Y := \overline{\mathrm{lin}\{x_1, x_2, ...\}}$. Then Y is separable and reflexive. As proved above, there exists a subsequence $(\xi_n)_{n\in\mathbb{N}}$ and an element $\xi \in Y$ such that $y^*(\xi_n) \to y^*(\xi)$ $(n \to +\infty)$ for all $y^* \in Y^*$. Let $x^* \in X^*$. Then $x^*|_Y \in Y^*$, which means that $x^*(\xi_n) \to x^*(\xi)$ $(n \to +\infty)$. This shows that $\xi_n \overset{w}{\to} \xi$ $(n \to +\infty)$. ■

In ℓ^2 we have that $e_n \overset{w}{\to} 0$ as $n \to +\infty$ and $(e_n)_{n\in\mathbb{N}} \subseteq S_{\ell^2} = \{x \in \ell^2 \mid \|x\| = 1\}$, which is a closed set. This shows that closed sets are not necessarily "weakly closed". However, the two notions are equivalent for convex sets. This is another consequence of the Hahn-Banach Theorem.

Theorem 8.9 *Let X be a normed space and $U \subseteq X$ a closed convex set. If $(x_n)_{n\in N} \subseteq U$ is weakly convergent to x, then $x \in U$.*

Proof. Assume that $x \notin U$. According to the Hahn-Banach Strong Separation Theorem, there exist $\varepsilon > 0$ and $x^* \in X^*$ such that $\varepsilon < \operatorname{Re} x^*(u) - \operatorname{Re} x^*(x) = \operatorname{Re} x^*(u-x)$ for all $u \in U$. Thus $\operatorname{Re} x^*(x_n - x) > \varepsilon$ for all $n \in \mathbb{N}$, which contradicts $\lim_{n\to+\infty} x^*(x_n) = x^*(x)$. ■

Corollary 8.10 *If $x_n \overset{w}{\to} x$ $(n \to +\infty)$, then there exists a sequence of convex combinations*

$$y_n = \sum_{i=n}^{N_n} \lambda_i^n x_i, \ \text{where } \lambda_i^n \ge 0, i = n, ..., N_n, \sum_{i=n}^{N_n} \lambda_i^n = 1,$$

such that $\lim_{n\to+\infty} \|y_n - x\| = 0$.

Proof. For all $n \in \mathbb{N}$, define $U_n := \overline{\mathrm{co}\{x_n, x_{n+1}, ...\}}$, where $\mathrm{co}\, A := \cap\{C \subseteq X \mid A \subseteq C, C \text{ is convex}\}$ denotes the convex hull of a set A. According to Theorem 8.9, $x \in U_n$ for all $n \in \mathbb{N}$. Thus, for all $n \in \mathbb{N}$, there exists an element $y_n \in \mathrm{co}\{x_n, x_{n+1}, ...\}$ such that $\|y_n - x\| < \frac{1}{n}$. ■

9 Adjoint operators

Adjoint operators are dual objects which are associated to continuous linear operators.

Definition 9.1 (adjoint operator) Let X and Y be normed spaces and $T \in L(X,Y)$. The operator $T^* : Y^* \to X^*$ defined through

$$(T^*y^*)(x) := y^*(Tx) \quad \forall x \in X \ \forall y^* \in Y^*$$

is called the adjoint operator of T.

It is easy to see that $T^*y^* \in X^*$ for every $y^* \in Y^*$ and that $T^* \in L(Y^*, X^*)$.

Example 9.2 (a) Let $1 \le p < \infty$ and $T : \ell^p \to \ell^p$ be the left shift operator

$$T(s_1, s_2, ...) = (s_2, s_3, ...).$$

In order to find its adjoint operator $T^* : \ell^q \to \ell^q$, where $\frac{1}{p} + \frac{1}{q} = 1$, we will identify $y^* \in (\ell^p)^*$ with an element $(t_n)_{n\in\mathbb{N}} \in \ell^q$. For $x = (s_n)_{n\in\mathbb{N}}$ we have $y^*(Tx) = \sum_{n=1}^{+\infty} s_{n+1}t_n = \sum_{n=2}^{+\infty} s_n t_{n-1}$. This shows that T^* is the right shift operator

$$T^*(t_1, t_2, ...) = (0, t_1, t_2, ...).$$

For $p = q = 2$ we have $TT^* = \mathrm{Id}$, however, $T^*T \neq \mathrm{Id}$.

(b) Let $1 \le p < \infty$ and $T_{(p)} : L^p[0,1] \to L^p[0,1]$ be the multiplication operator $T_{(p)}f = hf$, where $h \in L^\infty[0,1]$. Its adjoint operator is the multiplication operator $T_{(q)} : L^q[0,1] \to L^q[0,1]$ on $L^q[0,1]$, where $\frac{1}{p} + \frac{1}{q} = 1$. Indeed, for all $f \in L^p[0,1]$ and $g \in L^q[0,1]$ it holds

$$\begin{aligned}(T^*_{(p)}g)(f) &= \int_0^1 (T_{(p)}f)(t)g(t)dt = \int_0^1 h(t)f(t)g(t)dt \\ &= \int_0^1 (T_{(q)}g)(t)f(t)dt = (T_{(q)}g)(f).\end{aligned}$$

(c) The adjoint operator of the canonical embedding $i_X : X \to X^{**}$, $(i_X)^* : X^{***} \to X^*$, fulfils $((i_X)^*(x^{***}))(x) = x^{***}(i_X(x))$ for all $x \in X$ and all $x^{***} \in X^{***}$, thus it is nothing else than the restriction mapping $x^{***} \mapsto x^{***}|_{i_X(X)}$.

Theorem 9.3 *(a) The mapping $T \mapsto T^*$ from $L(X,Y)$ to $L(Y^*,X^*)$ is linear and isometric, namely, $\|T\| = \|T^*\|$.*

(b) For $T \in L(X,Y)$ and $S \in L(Y,Z)$ it holds $(ST)^* = T^*S^*$.

Proof. (a) The linearity of the mapping is clear. By using Corollary 6.8 we get

$$\|T\| = \sup_{x\in B_X} \|Tx\| = \sup_{x\in B_X} \sup_{y^*\in B_{Y^*}} |y^*(Tx)| = \sup_{y^*\in B_{Y^*}} \|T^*y^*\| = \|T^*\|.$$

(b) Follows by straightforward calculations. ∎

Lemma 9.4 *For $T \in L(X,Y)$ we have*

$$T^{**} \circ i_X = i_Y \circ T.$$

Proof. For every $x \in X$ and every $y^* \in Y^*$ it holds

$$[T^{**}(i_X(x))](y^*) = (i_X(x))(T^*y^*) = (T^*y^*)(x) = y^*(Tx) = [i_Y(Tx)](y^*).$$

∎

By using Lemma 9.4 one can easily see that $S \in L(Y^*, X^*)$ is an adjoint operator if and only if $S^*(X) \subseteq Y$, where X and Y are identified with $i_X(X)$ and $i_Y(Y)$, respectively. This remark is helpful when showing that the mapping in Theorem 9.3(a) is in general not surjective (see Übungsbeispiel 37).

Theorem 9.5 (Schauder Theorem) *Let X and Y be Banach spaces and $T : X \to Y$ a continuous linear operator. T is compact if and only if T^* is compact.*

Proof. "⇒" Let T be compact and $(y_n^*)_{n\in\mathbb{N}}$ a bounded sequence in Y^*. Then $K := \overline{T(B_X)} \subseteq Y$ is a compact metric space. The sequence $(f_n := y_n^*|_K)_{n\in\mathbb{N}}$ is bounded and equicontinuous, since

$$|f_n(u) - f_n(v)| \leq \sup_{n\in\mathbb{N}} \|y_n^*\| \|u - v\| \quad \forall n \in \mathbb{N}.$$

According to the Arzelá-Ascoli Theorem, $\{f_n \mid n \in \mathbb{N}\}$ is relatively compact, thus there exists a convergent subsequence $(f_{n_k})_{k\in\mathbb{N}}$. For all $k, l \in \mathbb{N}$ it holds

$$\|T^*y_{n_k}^* - T^*y_{n_l}^*\| = \sup_{x\in B_x} |(y_{n_k}^* - y_{n_l}^*)(Tx)| = \sup_{y\in T(B_x)} |(y_{n_k}^* - y_{n_l}^*)(y)| = \|y_{n_k}^* - y_{n_l}^*\|_\infty,$$

since $T(B_X)$ is dense in K. Thus $(T^*y_{n_k}^*)_{k\in\mathbb{N}}$ is convergent, which proves that T^* is compact.

"⇐" Let T^* be compact. From the first part of the proof it follows that T^{**} is also compact, thus $T^{**} \circ i_X$ is compact and, according to Lemma 9.4, $i_Y \circ T : X \to Y^{**}$ is also compact. Since $i_Y(Y)$ is a closed linear subspace of Y^{**}, it follows that T is compact. ∎

For $T : X \to Y$ a given linear operator, we have seen that $\ker T = \{x \in X : Tx = 0\}$ and $\operatorname{ran} T = T(X)$ are linear subspaces. If T is continuous, then $\ker T$ is also closed; on the other hand, $\operatorname{ran} T$ is in general not closed. To see this, consider for instance the identity operator from $(P[a,b], \|\cdot\|_\infty)$ to $(C[a,b], \|\cdot\|_\infty)$ (see Übungsbeispiel 8). We have the following characterization of the closure of the range of a continuous linear operator.

Theorem 9.6 *It holds*

$$\overline{\operatorname{ran} T} = (\ker T^*)_\perp.$$

Proof. "$\subseteq$" Let $y = Tx \in \operatorname{ran} T$. For all $y^* \in \ker T^*$ we have $y^*(y) = y^*(Tx) = T^*y^*(x) = 0$, thus $\operatorname{ran} T \subseteq (\ker T^*)_\perp$, and the inclusion follows from the fact that the anihilator is closed.

"$\supseteq$" Let $y \notin \overline{\operatorname{ran} T}$. According to Corollary 6.9, there exists $y^* \in Y^*$ such that $y^*|_{\overline{\operatorname{ran} T}} = 0$ and $y^*(y) \neq 0$. Thus $T^*y^*(x) = y^*(Tx) = 0$ for all $x \in X$, which shows that $y^* \in \ker T^*$. Since $y^*(y) \neq 0$, $y \notin (\ker T^*)_\perp$. ■

The following corollary is a direct consequence of Theorem 9.6.

Corollary 9.7 *Let $T : X \to Y$ a continuous linear operator with closed range and $y \in Y$. The operator equation*

$$Tx = y$$

has a solution if and only if the following implication holds

$$T^*y^* = 0 \quad \Longrightarrow \quad y^*(y) = 0.$$

Chapter IV

The fundamental theorems for operators on Banach spaces

In this chapter we will study the fundamental theorems for operators on Banach spaces. The main tool in the derivation of these results is the Baire Category Theorem.

10 An essential tool: the Baire Category Theorem

Theorem 10.1 (Baire Theorem) *Let (T, d) be a complete metric space and a sequence $(O_n)_{n\in\mathbb{N}}$ of open and dense sets in T. Then $\cap_{n\in\mathbb{N}} O_n$ is dense in T.*

Proof. Set $D := \cap_{n\in\mathbb{N}} O_n$. We will prove that every open ball $U_\varepsilon(x_0) = \{x \in T \mid d(x_0, x) < \varepsilon\}$, for $\varepsilon > 0$ and $x_0 \in T$, intersects D.

As O_1 is open and dense, $O_1 \cap U_\varepsilon(x_0)$ is open and nonempty. We choose $x_1 \in O_1$ and $0 < \varepsilon_1 < \frac{1}{2}\varepsilon$ such that

$$\overline{U_{\varepsilon_1}(x_1)} \subseteq O_1 \cap U_\varepsilon(x_0).$$

As O_2 is open and dense, $O_2 \cap U_{\varepsilon_1}(x_1)$ is open and nonempty. Then there exist $x_2 \in O_2$ and $0 < \varepsilon_2 < \frac{1}{2}\varepsilon_1$ such that

$$\overline{U_{\varepsilon_2}(x_2)} \subseteq O_2 \cap U_{\varepsilon_1}(x_1) \subseteq O_1 \cap O_2 \cap U_\varepsilon(x_0).$$

By repeating the above procedure, we can inductively construct the sequences $(\varepsilon_n)_{n\in\mathbb{N}}$ and $(x_n)_{n\in\mathbb{N}}$ such that for all $n \geq 2$

$$0 < \varepsilon_n < \frac{1}{2}\varepsilon_{n-1}, \text{ thus, } \varepsilon_n < \frac{\varepsilon}{2^n},$$

and

$$\overline{U_{\varepsilon_n}(x_n)} \subseteq O_n \cap U_{\varepsilon_{n-1}}(x_{n-1}) \subseteq ... \subseteq O_1 \cap ... \cap O_n \cap U_\varepsilon(x_0).$$

Let $N \in \mathbb{N}$. For all $n > N$ it holds $x_n \in U_{\varepsilon_N}(x_N) \subseteq U_{\frac{1}{2^N}\varepsilon}(x_N)$, thus $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Since T is complete, there exists $x := \lim_{n\to+\infty} x_n \in \overline{U_{\varepsilon_N}(x_N)}$ for all $N \in \mathbb{N}$, thus

$$x \in \cap_{n\in\mathbb{N}} O_n \cap U_\varepsilon(x_0) = D \cap U_\varepsilon(x_0).$$

■

Notice that Theorem 10.1 does not claim that $\cap_{n\in\mathbb{N}} O_n$ is open. Notice also that the completeness of the metric space T is an essential hypothesis. Indeed, for $T = \mathbb{Q}$, and assuming that $\mathbb{Q} = \{q_1, q_2, ...\}$, the sets $O_n := \mathbb{Q} \setminus \{q_n\}$ are open and dense in $\mathbb{Q}$, while $\cap_{n\in\mathbb{N}} O_n = \emptyset$.

A countable intersection of open sets is called a G_δ set (the notation originated in Germany: G means "Gebiet", δ means "Durchschnitt"). Theorem 10.1 says that in a complete metric space the countable intersection of dense G_δ sets is a dense G_δ set.

Definition 10.2 (a) A subset D of a metric space is called nowhere dense if $\mathrm{int}(\overline{D}) = \emptyset$.

(b) The set D is said to be of the first category if there exists a sequence $(D_n)_{n\in\mathbb{N}}$ of nowhere dense sets with $D = \cup_{n\in\mathbb{N}} D_n$.

(c) The set D is said to be of the second category if D is not of the first category.

The set $\{1/n \mid n \in \mathbb{N}\}$ is nowhere dense in $\mathbb{R}$. The set $\mathbb{Q}$ is of the first category in $\mathbb{R}$.

Corollary 10.3 (Baire Category Theorem) *The complement of a subset of the first category of a complete metric space is dense.*

Proof. Let $(D_n)_{n\in\mathbb{N}}$ be a sequence of nowhere dense sets with $D = \cup_{n\in\mathbb{N}} D_n$. Then $X \setminus D = \cap_{n\in\mathbb{N}}(X \setminus D_n) \supseteq \cap_{n\in\mathbb{N}}(X \setminus \overline{D_n})$. The conclusion follows from Theorem 10.1, by noticing that $X \setminus \overline{D_n}$ are open and dense sets, since $\overline{X \setminus \overline{D_n}} = X \setminus \mathrm{int}(\overline{D_n}) = X$ for all $n \in \mathbb{N}$. ■

A useful consequence of Corollary 10.3 follows.

Corollary 10.4 *A nonempty complete metric space is a set of the second category in itself.*

The Baire Category Theorem is a very helpful instrument when proving existence results. We will illustrate this in Übungsbeispiel 38, Gruppe 1, where we will show that there exists continuous functions on $[0, 1]$ that are nowhere differentiable.

11 The Uniform Boundedness Principle

Theorem 11.1 (Banach-Steinhaus Theorem) *Let X be a Banach space, Y a normed space, I an index set, and $T_i \in L(X, Y)$, for $i \in I$, a family of continuous linear operators. If*

$$\sup_{i\in I} \|T_i x\| < +\infty \quad \forall x \in X,$$

then

$$\sup_{i\in I} \|T_i\| < +\infty.$$

Proof. Define $E_n := \{x \in X \mid \sup_{i\in I} \|T_i x\| \leq n\}$ for all $n \in \mathbb{N}$. It holds $X = \cup_{n\in\mathbb{N}} E_n$. Since T_i is continuous, for all $i \in I$, the set $E_n = \cap_{i\in I} \|T_i\|^{-1}([0, n])$, for $n \in \mathbb{N}$, is closed. According to Corollary 10.4, X is of second category, thus at least one of the sets $(E_n)_{n\in\mathbb{N}}$ has an interior point.

Let $N \in \mathbb{N}$, $y_N \in E_N$ and $\varepsilon > 0$ be such that $U_\varepsilon(y_N) = \{x \in X \mid \|x - y_N\| < \varepsilon\} \subseteq E_N$. For every $u \in X$ with $\|u\| < \varepsilon$ it holds $u + y_N, u - y_N \in E_N$, thus $u = \frac{1}{2}(u + y_N) + \frac{1}{2}(u - y_N) \in \frac{1}{2}E_N + \frac{1}{2}E_N \subseteq E_N$, since E_N is convex.

Let $i \in I$. For all $\|u\| < \varepsilon$ it holds $\|T_i u\| \leq N$, thus $\|T_i\| \leq \frac{N}{\varepsilon}$. In conclusion,

$$\sup_{i\in I} \|T_i\| \leq \frac{N}{\varepsilon}.$$

■

The completeness of X is an essential hypothesis for the Uniform Boundedness Principle. Indeed, consider the family of operators $(T_n)_{n\in\mathbb{N}} \subseteq L(c_{00}, \mathbb{K})$ defined by $T_n((t_m)_{m\in\mathbb{N}}) = nt_n$ for all $n \in \mathbb{N}$. Obviously, we have $\sup_{n\in\mathbb{N}} |T_n((t_m)_{m\in\mathbb{N}})| < +\infty$ for all $(t_m)_{m\in\mathbb{N}} \in c_{00}$, however, $\sup_{n\in\mathbb{N}} \|T_n\| = \sup_{n\in\mathbb{N}} n = +\infty$.

Corollary 11.2 *Let U be a subset of a normed space X. The following statements are equivalent:*

(a) U is bounded.

(b) $x^*(U) \subseteq \mathbb{K}$ is bounded for all $x^* \in X^*$.

Proof. "$(a) \Rightarrow (b)$" Clear.

"$(b) \Rightarrow (a)$". Consider the family of operators $(i_X(x))_{x\in U} \subseteq L(X^*, \mathbb{K})$. For every $x^* \in X^*$ it holds

$$\sup_{x\in U} |(i_X(x))(x^*)| = \sup_{x\in U} |x^*(x)| < +\infty,$$

thus

$$\sup_{x\in U} \|(i_X(x)\| = \sup_{x\in U} \|x\| < +\infty,$$

which proves that U is bounded. ■

Corollary 11.3 *Weak convergent sequences are bounded.*

Proof. Let $(x_n)_{n\in\mathbb{N}}$ converge weakly to x. The statement follows from Corollary 11.2, by noticing that $(x^*(x_n))_{n\in\mathbb{N}}$ is convergent, thus bounded, for all $x^* \in X^*$. ■

The following corollary is a dual statement to Corollary 11.2 and is a direct consequence of Theorem 11.1.

Corollary 11.4 *Let X be a Banach space and U be a subset of its dual space X^*. The following statements are equivalent:*

(a) U is bounded.

(b) The set $\{x^*(x) \mid x^* \in U\}$ is bounded for all $x \in X$.

The following corollary shows that operators defined as pointwise limits of continuous linear operators are also continuous and linear.

Corollary 11.5 *Let X be a Banach space, Y a normed space and $(T_n)_{n\in\mathbb{N}} \subseteq L(X,Y)$ such that for all $x \in X$ the limit $Tx := \lim_{n\to+\infty} T_n x$ exists. Then it holds $T \in L(X,Y)$.*

Proof. The linearity of T is clear. For all $x \in X$ we have that $\sup_{n\in\mathbb{N}} \|T_n x\| < +\infty$, thus $M := \sup_{n\in\mathbb{N}} \|T_n\| < +\infty$. From here we get

$$\|Tx\| = \lim_{n\to+\infty} \|T_n x\| \leq M\|x\| \quad \forall x \in X.$$

■

12 The Open Mapping Theorem

Definition 12.1 (open mapping) A mapping between two metric spaces that maps open sets to open sets is called open.

An open mapping does not necessarily map closed sets to closed sets. The projection mapping $p : \mathbb{R}^2 \to \mathbb{R}, p(u,v) = u$ is open, however, $p(\{(u,v) \in \mathbb{R}^2 \mid u > 0, v > 0, uv \geq 1\}) = (0,+\infty)$. Obviously, a bijective mapping is open if and only if its inverse is continuous.

Lemma 12.2 *Let $T : X \to Y$ be a linear operator between two normed spaces X and Y. The following statements are equivalent:*

(i) T is open.

(ii) *T maps open balls around zero to zero neighbourhoods, namely,*

$$\forall r > 0 \; \exists \varepsilon > 0 \text{ such that } V_\varepsilon \subseteq T(U_r),$$

where $U_r := \{x \in X \mid \|x\| < r\}$ *and* $V_\varepsilon := \{y \in Y \mid \|y\| < \varepsilon\}$.

(iii)

$$\exists \varepsilon > 0 \text{ such that } V_\varepsilon \subseteq T(U_1).$$

Proof. "$(ii) \Leftrightarrow (iii)$" Clear, since $U_r = rU_1$ and T is linear.

"$(i) \Rightarrow (ii)$" Also clear, since $0 \in T(U_r)$ and $T(U_r)$ is open.

"$(ii) \Rightarrow (i)$" Let O be an open set in X, $y \in T(O)$ and $x \in O$ such that $Tx = y$. Let $r > 0$ be such that $x + U_r \subseteq O$ and, according to (ii), let $\varepsilon > 0$ be such that $V_\varepsilon \subseteq T(U_r)$. Then it holds $y + V_\varepsilon \subseteq Tx + T(U_r) \subseteq T(O)$, which proves that $T(O)$ is an open neighbourhood of y. This shows that $T(O)$ is open. ■

Example 12.3 (a) Every quotient operator (see Example 3.16) is open.

(b) The operator $T : \ell^\infty \to c_0, T((t_n)_{n\in\mathbb{N}}) = (\frac{1}{n}t_n)_{n\in\mathbb{N}}$, is not open, since the set $T(U_1) = \left\{(s_n)_{n\in\mathbb{N}} \in c_0 \mid |s_n| < \frac{1}{n} \; \forall n \in \mathbb{N}\right\}$ is not open.

Every open linear operator is surjective. The following remarkable result, which goes back to Stefan Banach, asserts that in complete normed spaces the opposite statement is also true.

Theorem 12.4 (Open Mapping Theorem) *Let X and Y be Banach spaces and $T \in L(X, Y)$ a surjective operator. Then T is open.*

Proof. We will show that statement (iii) in Lemma 12.2 holds.

Step 1. We will show that there exists $\varepsilon_0 > 0$ such that $V_{\varepsilon_0} \subseteq \overline{T(U_1)}$.

Since T is surjective, we have $Y = \cup_{n\in\mathbb{N}} T(U_n)$. According to the Baire Category Theorem, there exists $N \in \mathbb{N}$ such that $\operatorname{int}(\overline{T(U_N)}) \neq \emptyset$. Let $y_0 \in \overline{T(U_N)}$ and $\varepsilon > 0$ such that

$$\|z - y_0\| < \varepsilon \quad \Rightarrow \quad z \in \overline{T(U_N)}.$$

Since $\overline{T(U_N)}$ is symmetric ($z \in \overline{T(U_N)} \Rightarrow -z \in \overline{T(U_N)}$), it holds $-y_0 \in \overline{T(U_N)}$ and

$$\|z + y_0\| < \varepsilon \quad \Rightarrow \quad z \in \overline{T(U_N)}.$$

Let $y \in Y$ with $\|y\| < \varepsilon$. Then $y + y_0, y - y_0 \in \overline{T(U_N)}$, thus (since $\overline{T(U_N)}$ is convex) $y = \frac{1}{2}(y + y_0) + \frac{1}{2}(y - y_0) \in \overline{T(U_N)}$, which shows that $V_\varepsilon \subseteq \overline{T(U_N)}$ and, consequently, $V_{\frac{\varepsilon}{N}} \subseteq \overline{T(U_1)}$.

Step 2. Let $\varepsilon_0 > 0$ with $V_{\varepsilon_0} \subseteq \overline{T(U_1)}$. We will prove that in fact it holds $V_{\varepsilon_0} \subseteq T(U_1)$. This will lead to the desired conclusion.

Let $y \in V_{\varepsilon_0}$. We choose $\varepsilon > 0$ such that $\|y\| < \varepsilon < \varepsilon_0$. Let $\overline{y} := \frac{\varepsilon_0}{\varepsilon} y$. Then $\|\overline{y}\| < \varepsilon_0$, thus, according to **Step 1**, $\overline{y} \in \overline{T(U_1)}$.

We choose $0 < \alpha < 1$ such that

$$\frac{\varepsilon}{\varepsilon_0} \frac{1}{1-\alpha} < 1.$$

Then there exists $y_0 = Tx_0 \in T(U_1)$ such that

$$\|\overline{y} - y_0\| < \alpha\varepsilon_0.$$

Since $\frac{\overline{y}-y_0}{\alpha} \in V_{\varepsilon_0} \subseteq \overline{T(U_1)}$, there exists $y_1 = Tx_1 \in T(U_1)$ such that

$$\left\| \frac{\overline{y} - y_0}{\alpha} - y_1 \right\| < \alpha\varepsilon_0$$

or, equivalently, $\|\overline{y} - (y_0 + \alpha y_1)\| < \alpha^2\varepsilon_0$.

Since $\frac{\overline{y}-(y_0+\alpha y_1)}{\alpha^2} \in V_{\varepsilon_0} \subseteq \overline{T(U_1)}$, there exists $y_2 = Tx_2 \in T(U_1)$ such that

$$\left\| \frac{\overline{y} - (y_0 + \alpha y_1)}{\alpha^2} - y_2 \right\| < \alpha\varepsilon_0$$

or, equivalently, $\|\overline{y} - (y_0 + \alpha y_1 + \alpha^2 y_2)\| < \alpha^3\varepsilon_0$.

In this way we can inductively construct a sequence $(x_n)_{n\geq 0} \subseteq U_1$ with

$$\left\| \overline{y} - T\left(\sum_{i=0}^{n} \alpha^i x_i \right) \right\| < \alpha^{n+1}\varepsilon_0 \quad \forall n \geq 0.$$

Since $\alpha < 1$, the series $\sum_{i=0}^{+\infty} \alpha^i x_i$ is absolutely convergent, thus there exists $\overline{x} := \sum_{i=0}^{+\infty} \alpha^i x_i \in X$. Obviously, $T\overline{x} = \overline{y}$.

For $x := \frac{\varepsilon}{\varepsilon_0}\overline{x}$ we have $Tx = y$ and

$$\|x\| \leq \frac{\varepsilon}{\varepsilon_0} \sum_{i=0}^{+\infty} \alpha^i \|x_i\| \leq \frac{\varepsilon}{\varepsilon_0} \sum_{i=0}^{+\infty} \alpha^i = \frac{\varepsilon}{\varepsilon_0} \frac{1}{1-\alpha} < 1,$$

which proves that $y \in T(U_1)$. ■

One can notice that the statement of the Open Mapping Theorem remains valid if instead of Y is a Banach space and T is surjective we assume that Y is a normed space and $\operatorname{ran} T$ is not of the first category in Y.

The Open Mapping Theorem has a series of important consequences.

Corollary 12.5 *Let X and Y be Banach spaces and $T \in L(X, Y)$ a bijective operator. Then T^{-1} is continuous.*

If $\|\cdot\|$ and $|||\cdot|||$ are two norms defined on a vector space X, such that both norms make X a Banach space, and there exists $M > 0$ with

$$\|x\| \leq M|||x||| \quad \forall x \in X,$$

then the identity operator from $(X, |||\cdot|||)$ to $(X, \|\cdot\|)$ has a continuous inverse, thus the two norms are equivalent.

Corollary 12.6 *Let X and Y be Banach spaces and $T \in L(X, Y)$ an injective operator. Then $\operatorname{ran} T$ is closed if and only if $T^{-1} : \operatorname{ran} T \to X$ is continuous.*

Proof. "$\Rightarrow$" If $\operatorname{ran} T$ is closed, then $\operatorname{ran} T$ is a Banach space. Since $T : X \to \operatorname{ran} T$ is a continuous and bijective operator, according to Corollary 12.5, $T^{-1} : \operatorname{ran} T \to X$ is continuous.

"$\Leftarrow$" If $T^{-1} : \operatorname{ran} T \to X$ is continuous, then $\operatorname{ran} T \simeq X$, which implies that $\operatorname{ran} T$ is complete. According to Lemma 1.6 (b), $\operatorname{ran} T$ is closed. ■

13 The Closed Graph Theorem

Definition 13.1 (closed graph operator) Let X and Y be normed spaces, $M \subseteq X$ a linear subspace and $T : M \to Y$ a linear operator. The operator T is called closed graph if for every sequence $(x_n)_{n\in\mathbb{N}} \subseteq M$ such that $(x_n)_{n\in\mathbb{N}}$ converges to $x \in X$ and $(Tx_n)_{n\in\mathbb{N}}$ converges to $y \in Y$, it holds $x \in M$ and $Tx = y$.

For an operator T defined on a linear subspace $M \subseteq X$ we usually write $\operatorname{dom} T = M$ and $T : \operatorname{dom} T \subseteq X \to Y$. The graph of a linear operator $T : \operatorname{dom} T \subseteq X \to Y$ is defined as

$$\operatorname{gr} T := \{(x, Tx) \mid x \in M\} \subseteq X \times Y.$$

Obviously, $\operatorname{gr} T$ is a linear subspace of $X \times Y$. The operator T is closed graph if and only if $\operatorname{gr} T$ is closed in $X \oplus_1 Y$ (see Übungsbeispiel 6).

Example 13.2 (a) Let $C[-1, 1]$ be endowed with the supremum norm. The differential operator $D : C^1[-1, 1] \subseteq C[-1, 1] \to C[-1, 1], Dx = x'$, is graph closed. Let $(x_n)_{n\in\mathbb{N}} \subseteq C^1[-1, 1]$ be such that $(x_n)_{n\in\mathbb{N}}$ converges (uniformly) to $x \in C[-1, 1]$ and $(x'_n)_{n\in\mathbb{N}}$ converges (uniformly) to $y \in C[-1, 1]$. Thus x is differentiable and $x' = y$, which proves that $x \in C^1[-1, 1]$ and $Dx = y$.

Notice that D is not continuous (see Example 3.8(c)).

(b) The differential operator $D : C^1[-1, 1] \subseteq L^2[-1, 1] \to L^2[-1, 1], Dx = x'$, is not graph closed. Let $x_n : [-1, 1] \to \mathbb{R}, x_n(t) = \left(t^2 + \frac{1}{n}\right)^{\frac{1}{2}}$, for all $n \in \mathbb{N}$, $x = |\cdot|$ and $y = \operatorname{sgn}$. Then $(x_n)_{n\in\mathbb{N}} \subseteq C^1[-1, 1]$, $x_n \to x$ and $Dx_n \to y$ in $L^2[-1, 1]$ as $n \to +\infty$. However, $x \notin C^1[-1, 1]$.

Lemma 13.3 *Let X and Y be Banach spaces, $M \subseteq X$ a linear subspace and $T : M \subseteq X \to Y$ a closed graph operator. The following statements are true:*

(a) M embedded with the norm $|||x||| := \|x\| + \|Tx\|, x \in X$, is a Banach space; $|||\cdot|||$ is called graph norm.

(b) $T : (M, |||\cdot|||) \to Y$ is continuous.

Proof. (a) It is easy to see that $|||\cdot|||$ is a norm. Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in M. This means that $(x_n)_{n\in\mathbb{N}}$ and $(Tx_n)_{n\in\mathbb{N}}$ are Cauchy sequences in X and Y, respectively. Thus there exist $x := \lim_{n\to+\infty} x_n \in X$ and $y := \lim_{n\to+\infty} Tx_n \in Y$. Since T is closed graph, $x \in M$ and $Tx = y$. From here it follows that $|||x_n - x||| \to 0$ as $n \to +\infty$, which proves that $(M, |||\cdot|||)$ is complete.

(b) The continuity follows from the fact that $\|Tx\| \le |||x|||$ for all $x \in M$. ■

The following result characterizes the continuity of the inverse of a closed graph operator.

Theorem 13.4 *Let X and Y be Banach spaces, $M \subseteq X$ a linear subspace and $T : M \subseteq X \to Y$ a closed graph and surjective operator. Then T is open. If, in addition, T is injective, then T^{-1} is continuous.*

Proof. According to Lemma 13.3, $T : (M, |||\cdot|||) \to Y$ is continuous. The Open Mapping Theorem guarantees that T is open. Let $O \subseteq M$ be an open set with respect to $\|\cdot\|$, the original norm on X. Since $\|\cdot\| \le |||\cdot|||$ on M, O is open with respect to $|||\cdot|||$, thus $T(O)$ is open. This shows that T is open with respect to the original norms or, equivalently, T^{-1} is continuous. ■

The main theorem of this section follows.

Theorem 13.5 (Closed Graph Theorem) *Let X and Y be Banach spaces and $T : X \to Y$ a linear and closed graph operator. Then T is continuous.*

Proof. Let $\|\cdot\|$ be the original norm on X, and $|||\cdot|||$ the graph norm on X. Since $\|\cdot\| \le |||\cdot|||$ on X and X endowed with each of the two norms is a Banach space, the two norms are equivalent. However, according to Lemma 13.3, T is continuous with respect to $|||\cdot|||$, thus T is continuous with respect to $\|\cdot\|$, too. ■

14 The Closed Range Theorem

We introduced in Corollary 9.7 a criterion which characterizes the solvability of operator equations governed by continuous linear operators with closed range. In this section we will provide useful necessary and sufficient conditions for an operator to have a closed range.

We start with two preparatory lemmas.

Lemma 14.1 *Let X and Y be Banach spaces and $T \in L(X, Y)$ an operator with closed range. Then there exists $K \geq 0$ such that*

$$\forall y \in \operatorname{ran} T \ \exists x \in X \ \textit{with } Tx = y \textit{ and } \|x\| \leq K\|y\|.$$

Proof. Notice that $\ker T$ is a closed linear subspace, thus $X/\ker T$ and $\operatorname{ran} T$ are Banach spaces. Consider the operator $\widehat{T} : X/\ker T \to \operatorname{ran} T, \widehat{T}[x] := Tx$. Then $\widehat{T}$ is well-defined, linear, continuous and bijective. According to Corollary 12.5, $\widehat{T}^{-1}$ is continuous, thus,

$$\|\widehat{T}^{-1}y\| \leq \|\widehat{T}^{-1}\|\|y\| \quad \forall y \in \operatorname{ran} T.$$

Let $K > \|\widehat{T}^{-1}\|$ and $y \in \operatorname{ran} T, y \neq 0$. Then there exists $x' \in X$ such that $Tx' = y$ and $\|[x']\| < K\|y\|$. Consequently, there exists $m \in \ker T$ such that $x := x' - m \in X$ fulfills $\|x\| \leq K\|y\|$ and $Tx = T(x' - m) = Tx' = y$. ■

Lemma 14.2 *Let X and Y be Banach spaces and $T \in L(X, Y)$. If there exists $c > 0$ such that*

$$c\|y^*\| \leq \|T^*y^*\| \quad \forall y^* \in Y^*,$$

then T is open and surjective.

Proof. We will prove that $\{y \in Y \mid \|y\| < c\} = V_c \subseteq T(U_1) = T(\{x \in X \mid \|x\| < 1\})$. To this end it is enough to prove that $V_c \subseteq \overline{T(U_1)}$ and further to argue similarly as in **Step 2** of the proof of Theorem 12.4.

Let $y_0 \in V_c$ and assume that $y_0 \notin \overline{T(U_1)}$. According to the Hahn-Banach Strong Separation Theorem, there exist $y^* \in Y^*$ and $\gamma \in \mathbb{R}$ such that

$$\operatorname{Re} y^*(y) \leq \gamma < \operatorname{Re} y^*(y_0) \quad \forall y \in \overline{T(U_1)}.$$

For all $x \in U_1$ it holds $\operatorname{Re}(T^*y^*)(x) = \operatorname{Re} y^*(Tx) \leq \gamma < \operatorname{Re} y^*(y_0)$ and, consequently,

$$\begin{aligned} \|T^*y^*\| = \|\operatorname{Re} T^*y^*\| &= \sup_{x \in U_1} \operatorname{Re}(T^*y^*)(x) < \operatorname{Re} y^*(y_0) \\ &\leq |y^*(y_0)| \leq \|y^*\|\|y_0\| < c\|y^*\|. \end{aligned}$$

Contradiction. ■

The main theorem of this section follows.

Theorem 14.3 (Closed Range Theorem) *Let X and Y be Banach spaces and $T \in L(X, Y)$. The following statements are equivalent:*

(i) $\operatorname{ran} T$ *is closed.*

(ii) $\operatorname{ran} T = (\ker T^*)_\perp$.

(iii) $\operatorname{ran} T^*$ *is closed.*

(iv) $\operatorname{ran} T^* = (\ker T)^\perp$.

Proof. "$(i) \Rightarrow (ii)$" This is Theorem 9.6.

"$(ii) \Rightarrow (i)$" Follows since the annihilator is closed.

"$(i) \Rightarrow (iv)$" For all $y^* \in Y^*$ and all $x \in \ker T$ it holds $(T^*y^*)(x) = y^*(Tx) = 0$, thus $\operatorname{ran} T^* \subseteq (\ker T)^\perp$. Let $x^* \in (\ker T)^\perp$ and define

$$z^* : \operatorname{ran} T \to \mathbb{K}, z^*(y) := x^*(x), \text{ where } y = Tx.$$

z^* is well-defined. Indeed, let $y = Tx = Tu$. Then $x - u \in \ker T$ and, so, $x^*(x - u) = 0$. It is easy to see that z^* is linear. Since $\operatorname{ran} T$ is closed, according to Lemma 14.1, there exists $K \geq 0$ such that

$$\forall y \in \operatorname{ran} T \ \exists x \in X \text{ with } Tx = y \text{ and } \|x\| \leq K\|y\|.$$

Then for all $y \in \operatorname{ran} T$ and such an element $x \in X$ it holds

$$|z^*(y)| = |x^*(x)| \leq \|x^*\|\|x\| \leq \|x^*\|K\|y\|,$$

which proves that z^* is continuous. Let $y^* \in Y^*$ be a Hahn-Banach extension of z^*. For all $x \in X$ we have

$$x^*(x) = z^*(Tx) = y^*(Tx) = (T^*y^*)(x),$$

which proves that $x^* = T^*y^* \in \operatorname{ran} T^*$.

"$(iv) \Rightarrow (iii)$" Follows since the annihilator is closed.

"$(iii) \Rightarrow (i)$" Let $Z := \overline{\operatorname{ran} T}$ and $S \in L(X, Z), Sx := Tx$ for all $x \in X$. Notice that Z is a closed linear subspace of Y, thus a Banach space. For all $y^* \in Y^*$ and all $x \in X$ it holds

$$(T^*y^*)(x) = y^*(Tx) = y^*|_Z(Sx) = (S^*y^*|_Z)(x),$$

thus $T^*y^* = S^*y^*|_Z$, which proves that $\operatorname{ran} T^* \subseteq \operatorname{ran} S^*$. Let now $S^*z^* \in \operatorname{ran} S^*$ for $z^* \in Z^*$ and let $y^* \in Y^*$ be a Hahn-Banach extension of z^*. Then $T^*y^* = S^*z^*$, which proves that $S^*z^* \in \operatorname{ran} T^*$. From here we conclude that $\operatorname{ran} T^* = \operatorname{ran} S^*$, and this is by assumption a closed set.

Since $\operatorname{ran} S$ is dense in Z, by Theorem 9.6 we have that $(\ker S^*)_\perp = Z$. This means that $\ker S^* = \{0\}$, which implies that S^* is injective. Consequently, S^* is a continuous linear bijective operator from Z^* to $\operatorname{ran} S^*$. Corollary 12.5 implies that $(S^*)^{-1}$ is continuous, which means that there exists $c > 0$ with

$$c\|z^*\| \leq \|S^*z^*\| \quad \forall z^* \in Z^*.$$

Lemma 14.2 guarantees that S is surjective, thus $\operatorname{ran} S = Z$ or, equivalently, $\operatorname{ran} T = \overline{\operatorname{ran} T}$, which leads to the desired conclusion. ■

Chapter V

Hilbert spaces

15 Definitions and examples

Definition 15.1 (inner product) Let X be a vector space over $\mathbb{K}$. A mapping $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{K}$ is called inner product (scalar product) if

(a) $\langle x_1 + x_2, y\rangle = \langle x_1, y\rangle + \langle x_2, y\rangle \quad \forall x_1, x_2, y \in X$;

(b) $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle \quad \forall x, y \in X\ \forall \lambda \in \mathbb{K}$;

(c) $\langle x, y\rangle = \overline{\langle y, x\rangle} \quad \forall x, y \in X$;

(d) $\langle x, x\rangle \geq 0 \quad \forall x \in X$;

(e) $\langle x, x\rangle = 0 \Leftrightarrow x = 0$.

The properties (a)-(c) imply

(a') $\langle x, y_1 + y_2\rangle = \langle x, y_1\rangle + \langle x, y_2\rangle \quad \forall x, y_1, y_2 \in X$;

(b') $\langle x, \lambda y\rangle = \overline{\lambda}\langle x, y\rangle \quad \forall x, y \in X\ \forall \lambda \in \mathbb{K}$.

Notice that the mapping $\langle\cdot,\cdot\rangle$ is bilinear for $\mathbb{K} = \mathbb{R}$, and it is sesquilinear for $\mathbb{K} = \mathbb{C}$. A mapping $\langle\cdot,\cdot\rangle$ fulfilling (d) and (e) is called positive definite.

Theorem 15.2 (Cauchy-Schwarz inequality) *Let X be a vector space endowed with a scalar product $\langle\cdot,\cdot\rangle$. It holds*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle \quad \forall x, y \in X.$$

The equality holds if and only if x and y are linearly dependent.

Proof. Take $x, y \in X$. For all $\lambda \in \mathbb{K}$ it holds

$$0 \leq \langle x + \lambda y, x + \lambda y \rangle = \langle x, x \rangle + \lambda \overline{\langle x, y \rangle} + \overline{\lambda} \langle x, y \rangle + |\lambda|^2 \langle y, y \rangle.$$

If $y = 0$, then the inequality holds. Assume that $y \neq 0$ and choose $\lambda := -\frac{\langle x,y \rangle}{\langle y,y \rangle}$. Then

$$0 \leq \left\langle x - \frac{\langle x, y \rangle}{\langle y, y \rangle} y, x - \frac{\langle x, y \rangle}{\langle y, y \rangle} y \right\rangle = \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} + \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle},$$

which gives the desired inequality.

If x and y are linearly dependent, then obviously the equality holds. Viceversa, if $y \neq 0$ and the equality holds, then $x = \frac{\langle x,y \rangle}{\langle y,y \rangle} y$. ■

Defining

$$\|x\| := \sqrt{\langle x, x \rangle} \quad \forall x \in X,$$

we have, according to Theorem 15.2,

$$|\langle x, y \rangle| \leq \|x\| \|y\| \quad \forall x, y \in X.$$

Lemma 15.3 *The mapping $x \mapsto \|x\| := \sqrt{\langle x, x \rangle}$ defines a norm on X.*

Proof. According to (b) and (c) in Definition 15.1, it holds $\|\lambda x\| = |\lambda| \|x\|$ for all $\lambda \in \mathbb{K}$ and all $x \in X$. According to (e) in Definition 15.1 we have that $\|x\| = 0$ if and only if $x = 0$. The triangle inequality is a consequence of the Cauchy-Schwarz inequality:

$$\begin{aligned} \|x + y\|^2 = \langle x + y, x + y \rangle &= \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2 \leq \|x\|^2 + 2 |\langle x, y \rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \quad \forall x, y \in X. \end{aligned}$$

■

Definition 15.4 (pre-Hilbert space) A normed space X is called pre-Hilbert space (inner product space) if there exists an inner product $\langle \cdot, \cdot \rangle$ on X such that $\sqrt{\langle x, x \rangle} = \|x\|$ for all $x \in X$. A complete pre-Hilbert space is called Hilbert space.

The Cauchy-Schwarz inequality guarantees that in a pre-Hilbert space X the mappings $x \mapsto \langle x, y \rangle$, for $y \in X$, and $y \mapsto \langle x, y \rangle$, for $x \in X$, are continuous. If U is a dense linear subspace of a pre-Hilbert space X such that $\langle x, u \rangle = 0$ for all $u \in U$, then $x = 0$.

One can express the inner product of a pre-Hilbert space in terms of the norm as follows:

- for $\mathbb{K} = \mathbb{R}$:

$$\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2) \quad \forall x, y \in X;$$

- for $\mathbb{K} = \mathbb{C}$:

$$\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2 + i \|x + iy\|^2 - i \|x - iy\|^2) \quad \forall x, y \in X.$$

Lemma 15.5 *The inner product* $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{K}$ *of a pre-Hilbert space* X *is a continuous mapping.*

Proof. The statement follows from the following inequality, which holds for all $x_1, x_2, y_1, y_2 \in X$:

$$\begin{aligned} |\langle x_1, y_1 \rangle - \langle x_2, y_2 \rangle| &= |\langle x_1 - x_2, y_1 \rangle + \langle x_2, y_1 - y_2 \rangle| \\ &\le \|x_1 - x_2\| \|y_1\| + \|x_2\| \|y_1 - y_2\|. \end{aligned}$$

■

Theorem 15.6 (parallelogram law) *A normed space* X *is a pre-Hilbert space if and only if*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad \forall x, y \in X.$$

Proof. "$\Rightarrow$" Easy.

"$\Leftarrow$" We consider the case $\mathbb{K} = \mathbb{R}$. Define

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

Notice that $\|x\| = \sqrt{\langle x, x \rangle}$ for all $x \in X$.

We will show that the mapping $\langle \cdot, \cdot \rangle$ fulfills the statements (a)-(e) in Definition 15.1.

(a) Let $x_1, x_2, y \in X$. We have

$$\|x_1 + x_2 + y\|^2 = 2\|x_1 + y\|^2 + 2\|x_2\|^2 - \|x_1 - x_2 + y\|^2$$

$$\|x_1 + x_2 + y\|^2 = 2\|x_2 + y\|^2 + 2\|x_1\|^2 - \|x_2 - x_1 + y\|^2,$$

thus,

$$\begin{aligned} \|x_1 + x_2 + y\|^2 = \|x_1 + y\|^2 + \|x_1\|^2 + \|x_2 + y\|^2 + \|x_2\|^2 \\ - \frac{1}{2}(\|x_1 - x_2 + y\|^2 + \|x_2 - x_1 + y\|^2). \end{aligned}$$

Similarly,

$$\begin{aligned} \|x_1 + x_2 - y\|^2 = \|x_1 - y\|^2 + \|x_1\|^2 + \|x_2 - y\|^2 + \|x_2\|^2 \\ - \frac{1}{2}(\|x_1 - x_2 - y\|^2 + \|x_2 - x_1 - y\|^2). \end{aligned}$$

Consequently,

$$\begin{aligned} \langle x_1 + x_2, y \rangle &= \frac{1}{4}(\|x_1 + x_2 + y\|^2 - \|x_1 + x_2 - y\|^2) \\ &= \frac{1}{4}(\|x_1 + y\|^2 + \|x_2 + y\|^2 - \|x_1 - y\|^2 - \|x_2 - y\|^2) \\ &= \langle x_1, y \rangle + \langle x_2, y \rangle. \end{aligned}$$

(b) Let $x, y \in X$. Statement (a) implies that (b) holds for all $\lambda \in \mathbb{N}$. By definition, (b) holds also for $\lambda = 0$ and $\lambda = -1$, thus for $\lambda \in \mathbb{Z}$. For $\lambda = \frac{m}{n} \in \mathbb{Q}$ we have

$$n\langle \lambda x, y\rangle = n\left\langle m\frac{x}{n}, y\right\rangle = m\langle x, y\rangle = n\lambda\langle x, y\rangle,$$

which shows that (b) holds for all $\lambda \in \mathbb{Q}$. The continuous functions $\lambda \mapsto \langle \lambda x, y\rangle$ and $\lambda \mapsto \lambda\langle x, y\rangle$ from $\mathbb{R}$ to $\mathbb{R}$ are equal on $\mathbb{Q}$, which shows that they are also equal on $\mathbb{R}$.

(c), (d) and (e) are clear.

In case $\mathbb{K} = \mathbb{C}$, one only has to repeat the above arguments for

$$\langle x, y\rangle := \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2) \quad \forall x, y \in X.$$

■

Theorem 15.7 *(a) Every linear subspace of a pre-Hilbert space is a pre-Hilbert space.*

(b) A normed space is a pre-Hilbert space if and only if every two-dimensional linear subspace is a pre-Hilbert space.

(c) The completion (the Banach space which contains the original space as a dense linear subspace) of a pre-Hilbert space is a Hilbert space.

Proof. (a) It follows by restricting the inner product to the linear subspace.

(b) Clear.

(c) Let X be a pre-Hilbert space and $\widehat{X}$ its completion (see Übungsbeispiel 4, Gruppe 1 or Section 8). Using the continuity of the norm, we have that the parallelogram law holds on $\widehat{X}$. This means that $\widehat{X}$ is a complete pre-Hilbert space, thus it is a Hilbert space. ■

Example 15.8 (a) $\mathbb{C}^n$ endowed with the scalar product

$$\langle (t_1, ..., t_n), (s_1, ..., s_n)\rangle := \sum_{i=1}^{n} t_i \overline{s_i},$$

for $(t_1, ..., t_n), (s_1, ..., s_n) \in \mathbb{C}^n$, is a Hilbert space.

(b) ℓ^2 endowed with the scalar product

$$\langle (t_k)_{k\in\mathbb{N}}, (s_k)_{k\in\mathbb{N}}\rangle := \sum_{k=1}^{+\infty} t_k \overline{s_k},$$

for $(t_k)_{k\in\mathbb{N}}, (s_k)_{k\in\mathbb{N}} \in \ell^2$, is a Hilbert space. The Hölder inequality guarantees that the inner product is well-defined.

(c) $L^2(I)$, for $I \subseteq \mathbb{R}$ an interval, is a Hilbert space. The inner product is defined by

$$\langle f, g\rangle = \int_I f\overline{g}d\lambda,$$

for $f, g \in L^2(I)$. $L^2(\Omega, \Sigma, \mu)$, for (Ω, Σ, μ) a measure space, is a Hilbert space.

(d) Let I be an index set and

$$\ell^2(I) := \left\{ f : I \to \mathbb{K} \mid f(s) \neq 0 \text{ for at most countable many } s, \sum_{s\in I} |f(s)|^2 < +\infty \right\}.$$

The sum over all $s \in I$ is to be understand as follows: let $\{s_1, s_2, ...\}$ be an enumeration of the set $\{s \mid f(s) \neq 0\}$. We set

$$\sum_{s\in I} |f(s)|^2 := \sum_{i=1}^{+\infty} |f(s_i)|^2$$

and notice that, due to the absolute convergence of the series, the summation order does not affect its convergence. Obviously, $\ell^2(\mathbb{N}) = \ell^2$.

For $f, g \in \ell^2(I)$,

$$\langle f, g\rangle = \sum_{s\in I} f(s)\overline{g(s)},$$

defines an inner product with induced norm

$$\|f\| = \left(\sum_{s\in I} |f(s)|^2\right)^{\frac{1}{2}}.$$

In order to see that $\ell^2(I)$ is complete one has to repeat the arguments used in Example 1.10.

(e) For $\lambda \in \mathbb{R}$, define $f_\lambda : \mathbb{R} \to \mathbb{C}$, $f_\lambda(t) = e^{i\lambda t}$. Define $X := \mathrm{lin}\{f_\lambda \mid \lambda \in \mathbb{R}\}$. We define on X the scalar product

$$\langle f, g\rangle = \lim_{T\to+\infty} \frac{1}{2T} \int_{-T}^{T} f(s)\overline{g(s)}ds,$$

for $f, g \in X$. The completion of X with respect to the norm $\|\cdot\| := \sqrt{\langle\cdot,\cdot\rangle}$ is a Hilbert space, which we denote by $AP^2(\mathbb{R})$ ("AP" stands for almost periodic).

Since

$$\|f_\lambda - f_\mu\| = \sqrt{2} \quad \forall \lambda \neq \mu,$$

one can show in a similar way as we did it for ℓ^∞ the $AP^2(\mathbb{R})$ is not separable.

16 Orthogonality

Definition 16.1 Let X be a pre-Hilbert space. Two vectors $x, y \in X$ are called orthogonal if $\langle x, y\rangle = 0$. In this case we write $x \perp y$. Two sets $A, B \subseteq X$ are called orthogonal (we write $A \perp B$) if $x \perp y$ for all $x \in A$ and all $y \in B$. The set $A^\perp := \{y \in X \mid x \perp y \; \forall x \in A\}$ is called the orthogonal complement of the set A.

We will see later that the notation $A^\perp$ is consistent with the one we used for the anihilator of a linear subspace. It is clear that $A^\perp$ is a closed linear subspace, $A \subseteq (A^\perp)^\perp$ and $A^\perp = (\overline{\operatorname{lin} A})^\perp$. Furthermore, we have the Pythagoras Theorem

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

Theorem 16.2 (projection operator) *Let H be a Hilbert space, $K \subseteq H$ a nonempty convex and closed set and $x_0 \in H$. Then there exists a unique element $x \in K$ such that*

$$\|x_0 - x\| = \inf_{k\in K} \|x_0 - k\|.$$

This defines a mapping $P_K : H \to K, P_K(x_0) := x$, which is the so-called projection operator onto the set K.

Proof. Let $d := \inf_{k\in K} \|x_0 - k\| \geq 0$. Then there exists a sequence $(k_n)_{n\in\mathbb{N}}$ such that $\|x_0 - k_n\| \to d$ as $n \to +\infty$. According to the parallelogram law, we have for all $m, n \in \mathbb{N}$

$$\begin{aligned} 0 \leq \left\|\frac{k_n - k_m}{2}\right\|^2 &= \frac{1}{2}(\|x_0 - k_n\|^2 + \|x_0 - k_m\|^2) - \left\|x_0 - \frac{k_n + k_m}{2}\right\|^2 \\ &\leq \frac{1}{2}(\|x_0 - k_n\|^2 + \|x_0 - k_m\|^2) - \|d\|^2, \end{aligned}$$

where we used that $\frac{k_n+k_m}{2} \in K$. This implies that $\lim_{m,n\to+\infty} \|k_n - k_m\| = 0$, therefore, $(k_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Since H is complete and K closed, there exists $x \in K$ such that $k_n \to x$ as $n \to +\infty$. The continuity of the norm implies that $\|x_0 - x\| = d$.

Assuming that there exist $x' \in K, x \neq x'$, with $\|x_0 - x\| = \|x_0 - x'\| = d$, by using again the parallelogram law, we get

$$\left\|x_0 - \frac{x + x'}{2}\right\|^2 = \frac{1}{2}(d^2 + d^2) - \left\|\frac{x' - x''}{2}\right\|^2 < d^2.$$

Contradiction. This proves the uniqueness of x. ■

Lemma 16.3 *Let H be a Hilbert space, $K \subseteq H$ a nonempty, convex and closed set, $x_0 \in H$ and $x \in K$. The following statements are equivalent:*

(i) $x = P_K(x_0)$;

(ii) $\operatorname{Re}\langle x_0 - x, k - x\rangle \leq 0 \quad \forall k \in K$.

Proof. "$(ii) \Rightarrow (i)$" For every $k \in K$ it holds

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

"$(i) \Rightarrow (ii)$" Let $k \in K$ and $k_t := (1-t)x + tk \in K$, for $t \in [0,1]$. According to (i), for all $t \in [0,1]$ it holds

$$\begin{aligned}\|x_0 - x\|^2 \leq \|x_0 - k_t\|^2 &= \langle x_0 - x + t(x-k), x_0 - x + t(x-k)\rangle\\ &= \|x_0 - x\|^2 + 2\operatorname{Re}\langle x_0 - x, t(x-k)\rangle + t^2\|x-k\|^2.\end{aligned}$$

From here we have

$$\operatorname{Re}\langle x_0 - x, k - x\rangle \leq \frac{t}{2}\|x-k\|^2 \quad \forall t \in (0,1],$$

which provides (by letting $t \to 0$) (ii). ■

Theorem 16.4 (orthogonal projection) *Let $M \neq \{0\}$ be a closed linear subspace of a Hilbert space H. The projection operator $P_M : H \to M$ onto the set M is linear and continuous, and it fulfills $\|P_M\| = 1$, $\ker P_M = M^\perp$ and $\operatorname{Id} - P_M = P_{M^\perp}$. It is also called the orthogonal projection onto M. It holds $H = M \oplus_2 M^\perp$.*

Proof. Let $x_0 \in H$ and $x \in M$. According to Lemma 16.3, $x = P_M(x_0)$ if and only if

$$\operatorname{Re}\langle x_0 - P_M(x_0), m - P_M(x_0)\rangle \leq 0 \quad \forall m \in M.$$

This is equivalent to (since $m \in M \Leftrightarrow m - P_M(x_0) \in M$)

$$\operatorname{Re}\langle x_0 - P_M(x_0), m\rangle \leq 0 \quad \forall m \in M$$

and further to (since $y \in M \Leftrightarrow -y \in M \Leftrightarrow iy \in M$)

$$\langle x_0 - P_M(x_0), m\rangle = 0 \quad \forall m \in M.$$

This shows that $P_M(x_0)$ is the uniquely determined element $x \in M$ which fulfills

$$x_0 - x \in M^\perp.$$

Let $x_1, x_2 \in H$ and $\lambda_1, \lambda_2 \in \mathbb{K}$. Since $M^\perp$ is a linear subspace, it holds

$$\lambda_1 x_1 + \lambda_2 x_2 - (\lambda_1 P_M(x_1) + \lambda_2 P_M(x_2)) \in M^\perp.$$

Thus $P_M(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 P_M(x_1) + \lambda_2 P_M(x_2)$, which means that P_M is linear.

In the light of the above characterization we have that $P_M(x_0) = 0$ if and only if $x_0 \in M^\perp$, which shows that $\ker P_M = M^\perp$.

For $x_0 \in H$, we have seen that $x_0 - P_M(x_0) \in M^\perp$. In addition, for all $m \in M^\perp$ it holds (according to Lemma 16.3)

$$\begin{aligned}\mathrm{Re}\langle x_0 - (x_0 - P_M(x_0)), m - (x_0 - P_M(x_0))\rangle &= \mathrm{Re}\langle P_M(x_0), m - x_0 + P_M(x_0)\rangle \\ &= \mathrm{Re}\langle -P_M(x_0), x_0 - P_M(x_0)\rangle \leq 0.\end{aligned}$$

By using again Lemma 16.3 it follows that $x_0 - P_M(x_0) = P_{M^\perp}(x_0)$.

Since, for every $x_0 \in H$, $P_M(x_0) \perp (x_0 - P_M(x_0))$, the Pythagoras Theorem leads to

$$\|P_M(x_0)\|^2 + \|x_0 - P_M(x_0)\|^2 = \|x_0\|^2,$$

which shows that $\|P_M\| \leq 1$. On the other hand, since $P_M = P_M^2$, it yields $\|P_M\| = \|P_M^2\| \leq \|P_M\|^2$, thus $1 \leq \|P_M\|$. This proves that $\|P_M\| = 1$.

Pythagoras Theorem proves that $H = M \oplus_2 M^\perp$. ■

Corollary 16.5 *Let M be a linear subspace of a Hilbert space H. It holds $\overline{M} = M^{\perp\perp}$.*

Proof. Let $N := \overline{M}$, which is a closed linear subspace. Then $P_N = \mathrm{Id} - P_{N^\perp} = P_{(N^\perp)^\perp}$, which proves that $\overline{M} = N = (N^\perp)^\perp = \overline{M}^{\perp\perp} = M^{\perp\perp}$. ■

Theorem 16.6 (Fréchet-Riesz Representation Theorem) *Let H be a Hilbert space. The mapping*

$$\Phi : H \to H^*, \quad y \mapsto \langle \cdot, y\rangle,$$

is bijective, isometric and conjugate linear. This means that for every element $x^* \in H^*$ there exists an unique element $u \in H$ such that $x^*(x) = \langle x, u\rangle$ for all $x \in H$, and $\|x^*\| = \|u\|$.

Proof. It is easy to see that $\Phi(y) \in H^*$ for all $y \in H$. Obviously, $\Phi(\lambda_1 y_1 + \lambda_2 y_2) = \overline{\lambda}_1\Phi(y_1) + \overline{\lambda}_2\Phi(y_2)$ for all $\lambda_1, \lambda_2 \in \mathbb{K}$ and $y_1, y_2 \in H$. According to the Cauchy-Schwarz inequality we have $\|\Phi(y)\| \leq \|y\|$ for all $y \in H$. For $y \in H, y \neq 0$, it holds $\Phi(y)(y/\|y\|) = \langle y, y\rangle/\|y\| = \|y\|$, thus $\|\Phi(y)\| = \|y\|$, which shows that Φ is a isometry. Consequently, Φ is injective.

We will show that Φ is surjective. Let $x^* \in H^*, x^* \neq 0$. Then $M := \ker x^*$ is a proper closed linear subspace of H. Let $y \in M^\perp, y \neq 0$ (see Theorem 16.4) be such that $x^*(y) = 1$. Then $H = M \oplus_2 \mathrm{lin}\{y\}$ (which actually shows that $M^\perp = \mathrm{lin}\{y\}$). Indeed, for $x \in H$ and $\lambda := x^*(x)$ we have $x - \lambda y \in M$. Every $x \in H$ can be represented as $x = m + \lambda y$, for $m \in M$ and $\lambda \in \mathbb{K}$. Then $x^*(x) = \lambda = \langle x, y\rangle/\|y\|^2$, thus $\Phi(y/\|y\|^2) = x^*$. This shows that Φ is surjective.

If for $u \in H$, $x^*(x) = \langle x, u\rangle = \Phi(u)(x)$ for all $x \in H$, we obviously have that $\|x^*\| = \|\Phi(u)\| = \|u\|$. ■

Corollary 16.7 *Let H be a Hilbert space. The following statements are true:*

(a) A sequence $(x_n)_{n\in\mathbb{N}} \subseteq H$ converges weakly to $x \in H$ if and only of

$$\langle x_n - x, y\rangle \to 0 \ (n \to +\infty) \quad \forall y \in H.$$

(b) H is reflexive.

(c) Every bounded sequence in H has a weakly convergent subsequence.

Proof. (a) Follows from the Fréchet-Riesz Representation Theorem.

(b) Let $\Phi : H \to H^*$ be the mapping in the Fréchet-Riesz Representation Theorem. Then

$$\langle \Phi(x), \Phi(y)\rangle_{H^*} := \langle y, x\rangle_H$$

defines an inner product on H^*, which is a Hilbert space. Let $\Psi : H^* \to H^{**}$ be the mapping in the Fréchet-Riesz Representation Theorem, this time from H^* to its bidual. It is easy to see that $\Psi \circ \Phi = i_H : H \to H^{**}$, which shows that the canonical embedding is surjective.

(c) Follows from (b) and Theorem 8.8. ■

The following results strenghtens statement (c) of the above corollary.

Theorem 16.8 (Banach-Saks Theorem) *Let H be a Hilbert space and $(x_n)_{n\in\mathbb{N}} \subseteq H$ a bounded sequence. Then it has a weakly convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that the sequence of arithmetic means*

$$\left(\frac{1}{n}\sum_{k=1}^{n} x_{n_k}\right)_{n\in\mathbb{N}}$$

converges in norm.

Proof. Without loss of generality we can that assume that $(x_n)_{n\in\mathbb{N}}$ is weakly convergent and that its weak limit is equal to 0. Let $M \geq 0$ such that $\|x_n\| \leq M$ for all $n \in \mathbb{N}$. Choose $x_{n_1} := x_1$. Since $\langle x_n, x_{n_1}\rangle$ converges to 0 as $n \to +\infty$, there exists $n_2 > 1$ such that

$$|\langle x_{n_2}, x_{n_1}\rangle| \leq 1.$$

Further, there exists $n_3 > n_2$ such that

$$|\langle x_{n_3}, x_{n_1}\rangle| \leq \frac{1}{2} \quad \text{and} \quad |\langle x_{n_3}, x_{n_2}\rangle| \leq \frac{1}{2}.$$

In this way we can inductively construct a sequence $(x_{n_k})_{k\in\mathbb{N}}$ such that

$$|\langle x_{n_{k+1}}, x_{n_i}\rangle| \leq \frac{1}{k} \quad \forall i = 1, ..., k.$$

This implies that $\sum_{i=1}^{k} |\langle x_{n_{k+1}}, x_{n_i}\rangle| \leq 1$ for all $k \in \mathbb{N}$ and $\sum_{k=2}^{n}\sum_{i=1}^{k-1} |\langle x_{n_k}, x_{n_i}\rangle| \leq n-1$ for all $n \in \mathbb{N}$. From here we have

$$\left\|\frac{1}{n}\sum_{k=1}^{n} x_{n_k}\right\|^2 \leq \frac{1}{n^2}\sum_{k=1}^{n}\sum_{i=1}^{n} |\langle x_{n_k}, x_{n_i}\rangle| = \frac{1}{n^2}\left(\sum_{k=1}^{n}\|x_{n_k}\|^2 + 2\sum_{k=2}^{n}\sum_{i=1}^{k-1}|\langle x_{n_k}, x_{n_i}\rangle|\right)$$
$$\leq \frac{nM^2+2n-2}{n^2} \longrightarrow 0 \quad (n \to +\infty).$$

■

17 Orthonormal bases

Definition 17.1 A subset S of a Hilbert space H is called orthonormal system if $\|e\| = 1$ for all $e \in S$ and $\langle e, f\rangle = 0$ for all $e, f \in S, e \neq f$. An orthonormal system S is called orthonormal basis if

$$S \subseteq T,\ T \text{ is an orthonormal system } \Rightarrow S = T.$$

Example 17.2 (a) For $H = \ell^2$, $S = \{e_n \mid n \in \mathbb{N}\}$ is an orthonormal system.

(b) For $H = L^2[0, 2\pi]$,

$$S = \left\{\frac{1}{\sqrt{2\pi}}\mathbb{1}\right\} \bigcup \left\{\frac{1}{\sqrt{\pi}}\cos(n\cdot) \mid n \in \mathbb{N}\right\} \bigcup \left\{\frac{1}{\sqrt{\pi}}\sin(n\cdot) \mid n \in \mathbb{N}\right\}$$

is an orthonormal system.

(c) For $H = L^2([0, 2\pi]; \mathbb{C})$ (which denotes the normed space $L^2[0, 2\pi]$ over $\mathbb{C}$), $S = \left\{\frac{1}{\sqrt{2\pi}}e^{i(n\cdot)} \mid n \in \mathbb{Z}\right\}$ is an orthonormal system.

(d) For $H = AP^2(\mathbb{R})$, $S = \{f_\lambda \mid \lambda \in \mathbb{R}\}$ is an orthonormal system.

Later we will show that the orthonormal systems in the above example are in fact orthonormal bases.

Theorem 17.3 (Gram-Schmidt Algorithm) *Let $(x_n)_{n\in\mathbb{N}}$ be a linearly independent subset of a Hilbert space H. Then there exists an orthonormal system S with $\operatorname{lin} S = \operatorname{lin}\{x_n \mid n \in \mathbb{N}\}$.*

Proof. Set $e_1 := \frac{x_1}{\|x_1\|}$. Define $f_2 := x_2 - \langle x_2, e_1\rangle e_1$. Since $\{x_1, x_2\}$ is linearly independent, $f_2 \neq 0$. Set $e_2 := \frac{f_2}{\|f_2\|}$. Then $e_1 \perp e_2$.

We define the sequence $(e_n)_{n\in\mathbb{N}}$ by means of the following iterative scheme

$$f_{n+1} := x_{n+1} - \sum_{i=1}^{n}\langle x_{n+1}, e_i\rangle e_i \neq 0 \text{ and } e_{n+1} := \frac{f_{n+1}}{\|f_{n+1}\|} \quad \forall n \geq 2.$$

Then $S := \{e_1, e_2, ...\}$ is an orthonormal system and for all $n \in \mathbb{N}$ it holds $x_n \in \operatorname{lin} S$ and $e_n \in \operatorname{lin}\{x_1, x_2, ..., x_n\}$. Consequently, $\operatorname{lin} S = \operatorname{lin}\{x_n \mid n \in \mathbb{N}\}$. ■

Example 17.4 For $H = L^2[-1,1]$ and the linearly independent set $\{x_n \mid n \geq 0\}$, where $x_n(t) = t^n, n \geq 0$, the Gram-Schmidt Algorithm furnishes the orthonormal system

$$e_n(t) = \sqrt{n + \frac{1}{2}} P_n(t), \quad n \geq 0,$$

where $P_n(t) = \frac{1}{2^n n!} \left(\frac{d}{dt}\right)^n (t^2 - 1)^n$ denotes the Legendre polynomial of order n.

Theorem 17.5 (Bessel's inequality) *Let $\{e_n \mid n \in \mathbb{N}\}$ be an orthonormal system of a Hilbert space H and $x \in H$. It holds*

$$\sum_{n=1}^{+\infty} |\langle x, e_n \rangle|^2 \leq \|x\|^2.$$

Proof. Let $N \in \mathbb{N}$ be fixed. Set $x_N := x - \sum_{n=1}^{N} \langle x, e_n \rangle e_n$. Then $x_N \perp e_n$ for all $n = 1, ..., N$. Pythagoras Theorem leads to

$$\|x\|^2 = \|x_N\|^2 + \left\| \sum_{n=1}^{N} \langle x, e_n \rangle e_n \right\|^2 = \|x_N\|^2 + \sum_{n=1}^{N} |\langle x, e_n \rangle|^2 \geq \sum_{n=1}^{N} |\langle x, e_n \rangle|^2.$$

The statement follows by letting $N \to +\infty$. ■

The following corollary follows by combining Theorem 17.5 with the Hölder inequality.

Corollary 17.6 *Let $\{e_n \mid n \in \mathbb{N}\}$ be an orthonormal system of a Hilbert space H and $x, y \in H$. It holds*

$$\sum_{n=1}^{+\infty} |\langle x, e_n \rangle \langle e_n, y \rangle| < +\infty.$$

Lemma 17.7 *Let S be an orthonormal system of a Hilbert space H and $x \in H$. The set $S_x := \{e \in S \mid \langle x, e \rangle \neq 0\}$ is countable.*

Proof. According to Bessel's inequality, the set $S_{x,n} := \{e \in S \mid |\langle x, e \rangle| \geq \frac{1}{n}\}$ is finite for all $n \in \mathbb{N}$. It holds $S_x = \cup_{n \in \mathbb{N}} S_{x,n}$. ■

The following notion is needed in order to handle non-separable Hilbert spaces.

Definition 17.8 (unconditional convergence) Let X be a normed space, I an infinite index set and $x_i \in X$ for all $i \in I$. The series $\sum_{i \in I} x_i$ is called unconditionally convergent to $x \in X$ if

(a) $I_0 = \{i \in I \mid x_i \neq 0\}$ is countable;

(b) For every enumeration $\{i_1, i_2, ...\}$ of I_0 it holds $\sum_{n=i}^{+\infty} x_{i_n} = x$.

Since $\sum_{n=1}^{+\infty} x_{i_n}$ does not depend on the order of the summands, we will write $\sum_{i\in I} x_i = x$. Even if $I = \mathbb{N}$, we will differentiate between $\sum_{n\in\mathbb{N}}$ and $\sum_{n=1}^{+\infty}$. If X is a Banach space, then every absolutely convergent series is unconditionally convergent. For $X = \mathbb{K}^n$, according to the Riemann Rearrangement Theorem, a series is unconditionally convergent if and only if it is absolutely convergent. On the other hand, in infinite-dimensional Banach spaces there exist unconditionally convergent series that are not absolutely convergent (Theorem of Dvoretzky-Rogers).

Bessel's inequality and Lemma 17.7 lead to the following result.

Corollary 17.9 (Bessel's inequality for orthonormal systems) *If S is an orthonormal system of a Hilbert space H and $x \in H$, then*

$$\sum_{e\in S} |\langle x, e\rangle|^2 \leq \|x\|^2.$$

Theorem 17.10 *Let S be an orthonormal system of a Hilbert space H.*

(a) For every $x \in H$, the series $\sum_{e\in S}\langle x, e\rangle e$ is unconditionally convergent.

(b) The mapping $P : x \mapsto \sum_{e\in S}\langle x, e\rangle e$ is the orthogonal projection onto $\overline{\operatorname{lin} S}$.

Proof. (a) Let $\{e_1, e_2, ...\}$ be an enumeration of $\{e \in S \mid \langle x, e\rangle \neq 0\}$. Let $N, M \in \mathbb{N}, M > N$. According to the Pythagoras Theorem we have

$$\left\|\sum_{n=N}^{M}\langle x, e_n\rangle e_n\right\|^2 = \sum_{n=N}^{M} \|\langle x, e_n\rangle\|^2 \to 0 \quad (N, M \to +\infty).$$

This shows that $\sum_{n=1}^{+\infty}\langle x, e_n\rangle e_n$ is a Cauchy series, thus there exists $y \in H$ such that $y = \sum_{n=1}^{+\infty}\langle x, e_n\rangle e_n$. Let $\pi : \mathbb{N} \to \mathbb{N}$ be a permutation and $y_\pi \in H$ such that $y_\pi = \sum_{n=1}^{+\infty}\langle x, e_{\pi(n)}\rangle e_{\pi(n)}$. We will prove that $y = y_\pi$.

For all $z \in H$ it holds (we use Corollary 17.6 and the fact that absolute convergence implies unconditional convergence)

$$\langle y, z\rangle = \sum_{n=1}^{+\infty}\langle x, e_n\rangle\langle e_n, z\rangle = \sum_{n=1}^{+\infty}\langle x, e_{\pi(n)}\rangle\langle e_{\pi(n)}, z\rangle = \langle y_\pi, z\rangle.$$

Consequently, $y - y_\pi \in H^\perp = \{0\}$.

(b) Let $x \in H$ and $\{e_1, e_2, ...\}$ be an enumeration of $\{e \in S \mid \langle x, e\rangle \neq 0\}$. According to Theorem 16.4 it suffices to prove that $x - Px \in (\overline{\operatorname{lin} S})^\perp = S^\perp$ or, in other words, that

$$\left\langle x - \sum_{n=1}^{+\infty}\langle x, e_n\rangle e_n, e\right\rangle = 0 \quad \forall e \in S.$$

This is clear, both when $\langle x, e\rangle = 0$ and when e is an element of the enumeration $\{e_1, e_2, ...\}$. ■

The following theorem provides a helpful characterization of the orthonormal bases.

Theorem 17.11 *If S is an orthonormal system of a Hilbert space H, then there exists an orthonormal base S' such that $S \subseteq S'$. In addition, the following statements are equivalent:*

(i) *S is an orthonormal base.*

(ii) *If $x \in H$ and $x \perp S$, then $x = 0$.*

(iii) *It holds $H = \overline{\operatorname{lin} S}$.*

(iv) *$x = \sum_{e \in S} \langle x, e\rangle e$ for all $x \in H$.*

(v) *$\langle x, y\rangle = \sum_{e \in S} \langle x, e\rangle\langle e, y\rangle$ for all $x, y \in H$.*

(vi) *(Parseval's identity)*

$$\|x\|^2 = \sum_{e \in S} |\langle x, e\rangle|^2 \quad \forall x \in H.$$

Proof. The existence is a consequence of the Zorn Lemma applied to the set $A := \{S' \mid S \subseteq S', S' \text{ is an orthonormal system in } H\}$ and the partial order $S_1 \leq S_2 \Leftrightarrow S_1 \subseteq S_2$.

"$(i) \Rightarrow (ii)$" If $x \neq 0$, then $S \cup \{\frac{x}{\|x\|}\}$ is an orthonormal system.

"$(ii) \Rightarrow (iii)$" Follows from Corollary 16.5.

"$(iii) \Rightarrow (iv)$" Follows from Theorem 17.10(b).

"$(iv) \Rightarrow (v)$" Follows by direct verification (see also Corollary 17.6).

"$(v) \Rightarrow (vi)$" Just set $x = y$.

"$(vi) \Rightarrow (i)$" If S is not an orthonormal base, then there exists $x \in H$, $\|x\| = 1$, such that $S \cup \{x\}$ is an orthonormal system. This gives $\sum_{e \in S} |\langle x, e\rangle|^2 = 0$. Contradiction. ■

Example 17.12 (a) For $H = \ell^2$, $S = \{e_n \mid n \in \mathbb{N}\}$ is an orthonormal basis, since $\ell^2 = \overline{\operatorname{lin} S}$.

(b) For $H = L^2[0, 2\pi]$, the orthonormal system S in Example 17.2(b) is an orthonormal basis, since $L^2[0, 2\pi] = \overline{\operatorname{lin} S}$. We have that $\operatorname{lin} S$ is dense in $M = \{f \in C[0, 2\pi] \mid f(0) = f(2\pi)\}$ with respect to $\|\cdot\|_\infty$ and, consequently, with respect to $\|\cdot\|_{L^2}$ and that M is dense in $L^2[0, 2\pi]$.

(c) For $H = L^2([0,2\pi];\mathbb{C})$, the orthonormal system S in Example 17.2(c) is an orthonormal basis. Indeed, $e^{int} = \cos nt + i \sin nt$, thus $\operatorname{lin} S = \{f \mid f \text{ is a complex-valued trigonometric polynom}\}$. It follows from (b) that $\operatorname{lin} S$ is dense in H.

(d) For $H = AP^2(\mathbb{R})$, the orthonormal system S in Example 17.2(d) is an orthonormal basis, since, $H = \overline{\operatorname{lin} S}$.

(e) For $H = L^2[-1,1]$, the orthonormal system $\{e_n \mid n \in \mathbb{N}\}$ in Example 17.4 is an orthonormal basis, since (see Example 2.12 (a), (b)) $L^2[-1,1] = \overline{\operatorname{lin}\{x_n \mid n \in \mathbb{N}\}} = \overline{\operatorname{lin}\{e_n \mid n \in \mathbb{N}\}}$.

Corollary 17.13 *Let H be an infinite-dimensional Hilbert space. The following statements are equivalent:*

(i) H is separable.

(ii) All orthonormal bases of H are countable.

(iii) H has a countable orthonormal basis.

Proof. "$(i) \Rightarrow (ii)$" Let S be an orthonormal basis of H. Since $\|e - f\| = \sqrt{2}$ for all $e, f \in S, e \neq f$, we have that S cannot be uncountable (see the proof of the statement that ℓ^∞ is not separable in Example 2.11(c)).

"$(ii) \Rightarrow (iii)$" Clear (see also Theorem 17.11).

"$(iii) \Rightarrow (i)$" According to Theorem 17.11 we have that $H = \overline{\operatorname{lin} S}$, where S is a countable orthonormal basis. Lemma 2.10 guarantees that H is separable. ∎

Lemma 17.14 *If S and T are orthonormal bases of the Hilbert space H, then $|S| = |T|$.*

Proof. For S finite, this is known from the linear algebra. Let $|S| \geq |\mathbb{N}|$. For $x \in S$, let $T_x := \{y \in T \mid \langle x, y\rangle \neq 0\}$. According to Lemma 17.7 we have that $|T_x| \leq |\mathbb{N}|$. According to Theorem 17.11 ($(i) \Leftrightarrow (ii)$), we have that $T \subseteq \cup_{x\in S} T_x$, thus $|T| \leq |S||\mathbb{N}| = |S|$. Similary, $|S| \leq |T|$, which implies that $|S| = |T|$ (Schröder-Bernstein Theorem). ∎

The cardinal number of a orthonormal basis of a Hilbert space is called the dimension of the Hilbert space.

Theorem 17.15 *If S is an orthonormal basis of a Hilbert space H, then $H \cong \ell^2(S)$.*

Proof. Let $x \in H$ and define $Tx : S \to \mathbb{K}, (Tx)(e) = \langle x, e\rangle$. The Bessel inequality guarantees that $Tx \in \ell^2(S)$. The operator $T : H \to \ell^2(S), x \mapsto Tx$, is linear and, due to Parseval's identity, isometric. Let $(f_e)_{e\in S} \in \ell^2(S)$. As in the proof of Theorem 17.10, one can see that $x := \sum_{e\in S} f_e e$ is an element of H and, obviously, $Tx = (f_e)_{e\in S}$. This shows that T is an isometric isomorphism. ∎

Corollary 17.16 *If H is a separable infinite-dimensional Hilbert space, then $H \cong \ell^2$.*

Theorem 17.17 (Fischer-Riesz Theorem)

$$L^2[0,1] \cong \ell^2.$$

18 Operators on Hilbert spaces

Definition 18.1 Let H and G be two Hilbert spaces and $T \in L(H,G)$. The adjoint operator (in Hilbert space sense) of T is defined as

$$T^* : G \to H, \quad \langle x, T^*y\rangle_H = \langle Tx, y\rangle_G \quad \forall x \in H\ \forall y \in G.$$

If $\Phi_H : H \to H^*$ and $\Phi_G : G \to G^*$, are the conjugate linear isometric isomorphism introduced in the Fréchet-Riesz Representation Theorem and $\widetilde{T}^* \in L(G^*, H^*)$ is the "adjoint operator" of T introduced in Section 9, then $T^* = \Phi_H^{-1} \circ \widetilde{T}^* \circ \Phi_G$.

Theorem 18.2 *Let H, G and K be Hilbert spaces, $S, T \in L(H,G)$, $R \in L(G,K)$, and $\lambda \in \mathbb{K}$. It holds:*

(a) $(S+T)^* = S^* + T^*$.

(b) $(\lambda S)^* = \overline{\lambda} S^*$.

(c) $(RS)^* = S^* R^*$.

(d) $S^* \in L(G,H)$ *and* $\|S\| = \|S^*\|$.

(e) $S^{**} = S$.

(f) $\|SS^*\| = \|S^*S\| = \|S\|^2$.

(g) $\ker S = (\operatorname{ran} S^*)^\perp$, $\ker S^* = (\operatorname{ran} S)^\perp$, *thus, S is injective if and only if* $\operatorname{ran} S^*$ *is dense in H.*

The mapping $S \mapsto S^*$ is a conjugate linear isometric isomorphism between $L(H,G)$ and $L(G,H)$.

Proof. (a)-(e) The statements follow directly from the definition.

(f) For all $x \in H$ it holds

$$\|Sx\|^2 = \langle Sx, Sx\rangle = \langle x, S^*Sx\rangle \le \|x\| \|S^*Sx\|,$$

thus

$$\|S\|^2 = \sup_{\|x\|\le 1} \|Sx\|^2 \le \sup_{\|x\|\le 1} \|x\|\|S^*Sx\| \le \|S^*S\| \le \|S^*\|\|S\| = \|S\|^2.$$

We get $\|S\|^2 = \|S^*S\|$ and, from here,

$$\|S\|^2 = \|S^*\|^2 = \|S^{**}S^*\| = \|SS^*\|.$$

(g) We have

$$Sx = 0 \Leftrightarrow \langle Sx, y\rangle = 0 \ \forall y \in G \Leftrightarrow \langle x, S^*y\rangle = 0 \ \forall y \in G \Leftrightarrow x \in (\operatorname{ran} S^*)^\perp,$$

thus $\ker S = (\operatorname{ran} S^*)^\perp$ and, from here, $\ker S^* = (\operatorname{ran} S^{**})^\perp = (\operatorname{ran} S)^\perp$. ■

The following definition introduces some important classes of operators defined on Hilbert spaces.

Definition 18.3 Let $T \in L(H, G)$.

(a) T is called unitary if $T^*T = \mathrm{Id}_H$ and $TT^* = \mathrm{Id}_G$, in other words, if T is invertible with $T^{-1} = T^*$.

(b) Let $H = G$. T is called self-adjoint if $T = T^*$.

(c) Let $H = G$. T is called normal if $TT^* = T^*T$.

It is easy to see that T is unitary if and only if it is surjective and $\langle Tx, Ty\rangle = \langle x, y\rangle$ for all $x, y \in H$. Let $H = G$. Then T is self-adjoint if and only if $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in H$. Furthermore, T is normal if and only if $\langle Tx, Ty\rangle = \langle T^*x, T^*y\rangle$ for all $x, y \in H$. Consequently, for a normal operator we have $\|Tx\| = \|T^*x\|$ for all $x \in H$, which implies that $\ker T = \ker T^*$.

Example 18.4 (a) Let $H = \mathbb{K}^n$. If $T \in L(H)$ is represented by a matrix $(a_{ij})_{i,j}$, then T^* is represented by its complex-conjugate transpose $(\overline{a_{ji}})_{i,j}$.

(b) Let $k \in L^2([0,1]^2)$. The operator

$$T_k : L^2[0,1] \to L^2[0,1], \quad (T_k x)(s) = \int_0^1 k(s,t)x(t)dt,$$

is linear and continuous with $\|T_k\| \le \|k\|_{L^2}$. Then $T_k^* = T_{\overline{k}}$, where $\overline{k}(s,t) = \overline{k(s,t)}$ for all $(s,t) \in [0,1]^2$. T_k is self-adjoint if and only if $k(s,t) = \overline{k(s,t)}$ almost everywhere; in this case k is called symmetric kernel.

(c) The left shift operator $T : \ell^2 \to \ell^2, T(s_1, s_2, ...) = (s_2, s_3, ...)$, is not normal (see Example 9.2).

(d) T^*T and TT^* are always self-adjoint.

Lemma 18.5 *Let H and G be Hilbert spaces and $T \in L(H,G)$. Then T is an isometry if and only if*

$$\langle Tx, Ty \rangle = \langle x, y \rangle \quad \forall x, y \in H.$$

In the last part of this section we will provide some characterizations of self-adjoint operators.

Theorem 18.6 (Hellinger-Toeplitz Theorem) *Let H be a Hilbert space and $T : H \to H$ a linear mapping which fulfills the symmetry condition*

$$\langle Tx, y \rangle = \langle x, Ty \rangle \quad \forall x, y \in H.$$

Then T is continuous and, consequentely, self-adjoint.

Proof. We will show that T is graphclosed and the Closed Graph Theorem will imply that T is continuous, consequently, self-adjoint.

Let $(x_n)_{n\in\mathbb{N}} \subseteq H$ such that $x_n \to 0$ and $Tx_n \to y \in H$ as $n \to +\infty$. From

$$\|y^2\| = \langle \lim_{n\to+\infty} Tx_n, y \rangle = \lim_{n\to+\infty} \langle Tx_n, y \rangle = \lim_{n\to+\infty} \langle x_n, Ty \rangle = 0$$

we have $y = 0$, and the conclusion follows. ■

Theorem 18.7 *Let H be a Hilbert space over $\mathbb{K} = \mathbb{C}$ and $T \in L(H)$. The following statements are equivalent:*

(i) T is self-adjoint.

(ii) $\langle Tx, x \rangle \in \mathbb{R}$ for all $x \in H$.

Proof. "$(i) \Rightarrow (ii)$" For all $x \in H$ it holds

$$\langle Tx, x \rangle = \langle x, T^*x \rangle = \langle x, Tx \rangle = \overline{\langle Tx, x \rangle}.$$

"$(ii) \Rightarrow (i)$" Let $\lambda \in \mathbb{C}$. Then

$$\langle T(x + \lambda y), x + \lambda y \rangle = \langle Tx, x \rangle + \overline{\lambda} \langle Tx, y \rangle + \lambda \langle Ty, x \rangle + |\lambda|^2 \langle Ty, y \rangle \in \mathbb{R},$$

thus, by taking the complex conjugate,

$$\langle T(x + \lambda y), x + \lambda y \rangle = \langle Tx, x \rangle + \lambda \langle y, Tx \rangle + \overline{\lambda} \langle x, Ty \rangle + |\lambda|^2 \langle Ty, y \rangle.$$

This yields

$$\overline{\lambda} \langle Tx, y \rangle + \lambda \langle Ty, x \rangle = \lambda \langle y, Tx \rangle + \overline{\lambda} \langle x, Ty \rangle.$$

This gives for $\lambda = 1$

$$\langle Tx, y \rangle + \langle Ty, x \rangle = \langle y, Tx \rangle + \langle x, Ty \rangle$$

and for $\lambda = -i$

$$\langle Tx, y \rangle - \langle Ty, x \rangle = -\langle y, Tx \rangle + \langle x, Ty \rangle.$$

Consequently, $\langle Tx, y \rangle = \langle x, Ty \rangle$. ■

Theorem 18.8 *Let H be a Hilbert space and $T \in L(H)$ self-adjoint. Then*

$$\|T\| = \sup_{\|x\|\leq 1} |\langle Tx, x\rangle|.$$

Proof. We have

$$\sup_{\|x\|\leq 1} |\langle Tx, x\rangle| \leq \sup_{\|x\|\leq 1} \|Tx\|\|x\| \leq \sup_{\|x\|\leq 1} \|T\|\|x\|^2 = \|T\|.$$

In order to prove the reverse inequality we denote $M := \sup_{\|x\|\leq 1} |\langle Tx, x\rangle|$. For all $x, y \in H$ it holds

$$\begin{aligned}\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle =& 2\langle Tx, y\rangle + 2\langle Ty, x\rangle \\ =& 2\langle Tx, y\rangle + 2\overline{\langle Tx, y\rangle} = 4\operatorname{Re}\langle Tx, y\rangle.\end{aligned}$$

The parallelogram law gives

$$4\operatorname{Re}\langle Tx, y\rangle \leq M(\|x+y\|^2 + \|x-y\|^2) = 2M(\|x\|^2 + \|y\|^2),$$

thus

$$\operatorname{Re}\langle Tx, y\rangle \leq M \quad \forall \|x\|, \|y\| \leq 1.$$

Let $\|x\|, \|y\| \leq 1$ and $\lambda \in \mathbb{K}, |\lambda| = 1$, be such that $\lambda |\langle Tx, y\rangle| = \langle Tx, y\rangle$. Then

$$|\langle Tx, y\rangle| = \lambda^{-1}\langle Tx, y\rangle = \langle T(\lambda^{-1}x), y\rangle \leq M.$$

This implies that $\|T\| \leq M$. ■

Corollary 18.9 *If H ia a Hilbert space and $T \in L(H)$ a self-adjoint operator such that $\langle Tx, x\rangle = 0$ for all $x \in H$, then $T = 0$.*

For $\mathbb{K} = \mathbb{C}$, if $\langle Tx, x\rangle = 0$ for all $x \in H$, then T is self-adjoint. This is in general not the case for $\mathbb{K} = \mathbb{R}$. Indeed, for $H = \mathbb{R}^2$,

$$T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

fulfills $\langle Tx, x\rangle = 0$ for all $x \in H$, while $T \neq T^*$.

Theorem 18.10 *Let H be a Hilbert space and $P \in L(H), P \neq 0$, a projection, namely, an operator such that $P^2 = P$. The following statements are equivalent:*

(i) $P : H \to \operatorname{ran} P$ *is an orthogonal projection, namely,* $\ker P \perp \operatorname{ran} P$.

(ii) $\|P\| = 1$.

(iii) P *is self-adjoint.*

(iv) P is normal.

(v) P is positive, this means, $\langle Px, x\rangle \geq 0$ for all $x \in H$.

Proof. Notice that $\mathrm{Id} - P$ is also a projection and $\operatorname{ran} P = \ker(\mathrm{Id} - P)$, which means that P has closed range. In addition, $\ker P = \{x - Px \mid x \in H\}$.

"$(i) \Rightarrow (ii)$" Follows from Theorem 16.4.

"$(ii) \Rightarrow (i)$" It is enough to prove that $x - Px \in (\operatorname{ran} P)^{\perp}$ for all $x \in H$ (see the proof of Theorem 16.4) or, equivalently, that $\ker P \perp \operatorname{ran} P$.

Let $x \in \ker P, y \in \operatorname{ran} P$ and $\lambda \in \mathbb{K}$. Then $P(x + \lambda y) = \lambda y$, thus

$$\|\lambda y\|^2 \leq \|x + \lambda y\|^2 = \|x\|^2 + 2 \operatorname{Re} \overline{\lambda}\langle x, y\rangle + \|\lambda y\|^2,$$

so,

$$-2 \operatorname{Re} \overline{\lambda}\langle x, y\rangle \leq \|x\|^2.$$

Since this inequality holds for every $\lambda \in \mathbb{R}$ and every $\lambda \in i\mathbb{R}$, it yields $\operatorname{Re}\langle x, y\rangle = 0$ and $\operatorname{Im}\langle x, y\rangle = 0$, thus $\langle x, y\rangle = 0$.

"$(i) \Rightarrow (iii)$" For all $x, y \in H$ it holds

$$\langle Px, y\rangle = \langle Px, Py + (y - Py)\rangle = \langle Px, Py\rangle = \langle Px + (x - Px), Py\rangle = \langle x, Py\rangle.$$

"$(iii) \Rightarrow (iv)$" Clear.

"$(iv) \Rightarrow (i)$" Since $0 = \langle (P^*P - PP^*)x, x\rangle = \|Px\|^2 - \|P^*x\|^2$ for all $x \in H$, we get that $\ker P = \ker P^* = (\operatorname{ran} P)^{\perp}$ (see Theorem 18.2(g)).

"$(iii) \Rightarrow (v)$" For all $x \in H$ we have $\langle Px, x\rangle = \langle P^2x, x\rangle = \langle Px, Px\rangle = \|Px\|^2 \geq 0$.

"$(v) \Rightarrow (i)$" For $x \in \ker P, y \in \operatorname{ran} P$ and $\lambda \in \mathbb{R}$ it holds

$$0 \leq \langle P(x + \lambda y), x + \lambda y\rangle = \langle \lambda y, x + \lambda y\rangle = \lambda^2\|y\| + \lambda\langle y, x\rangle,$$

which implies that

$$\langle y, x\rangle \geq -\lambda\|y\|^2 \quad \forall \lambda > 0 \text{ and } \langle y, x\rangle \leq -\lambda\|y\|^2 \quad \forall \lambda < 0.$$

Thus $\langle x, y\rangle = 0$. ■

Chapter VI

Spectral theory

19 The spectrum of a bounded operator

The spectral theory generalizes the eigenvalue theory for matrices to bounded operators on infinite-dimensional Banach spaces. An eigenvalue λ of a linear operator T defined on a finite-dimensional space (which can be identified with a matrix) is characterized through the fact that $\lambda \operatorname{Id} - T$ is not injective or, equivalently, $\lambda \operatorname{Id} - T$ is not surjective. This equivalence is in infinite-dimensional spaces in general not true, which means that a more general concept is needed.

In the following we assume that X is a Banach space and $T \in L(X)$ is a continuous linear operator. We write $\lambda - T$ instead of $\lambda \operatorname{Id} - T$.

Definition 19.1 Let $T \in L(X)$.

(a) The resolvent set of T is defined as

$$\rho(T) := \{\lambda \in \mathbb{K} \mid (\lambda - T)^{-1} \text{ exists in } L(X)\}.$$

(b) The mapping

$$\rho(T) \mapsto L(X), \quad \lambda \mapsto R_\lambda := R_\lambda(T) := (\lambda - T)^{-1},$$

is called resolvent mapping.

(c) The spectrum of T is defined as

$$\sigma(T) := \mathbb{K} \setminus \rho(T).$$

The spectrum is decomposed into:

- point spectrum: $\sigma_p(T) := \{\lambda \mid \lambda - T \text{ is not injective}\}$;

- continuous spectrum:

$$\sigma_c(T) := \{\lambda \mid \lambda - T \text{ is injective, not surjective, and has dense range}\};$$

- residual spectrum: $\sigma_r(T) := \{\lambda \mid \lambda - T \text{ is injective and has no dense range}\}$.

Notice that according to Corollary 12.5, if $\lambda - T$ is bijective, then $(\lambda - T)^{-1}$ is continuous, consequently, $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$. The elements of $\sigma_p(T)$ are called eigenvalues; an element $x \neq 0$ with the property that $Tx = \lambda x$ is called eigenvector.

Theorem 19.2 *It holds $\sigma(T) = \sigma(T^*)$. If X is a Hilbert space, then $\sigma(T^*) = \{\overline{\lambda} : \lambda \in \sigma(T)\}$, whereby T^* is the adjoint operator of T in Hilbert space sense (see Definition 19.1).*

Proof. The first statement follows from the fact that a continuous linear operator between Banach spaces is an isomorphism if and only if its adjoint operator is an isomorphism (see Übungsbeispiel 38(a)). If X is an Hilbert space, then

$$\left((\lambda - T)^{-1}\right)^* = ((\lambda - T)^*)^{-1} = \left(\overline{\lambda} - T^*\right)^{-1}.$$

■

Example 19.3 (a) If X is finite-dimensional, then $\sigma(T) = \sigma_p(T)$. For $\mathbb{K} = \mathbb{R}$ this set can be empty.

(b) Let $X = C[0,1]$ and $(Tx)(s) = \int_0^s x(t)dt$. Then $\sigma(T) = \sigma_r(T) = \{0\}$. It is easy to see hat $0 \in \sigma_r(T)$, since T is injective (the fundamental theorem of calculus) and it has no dense range, since $(Tx)(0) = 0$ for all $x \in C[0,1]$.

Let $\lambda \neq 0$. We will prove that the equation

$$\lambda x - Tx = y$$

has a unique solution, which will imply that $\lambda - T$ is bijective. The continuity of its inverse will follow as a consequence of the Open Mapping Theorem.

Denoting $z := Tx$, the equation gives rise to the following ordinary differential equation

$$\dot{z}(t) - \frac{1}{\lambda}z(t) = \frac{1}{\lambda}y(t), \quad z(0) = 0.$$

Its unique solution reads

$$z(t) = e^{\frac{t}{\lambda}}\frac{1}{\lambda}\int_0^t e^{-\frac{s}{\lambda}}y(s)ds,$$

thus

$$x(t) = \dot{z}(t) = \frac{1}{\lambda}(z(t) + y(t)) = \frac{1}{\lambda^2}\int_0^t e^{\frac{t-s}{\lambda}} y(s)ds + \frac{1}{\lambda}y(t).$$

This proves that $\lambda - T$ is bijective.

(c) Let $X = \{x \in C[0,1] \mid x(0) = 0\}$ and $(Tx)(s) = \int_0^s x(t)dt$. Then $\sigma(T) = \sigma_c(T) = \{0\}$. One can prove as in (b) that $\lambda \in \rho(T)$ for all $\lambda \neq 0$. On the other hand, T is injective and, since $\operatorname{ran} T = \{y \in C^1[0,1] \mid y(0) = 0, \dot{y}(0) = 0\}$, it has a dense range in X.

Theorem 19.4 *(a) $\rho(T)$ is open.*

(b) The resolvent mapping is analytic, which means that it can be locally represented by a power series with coefficients in $L(X)$.

(c) $\sigma(T)$ is compact, more precisely, $|\lambda| \leq \|T\|$ for all $\lambda \in \sigma(T)$.

(d) If $\mathbb{K} = \mathbb{C}$, then $\sigma(T) \neq \emptyset$.

Proof. (a) Let $\lambda_0 \in \rho(T)$ and $|\lambda - \lambda_0| < \|(\lambda_0 - T)^{-1}\|^{-1}$. Then

$$\lambda - T = (\lambda_0 - T) + (\lambda - \lambda_0) = (\lambda_0 - T)\left(\operatorname{Id} - (\lambda_0 - \lambda)(\lambda_0 - T)^{-1}\right).$$

According to Theorem 3.17, the Neumann series $\sum_{n=0}^{+\infty}\left((\lambda_0 - \lambda)(\lambda_0 - T)^{-1}\right)^n$ converges and $\operatorname{Id} - (\lambda_0 - \lambda)(\lambda_0 - T)^{-1}$ is invertible. Consequently, $\lambda - T$ is also invertible.

(b) As we have seen above, we can represent the resolvent mapping for $|\lambda - \lambda_0| < \|(\lambda_0 - T)^{-1}\|^{-1}$ as

$$R_\lambda = \left(\operatorname{Id} - (\lambda_0 - \lambda)(\lambda_0 - T)^{-1}\right)^{-1}(\lambda_0 - T)^{-1} = \sum_{n=0}^{+\infty}(\lambda_0 - \lambda)^n \left((\lambda_0 - T)^{-1}\right)^{n+1},$$

which gives the desired local representation as power series with coefficients $\left((\lambda_0 - T)^{-1}\right)^{n+1}$.

(c) According to (a), $\sigma(T)$ is closed. Let $|\lambda| > \|T\|$. The series

$$(\lambda - T)^{-1} = \lambda^{-1}\left(\operatorname{Id} - \frac{1}{\lambda}T\right)^{-1} = \lambda^{-1}\sum_{n=0}^{+\infty}\lambda^{-n}T^n$$

is convergent. This implies $\sigma(T) \subseteq \{\lambda \in \mathbb{K} \mid |\lambda| \leq \|T\|\}$, which shows that $\sigma(T)$ is bounded, consequently, compact.

(d) Assume that $\sigma(T) = \emptyset$. The resolvent mapping is defined on the whole space $\mathbb{C}$ and is analytic. Let $\ell \in (L(X))^*$. The mapping $\lambda \mapsto \ell(R_\lambda)$ can for every $\lambda_0 \in \mathbb{C}$ be locally represented as

$$\ell(R_\lambda) = \sum_{n=0}^{+\infty} (-1)^n (\lambda - \lambda_0)^n \ell(R_{\lambda_0}^{n+1}),$$

which means that it is analytic. For $|\lambda| > 2\|T\|$ it holds

$$|\ell(R_\lambda)| \le \|\ell\| |\lambda^{-1}| \sum_{n=0}^{+\infty} \left\| \frac{T}{\lambda} \right\|^n \le \|\ell\| \frac{1}{2\|T\|} 2 = \frac{\|\ell\|}{\|T\|}.$$

On the other hand, the continuous mapping $\lambda \mapsto \ell(R_\lambda)$ is bounded on the compact set $\{\lambda \mid |\lambda| \le 2\|T\|\}$, thus it is bounded on $\mathbb{C}$. According to the Liouville Theorem, an analytic function which is bounded on $\mathbb{C}$ must be constant. We take $\lambda_0 := 0$ and notice that $\ell(R_0^{n+1})$ must be equal to zero for all $n \ge 1$. In particular, $\ell(R_0^2) = \ell(T^{-2}) = 0$. Since this is true for all $\ell \in (L(X))^*$, according to the Hahn-Banach Theorem, we must have $T^{-2} = 0$, which is a contradiction to the fact that this operator is the inverse of T^2. This shows that $\sigma(T)$ cannot be empty. ■

The estimation in statement (c) of the above theorem can be strengthen. To this end we will use the following convergence statement for sequences of real numbers.

Lemma 19.5 *Let $(t_n)_{n\in\mathbb{N}}$ a sequence of real numbers such that $0 \le t_{n+m} \le t_n t_m$ for all $n, m \in \mathbb{N}$. Then $(\sqrt[n]{t_n})_{n\in\mathbb{N}}$ converges to $t := \inf_{n\in\mathbb{N}} \sqrt[n]{t_n}$.*

Proof. Let $\varepsilon > 0$ and $N \in \mathbb{N}$ be such that $\sqrt[N]{t_N} < t+\varepsilon$. Set $M := \max\{t_1, ..., t_N\}$. Every $n \ge N$ can be written as $n = kN + r$ for $1 \le r \le N$. We have

$$\sqrt[n]{t_n} = t_{kN+r}^{\frac{1}{n}} \le (t_N^k t_r)^{\frac{1}{n}} \le (t+\varepsilon)^{\frac{kN}{n}} M^{\frac{1}{n}} = (t+\varepsilon)(t+\varepsilon)^{-\frac{r}{n}} M^{\frac{1}{n}} < t + 2\varepsilon$$

for n big enough. This proves the statement. ■

Lemma 19.5 shows that the limit in the following definition exists and that it coincides with the infimum of the sequence.

Definition 19.6 The spectral radius of $T \in L(X)$ is defined as

$$r(T) := \inf \|T^n\|^{\frac{1}{n}} = \lim_{n\to+\infty} \|T^n\|^{\frac{1}{n}}.$$

Theorem 19.7 *It holds*

(a) $|\lambda| \le r(T)$ *for all* $\lambda \in \sigma(T)$*;*

(b) if $\mathbb{K} = \mathbb{C}$*, then there exists* $\lambda \in \sigma(T)$ *such that* $|\lambda| = r(T)$*.*

Thus, for $\mathbb{K} = \mathbb{C}$ *we have*

$$r(T) = \max\{|\lambda| \mid \lambda \in \sigma(T)\},$$

which explains the name of spectral radius of the operator T*.*

Proof. (a) The statement follows if we prove that for $|\lambda| > r(T)$ the series $\lambda^{-1}\sum_{n=0}^{+\infty}(T/\lambda)^n$ converges. This is indeed the case if $\limsup_{n\to+\infty}\|(T/\lambda)^n\|^{\frac{1}{n}} < 1$, which is true, since

$$\limsup_{n\to+\infty}\|(T/\lambda)^n\|^{\frac{1}{n}} = \lim_{n\to+\infty}|\lambda^{-1}|\|T^n\|^{\frac{1}{n}} = \frac{r(T)}{|\lambda|} < 1.$$

(b) Let $r_0 := \sup\{|\lambda| \mid \lambda \in \sigma(T)\}$. According to (a), $r_0 \le r(T)$. Let $|\mu| > r_0$. We will show that $|\mu| \ge r(T)$, which will imply that $r_0 = r(T)$. The statement will follow by using that the spectrum is compact.

Let $\ell \in L(X)^*$. On $\{\lambda \mid |\lambda| > r_0\} \subseteq \rho(T)$ we consider the analytic function $f_\ell(\lambda) = \ell((\lambda - T)^{-1})$. As seen in the proof of statement (a), this function can be represented for $|\lambda| > r(T)$ by the convergent series

$$f_\ell(\lambda) = \sum_{n=0}^{+\infty}\ell(T^n)\lambda^{-(n+1)}.$$

The series converges in the largest open annulus on which f_ℓ is analytic. In particular, it converges at μ, which implies that $\lim_{n\to+\infty}\ell(T^n/\mu^{n+1}) = 0$. This shows that $(T^n/\mu^{n+1})_{n\in\mathbb{N}}$ converges weakly to 0, consequently, it is bounded. Then there exists $K > 0$ such that $\|T^n\|^{\frac{1}{n}} \le K^{\frac{1}{n}}|\mu|^{\frac{n+1}{n}}$ for all $n \in \mathbb{N}$. Letting $n \to +\infty$, it yields $r(T) \le |\mu|$. ∎

We have in general that $r(T) < \|T\|$ (see Example 19.3(b), where $r(T) = 0$). However, for normal operators defined on Hilbert spaces equality holds.

Theorem 19.8 *Let* H *be a Hilbert space and* $T \in L(H)$ *a normal operator. It holds* $r(T) = \|T\|$*.*

Proof. According to Theorem 18.2 we have

$$\|T^2\|^2 = \|T^2(T^2)^*\| = \|(TT^*)^2\| = \|TT^*\|^2 = \|T\|^4,$$

thus $\|T^2\| = \|T\|^2$. Since T^n is also normal for all $n \in \mathbb{N}$, it yields by induction that $\|T^{2^k}\| = \|T\|^{2^k}$ for all $k \in \mathbb{N}$, so

$$r(T) = \lim_{k\to+\infty}\|T^{2^k}\|^{\frac{1}{2^k}} = \|T\|.$$

∎

20 The Riesz theory

In this section we will provide a complete description of the spectrum of a compact operator defined on a Banach space X.

Theorem 20.1 (Riesz-Schauder Theorem) *Let X be a Banach space, $T \in K(X)$ and $S = \operatorname{Id} - T$.*

(a) The kernel of S has finite dimension.

(b) S has closed range and $X/\operatorname{ran} S$ has finite dimension.

(c) $\dim(X/\operatorname{ran} S) = \dim(\ker S) = \dim(X^*/\operatorname{ran} S^*) = \dim(\ker S^*)$.

An operator fulfilling (a) and (b) is called Fredholm operator and the number

$$\operatorname{ind}(T) := \dim(\ker S) - \dim(X/\operatorname{ran} S)$$

is called its index. According to the Riesz-Schauder Theorem, operators of the form $S = \operatorname{Id} - T$, where T is a compact operator, are Fredholm operators with index 0. Such operators are in particular surjective if and only if they are injective. This also means that the integral equation $Sx = y$ is for every right-hand side y solvable if and only if the equation $Sx = 0$ has only the trivial solution $x = 0$, which is much easier to verify. If so, then $Sx = y$ has a unique solution.

Proof. (a) Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\ker S$. Then there exists a convergent subsequence $(Tx_{n_k})_{k\in\mathbb{N}}$. Since $0 = Sx_{n_k} = x_{n_k} - Tx_{n_k}$ for all $k \in \mathbb{N}$, $(x_{n_k})_{k\in\mathbb{N}}$ is convergent in X. Using that $\ker S$ is closed we have that $(x_{n_k})_{k\in\mathbb{N}}$ is convergent also in $\ker S$. According to Theorem 2.8 we obtain that $\ker S$ is finite-dimensional.

(b) Consider $\widehat{S} : X/\ker S \to \operatorname{ran} S, \widehat{S}([x]) = Sx$. It is easy to see that $\widehat{S}$ is linear, continuous and bijective. We will prove that $\widehat{S}^{-1}$ is continuous. Assuming that this is not true, by Theorem 3.12 we obtain a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ such that

$$\operatorname{dist}(x_n, \ker S) = \|[x_n]\|_{X/\ker S} = 1 \;\forall n \in \mathbb{N} \text{ and } Sx_n \to 0 \;(n \to +\infty).$$

This gives a sequence $(y_n)_{n\in\mathbb{N}} \subseteq X$ such that $\|y_n\| < 2$ for all $n \in \mathbb{N}$ and $Sy_n \to 0$ as $n \to +\infty$. Then there exists a subsequence $(y_{n_k})_{k\in\mathbb{N}}$ such that $(Ty_{n_k})_{k\in\mathbb{N}}$ is convergent, which implies that $(y_{n_k})_{k\in\mathbb{N}} = (Ty_{n_k} + Sy_{n_k})_{k\in\mathbb{N}}$ is convergent. We denote its limit by x. Thus $Sx = 0$ and $\|[x]\|_{X/\ker S} = 0$, which is a contradiction. This shows that $\widehat{S}^{-1}$ is an isomorphismus from $X/\ker S$ to $\operatorname{ran} S$ and, since $X/\ker S$ is complete, it follows that $\operatorname{ran} S$ is complete. According to Lemma 1.6(b), $\operatorname{ran} S$ is closed.

According to Theorem 6.11 we have that $(X/\operatorname{ran} S)^* \cong (\operatorname{ran} S)^\perp$. Obviously, $(\operatorname{ran} S)^\perp = \ker S^*$, which gives $(X/\operatorname{ran} S)^* \cong \ker S^*$. Thus $\dim(X/\operatorname{ran} S)^* = \dim(\ker S^*) < +\infty$, since, according to (a), the kernel of S^* is finite-dimensional (we use here that $T^* = \operatorname{Id} - S^*$ is also compact; see Schauder Theorem). Consequently, $\dim(X/\operatorname{ran} S) = \dim(X/\operatorname{ran} S)^* < +\infty$.

(c) From Theorem 6.11 and Theorem 14.3 we have that $(\ker S)^* \cong X^*/(\ker S)^\perp = X^*/\operatorname{ran} S^*$, thus

$$\dim(X^*/\operatorname{ran} S^*) = \dim(\ker S)^* < +\infty.$$

It remains to prove that

$$\dim(X/\operatorname{ran} S) = \dim(\ker S)$$

and from here, taking also into account the proof of (b), the conclusion will follow. In order to prove this statement we will need some intermediate results. ■

For $m \in \mathbb{N}$ we set

$$N_m := \ker S^m, \quad N_0 := \{0\}, \quad R_m := \operatorname{ran} S^m, \quad R_0 := X.$$

We have

$$N_0 \subseteq N_1 \subseteq N_2 \subseteq \ldots \quad \text{and} \quad R_0 \supseteq R_1 \supseteq R_2 \supseteq \ldots.$$

Since for every $m \in \mathbb{N}$ it holds

$$S^m = (\operatorname{Id} - T)^m = \operatorname{Id} - \sum_{k=1}^{m} \binom{m}{k} (-1)^{k+1} T^k := \operatorname{Id} - \widetilde{T},$$

where $\widetilde{T}$ is compact, N_m is finite-dimensional, R_m is closed and $\dim X/R_m < +\infty$.

Lemma 20.2 *(a) There exists a minimal number $p \in \mathbb{N} \cup \{0\}$ such that $N_p = N_{p+1}$.*

(b) It holds $N_p = N_{p+r}$ for all $r > 0$.

(c) $N_p \cap R_p = \{0\}$.

(d) There exists a minimal number $q \in \mathbb{N} \cup \{0\}$ such that $R_q = R_{q+1}$.

(e) It holds $R_q = R_{q+r}$ for all $r > 0$.

(f) $N_q + R_q = X$.

(g) p=q.

Proof. (a) Assume that such a number p does not exist, namely,

$$N_0 \subsetneq N_1 \subsetneq N_2 \subsetneq \ldots.$$

Since N_n is closed, for every $n \in \mathbb{N}$ there exists due to the Riesz Lemma (Lemma 2.6) an element $x_n \in N_n$ such that $\operatorname{dist}(x_n, N_{n-1}) > \frac{1}{2}$ and $\|x_n\| = 1$. For every $n > m \geq 1$ it holds

$$\|Tx_n - Tx_m\| = \|x_n - (Sx_n + x_m - Sx_m)\| > \frac{1}{2},$$

since $S(N_n) \subseteq N_{n-1}$, which means that $Sx_n + x_m - Sx_m \in N_{n-1}$. This implies that $(Tx_n)_{n\in\mathbb{N}}$ does not have a convergent subsequence. This is a contradiction to $T \in K(X)$.

(b) Let $r > 0$. We will show that $N_{p+r} \subseteq N_p$. Let $x \in N_{p+r}$. Then $S^{r-1}(x) \in N_{p+1} = N_p$, consequently, $x \in N_{p+r-1}$. By repeating this argument we finally get that $x \in N_p$.

(c) Let $x \in N_p \cap R_p$. It holds $S^p x = 0$ and $x = S^p y$, for $y \in X$. This yields $S^{2p} y = 0$, thus $y \in N_{2p} = N_p$. Consequently, $S^p y = 0$, thus $x = 0$.

(d) Assume that such a number q does not exist, namely,

$$R_0 \supsetneq R_1 \supsetneq R_2 \supsetneq \ldots.$$

Since R_n is closed, for every $n \in \mathbb{N}$ there exists due to the Riesz Lemma (Lemma 2.6) an element $x_n \in R_n$ such that $\operatorname{dist}(x_n, R_{n+1}) > \frac{1}{2}$ and $\|x_n\| = 1$. For every $m > n \geq 1$ it holds

$$\|Tx_n - Tx_m\| = \|x_n - (Sx_n + x_m - Sx_m)\| > \frac{1}{2},$$

since $S(R_n) = R_{n+1}$, which means that $Sx_n + x_m - Sx_m \in R_{n+1}$. This implies that $(Tx_n)_{n\in\mathbb{N}}$ does not have a convergent subsequence. This is a contradiction to $T \in K(X)$.

(e) Let $r > 0$. We will show that $R_{q+r} \supseteq R_q$. Let $x \in R_q$, which means that $x = S^q y$, for $y \in X$. Since $S^q y \in R_q = R_{q+1}$, it yields $x = S^q y = R_{q+1}$. By repeating this argument we finally get that $x \in R_{q+r}$.

(f) Let $x \in X$. According to (e) we have $S^q x \in R_q = R_{2q}$, which means that $S^q x = S^{2q} y$, for $y \in X$. Then we can write x as $x = (x - S^q y) + S^q y \in N_q + R_q$.

(g) Assume that $p > q$. Then $R_p = R_q$ and there exists $x \in N_p \setminus N_q$. We represent x as $x = y + z \in N_q + R_q$. Then $z = x - y \in N_p + N_q = N_p$ and $z \in R_q = R_p$. Consequently, $z = 0$, and so $x \in N_q$. Contradiction.

Assume that $p < q$. Then $N_p = N_q$ and there exists $x \in R_p \setminus R_q$. We represent x as $x = y + z \in N_q + R_q$. Then $y = x - z \in R_p + R_q = R_p$ and $y \in N_q = N_p$. Consequently, $y = 0$, and so $x \in R_q$. Contradiction.

■

Corollary 20.3 *Let X be a Banach space, $T \in K(X)$ and $S = \mathrm{Id} - T$. Then there exist closed linear subspaces $\widehat{N}$ and $\widehat{R}$ such that*

(a) $\dim \widehat{N} < +\infty$.

(b) $X = \widehat{N} \oplus \widehat{R}$ *is a topological direct sum, which means that X is the algebraic direct sum of $\widehat{N}$ and $\widehat{R}$ and that X is isomorphic to the space $\widehat{N} \oplus \widehat{R}$ (endowed with one of the norms introduced in Übungsbeispiel 6).*

(c) $S(\widehat{N}) \subseteq \widehat{N}, S(\widehat{R}) \subseteq \widehat{R}$.

(d) $S|_{\widehat{R}} : \widehat{R} \to \widehat{R}$ *is an isomorphism.*

Proof. Choose p and q as in the above lemma and set $\widehat{N} := N_p$ and $\widehat{R} := R_q$. Then $\widehat{N}$ is finite-dimensional and $\widehat{R}$ is closed. Lemma 20.2(c) and (b) guarantees that $X = \widehat{N} \oplus \widehat{R}$ is an algebraic direct sum. Since $\widehat{N}$ and $\widehat{R}$ are closed, they are Banach spaces, thus $\widehat{N} \oplus \widehat{R}$ is a Banach space. The mapping $\widehat{N} \oplus \widehat{R} \mapsto X, (x_1, x_2) \mapsto x_1 + x_2$, is linear, continuous and bijective, thus, according to Corollary 12.5, it is an isomorphism. This proves (b).

On the one hand we have $S(\widehat{N}) = S(N_p) \subseteq N_{p-1} \subseteq N_p = \widehat{N}$, for $p \geq 1$, and $S(\widehat{N}) = S(N_0) = N_0 = \widehat{N}$, for $p = 0$. On the other hand, $S(\widehat{R}) = S(R_p) \subseteq R_{p+1} = R_p = \widehat{R}$, which proves (c).

We have seen above that the mapping $x \mapsto Sx$ from $\widehat{R}$ to $\widehat{R}$ is surjective. If $Sy = 0$ for $y \in \widehat{R}$, where $y = S^p x$, then $S^{p+1}x = 0$, which means that $x \in N_{p+1} = N_p = \widehat{N}$. This proves that $y = S^p x = 0$, consequently, the mapping is bijective. It is also continuous and, according to Corollary 12.5, it is an isomorphism. This proves statement (d). ■

We will continue with the proof of the Riesz-Schauder Theorem.

Proof. (Continuation of the proof of Theorem 20.1)

We consider the decomposition $X = \widehat{N} \oplus \widehat{R}$ in Corollary 20.3 and set $\widehat{S} := S|_{\widehat{N}} : \widehat{N} \to \widehat{N}$. Let $\Phi : \widehat{N}/\operatorname{ran} \widehat{S} \to X/\operatorname{ran} S, x + \operatorname{ran} \widehat{S} \mapsto x + \operatorname{ran} S$. Since $\operatorname{ran} \widehat{S} \subseteq \operatorname{ran} S$, Φ is well-defined and linear.

We will prove that Φ is bijective and this will lead to $\dim(X/\operatorname{ran} S) = \dim(\ker S)$, since $\dim(\widehat{N}/\operatorname{ran} \widehat{S}) = \dim(\ker \widehat{S})$, which is known from the linear algebra (notice that $\widehat{N}$ is finite-dimensional), and $\ker S = \ker \widehat{S}$, which follows from the following equivalence:

$$
\begin{aligned}
x \in \ker S &\Leftrightarrow x = x_1 + x_2 \in \widehat{N} \oplus \widehat{S} \text{ and } Sx = 0 \\
&\Leftrightarrow x = x_1 + x_2 \text{ and } Sx_2 = -Sx_1 \in \widehat{N} \cap \widehat{R} \\
&\Leftrightarrow x = x_1 + x_2 \text{ and } Sx_2 = Sx_1 = 0 \\
&\Leftrightarrow x = x_1 + x_2 \text{ and } x_2 = Sx_1 = 0 \Leftrightarrow x \in \widehat{N} \text{ and } Sx = 0 \Leftrightarrow x \in \ker \widehat{S}.
\end{aligned}
$$

In order to prove that Φ is injective, let $x \in \widehat{N}$ with $\Phi([x]) = 0$, in other words, $x \in \operatorname{ran} S$. We will prove that $x \in \operatorname{ran} \widehat{S}$. Let $y \in X, y = y_1 + y_2 \in \widehat{N} \oplus \widehat{R}$, such that $x = Sy$. We have $Sy_2 = Sy - Sy_1 = x - Sy_1 \in \widehat{N}$. On the other hand, $Sy_2 \in \widehat{R}$ and so $Sy_2 = 0$. Since $y_2 \in \widehat{R}$, it holds $y_2 = 0$, consequently, $x = Sy_1 \in S(\widehat{N}) = \operatorname{ran} \widehat{S}$.

In order to prove that Φ is surjective, let $x = x_1 + x_2 \in X = \widehat{N} \oplus \widehat{R} = \widehat{N} \oplus S(\widehat{R})$. It holds $x + \operatorname{ran} S = x_1 + \operatorname{ran} \widehat{S}$, thus $\Phi(x_1 + \operatorname{ran} \widehat{S}) = x + \operatorname{ran} S$. ■

The following result is an important consequence of the Riesz-Schauder Theorem.

Theorem 20.4 (The Fredholm Alternative) *Let X be a Banach space, $T \in K(X)$ and $\lambda \in \mathbb{K}, \lambda \neq 0$. Either the equation*

$$\lambda x - Tx = 0$$

has only the trivial solution; in this case the inhomogeneous equation

$$\lambda x - Tx = y$$

is uniquely solvable for every $y \in X$; or there exists $n := \dim \ker(\lambda - T) < +\infty$ linearly independent solutions of the homogeneous equation, and the adjoint equation

$$\lambda x^* - T^* x^* = 0$$

has precisely n linearly independent solutions; in this case, the inhomogeneous equation is solvable if and only if $y \in (\ker(\lambda - T^*))_\perp$.

Proof. We only have to use Theorem 20.1 and also Theorem 9.6; notice that $\lambda - T = \lambda(\mathrm{Id} - T/\lambda)$. ■

Example 20.5 Let $T : C[0,1] \to C[0,1], (Tx)(s) = \int_0^s k(s,t)x(t)dt$, be the Volterra integral operator with continuous kernel $k : [0,1] \times [0,1] \to \mathbb{K}$. We have seen in Übungsbeispiel 28 that T is compact. For $\lambda \neq 0$ we consider the operator equation

$$\lambda x - Tx = 0.$$

We will prove that $\lambda - T$ is injective. Without loss of generality we assume that $\lambda = 1$. Let $x \in C[0,1]$ such that $Tx = x$. For every $s \in [0,1]$ it holds

$$|x(s)| = |Tx(s)| \leq \int_0^s |k(s,t)||x(t)|dt \leq s\|k\|_\infty \|x\|_\infty$$

and, from here,

$$|x(s)| = |Tx(s)| \leq \int_0^s |k(s,t)|t\|k\|_\infty \|x\|_\infty dt \leq \frac{s^2}{2}\|k\|_\infty \|x\|_\infty.$$

By repeating this approach it yields for every $s \in [0,1]$

$$|x(s)| \leq \frac{s^n \|k\|_\infty^n}{n!} \|x\|_\infty \to 0 \quad (n \to +\infty),$$

thus $x = 0$.

The uniqueness of the solution of the homogeneous equation $\lambda x - Tx = 0$ yields via the Fredholm Alternative the existence of the solution of the inhomogeneous equation

$$\lambda x - Tx = y$$

for every $y \in C[0,1]$.

We state now the main theorem of the spectral theory of compact operators on Banach spaces.

Theorem 20.6 *Let X be a Banach space and $T \in K(X)$.*

(a) If X is infinite-dimensional, then $0 \in \sigma(T)$.

(b) The (possibly empty) set $\sigma(T) \setminus \{0\}$ is at most countable.

(c) Every $\lambda \in \sigma(T) \setminus \{0\}$ is an eigenvalue of T and the corresponding eigenspace $\ker(\lambda - T)$ is finite-dimensional. In addition, there exists a topological direct sum $X = N(\lambda) \oplus R(\lambda)$ with $T(N(\lambda)) \subseteq N(\lambda)$ and $T(R(\lambda)) \subseteq R(\lambda)$, $N(\lambda)$ finite-dimensional, $\ker(\lambda - T) \subseteq N(\lambda)$, and $(\lambda - T)|_{R(\lambda)}$ an isomorphism from $R(\lambda)$ to $R(\lambda)$.

(d) $\sigma(T)$ has no nonzero accumulation points.

Proof. (a) If $0 \in \rho(T)$, then T is invertible with continuous inverse and, so, $\mathrm{Id} = TT^{-1}$ is compact. According to Theorem 2.8 we obtain that $\dim X < +\infty$. Contradiction.

(c) The statement follows from Theorem 20.1, Corollary 20.3 and Theorem 20.4.

(b) + (d) The statement follows by proving that for every $\varepsilon > 0$ the set $\{\lambda \in \sigma(T) \mid |\lambda| \geq \varepsilon\}$ is finite.

Assume that there exist $\varepsilon > 0$, a sequence $(\lambda_n)_{n \in \mathbb{N}} \subseteq \mathbb{K}$ and a sequence $(x_n)_{n \in \mathbb{N}} \subseteq X$ such that $|\lambda_n| \geq \varepsilon, x_n \neq 0, Tx_n = \lambda_n x_n$ for all $n \in \mathbb{N}$ and $\lambda_n \neq \lambda_m$ for all $n \neq m$.

The set $\{x_n \mid n \in \mathbb{N}\}$ is linear independent. Assuming that there exists $N \in \mathbb{N}$, linearly independent vectors $x_1, ..., x_N$ and scalars $\alpha_1, ..., \alpha_N$, not all equal to zero, with $x_{N+1} = \sum_{i=1}^N \alpha_i x_i$, it yields

$$Tx_{N+1} = \sum_{i=1}^N \alpha_i T x_i = \sum_{i=1}^N \lambda_i \alpha_i x_i = \lambda_{N+1} x_{N+1} = \sum_{i=1}^N \lambda_{N+1} \alpha_i x_i.$$

This means that for at least one i it must hold $\lambda_i = \lambda_{N+1}$. Contradiction.

We define $E_n := \operatorname{lin}\{x_1, ..., x_n\}$ for all $n \in \mathbb{N}$. Consequently,

$$E_1 \subsetneqq E_2 \subsetneqq E_3 \subsetneqq ....$$

For every $n \in \mathbb{N}$ it holds $T(E_n) \subseteq E_n$. Let $n \geq 2$. According to the Riesz Lemma (Lemma 2.6) there exists $y_n = \sum_{i=1}^n \alpha_i^{(n)} x_i \in E_n$ such that $\operatorname{dist}(y_n, E_{n-1}) > \frac{1}{2}$ and $\|x_n\| = 1$. For every $n > m > 1$ it holds

$$\begin{aligned}\|Ty_n - Ty_m\| &= \|\lambda_n y_n - (Ty_m + \lambda_n y_n - Ty_n)\| \\ &= |\lambda_n| \left\|y_n - \lambda_n^{-1}(y_m + \lambda_n y_n - Ty_n)\right\| \geq |\lambda_n| \operatorname{dist}(y_n, E_{n-1}) \geq \frac{\varepsilon}{2},\end{aligned}$$

since $Ty_m \in E_m \subseteq E_{n-1}$ and $\lambda_n y_n - Ty_n = \sum_{i=1}^n (\lambda_n - \lambda_i)\alpha_i^{(n)} x_i \in E_{n-1}$.

This means that $(Ty_n)_{n\in\mathbb{N}}$ cannot be a Cauchy sequence, which contradicts the compactness of the operator T. ∎

Theorem 20.6 states that the spectrum of a compact operator T consists, excepting 0, of a zero sequence (or a finite set) of eigenvalues λ. The operator $\lambda - T$ can be decomposed in a finite-dimensional component (which can be represented by a matrix) and an isomorphism. However, as seen in Example 19.3 (c) and (d), not every compact operators has nonzero spectral values.

Bibliography

[1] M. Fabian, P. Habala, P. Hájek, V. Montesinos, V. Zizler: *Banach Space Theory*, Springer, 2011

[2] W. Kaballo: *Grundkurs Funktionalanalysis*, Springer Spektrum, 2018

[3] W. Rudin: *Functional Analysis*, McGraw-Hill, 1991

[4] W. Rudin: *Principles of Mathematical Analysis*, McGraw-Hill, 1976

[5] W. Rudin: *Real and Complex Analysis*, McGraw-Hill, 1986

[6] D. Werner: *Funktionalanalysis*, Springer, 2011

[7] C. Zălinescu: *Convex Analysis in General Vector Spaces*, World Scientific, 2002