Functional Analysis *

Martin Brokate **

Contents

*Lecture Notes, WS 2017/18

**Zentrum Mathematik, TU München

1 Normed Spaces

Functional analysis is concerned with normed spaces and with operators between normed spaces.

As an example we consider the initial value problem on $I = [t_0, t_1]$,

$$y' = f(t, y)\,, \quad y(t_0) = y_0\,, \tag{1.1}$$

with the right side $f : I \times \mathbb{R}^n \to \mathbb{R}^n$. We transform (1.1) into the integral equation

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds\,. \tag{1.2}$$

We define an operator

$$T : C(I; \mathbb{R}^n) \to C(I; \mathbb{R}^n)\,, \quad (Ty)(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds\,. \tag{1.3}$$

Now (1.2) is equivalent to the operator equation

$$y = Ty\,. \tag{1.4}$$

In this manner we have transformed our original problem into an equation whose unknown variable is a function (in this case the function y), not a number nor a vector with finitely many components. In this equation there appears a mapping (here T) between function spaces (here $C(I; \mathbb{R}^n)$). Such mappings are usually called “operators”. The function spaces are typically infinite-dimensional Banach or Hilbert spaces.

In the following we write $\mathbb{K}$ for $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.1 (Norm, normed space)
*Let X be a vector space over $\mathbb{K}$. A mapping $\|\cdot\| : X \to [0, \infty)$ is called a **norm** on X, if*

$$\|x\| = 0 \qquad \Leftrightarrow \qquad x = 0\,, \tag{1.5}$$

$$\|\alpha x\| = |\alpha|\, \|x\| \qquad \text{for all } \alpha \in \mathbb{K},\ x \in X, \tag{1.6}$$

$$\|x + y\| \le \|x\| + \|y\| \qquad \text{for all } x, y \in X. \tag{1.7}$$

*The pair $(X, \|\cdot\|)$ is called a **normed space**. (Often, one just writes X.)*

We repeat some basic notions from the first year analysis course. (See, for example, my lecture notes.)

Let X be a normed space. A sequence $(x_n)_{n\in\mathbb{N}}$ in X is called **convergent** to the **limit** $x \in X$, written as

$$\lim_{n\to\infty} x_n = x\,, \tag{1.8}$$

if

$$\lim_{n\to\infty} \|x_n - x\| = 0\,. \tag{1.9}$$

The sequence is called a **Cauchy sequence**, if for every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that

$$\|x_n - x_m\| \le \varepsilon\,, \quad \text{for all } n, m \ge N. \tag{1.10}$$

X is called **complete**, if every Cauchy sequence in X has a limit (which by the definition above has to be an element of X. If X is complete, X is called a **Banach space**. For $x \in X$ and $\varepsilon > 0$ we call

$$B(x, \varepsilon) = \{y : \, y \in X, \, \|y - x\| < \varepsilon\}\,, \tag{1.11}$$

$$K(x, \varepsilon) = \{y : \, y \in X, \, \|y - x\| \le \varepsilon\}\,, \tag{1.12}$$

the **open** and **closed** ε**-ball** around x, respectively. A subset O of X is called **open in** X, if for every $x \in O$ there exists an $\varepsilon > 0$ such that $B(x, \varepsilon) \subset O$. A subset A of X is called **closed in** X, if $X \setminus A$ is open in X, or equivalently, if for every convergent sequence $(x_n)_{n \in \mathbb{N}}$ in X whose elements x_n all belong to A, its limit also belongs to A. The **closure** $\overline{Y}$, the **interior** $\operatorname{int}(Y)$ and the **boundary** ∂Y of a subset Y of X are given by

$$\overline{Y} = \bigcap_{\substack{A \supset Y \\ A \text{ closed}}} A\,, \quad \operatorname{int}(Y) = \bigcup_{\substack{O \subset Y \\ O \text{ open}}} O\,, \quad \partial Y = \overline{Y} \setminus \operatorname{int}(Y)\,. \tag{1.13}$$

A subspace U of the vector space X becomes itself a normed space, if we define the norm on U to be the restriction of the given norm on X. If U is complete then U is closed in X; if X itself is complete, the converse also holds. The closure $\overline{U}$ of a subspace U of X is also a subspace of X.

In the basic analysis lectures we have already encountered some Banach spaces. The spaces $(\mathbb{K}^n, \|\cdot\|_p)$, $1 \le p \le \infty$, are Banach spaces with

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}\,, \quad 1 \le p < \infty\,, \quad \|x\|_\infty = \max_{1 \le i \le n} |x_i|\,. \tag{1.14}$$

Let D be an arbitrary set. Then $(B(D; \mathbb{K}), \|\cdot\|_\infty)$,

$$B(D; \mathbb{K}) = \{f | \, f : D \to \mathbb{K}, \; f \text{ bounded}\}\,, \quad \|f\|_\infty = \sup_{x \in D} |f(x)|\,, \tag{1.15}$$

is a Banach space. If moreover D is a compact metric space (for example, if D is a closed and bounded subset of $\mathbb{K}^n$), then $(C(D; \mathbb{K}), \|\cdot\|_\infty)$,

$$C(D; \mathbb{K}) = \{f | \, f : D \to \mathbb{K}, \; f \text{ stetig}\}\,, \tag{1.16}$$

is a closed subspace of $(B(D; \mathbb{K}), \|\cdot\|_\infty)$ and therefore itself a Banach space. If D is a Lebesgue measurable subset of $\mathbb{R}^n$ (this is the case, for example, if D is open or closed), then for $1 \le p < \infty$ the space $L^p(D; \mathbb{K})$ of functions which are Lebesgue integrable to the p-th power, that is, those measurable functions for which

$$\|f\|_p = \left(\int_D |f(x)|^p \, dx \right)^{\frac{1}{p}} < \infty\,, \tag{1.17}$$

becomes a Banach space for this norm $\|\cdot\|_p$, if we "identify" functions which are equal almost everywhere. ("Identify" means that we pass to the quotient space formed by equivalence classes of functions which are equal almost everywhere.) This is discussed in detail in the course on measure and integration theory.

Sequence spaces. Let us set $D = \mathbb{N}$. Then $B(D;\mathbb{K})$ coincides with the space of all bounded sequences, denoted as $\ell^\infty(\mathbb{K})$,

$$\ell^\infty(\mathbb{K}) = \{x : x = (x_k)_{k\in\mathbb{N}},\, x_k \in \mathbb{K},\, \sup_{k\in\mathbb{N}} |x_k| < \infty\}\,, \quad \|x\|_\infty = \sup_{k\in\mathbb{N}} |x_k|\,. \tag{1.18}$$

Because this is a special case of (1.15), the space $\ell^\infty(\mathbb{K})$ is a Banach space. Let us consider the subsets

$$c(\mathbb{K}) = \{x : x = (x_k)_{k\in\mathbb{N}} \text{ is a convergent sequence in } \mathbb{K}\}\,, \tag{1.19}$$

$$c_0(\mathbb{K}) = \{x : x = (x_k)_{k\in\mathbb{N}} \text{ converges to 0 in } \mathbb{K}\}\,. \tag{1.20}$$

Proposition 1.2 *We have $c_0(\mathbb{K}) \subset c(\mathbb{K}) \subset \ell^\infty(\mathbb{K})$. Endowed with the supremum norm, the spaces $c_0(\mathbb{K})$ are $c(\mathbb{K})$ Banach spaces.*

Proof: Exercise. It suffices to show that $c(\mathbb{K})$ is a closed subspace of $\ell^\infty(\mathbb{K})$, and that $c_0(\mathbb{K})$ is a closed subspace of $c(\mathbb{K})$. □

Moreover, we consider the space $c_e(\mathbb{K})$ of all finite sequences,

$$c_e(\mathbb{K}) = \{x : x = (x_k)_{k\in\mathbb{N}}, \text{ there exists } N \in \mathbb{N} \text{ with } x_k = 0 \text{ for all } k \geq N\}\,. \tag{1.21}$$

The space $c_e(\mathbb{K})$ is a subspace of $c_0(\mathbb{K})$; it is not closed in $c_0(\mathbb{K})$. Thus, it is not a Banach space; we have (exercise)

$$\overline{c_e(\mathbb{K})} = c_0(\mathbb{K})\,. \tag{1.22}$$

Let $x = (x_k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{K}$. We define

$$\|x\|_p = \left(\sum_{k=1}^\infty |x_k|^p\right)^{\frac{1}{p}}\,, \quad 1 \leq p < \infty\,, \tag{1.23}$$

and

$$\ell^p(\mathbb{K}) = \{x : x = (x_k)_{k\in\mathbb{N}},\, x_k \in \mathbb{K},\, \|x\|_p < \infty\}\,, \tag{1.24}$$

the space of sequences which are summable to the p-th power.

Proposition 1.3 *The space $(\ell^p(\mathbb{K}), \|\cdot\|_p)$, $1 \leq p < \infty$, is a Banach space.*

Proof: For $x \in \ell^p(\mathbb{K})$ we have

$$\|x\|_p = 0 \quad \Leftrightarrow \quad \sum_{k=1}^\infty |x_k|^p = 0 \quad \Leftrightarrow \quad |x_k| = 0 \text{ for all } k \quad \Leftrightarrow \quad x = 0\,.$$

If $x \in \ell^p(\mathbb{K})$, $\alpha \in \mathbb{K}$, then

$$\|\alpha x\|_p = \left(\sum_{k=1}^\infty |\alpha x_k|^p\right)^{\frac{1}{p}} = |\alpha| \left(\sum_{k=1}^\infty |x_k|^p\right)^{\frac{1}{p}} = |\alpha| \|x\|_p\,.$$

In order to prove the triangle inequality, let $x, y \in \ell^p(\mathbb{K})$. For arbitrary $N \in \mathbb{N}$ we have

$$\left(\sum_{k=1}^{N} |x_k + y_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{N} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{N} |y_k|^p\right)^{\frac{1}{p}},$$

this is just Minkowski's inequality in $\mathbb{K}^N$. It follows that

$$\sum_{k=1}^{N} |x_k + y_k|^p \leq (\|x\|_p + \|y\|_p)^p\,, \quad \text{for all } N \in \mathbb{N}.$$

Passing to the limit $N \to \infty$ yields

$$\sum_{k=1}^{\infty} |x_k + y_k|^p \leq (\|x\|_p + \|y\|_p)^p\,,$$

so

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p\,.$$

Thus, $\ell^p(\mathbb{K})$ is a normed space. In order to show that it is complete, let $(x^n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\ell^p(\mathbb{K})$. Then for all $k, n, m \in \mathbb{N}$ we have

$$|x_k^n - x_k^m|^p \leq \sum_{j=1}^{\infty} |x_j^n - x_j^m|^p = \|x^n - x^m\|_p^p\,.$$

Therefore, $(x_k^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{K}$ for all k. Since $\mathbb{K}$ is complete, there exists

$$x_k^\infty = \lim_{n\to\infty} x_k^n\,, \quad k \in \mathbb{N}\,.$$

For all k, n, m, N we have (Minkowski)

$$\begin{aligned}\left(\sum_{k=1}^{N} |x_k^n - x_k^\infty|^p\right)^{\frac{1}{p}} &\leq \left(\sum_{k=1}^{N} |x_k^n - x_k^m|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{N} |x_k^m - x_k^\infty|^p\right)^{\frac{1}{p}} \\ &\leq \|x^n - x^m\|_p + \left(\sum_{k=1}^{N} |x_k^m - x_k^\infty|^p\right)^{\frac{1}{p}}. \end{aligned} \tag{1.25}$$

Let $\varepsilon > 0$ be arbitrary. We choose M large enough such that

$$\|x^n - x^m\|_p \leq \varepsilon\,, \quad \text{for all } n, m \geq M. \tag{1.26}$$

Next, for every $N \in \mathbb{N}$ we choose $m(N) \in \mathbb{N}$ such that $m(N) \geq M$ and

$$\left(\sum_{k=1}^{N} |x_k^{m(N)} - x_k^\infty|^p\right)^{\frac{1}{p}} \leq \varepsilon\,. \tag{1.27}$$

We set $m = m(N)$ in (1.25). Then it follows from (1.25) – (1.27) that

$$\left(\sum_{k=1}^{N} |x_k^n - x_k^\infty|^p\right)^{\frac{1}{p}} \leq 2\varepsilon\,, \quad \text{for all } n \geq M \text{ and all } N \in \mathbb{N}.$$

Passing to the limit $N \to \infty$ we obtain

$$\|x^n - x^\infty\|_p \le 2\varepsilon \,, \quad \text{for all } n \ge M.$$

This implies $x^\infty = (x^\infty - x^n) + x^n \in \ell^p(\mathbb{K})$ and $x^n \to x^\infty$ in $\ell^p(\mathbb{K})$. □

Linear continuous mappings. Let X and Y be normed spaces. By definition, a mapping $f : X \to Y$ is continuous on X if and only if

$$f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n) \tag{1.28}$$

holds for all convergent sequences $(x_n)_{n\in\mathbb{N}}$ in X. From the analysis course we know that the assertions

$f^{-1}(O)$ is open for every open set $O \subset Y$,

and

$f^{-1}(A)$ is closed for every closed set $A \subset Y$

both are equivalent to (1.28).

f is continuous at a point $x \in X$, if (1.28) holds for all sequences $(x_n)_{n\in\mathbb{N}}$ which converge to x.

Proposition 1.4 *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed spaces, let $T : X \to Y$ be linear. Then the following are equivalent:*

(i) T is continuous on X.

(ii) T is continuous in 0.

(iii) There exists a $C > 0$ such that

$$\|T(x)\|_Y \le C\|x\|_X \,, \quad \textit{for all } x \in X. \tag{1.29}$$

(iv) T is Lipschitz continuous on X with Lipschitz constant C.

Proof: "(iii)⇒(iv)": For all $x, y \in X$ we have

$$\|T(x) - T(y)\|_Y = \|T(x-y)\|_Y \le C\|x - y\|_X \,.$$

"(iv)⇒(i)⇒(ii)": obvious.
"(ii)⇒(iii)": Contraposition. Assume that (iii) does not hold. We choose a sequence $(x_n)_{n\in\mathbb{N}}$ in X satisfying

$$\|T(x_n)\|_Y > n\|x_n\|_X \,. \tag{1.30}$$

We set

$$z_n = \frac{1}{n\|x_n\|_X} x_n \,,$$

this is possible, because $x_n \neq 0$ due to (1.30). It follows that $z_n \to 0$, but $\|T(z_n)\|_Y > 1$ and therefore $T(z_n)$ does not converge to 0. Consequently, (ii) does not hold. $\square$

Not all linear mappings are continuous. Here is an example: the unit vectors $\{e_k : k \in \mathbb{N}\}$ form a basis of $c_e(\mathbb{K})$. We define a linear mapping $T : c_e(\mathbb{K}) \to \mathbb{K}$ by $T(e_k) = k$. Then we have $\|e_k\|_\infty = 1$ and $|T(e_k)| = k$, so (iii) in Proposition 1.4 is not satisfied.

In the following we will write $\|x\|$ instead of $\|x\|_X$ if it is obvious which norm is meant. We will also write Tx instead of $T(x)$.

Definition 1.5 (Isomorphism)
Let X, Y be normed spaces. A mapping $T : X \to Y$ which is bijective, linear and continuous is called an ***isomorphism*** *between X and Y, if T^{-1}, too, is continuous. If moreover $\|Tx\| = \|x\|$ for all $x \in X$, T is called an* ***isometric isomorphism****. X and Y are called* ***(isometrically) isomorphic****, if there exists an (isometric) isomorphism between X and Y. In this case we write $X \simeq Y$ ($X \cong Y$).* $\square$

Obviously we have (and the same for "$\cong$")

$$X \simeq Y , \quad Y \simeq Z \qquad \Rightarrow \qquad X \simeq Z . \tag{1.31}$$

Let $T : X \to Y$ be an isomorphism. According to Proposition 1.4 there exist constants C_1 and C_2 such that

$$\|Tx\| \le C_1 \|x\| , \quad \|x\| = \|T^{-1}Tx\| \le C_2 \|Tx\| , \quad \text{for all } x \in X. \tag{1.32}$$

If $\|\cdot\|_1$ and $\|\cdot\|_2$ are two different norms on X, the identity mapping is an isomorphism between $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ if and only if there exist constants C_1 and C_2 such that

$$\|x\|_1 \le C_1 \|x\|_2 , \quad \|x\|_2 \le C_2 \|x\|_1 , \quad \text{for all } x \in X. \tag{1.33}$$

In this case the norms $\|\cdot\|_1$ und $\|\cdot\|_2$ are called **equivalent**.

We already know from the basic analysis course that on $\mathbb{K}^n$ all norms are equivalent. An immediate generalization of this result is the following.

Proposition 1.6 *Let X, Y be finite dimensional normed spaces with* $\dim(X) = \dim(Y)$. *Then $X \simeq Y$.*

Proof: Let $\dim(X) = n$, let $\{v_1, \dots, v_n\}$ be a basis of X, let $e_1, \dots, e_n$ be the unit vectors in $\mathbb{K}^n$. We define

$$T : \mathbb{K}^n \to X , \quad Tx = \sum_{i=1}^n x_i v_i \quad \text{if} \quad x = \sum_{i=1}^n x_i e_i .$$

One can verify immediately that

$$\|x\|_X = \|Tx\|$$

defines a norm on $\mathbb{K}^n$; T then becomes an isometric isomorphism. For Y, we proceed analogously. Then

$$X \cong (\mathbb{K}^n, \|\cdot\|_X) \simeq (\mathbb{K}^n, \|\cdot\|_Y) \cong Y .$$

□

If X, Y are finite dimensional vector spaces with $Y \subset X$, $Y \neq X$, then $\dim(Y) < \dim(X)$, and X und Y are not isomorphic, since bijective linear mappings leave the dimension invariant. In the infinite dimensional case, the situation is not that simple. For example, $c_0(\mathbb{K}) \subset c(\mathbb{K})$ and $c_0(\mathbb{K}) \neq c(\mathbb{K})$, but we have

$$c_0(\mathbb{K}) \simeq c(\mathbb{K}) \,. \tag{1.34}$$

Indeed, we claim that an isomorphism $T : c(\mathbb{K}) \to c_0(\mathbb{K})$ is given by

$$(Tx)_1 = \lim_{j\to\infty} x_j \,, \quad (Tx)_k = x_{k-1} - \lim_{j\to\infty} x_j \,,\, k \geq 2 \,, \tag{1.35}$$

and that the mapping $S : c_0(\mathbb{K}) \to c(\mathbb{K})$,

$$(Sy)_k = y_{k+1} + y_1 \,, \tag{1.36}$$

is the inverse of T. One may compute directly that $T \circ S$ and $S \circ T$ are equal to the identity on $c_0(\mathbb{K})$ and $c(\mathbb{K})$ resp., and that $\|Tx\|_\infty \leq 2\|x\|_\infty$ as well as $\|Sy\|_\infty \leq 2\|y\|_\infty$ hold for all $x \in c(\mathbb{K})$ and all $y \in c_0(\mathbb{K})$.

Proposition 1.7 *Let $(X_1, \|\cdot\|_1), \ldots, (X_m, \|\cdot\|_m)$ be normed spaces. On the product space*

$$X = \prod_{i=1}^m X_i = X_1 \times \cdots \times X_m \tag{1.37}$$

the expressions

$$\|x\|_\infty = \max_{1\leq i\leq m} \|x_i\|_i \,, \quad \|x\|_p = \left(\sum_{i=1}^m \|x_i\|_i^p \right)^{\frac{1}{p}} \,, \tag{1.38}$$

where $x = (x_1, \ldots, x_m) \in X$, define norms $\|\cdot\|_p$, $1 \leq p \leq \infty$. All these norms are equivalent. A sequence $(x^n)_{n\in\mathbb{N}}$ in X converges to an $x = (x_1, \ldots, x_m) \in X$ if and only if all component sequences $(x_i^n)_{n\in\mathbb{N}}$ converge to x_i. X is complete if and only if all X_i are complete.

Proof: Exercise. □

Corollary 1.8 *Let X be a normed space. The addition $+ : X \times X \to X$ and the scalar multiplication $\cdot : \mathbb{K} \times X \to X$ are continuous.*

Proof: If $x_n \to x$ and $y_n \to y$, then

$$0 \leq \|(x_n + y_n) - (x + y)\| \leq \|x_n - x\| + \|y_n - y\| \to 0 \,.$$

If $\alpha_n \to \alpha$ and $x_n \to x$, then

$$0 \leq \|\alpha_n x_n - \alpha x\| \leq |\alpha_n| \, \|x_n - x\| + |\alpha_n - \alpha| \, \|x\| \to 0 \,.$$

□

Definition 1.9 (Space of operators, dual space)
Let X, Y be normed spaces. We define

$$L(X;Y) = \{T|\, T : X \to Y,\, T \text{ is linear and continuous}\}\,. \tag{1.39}$$

*The space $L(X;\mathbb{K})$ is called the **dual space** of X, denoted X^*. The elements of X^* are called **functionals**.* □

From linear algebra and analysis it is known that $L(X;Y)$ is a vector space.

Proposition 1.10 (Operator norm)
Let X, Y be normed spaces. Then

$$\|T\| = \sup_{x\in X,\, x\neq 0} \frac{\|Tx\|}{\|x\|} \tag{1.40}$$

defines a norm on $L(X,Y)$, called the operator norm. We have

$$\|Tx\| \le \|T\|\,\|x\|\,, \quad \text{for all } x \in X, \tag{1.41}$$

and

$$\|T\| = \sup_{x\in X,\, \|x\|\le 1} \|Tx\| = \sup_{x\in X,\, \|x\|=1} \|Tx\|\,, \tag{1.42}$$

as well as

$$\|T\| = \inf\{C : C > 0,\, \|Tx\| \le C\|x\| \text{ for all } x \in X\}\,. \tag{1.43}$$

If Y is a Banach space, then so is $L(X,Y)$.

Proof: Let $T \in L(X;Y)$, let $C > 0$ with $\|Tx\| \le C\|x\|$ for all $x \in X$. Dividing by $\|x\|$ we see that $\|T\| \le C$, so $\|T\| \in \mathbb{R}_+$, and (1.41) as well as (1.43) hold. From (1.41) we obtain that

$$\sup_{x\in X,\, \|x\|=1} \|Tx\| \le \sup_{x\in X,\, \|x\|\le 1} \|Tx\| \le \|T\|\,.$$

Since

$$\frac{\|Tx\|}{\|x\|} = \left\|T(\frac{x}{\|x\|})\right\|,$$

(1.42) follows. We have

$$\|T\| = 0 \quad \Leftrightarrow \quad \|Tx\| = 0 \text{ for all } x \in X \;\Leftrightarrow\; \quad Tx = 0 \text{ for all } x \in X \quad \Leftrightarrow \quad T = 0\,.$$

For $\alpha \in \mathbb{K}$ we get

$$\|\alpha T\| = \sup_{\|x\|=1} \|\alpha Tx\| = |\alpha| \sup_{\|x\|=1} \|Tx\| = |\alpha|\,\|T\|\,,$$

and for $S, T \in L(X;Y)$ we get

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

Thus, the properties of a norm are satisfied. We now prove that $L(X, Y)$ is complete. Let $(T_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $L(X; Y)$. Since

$$\|T_n x - T_m x\| = \|(T_n - T_m)(x)\| \leq \|T_n - T_m\| \, \|x\|$$

the sequence $(T_n x)_{n\in\mathbb{N}}$ is a Cauchy sequence in Y for every $x \in X$. Then (as Y is complete)

$$Tx = \lim_{n\to\infty} T_n x$$

defines a mapping $T : X \to Y$. Let $x, z \in X$, $\alpha, \beta \in \mathbb{K}$, then

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

Therefore, T is linear. Since (as the norm is continuous)

$$\|Tx\| = \|\lim_{n\to\infty} T_n x\| = \lim_{n\to\infty} \|T_n x\| \leq (\sup_{n\in\mathbb{N}} \|T_n\|)\|x\|\,,$$

T is continuous. It remains to prove that $\|T_n - T\| \to 0$. Let $\varepsilon > 0$. We choose $N \in \mathbb{N}$ such that $\|T_n - T_m\| \leq \varepsilon$ for all $n, m \geq N$. For arbitrary $x \in X$ with $\|x\| = 1$ we get

$$\|(T_n - T)x\| \leq \|(T_n - T_m)x\| + \|(T_m - T)x\| \leq \|T_n - T_m\|\|x\| + \|T_m x - Tx\|\,.$$

Therefore, for $n \geq N$ it follows, choosing m sufficiently large,

$$\|(T_n - T)x\| \leq 2\varepsilon\,,$$

and thus $\|T_n - T\| \leq 2\varepsilon$ if $n \geq N$. $\square$

Example 1.11

1. D compact metric space, $X = C(D; \mathbb{K})$ with the supremum norm, $a \in D$, $T_a : X \to \mathbb{K}$, $T_a x = x(a)$. T_a is linear, and $|T_a x| = |x(a)| \leq \|x\|_\infty$ with equality, if x is a constant function. Therefore, T_a is continuous, and $\|T_a\| = 1$. The functional T_a is called the **Dirac functional** in a.

2. $X = C([a, b]; \mathbb{R})$ with the supremum norm, $T : X \to \mathbb{R}$,

   $$Tx = \int_a^b x(t)\,dt\,.$$

   T is linear,

   $$|Tx| = \left|\int_a^b x(t)\,dt\right| \leq (b - a)\|x\|_\infty$$

   with equality if x is constant. Therefore, T is continuous and $\|T\| = b - a$.

3. $X = L^1([a, b]; \mathbb{R})$ with the L^1 norm, T as before, then

   $$|Tx| = \left|\int_a^b x(t)\,dt\right| \leq \int_a^b |x(t)|\,dt = \|x\|_1$$

   with equality if x is constant. Therefore, T is continuous and $\|T\| = 1$.

4. X as before, $f \in C([a,b];\mathbb{R})$, $T : X \to \mathbb{R}$,

$$Tx = \int_a^b f(t)x(t)\,dt\,.$$

T is linear, and

$$|Tx| \le \int_a^b |f(t)|\,|x(t)|\,dt \le \|f\|_\infty \int_a^b |x(t)|\,dt = \|f\|_\infty\,\|x\|_1\,.$$

Therefore, T is continuous and $\|T\| \le \|f\|_\infty$. In order to prove that actually equality holds, we choose $t_* \in [a,b]$ with $f(t_*) = \|f\|_\infty$ (if the maximum of the absolute value is attained at a point where f is negative, we instead consider $-f$ resp. $-T$). Let now $\varepsilon > 0$, $\varepsilon < \|f\|_\infty$. We choose an interval I with $t_* \in I \subset [a,b]$, so that $f(t) \ge \|f\|_\infty - \varepsilon$ holds for all $t \in I$, and set ($|I|$ denotes the length of I)

$$x = \frac{1}{|I|}1_I\,, \quad \text{so} \quad x(t) = \begin{cases} \frac{1}{|I|}\,, & t \in I\,,\\ 0\,, & \text{sonst}\,. \end{cases}$$

Then $\|x\|_1 = 1$ and

$$|Tx| = \left|\int_a^b f(t)x(t)\,dt\right| = \int_I f(t)\frac{1}{|I|}\,dt \ge \int_I (\|f\|_\infty - \varepsilon)\frac{1}{|I|}\,dt = \|f\|_\infty - \varepsilon\,.$$

It follows that $\|T\| \ge \|f\|_\infty - \varepsilon$ and therefore $\|T\| = \|f\|_\infty$.

5. Let $D = (0,1)\times(0,1)$, $k \in L^2(D;\mathbb{R})$. On $X = L^2((0,1);\mathbb{R})$ we want to define an **integral operator** $T : X \to X$ by

$$(Tx)(s) = \int_0^1 k(s,t)x(t)\,dt\,, \quad s \in (0,1)\,.$$

Once we have proved that this integral is well defined, it is clear that T is linear. We have (the integrals are well defined as a consequence of Fubini's and Tonelli's theorem, the second inequality follows from Hölder's inequality)

$$\begin{aligned}
\|Tx\|^2_{L^2((0,1);\mathbb{R})} &= \int_0^1 \left(\int_0^1 k(s,t)x(t)\,dt\right)^2 ds \le \int_0^1 \left(\int_0^1 |k(s,t)|\,|x(t)|\,dt\right)^2 ds\\
&\le \int_0^1 \left(\int_0^1 |k(s,t)|^2\,dt\right)\cdot\left(\int_0^1 |x(t)|^2\,dt\right) ds\\
&= \|x\|^2_{L^2((0,1);\mathbb{R})} \int_0^1\int_0^1 |k(s,t)|^2\,dt\,ds\,.
\end{aligned}$$

Therefore, T is continuous and

$$\|T\| \le \left(\int_0^1\int_0^1 |k(s,t)|^2\,dt\,ds\right)^{\frac{1}{2}} = \|k\|_{L^2(D;\mathbb{R})}\,.$$

6. Let X, Y be normed spaces with $\dim(X) < \infty$. Then every linear mapping $T : X \to Y$ is continuous: Let $\{v_1, \dots, v_n\}$ be a basis of X. Then for
$$x = \sum_{i=1}^n x_i v_i \in X\,, \quad x_i \in \mathbb{K}\,,$$
we have that
$$\|Tx\| = \left\|\sum_{i=1}^n x_i T v_i\right\| \le \sum_{i=1}^n |x_i|\, \|Tv_i\| \le \max_{1\le i\le n} \|Tv_i\| \sum_{i=1}^n |x_i|\,.$$
Therefore,
$$\|T\| \le \max_{1\le i\le n} \|Tv_i\|\,, \quad \text{if} \quad \|x\| = \sum_{i=1}^n |x_i|\,.$$

7. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be linear, let $T(x) = Ax$ for the matrix $A \in \mathbb{R}^{(m,n)}$. Via this correspondence, we can identify the space $L(\mathbb{R}^n;\mathbb{R}^m)$ with the space $\mathbb{R}^{(m,n)}$ of all $m \times n$-matrices. The operator norm then becomes a so-called matrix norm whose explicit form depends on the choice of the norms in $\mathbb{R}^n$ and $\mathbb{R}^m$. Matrix norms play an important role in the construction and analysis of numerical algorithms, in particular in numerical linear algebra. $\Box$

Lemma 1.12 *Let X, Y, Z be normed spaces, let $T : X \to Y$ and $S : Y \to Z$ be linear and continuous. Then $S \circ T : X \to Z$ is linear and continuous, and*
$$\|S \circ T\| \le \|S\|\, \|T\|\,. \tag{1.44}$$

Proof: For all $x \in X$ we have $\|(S \circ T)x\| \le \|S\|\, \|Tx\| \le \|S\|\, \|T\|\, \|x\|$. The assertion now follows from Proposition 1.10. $\Box$

Seminorms and quotient spaces.

Definition 1.13 (Seminorm)
*Let X be a vector space over $\mathbb{K}$. A mapping $p : X \to [0,\infty)$ is called a **seminorm** on X, if*
$$p(\alpha x) = |\alpha|\, p(x) \qquad \text{for all } \alpha \in \mathbb{K},\ x \in X, \tag{1.45}$$
$$p(x+y) \le p(x) + p(y) \qquad \text{for all } x, y \in X. \tag{1.46}$$
In this case, (X, p) is called a seminormed space.

Obviously,(1.45) implies $p(0) = 0$, but $p(x) = 0$ does not imply $x = 0$. Every norm is a seminorm. If (Y, q) is a seminormed space, then every linear mapping $T : X \to Y$ generates a seminorm on X by
$$p(x) = q(Tx)\,. \tag{1.47}$$
Here are some examples of seminorms:
$$X = \mathbb{R}^n\,, \quad p(x) = |x_1|\,, \tag{1.48}$$
$$X = B(D;\mathbb{K})\,, \quad a \in D\,, \quad p(x) = |x(a)|\,, \tag{1.49}$$
$$X = L^1((0,1);\mathbb{R})\,, \quad p(x) = \left|\int_0^1 x(t)\,dt\right|\,, \tag{1.50}$$
$$X = C^1([0,1];\mathbb{R})\,, \quad p(x) = \|\dot{x}\|_\infty\,. \tag{1.51}$$

From linear algebra, we recall the concept of a quotient space. Let X be a vector space, let U be a subspace of X. Then

$$x \sim_U z \quad \Leftrightarrow \quad x - z \in U \tag{1.52}$$

defines an equivalence relation on X. Let

$$[x] = \{z : z \in X,\, x \sim_U z\} \tag{1.53}$$

denote the equivalence class of $x \in X$. The quotient space X/U is defined by

$$X/U = \{[x] : x \in X\}\,. \tag{1.54}$$

With the addition and scalar multiplication on X/U defined by

$$[x] + [y] = [x + y]\,, \quad \alpha[x] = [\alpha x]\,, \tag{1.55}$$

X/U becomes a vector space, and the mapping

$$Q : X \to X/U\,, \quad Qx = [x]\,, \tag{1.56}$$

is linear and surjective. We have $0 = [x] = Qx$ if and only if $x \in U$.

Proposition 1.14 (Quotient norm) *Let U be a subspace of a normed space X.*

(i) The formula

$$p([x]) = \operatorname{dist}(x, U) = \inf_{z \in U} \|x - z\| \tag{1.57}$$

defines a seminorm on X/U which satisfies

$$p([x]) \le \|x\|\,, \quad \text{for all } x \in X. \tag{1.58}$$

(ii) If U is closed, then p is a norm.

(iii) If moreover X is a Banach space and U is closed, then $(X/U, p)$ is a Banach space.

Proof: Part (i) is an exercise. Concerning (ii): Let

$$0 = p([x]) = \inf_{z \in U} \|x - z\|\,.$$

There exists a sequence $(z_n)_{n\in\mathbb{N}}$ in U satisfying $\|x - z_n\| \to 0$, thus $z_n \to x$ and therefore (if U is closed) $x \in U$, so finally $[x] = 0$. Concerning (iii): We begin by proving: If $x, y \in X$, then there exists a $\tilde{y} \in [y]$ with

$$\|\tilde{y} - x\| \le 2p([y - x])\,. \tag{1.59}$$

We obtain such a $\tilde{y}$ by first choosing a $z \in U$ with

$$\|y - x - z\| \le 2p([y - x])$$

and then setting $\tilde{y} = y - z$. Let now $([x_n])_{n\in\mathbb{N}}$ be a Cauchy sequence in X/U. Passing to a subsequence if necessary we may assume that

$$p([x_{n+1}] - [x_n]) \leq 2^{-n}\,.$$

We now choose according to (1.59) an $\tilde{x}_1 \in [x_1]$ and for $n > 1$ an $\tilde{x}_n \in [x_n]$ mit

$$\|\tilde{x}_n - \tilde{x}_{n-1}\| \leq 2p([x_n] - [x_{n-1}])\,. \tag{1.60}$$

Then for all $n, p \in \mathbb{N}$ we have

$$\|\tilde{x}_{n+p} - \tilde{x}_n\| \leq \sum_{j=1}^{p} \|\tilde{x}_{n+j} - \tilde{x}_{n+j-1}\| \leq \sum_{j=1}^{p} 2 \cdot 2^{-n-j+1} \leq 2^{-n+2}\,.$$

Therefore, $(\tilde{x}_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X. As X is complete, there exists $x = \lim_{n\in\mathbb{N}} \tilde{x}_n$. From (1.58) it follows that

$$0 \leq p([x_n] - [x]) = p([\tilde{x}_n] - [x]) \leq \|\tilde{x}_n - x\| \to 0\,,$$

so $[x_n] \to [x]$ in X/U. $\square$

Here is an example: $X = C([0,1])$, $U = \{x : x \in X,\, x(0) = 0\}$. We have

$$x \sim_U y \quad \Leftrightarrow \quad x(0) = y(0)\,.$$

One directly checks that

$$p([x]) = |x(0)|\,, \quad X/U \cong \mathbb{R}\,.$$

The mapping $T : X/U \to \mathbb{R}$, $T([x]) = x(0)$, is an isometric isomorphism

Dense subsets. A subset A of a metric space (X, d) is called **dense** in X, if

$$\overline{A} = X \tag{1.61}$$

holds. If A is dense in (X, d) and B is dense in (A, d_A), then B is dense in (X, d).

Definition 1.15 (Separable space)
A metric space (X, d) is called ***separable*** *if there exists a finite or countably infinite subset A of X which is dense in X.* $\square$

Example: $\mathbb{Q}^n$ is a dense subset of $\mathbb{R}^n$, therefore $\mathbb{R}^n$ is separable. Analogous for $\mathbb{C}^n$.

Proposition 1.16 *The space $(C([a,b];\mathbb{K}), \|\cdot\|_\infty)$ is separable.*

Proof: By the approximation theorem of Weierstra"s, the set P of all polynomials is dense in $C([a,b];\mathbb{K})$. Moreover, the set of all polynomials with rational coefficients is countable. It is dense in P, and therefore in $C([a,b];\mathbb{K})$, too. $\square$

Proposition 1.17 *Let X be a normed space, let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X with $X = \overline{\operatorname{span}\{x_n : n \in \mathbb{N}\}}$. Then X is separable.*

Proof: Exercise. $\square$

Proposition 1.18 *The space $\ell^p(\mathbb{K})$ is separable for $1 \leq p < \infty$. The space $\ell^\infty(\mathbb{K})$ is not separable.*

Proof: For $p < \infty$ we have $\ell^p(\mathbb{K}) = \overline{\operatorname{span}\{e_n : n \in \mathbb{N}\}}$. The assertion then follows from 1.17. Let now $p = \infty$. For arbitrary $M \subset \mathbb{N}$ we define $x^M \in \ell^\infty(\mathbb{K})$ by

$$x_k^M = \begin{cases} 1\,, & k \in M\,, \\ 0\,, & k \notin M\,. \end{cases}$$

If $M, N \subset \mathbb{N}$ with $M \neq N$, then $\|x^M - x^N\|_\infty = 1$. Therefore,

$$\{B(x^M, \frac{1}{2}) : M \subset \mathbb{N}\}$$

is an uncountable set whose elements are disjoint open balls. If A is a dense subset of $\ell^\infty(\mathbb{K})$, it has to be dense in each of those balls. Therefore A cannot be countable. $\square$

Analogously it holds that, for $D \subset \mathbb{R}^n$ open, the space $(L^p(D;\mathbb{K}), \|\cdot\|_p)$ is separable for $1 \leq p < \infty$, but not separable for $p = \infty$. If $p < \infty$ and moreover D is bounded, this follows from the result that $C(\overline{D})$ is dense in $(L^p(D), \|\cdot\|_p)$ and that (in analogy to Proposition 1.16) the polynomials with rational coefficients form a dense subset of $C(\overline{D})$. For D unbounded the claim follows from the representation

$$L^p(D;\mathbb{K}) = \overline{\bigcup_{n\in\mathbb{N}} L^p(D \cap B_n;\mathbb{K})}$$

with $B_n = \{x : \|x\|_p < n\}$.

Proposition 1.19 *Let X be a normed space, Y a Banach space, U a dense subspace of X and $S : U \to Y$ linear and continuous. Then there exists a unique linear and continuous mapping $T : X \to Y$ with $T|U = S$, and $\|T\| = \|S\|$.*

Proof: Given $x \in X$, we choose a sequence in $(x_n)_{n\in\mathbb{N}}$ in U with $x_n \to x$ and define $Tx = \lim_{n\to\infty} Sx_n$. The limit exists since $(x_n)_{n\in\mathbb{N}}$ and consequently $(Sx_n)_{n\in\mathbb{N}}$ are Cauchy sequences, and since Y is complete. The assertions now follow directly from the definitions and elementary properties of the operator norm and of convergent sequences; we do not write down the details. $\square$

Dual spaces. Let a normed space X be given. According to Proposition 1.10, the dual space $X^* = L(X;\mathbb{K})$ is a Banach space.

Proposition 1.20 *Let $p, q \in (1, \infty)$, let*

$$\frac{1}{p} + \frac{1}{q} = 1\,. \tag{1.62}$$

Then

$$(\ell^p(\mathbb{K}))^* \cong \ell^q(\mathbb{K})\,. \tag{1.63}$$

Proof: We want to define an isometric isomorphism $T : \ell^q(\mathbb{K}) \to (\ell^p(\mathbb{K}))^*$ by

$$(Tx)(y) = \sum_{k=1}^{\infty} x_k y_k \,, \quad x \in \ell^q(\mathbb{K}) \,, \quad y \in \ell^p(\mathbb{K}) \,. \tag{1.64}$$

By virtue of Hölder's inequality,

$$\sum_{k=1}^{N} |x_k|\,|y_k| \le \left(\sum_{k=1}^{N} |x_k|^q\right)^{\frac{1}{q}} \cdot \left(\sum_{k=1}^{N} |y_k|^p\right)^{\frac{1}{p}} \le \|x\|_q \, \|y\|_p \,.$$

Therefore, the series $\sum_k x_k y_k$ converges absolutely, and

$$|(Tx)(y)| \le \sum_{k=1}^{\infty} |x_k y_k| \le \|x\|_q \, \|y\|_p \,.$$

Thus, (1.64) defines for any given $x \in \ell^q(\mathbb{K})$ a linear continuous mapping $Tx : \ell^p(\mathbb{K}) \to \mathbb{K}$ which satisfies

$$\|Tx\| \le \|x\|_q \,. \tag{1.65}$$

Therefore $T : \ell^q(\mathbb{K}) \to (\ell^p(\mathbb{K}))^*$ is well-defined. It follows from (1.64) that T is linear and from (1.65) that T is continuous. T is injective, as $(Tx)(e_k) = x_k$, and consequently $Tx = 0$ implies that $x_k = 0$ for all k. We now prove that T is surjective. Let $y^* \in (\ell^p(\mathbb{K}))^*$ be arbitrary. We want to find an $x = (x_k)_{k\in\mathbb{N}}$ in $\ell^q(\mathbb{K})$ such that $Tx = y^*$. Since $(Tx)(e_k) = x_k$ for such an x, we have to define

$$x_k = y^*(e_k) \,. \tag{1.66}$$

Let $y \in c_e(\mathbb{K})$,

$$y = \sum_{k=1}^{N} y_k e_k \,, \quad y_k \in \mathbb{K} \,, \quad N \in \mathbb{N} \,, \tag{1.67}$$

be given. We have

$$y^*(y) = y^*\left(\sum_{k=1}^{N} y_k e_k\right) = \sum_{k=1}^{N} y_k y^*(e_k) = \sum_{k=1}^{N} y_k x_k \,. \tag{1.68}$$

We choose

$$y_k = |x_k|^{q-1} \operatorname{sign}(x_k) \,, \quad 1 \le k \le N \,. \tag{1.69}$$

Then

$$|y_k|^p = |x_k|^{p(q-1)} = |x_k|^q = x_k y_k \,,$$

and it follows from (1.68) that

$$\sum_{k=1}^{N} |x_k|^q = \sum_{k=1}^{N} x_k y_k = y^*(y) \le \|y^*\| \, \|y\|_p = \|y^*\| \left(\sum_{k=1}^{N} |y_k|^p\right)^{\frac{1}{p}} = \|y^*\| \left(\sum_{k=1}^{N} |x_k|^q\right)^{\frac{1}{p}} .$$

Thus,

$$\left(\sum_{k=1}^{N} |x_k|^q\right)^{\frac{1}{q}} \le \|y^*\| \,. \tag{1.70}$$

Passing to the limit $N \to \infty$ yields $x \in \ell^q(\mathbb{K})$ and

$$\|x\|_q \le \|y^*\| . \tag{1.71}$$

From (1.64) and (1.68) it follows that

$$(Tx)(y) = y^*(y) , \quad \text{for all } y \in c_e(\mathbb{K}).$$

We recall that $c_e(\mathbb{K})$ is dense in $(\ell^p(\mathbb{K}), \|\cdot\|_p)$. It follows from Proposition 1.19 that

$$Tx = y^* . \tag{1.72}$$

Therefore T is surjective, and from (1.65) and (1.71) we obtain that $\|Tx\| = \|x\|_q$. Therefore, T is an isometric isomorphism. $\square$

We mention some other results concerning the representation of dual spaces. We have

$$c_0(\mathbb{K})^* \cong \ell^1(\mathbb{K}) , \quad \ell^1(\mathbb{K})^* \cong \ell^\infty(\mathbb{K}) . \tag{1.73}$$

We do not present the proof here.

For function spaces, analogous results hold. For example, let $D \subset \mathbb{R}^n$ be open. Then

$$L^p(D;\mathbb{K})^* \cong L^q(D;\mathbb{K}) , \quad 1 < p < \infty , \quad \frac{1}{p} + \frac{1}{q} = 1 . \tag{1.74}$$

For the case $p = 2$, (1.74) is a consequence of a general result for Hilbert spaces, which will be treated in the next chapter. We sketch the proof for arbitrary $p \in (1, \infty)$; its general structure is the same as that for the sequence space $\ell^p(\mathbb{K})$. Setting

$$(Tx)(y) = \int_D x(t)y(t)\, dt , \quad x \in L^q(D;\mathbb{K}),\ y \in L^p(D;\mathbb{K}) , \tag{1.75}$$

one obtains a linear continuous mapping $T : L^q(D;\mathbb{K}) \to L^p(D;\mathbb{K})^*$ with $\|Tx\| = \|x\|_q$. Indeed, by virtue of Hölder's inequality,

$$|(Tx)(y)| \le \int_D |x(t)|\,|y(t)|\, dt \le \left(\int_D |x(t)|^q\, dt\right)^{\frac{1}{q}} \left(\int_D |y(t)|^p\, dt\right)^{\frac{1}{p}} = \|x\|_q\, \|y\|_p , \tag{1.76}$$

and for

$$y(t) = \operatorname{sign}(x(t))|x(t)|^{q-1}$$

we have $|y(t)|^p = |x(t)|^q$ and therefore

$$\begin{aligned}(Tx)(y) &= \int_D x(t)\operatorname{sign}(x(t))|x(t)|^{q-1}\, dt = \int_D |x(t)|^q\, dt \\ &= \left(\int_D |x(t)|^q\, dt\right)^{\frac{1}{q}} \left(\int_D |x(t)|^q\, dt\right)^{\frac{1}{p}} = \|x\|_q\, \|y\|_p .\end{aligned}$$

In order to prove that T is surjective, for a given $y^* \in L^p(D;\mathbb{K})^*$ one constructs an $x \in L^q(D;\mathbb{K})$ satisfying $Tx = y^*$ by employing the Radon-Nikodym theorem from measure and integration theory.

One also has the result that

$$L^1(D;\mathbb{K})^* \cong L^\infty(D;\mathbb{K})\,. \tag{1.77}$$

The representation theorem of Riesz states that, for compact sets $D \subset \mathbb{R}^n$, the space $C(D;\mathbb{R})^*$ is isometrically isomorphic to the space of all signed regular measures on the Borel σ-algebra on D. In particular, for every $y^* \in C(D;\mathbb{R})^*$ there exists a measure μ such that

$$y^*(y) = \int_D y\,d\mu\,. \tag{1.78}$$

Definition 1.21 (Series in normed spaces)
Let X be a normed space, let $(x_k)_{k\in\mathbb{N}}$ be a sequence in X. If the sequence

$$s_n = \sum_{k=1}^n x_k \tag{1.79}$$

of the partial sums converges to an element $s \in X$, we say that the corresponding series $\sum_{k=1}^\infty x_k$ converges, and we define

$$\sum_{k=1}^\infty x_k = s\,. \tag{1.80}$$

*The series $\sum_{k=1}^\infty x_k$ is called **absolutely convergent** if*

$$\sum_{k=1}^\infty \|x_k\| < \infty\,. \tag{1.81}$$

□

Because addition and scalar multiplication in normed spaces are continuous operations, we have the rules

$$\sum_{k=1}^\infty (x_k + y_k) = \sum_{k=1}^\infty x_k + \sum_{k=0}^\infty y_k\,, \quad \sum_{k=1}^\infty \alpha x_k = \alpha \sum_{k=1}^\infty x_k\,, \tag{1.82}$$

$\alpha \in \mathbb{K}$. They are valid when the limits on the right side exist.

Proposition 1.22 *Let X be a Banach space, assume that the series $\sum_{k=1}^\infty x_k$ converges absolutely. Then it also converges, and*

$$\left\| \sum_{k=1}^\infty x_k \right\| \le \sum_{k=1}^\infty \|x_k\|\,. \tag{1.83}$$

Moreover, every reordering of the series converges, and the limits are identical.

Proof: Let

$$\sigma_n = \sum_{k=1}^n \|x_k\|\,.$$

For the partial sums defined in (1.79) we have, if $n > m$,

$$\|s_n - s_m\| \leq \sum_{k=m+1}^{n} \|x_k\| = |\sigma_n - \sigma_m| \,.$$

Since (σ_n) is a Cauchy sequence in $\mathbb{K}$, also (s_n) is a Cauchy sequence in X, hence convergent to some $s \in X$. Due to $\|s_n\| \leq |\sigma_n|$, (1.83) follows from

$$\|s\| = \lim_{n\to\infty} \|s_n\| \leq \lim_{n\to\infty} |\sigma_n| = \sum_{k=1}^{\infty} \|x_k\| \,.$$

Let

$$\sum_{k=1}^{\infty} \tilde{x}_k \,, \quad \tilde{x}_k = x_{\pi(k)} \,, \quad \pi : \mathbb{N} \to \mathbb{N} \quad \text{bijective,}$$

be a reordering of $\sum x_k$ with the partial sums

$$\tilde{s}_n = \sum_{k=1}^{n} \tilde{x}_k \,.$$

Let $\varepsilon > 0$. We choose M large enough such that

$$\sum_{k=M+1}^{\infty} \|x_k\| \leq \varepsilon \,.$$

Next, we choose N such that $N \geq M$ and $\pi(\{1, \ldots, N\}) \supset \{1, \ldots, M\}$. Then we have for all $n > N$

$$\|\tilde{s}_n - s_n\| \leq \sum_{k=M+1}^{\infty} \|x_k\| \leq \varepsilon \,,$$

therefore $\|\tilde{s}_n - s_n\| \to 0$. $\square$

2 Hilbert Spaces

Definition 2.1 (Scalar product)
*Let X be a vector space over $\mathbb{K}$. A mapping $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{K}$ is called a **scalar product** on X, if*

$$\langle x, x\rangle > 0\,, \quad \text{for all } x \in X \text{ with } x \neq 0, \tag{2.1}$$

$$\langle x, y\rangle = \overline{\langle y, x\rangle}\,, \quad \text{for all } x, y \in X, \tag{2.2}$$

$$\langle \alpha x + \beta y, z\rangle = \alpha \langle x, z\rangle + \beta \langle y, z\rangle\ , \quad \text{for all } x, y, z \in X \text{ and all } \alpha, \beta \in \mathbb{K}. \tag{2.3}$$

$(X, \langle\cdot,\cdot\rangle)$ is called a prehilbert space. □

From (2.2) and (2.3) we immediately obtain

$$\langle x, \alpha y + \beta z\rangle = \overline{\alpha} \langle x, y\rangle + \overline{\beta} \langle x, z\rangle\ , \quad \text{for all } x, y, z \in X \text{ and all } \alpha, \beta \in \mathbb{K}. \tag{2.4}$$

In the case $\mathbb{K} = \mathbb{R}$, a scalar product is nothing else than a symmetric positive definite bilinear form.

Proposition 2.2 (Cauchy-Schwarz inequality, Hilbert space)
Let $(X, \langle\cdot,\cdot\rangle)$ be a prehilbert space. Then

$$\|x\| = \sqrt{\langle x, x\rangle} \tag{2.5}$$

defines a norm on X which satisfies the Cauchy-Schwarz inequality

$$|\langle x, y\rangle| \le \|x\|\,\|y\|\,, \quad \text{for all } x, y \in X. \tag{2.6}$$

*If X is complete, then X is called a **Hilbert space**.*

Proof: First, we prove (2.6). Let $x, y \in X$ with $y \neq 0$. For every $\alpha \in \mathbb{K}$ we have

$$0 \le \langle x + \alpha y, x + \alpha y\rangle = \langle x, x\rangle + \alpha\overline{\langle x, y\rangle} + \overline{\alpha}\langle x, y\rangle + |\alpha|^2 \langle y, y\rangle\ .$$

We set

$$\alpha = -\frac{\langle x, y\rangle}{\langle y, y\rangle}\,.$$

Then

$$0 \le \langle x, x\rangle - 2\frac{\langle x, y\rangle\,\overline{\langle x, y\rangle}}{\langle y, y\rangle} + \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle}\,,$$

and moreover

$$0 \le \langle x, x\rangle \langle y, y\rangle - |\langle x, y\rangle|^2\,.$$

This proves (2.6). Now we show that the properties of a norm are satisfied. From $\|x\| = 0$ it follows that $\langle x, x\rangle = 0$ and therefore $x = 0$ because of (2.1). For $\alpha \in \mathbb{K}$ and $x \in X$ we have

$$\|\alpha x\|^2 = \langle \alpha x, \alpha x\rangle = \alpha\overline{\alpha}\langle x, x\rangle = |\alpha|^2\|x\|^2\,.$$

The triangle inequality holds since

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2 \le \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2 .\end{aligned}$$

□

From $\langle x,y\rangle = \overline{\langle y,x\rangle}$ it follows that Re $\langle x,y\rangle =$ Re $\langle y,x\rangle$, and we have

$$\|x+y\|^2 = \|x\|^2 + 2\mathrm{Re}\,\langle x,y\rangle + \|y\|^2 , \quad \text{for all } x,y \in X. \tag{2.7}$$

The Cauchy-Schwarz inequality directly implies (see the exercises) that the scalar product $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{K}$ is continuous.

Examples for Hilbert spaces:

$$X = \mathbb{K}^n , \quad \langle x,y\rangle = \sum_{k=1}^{n} x_k \overline{y_k} , \tag{2.8}$$

also

$$X = \ell^2(\mathbb{K}) , \quad \langle x,y\rangle = \sum_{k=1}^{\infty} x_k \overline{y_k} , \tag{2.9}$$

and

$$X = L^2(D;\mathbb{K}) , \quad \langle x,y\rangle = \int_D x(t)\overline{y(t)}\,dt , \tag{2.10}$$

where $D \subset \mathbb{R}^n$ is open. For the examples (2.9) and (2.10), it follows from Hölder's inequality that the scalar product is well defined.

For these spaces, the norm defined in (2.5) is identical with those considered in the previous chapter.

Proposition 2.3 (Parallelogram identity)
Let X be a prehilbert space. Then

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) , \quad \textit{for all } x,y \in X. \tag{2.11}$$

Proof: A direct computation using definition (2.5) and the properties of a scalar product.
□

Similarly, one can compute directly that, in a prehilbert space, the scalar product can be expressed by the norm. Namely,

$$\langle x,y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) , \quad \mathbb{K} = \mathbb{R} , \tag{2.12}$$

and

$$\langle x,y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2) , \quad \mathbb{K} = \mathbb{C} . \tag{2.13}$$

These identities are called **polarization** identities.

Let X be a normed space in which the parallelogram identity (2.11) does not hold. Then there cannot exist a scalar product which generates this norm according to (2.5). Using this argument one can show (exercise) that $(C(D;\mathbb{K}), \|\cdot\|_\infty)$, D compact metric space, is not a prehilbert space.

If (2.11) holds in a given normed space X, one can prove that (2.12) resp. (2.13) defines a scalar product; thus, X then becomes a prehilbert space (we do not give the proof).

Proposition 2.4 (Projection)
Let X be a Hilbert space, let $K \subset X$ nonempty, convex and closed. Then for every $x \in X$ there exists a unique $y \in K$ with

$$\|x - y\| = \inf_{z \in K} \|x - z\| . \tag{2.14}$$

The association $x \mapsto y$ thus defines a mapping $P_K : X \to K$, called the ***projection on K****.*

Proof: Let $x \in X$, let $(y_n)_{n\in\mathbb{N}}$ be a sequence in K such that

$$\lim_{n\to\infty} \|x - y_n\| = \inf_{z\in K} \|x - z\| =: d .$$

With the aid of the parallelogram identity we get

$$2(\|x - y_n\|^2 + \|x - y_m\|^2) = \|2x - (y_n + y_m)\|^2 + \|y_n - y_m\|^2 ,$$

so

$$\|y_n - y_m\|^2 = 2(\|x - y_n\|^2 + \|x - y_m\|^2) - 4\left\|x - \frac{y_n + y_m}{2}\right\|^2 . \tag{2.15}$$

As K is convex,

$$\frac{y_n + y_m}{2} \in K ,$$

therefore

$$0 \le \|y_n - y_m\|^2 \le 2(\|x - y_n\|^2 + \|x - y_m\|^2) - 4d^2 . \tag{2.16}$$

Thus, $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X. As X is complete, there exists

$$y = \lim_{n\to\infty} y_n . \tag{2.17}$$

Since K is closed, we have $y \in K$. The continuity of the norm implies

$$\|x - y\| = \lim_{n\to\infty} \|x - y_n\| = d .$$

Therefore we have proved the existence of a y satisfying (2.14). Let now $\tilde{y} \in K$ with $\|x - \tilde{y}\| = d$. Then as in (2.15) we conclude that

$$\begin{aligned}\|y - \tilde{y}\|^2 &= 2(\|x - y\|^2 + \|x - \tilde{y}\|^2) - 4\left\|x - \frac{y + \tilde{y}}{2}\right\|^2 = 4d^2 - 4\left\|x - \frac{y + \tilde{y}}{2}\right\|^2 \\ &\le 0 , \quad \text{as} \quad \frac{y + \tilde{y}}{2} \in K ,\end{aligned}$$

so $y = \tilde{y}$. □

Proposition 2.5 (Variational inequality)
Let X be a Hilbert space, $K \subset X$ nonempty, convex and closed. Then for every $x \in X$ there exists a unique $y \in K$ with

$$\mathrm{Re}\,\langle x-y, z-y\rangle \le 0\,, \quad \text{for all } z \in K, \tag{2.18}$$

and we have $y = P_K x$.

Proof: Uniqueness: Let $y, \tilde{y} \in K$ be solutions of (2.18). Then

$$\begin{aligned} 0 &\ge \mathrm{Re}\,\langle x-y, \tilde{y}-y\rangle\,, \\ 0 &\ge \mathrm{Re}\,\langle x-\tilde{y}, y-\tilde{y}\rangle = \mathrm{Re}\,\langle \tilde{y}-x, \tilde{y}-y\rangle\,. \end{aligned}$$

Adding these inequalities yields

$$0 \ge \mathrm{Re}\,\langle \tilde{y}-y, \tilde{y}-y\rangle = \|\tilde{y}-y\|^2\,,$$

and therefore $y = \tilde{y}$. It remains to show that $y = P_K x$ is a solution. Let $z \in K$, $t \in [0,1]$ be arbitrary. Then $(1-t)P_K x + tz \in K$, so

$$\begin{aligned} \|x-P_Kx\|^2 &\le \|x-(1-t)P_Kx - tz\|^2 = \|(x-P_Kx)+t(P_Kx-z)\|^2 \\ &= \|x-P_Kx\|^2 + 2\,\mathrm{Re}\,\langle x-P_Kx, t(P_Kx-z)\rangle + t^2\|P_Kx-z\|^2\,. \end{aligned}$$

Dividing by t gives

$$0 \le 2\,\mathrm{Re}\,\langle x-P_Kx, P_Kx-z\rangle + t\|P_Kx-z\|^2\,.$$

Passing to the limit $t \to 0$ yields the assertion. □

When $\mathbb{K} = \mathbb{R}$, (2.18) becomes

$$\langle x-y, z-y\rangle \le 0\,, \quad \text{for all } z \in K. \tag{2.19}$$

Corollary 2.6 *Let X be a Hilbert space, $K \subset X$ nonempty, convex and closed. Then we have*

$$\|P_Kx - P_K\tilde{x}\| \le \|x-\tilde{x}\|\,, \quad \text{for all } x, \tilde{x} \in X. \tag{2.20}$$

Proof: Exercise. □

Corollary 2.7 *Let X be a Hilbert space, U a closed subspace of X. Then P_Ux is the unique element $y \in U$ which satisfies*

$$\langle x-y, z\rangle = 0\,, \quad \text{for all } z \in U. \tag{2.21}$$

Proof: If $z \in U$ then also $z + P_Ux \in U$. Then (2.18) implies

$$\mathrm{Re}\,\langle x-P_Ux, z\rangle \le 0\,, \quad \text{for all } z \in U. \tag{2.22}$$

Replacing z by $-z$ we see that $\mathrm{Re}\,\langle x-P_Ux, z\rangle = 0$; replacing z by iz yields $\mathrm{Im}\,\langle x-P_Ux, z\rangle = 0$, so (2.21) holds. Conversely every $y \in U$ which satisfies (2.21) also solves the variational inequality (2.18), since $z-y \in U$ for all $z \in U$. □

Definition 2.8 (Orthogonality)
Let X be a prehilbert space. If $x, y \in X$ satisfy

$$\langle x, y \rangle = 0 \,, \tag{2.23}$$

we say that x and y are ***orthogonal****, and write $x \perp y$. If $Y \subset X$, we define the* ***orthogonal complement*** *$Y^\perp$ of Y by*

$$Y^\perp = \{z : z \in X,\ z \perp y \text{ for all } y \in Y\} \,. \tag{2.24}$$

□

If $x \perp y$ then (Pythagoras)

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 \,.$$

Lemma 2.9 *Let X be a prehilbert space, $Y \subset X$. Then $Y^\perp$ is a closed subspace of X with $Y^\perp \cap Y = \{0\}$.*

Proof: This follows directly from the fact that the scalar product is bilinear, continuous and definite. □

Proposition 2.10 *Let X be a Hilbert space, U a closed subspace of X. Then the projection $P_U : X \to U$ is linear and continuous, $\ker(P_U) = U^\perp$, and*

$$P_{U^\perp} = id - P_U \,. \tag{2.25}$$

Proof: By Corollary 2.7 we have

$$\langle x - P_U x, z \rangle = 0 \,, \quad \text{for all } x \in X,\ z \in U. \tag{2.26}$$

Therefore $x - P_U x \in U^\perp$. For $x, \tilde{x} \in X$ and $\alpha, \beta \in \mathbb{K}$ we have

$$\alpha P_U x + \beta P_U \tilde{x} \in U \,,$$
$$\langle (\alpha x + \beta \tilde{x}) - (\alpha P_U x + \beta P_U \tilde{x}), z \rangle = 0 \,, \quad \text{for all } z \in U,$$

so, again using Corollary 2.7,

$$P_U(\alpha x + \beta \tilde{x}) = \alpha P_U x + \beta P_U \tilde{x} \,.$$

From (2.26) and Lemma 2.9 it follows that

$$x \in U^\perp \quad \Leftrightarrow \quad P_U x \in U^\perp \quad \Leftrightarrow \quad P_U x = 0 \,.$$

For $z \in U^\perp$ we have

$$\langle x - (id - P_U)x, z \rangle = \langle P_U x, z \rangle = 0 \,,$$

and because of $x - P_U x \in U^\perp$ we conclude that

$$(id - P_U)x = P_{U^\perp} x \,.$$

$\square$

From (2.25) we immediately obtain

$$\|x\|^2 = \|x - P_U x\|^2 + \|P_U x\|^2$$

and therefore

$$\|P_U\| = 1\,, \quad \text{if } U \neq \{0\}.$$

The mapping $x \mapsto (P_U x, x - P_U x)$ yields an isometric isomorphism

$$(X, \|\cdot\|) \cong (U \times U^\perp, \|\cdot\|_2)\,.$$

We may X also interpret as a direct sum

$$X = U \oplus U^\perp$$

in the sense of linear algebra.

Lemma 2.11 *Let X be a Hilbert space, $Y \subset X$. Then*

$$Y^{\perp\perp} = \overline{\operatorname{span} Y}\,. \tag{2.27}$$

In particular, we have $Y^{\perp\perp} = Y$ if Y is a closed subspace.

Proof: Exercise. $\square$

Proposition 2.12 (Riesz representation theorem)
Let X be a Hilbert space. Then for every $x^* \in X^*$ there exists a unique $x \in X$ satisfying

$$x^*(z) = \langle z, x \rangle\,, \quad \textit{for all } z \in X. \tag{2.28}$$

The mapping $J : X \to X^*$ defined by (2.28) is bijective, isometric and conjugate linear, that is,

$$J(\alpha x + \beta y) = \overline{\alpha} J(x) + \overline{\beta} J(y) \tag{2.29}$$

holds for all $x, y \in X$ and all $\alpha, \beta \in \mathbb{K}$.

Proof: Formula (2.28) defines a linear functional $x^* : X \to \mathbb{K}$. Since

$$|x^*(z)| \le \|x\|\,\|z\|\,, \quad |x^*(x)| = \|x\|^2\,,$$

we have $x^* \in X^*$. Thus J is well-defined, conjugate linear and satisfies $\|J(x)\| = \|x\|$. It remains to show that J is surjective. Let $x^* \in X^*$ with $x^* \neq 0$. We set $U = \ker(x^*)$. Since $U \neq X$, we get $U^\perp \neq \{0\}$. We choose an $x \in U^\perp$ with $x^*(x) = 1$. For arbitrary $z \in X$ we have

$$z = z - x^*(z)x + x^*(z)x\,, \quad z - x^*(z)x \in U\,.$$

This implies

$$\langle z, x \rangle = \underbrace{\langle z - x^*(z)x, x \rangle}_{=0} + \langle x^*(z)x, x \rangle = x^*(z)\,\langle x, x \rangle\,,$$

and therefore

$$x^* = J\left(\frac{x}{\|x\|^2}\right) .$$

$\square$

For the Hilbert space $X = L^2(D;\mathbb{K})$, Proposition 2.12 yields that for every functional $x^* \in X^*$ there exists a unique function $x \in L^2(D;\mathbb{K})$ such that

$$x^*(z) = \int_D z(t)\overline{x}(t)\,dt\,, \quad \text{for all } z \in L^2(D;\mathbb{K}).$$

Let U be a closed subspace of a Hilbert space X with $U \neq X$. Then there exists $x \in X$ satisfying

$$\operatorname{dist}(x,U) = \inf_{z\in U}\|x-z\| = 1\,, \quad \|x\| = 1\,, \tag{2.30}$$

because every $x \in U^\perp$ with $\|x\| = 1$ has this property.

In a Banach space which is not a Hilbert space, the projection theorem in general does not hold, not even when the convex set K is a subspace, and an $x \in X$ which satisfies (2.30) need not exist. Instead, the following weaker result holds.

Lemma 2.13 *Let X be a normed space, U a closed subspace of X with $U \neq X$. Then there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in X with*

$$\lim_{n\to\infty} \operatorname{dist}(x_n,U) = 1\,, \quad \|x_n\| = 1 \quad \textit{for all } n \in \mathbb{N}. \tag{2.31}$$

Proof: Let $\tilde{x} \in X \setminus U$, let $(y_n)_{n\in\mathbb{N}}$ be a sequence in U with

$$\lim_{n\to\infty}\|\tilde{x} - y_n\| = \operatorname{dist}(\tilde{x},U)\,.$$

Since U is closed, we have that $\operatorname{dist}(\tilde{x},U) > 0$. We define

$$x_n = \frac{\tilde{x} - y_n}{\|\tilde{x} - y_n\|}\,. \tag{2.32}$$

For arbitrary $z \in U$ we have

$$\|x_n - z\| = \frac{1}{\|\tilde{x}-y_n\|}\|\tilde{x} - \underbrace{(y_n + \|\tilde{x}-y_n\|z)}_{\in U}\| \geq \frac{1}{\|\tilde{x}-y_n\|}\operatorname{dist}(\tilde{x},U)\,,$$

therefore

$$1 = \|x_n\| \geq \operatorname{dist}(x_n,U) \geq \frac{1}{\|\tilde{x}-y_n\|}\operatorname{dist}(\tilde{x},U) \to 1 \quad \text{for } n \to \infty.$$

$\square$

Corollary 2.14 *Let X be a normed space with* $\dim(X) = \infty$. *Then the closed unit ball $K(0;1) = \{x : x \in X,\ \|x\| \leq 1\}$ is not compact.*

Proof: $K(0;1)$ is a subset of X and thus a metric space. It therefore suffices to construct a sequence $(x_n)_{n\in\mathbb{N}}$ in $K(0;1)$ which does not have a convergent subsequence. Let $x_1 \in K(0;1)$ with $\|x_1\| = 1$. Assume that $x_1, \ldots, x_n$ are already defined. Then choose x_{n+1} according to Lemma 2.13 such that $\|x_{n+1}\| = 1$ and

$$\operatorname{dist}(x_{n+1}, U_n) \geq \frac{1}{2}, \quad \text{where} \qquad U_n = \operatorname{span}\{x_1, \ldots, x_n\}.$$

Then we have

$$\|x_n - x_m\| \geq \frac{1}{2}, \quad n \neq m.$$

Therefore, $(x_n)_{n\in\mathbb{N}}$ has no convergent subsequence. □

If X has finite dimension, say $\dim(X) = n$, then $X \simeq \mathbb{K}^n$ by Proposition 1.6. Thus we have proved: **In a normed space, the unit ball is compact if and only if the space has finite dimension.**

Definition 2.15 (Orthonormal basis)
Let X be a Hilbert space. A set $S \subset X$ is called an ***orthonormal system*** *in X, if $\|e\| = 1$ for all $e \in S$ and $e \perp f$ for all $e, f \in S$ with $e \neq f$. An orthonormal system S is called an* ***orthonormal basis****, if there does not exist an orthonormal system $\tilde{S}$ with $S \subset \tilde{S}$ and $S \neq \tilde{S}$.* □

An orthonormal basis is also called a "complete orthonormal system".

We know from linear algebra that every vector space has a basis. In an analogous manner to that proof one can also prove that every Hilbert space has an orthonormal basis.

Proposition 2.16 *Let X be a Hilbert space, $X \neq \{0\}$. Then X has orthonormal basis. For every orthonormal system S_0 there exists an orthonormal basis S with $S_0 \subset S$.*

Proof: The first assertion follows from the second, since X has an orthonormal system; namely, $S_0 = \{x/\|x\|\}$ is an orthonormal system whenever $x \in X$, $x \neq 0$. Let now S_0 be an arbitrary orthonormal system. We consider the set

$$\mathcal{M} = \{S : S \supset S_0, S \text{ is an orthonormal system}\}.$$

The set $\mathcal{M}$ is not empty, and the set inclusion defines a partial order "$\leq$" on $\mathcal{M}$,

$$S_1 \leq S_2 \quad \Leftrightarrow \quad S_1 \subset S_2.$$

Let $\mathcal{V}$ be a completely ordered subset of $\mathcal{M}$ (that is, for arbitrary $S_1, S_2 \in \mathcal{V}$ we have $S_1 \leq S_2$ or $S_2 \leq S_1$). Then

$$T = \bigcup_{S\in\mathcal{V}} S$$

also is an element of $\mathcal{M}$ (because it follows from $e, f \in T$ that there exists an $S \in \mathcal{V}$ such that $e, f \in S$, so $\|e\| = \|f\| = 1$ and $\langle e, f\rangle = 0$). Therefore, T is an upper bound of $\mathcal{V}$ in $\mathcal{M}$. Consequently, every totally ordered subset of $\mathcal{M}$ has an upper bound. It now follows from Zorn's lemma that $\mathcal{M}$ possesses a maximal element S, that is, an element for which there does not exist a $\tilde{S} \in \mathcal{M}$ with $S \subset \tilde{S}$ and $S \neq \tilde{S}$. Therefore, S is an orthonormal basis. □

Proposition 2.17 *Let X be a separable Hilbert space with $\dim(X) = \infty$. Then every orthonormal basis S has countably infinitely many elemente.*

Proof: If S is an uncountable orthonormal basis, the space X cannot be separable, since $\|e - f\| = \sqrt{2}$ holds for elements $e, f \in S$ with $e \neq f$ (the reasoning is analogous to that in the proof that $\ell^\infty(\mathbb{K})$ is not separable). On the other hand, if S is a finite orthonormal system, and $x \notin U = \operatorname{span}(S)$, then

$$S \cup \left\{ \frac{x - P_U x}{\|x - P_U x\|} \right\}$$

is also an orthonormal system. Thus S is not an orthonormal basis. □

Lemma 2.18 *Let X be a Hilbert space, $S \subset X$ an orthonormal system. Let $e_1, \ldots, e_n \in S$ and $U = \operatorname{span}\{e_1, \ldots, e_n\}$. Then*

$$P_U x = \sum_{k=1}^{n} \langle x, e_k \rangle \, e_k \,, \quad \text{for all } x \in X. \tag{2.33}$$

Moreover, Bessel's inequality

$$\sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \leq \|x\|^2 \tag{2.34}$$

holds.

Proof: Setting

$$y = \sum_{k=1}^{n} \langle x, e_k \rangle \, e_k$$

we have $y \in U$ and

$$\langle x - y, e_j \rangle = \langle x, e_j \rangle - \sum_{k=1}^{n} \langle x, e_k \rangle \langle e_k, e_j \rangle = \langle x, e_j \rangle - \langle x, e_j \rangle = 0 \,.$$

Consequently, $\langle x - y, z \rangle = 0$ for all $z \in U$ and therefore $y = P_U x$ according to Corollary 2.7. Bessel's inequality follows because of

$$\begin{aligned} \|P_U x\|^2 &= \left\langle \sum_{k=1}^{n} \langle x, e_k \rangle \, e_k, \sum_{j=1}^{n} \langle x, e_j \rangle \, e_j \right\rangle = \sum_{j,k=1}^{n} \langle x, e_k \rangle \, \overline{\langle x, e_j \rangle} \, \langle e_k, e_j \rangle \\ &= \sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \end{aligned}$$

from the inequality $\|P_U x\| \leq \|x\|$. □

Corollary 2.19 *Let X be a Hilbert space, $S \subset X$ an orthonormal system, $x \in X$. Then the set*

$$S_x = \{e : e \in S, \, \langle x, e \rangle \neq 0\} \tag{2.35}$$

is finite or countably infinite.

Proof: Because of (2.34), for any given $m \in \mathbb{N}$ there cannot exist more than $m\|x\|^2$ different $e \in S$ with $|\langle x, e\rangle| > 1/m$. $\square$

An orthonormal system S in X is linearly independent and therefore a vector space basis of the subspace $\operatorname{span}(S)$ generated by S. Orthonormal systems and bases are important, because one also can represent elements in the closure $\overline{\operatorname{span}(S)}$, as limits of a series

$$x = \sum_{k=1}^{\infty} \langle x, e_k\rangle\, e_k\,, \quad e_k \in S\,. \tag{2.36}$$

One may extend an orthonormal system S to a vector space basis B of X (by the corresponding result from linear algebra) and thus represent every element of X as a finite linear combination of elements of B. But such a basis B is uncountably infinite (as we will see in the next chapter), and one cannot describe it in a "constructive" manner. For this reason, vector space bases are unsuitable for the analysis of function spaces.

Proposition 2.20 *Let X be a separable Hilbert space with* $\dim(X) = \infty$*, let $S = \{e_k : k \in \mathbb{N}\}$ be an orthonormal system. Then there are equivalent*

(i) S is an orthonormal basis.

(ii) $S^\perp = \{0\}$.

(iii) $X = \overline{\operatorname{span}(S)}$.

(iv) For all $x \in X$ we have

$$x = \sum_{k=1}^{\infty} \langle x, e_k\rangle\, e_k\,. \tag{2.37}$$

(v) For all $x \in X$ there holds ***Parseval's equality***

$$\|x\|^2 = \sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2\,. \tag{2.38}$$

Proof: (i)⇒(ii): Let $x \in S^\perp$, $x \neq 0$. Then

$$S \cup \left\{\frac{x}{\|x\|}\right\}$$

is an orthonormal system, so S is not an orthonormal basis.
(ii)⇒(iii): By Lemma 2.11,

$$\overline{\operatorname{span}(S)} = S^{\perp\perp} = \{0\}^\perp = X\,.$$

(iii)⇒(iv): Let $U_m = \operatorname{span}\{e_1, \dots, e_m\}$, $P_m = P_{U_m}$. Let $x \in X$ be arbitrary, let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\operatorname{span}(S)$ with $x_n \to x$, let $(m_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence in $\mathbb{N}$ with $x_n \in U_{m_n}$. Then we have

$$0 \le \|x - P_{m_n}x\| \le \|x - x_n\| \to 0\,,$$

and since $(\|x - P_n x\|)_{n\in\mathbb{N}}$ is decreasing, it follows from Proposition 2.18 that

$$0 \le \|x - \sum_{k=1}^{n} \langle x, e_k\rangle\, e_k\| = \|x - P_n x\| \to 0\,.$$

(iv)⇒(v): The partial sums $s_n = \sum_{k=1}^{n} \langle x, e_k\rangle\, e_k$ satisfy

$$\|s_n\|^2 = \langle s_n, s_n\rangle = \sum_{k=1}^{n} |\, \langle x, e_k\rangle\,|^2\,.$$

Passing to the limit $n \to \infty$ yields (2.38).
(v)⇒(i): If S is not an orthonormal basis, then there exists an $x \in X$ with $\|x\| = 1$ such that $S \cup \{x\}$ becomes an orthonormal system. For such an x, (2.38) does not hold. □

From Proposition 2.20 it follows immediately that in $\ell^2(\mathbb{K})$ the unit vectors e_k form an orthonormal basis $S = \{e_k\}$, because $\mathrm{span}\,(S) = c_e(\mathbb{K})$ and $c_e(\mathbb{K})$ is dense in $\ell^2(\mathbb{K})$.

The series in (2.37) in general is **not** absolutely convergent. Example: For

$$x_k = \frac{1}{k}$$

we have $x = (x_k)_{k\in\mathbb{N}} \in \ell^2(\mathbb{K})$, but $\|\,\langle x, e_k\rangle\, e_k\|_2 = \frac{1}{k}$. Nevertheless, (2.38) shows that, in the situation of Proposition 2.20, the limit x does not depend on the order in which the basis vectors are chosen.

When X is a nonseparable Hilbert space, one can prove an analogous characterization, because for every $x \in X$ the scalar product $\langle x, e\rangle$ can be nonzero for at most countably infinitely many elements $e \in S$, due to Lemma 2.19.

Proposition 2.21 *Let X be a separable Hilbert space with* $\dim(X) = \infty$. *Then*

$$X \cong \ell^2(\mathbb{K})\,. \tag{2.39}$$

Proof: Let $S = \{e_k : k \in \mathbb{N}\}$ be an orthonormal basis of X. For given $x \in X$ we define a sequence Tx by

$$(Tx)_k = \langle x, e_k\rangle\ , \quad x \in X\,.$$

Parseval's equality implies that $Tx \in \ell^2(\mathbb{K})$ and that $\|Tx\|_2 = \|x\|_X$. Obviously, T is linear and injective. We define $R : c_e(\mathbb{K}) \to X$ by

$$R(y) = \sum_{k=1}^{m} y_k e_k\,, \quad \text{if } y = (y_1, \ldots, y_m, 0, \ldots).$$

Then R is linear and continuous (even isometric). Therefore, by Proposition 1.19 we can uniquely extend R to a linear continuous mapping $R : \ell^2(\mathbb{K}) \to X$. Let now $y \in \ell^2(\mathbb{K})$, let $(y^n)_{n\in\mathbb{N}}$ be a sequence in $c_e(\mathbb{K})$ with $y^n \to y$. Then $y^n = TRy^n \to TRy$, so $y = TRy$. Thus T is surjective. □

From Proposition 2.21 it follows that all separable Hilbert spaces of infinite dimension are isometrically isomorphic. As a somewhat surprising consequence of this result we see that

$$L^2(D; \mathbb{K}) \cong \ell^2(\mathbb{K})\,,$$

if D is an open subset of $\mathbb{R}^n$.

In $L^2((-\pi,\pi);\mathbb{C})$ we now consider the functions

$$e_k(t) = \frac{1}{\sqrt{2\pi}} e^{ikt}, \quad k \in \mathbb{Z}, \tag{2.40}$$

It has been proved in the basic lecture on analysis that

$$S = \{e_k : k \in \mathbb{Z}\} \tag{2.41}$$

is an orthonormal basis. Consequently we conclude from Proposition 2.20 that for every $x \in L^2((-\pi,\pi);\mathbb{C})$ we have

$$x = \sum_{k\in\mathbb{Z}} c_k e_k, \quad c_k = \langle x, e_k\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} x(s) e^{-iks}\, ds, \tag{2.42}$$

where the limit has to be interpreted as a limit in L^2. The series (2.42) is called the **Fourier series** of x.

3 The Principle of Uniform Boundedness

Many results of functional analysis rest on the following proposition.

Proposition 3.1 (Baire)
Let (X, d) be a complete metric space, let $(U_n)_{n\in\mathbb{N}}$ be a sequence of open subsets of X, let U_n be dense in X for all $n \in \mathbb{N}$. Then

$$D = \bigcap_{n\in\mathbb{N}} U_n \tag{3.1}$$

is dense in X.

Proof: It suffices to show: if $V \subset X$ is open and not empty, then $V \cap D \neq \emptyset$. Let V be open, $V \neq \emptyset$. We choose $x_1 \in U_1 \cap V$ (this is possible as U_1 is dense in X) and an $\varepsilon_1 > 0$ such that

$$x_1 \in K(x_1; \varepsilon_1) \subset U_1 \cap V\,. \tag{3.2}$$

Assume that $(x_1, \varepsilon_1), \ldots, (x_{n-1}, \varepsilon_{n-1})$ are defined already. We choose

$$x_n \in U_n \cap B(x_{n-1}; \varepsilon_{n-1}) \qquad \text{(possible since } U_n \text{ is dense in } X) \tag{3.3}$$

and $\varepsilon_n > 0$ such that

$$x_n \in K(x_n; \varepsilon_n) \subset U_n \cap B(x_{n-1}; \varepsilon_{n-1})\,, \quad \varepsilon_n \leq \frac{1}{2}\varepsilon_{n-1}\,. \tag{3.4}$$

Then

$$d(x_{n+1}, x_n) \leq \varepsilon_n \leq 2^{-n+1}\varepsilon_1\,, \quad d(x_m, x_n) \leq 2 \cdot 2^{-n+1}\varepsilon_1 \quad \text{for all } m > n.$$

Therefore, $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Let $x = \lim_{n\to\infty} x_n$. We have $K(x_n; \varepsilon_n) \subset K(x_{n-1}; \varepsilon_{n-1})$ for all $n \in \mathbb{N}$ and therefore $x_m \in K(x_n; \varepsilon_n)$ for all $m > n$. Since $K(x_n; \varepsilon_n)$ is closed, it follows that $x \in K(x_n; \varepsilon_n) \subset V$ for all n. Consequently,

$$x \in \bigcap_{n\in\mathbb{N}} K(x_n; \varepsilon_n) \subset \bigcap_{n\in\mathbb{N}} U_n \cap V = D \cap V\,.$$

$\Box$

Definition 3.2 *Let (X, d) be a metric space, $M \subset X$. M is called **nowhere dense** in X, if $\operatorname{int}(\overline{M}) = \emptyset$ (that is, the closure of M has no interior points).* $\Box$

A closed set $A \subset X$ is nowhere dense in X if and only if its complement $X \setminus A$ is open and dense in X.

Corollary 3.3 *Let (X, d) be a complete metric space, let $(A_n)_{n\in\mathbb{N}}$ be a sequence of closed and nowhere dense subsets of X. Then*

$$X \setminus \bigcup_{n\in\mathbb{N}} A_n \tag{3.5}$$

is dense in X.

Proof: We apply Proposition 3.1, setting $U_n = X \setminus A_n$. $\square$

Proposition 3.1 and Corollary 3.3 imply in particular that $D = \cap_n U_n$ resp. $X \setminus \cup_n A_n$ are nonempty. Thus, they provide a way to prove existence of certain objects. For example one can show in this manner (see Werner p.139) that the set of all continuous but nowhere differentiable functions forms a dense subset of $(C([a,b]), \|\cdot\|_\infty)$.

We present another existence result.

Corollary 3.4 *Let (X,d) be a complete metric space, let $(A_n)_{n\in\mathbb{N}}$ be a sequence of closed subsets of X such that*

$$X = \bigcup_{n\in\mathbb{N}} A_n \,. \tag{3.6}$$

Then there exists a $k \in \mathbb{N}$ such that $\mathrm{int}\,(A_k) \neq \emptyset$.

Proof: This is a direct consequence of Corollary 3.3. $\square$

From Corollary 3.4 one can conclude (exercise) for example that there does not exist a Banach space which has a countably infinite vector space basis.

Subsets of X of the form

$$\bigcup_{n\in\mathbb{N}} M_n \,, \quad M_n \text{ nowhere dense in } X,$$

are also called sets **of first category** in X. Subsets of X which are not of first category are called **of second category** in X. When one uses this terminology, a complete metric space is always of second category (in itself).

The statement of the following proposition is called the principle of uniform boundedness.

Proposition 3.5 (Banach-Steinhaus)
Let X be a Banach space, Y a normed space, let $\mathcal{T} \subset L(X;Y)$. Assume that

$$\sup_{T\in\mathcal{T}} \|Tx\| < \infty \,, \quad \textit{for all } x \in X. \tag{3.7}$$

Then

$$\sup_{T\in\mathcal{T}} \|T\| < \infty \,. \tag{3.8}$$

Thus, if a family of linear and continuous operators is pointwise bounded, it is uniformly bounded.
Proof: For $n \in \mathbb{N}$ we define

$$A_n = \{x : x \in X, \sup_{T\in\mathcal{T}} \|Tx\| \le n\} \,. \tag{3.9}$$

A_n is closed, since for $f_T(x) := \|Tx\|$ we have

$$A_n = \bigcap_{T\in\mathcal{T}} f_T^{-1}([0,n]) \,.$$

By assumption (3.7) we have

$$X = \bigcup_{n \in \mathbb{N}} A_n \,.$$

Due to Corollary 3.4 we can choose a $k \in \mathbb{N}$ such that $\operatorname{int}(A_k) \neq \emptyset$. Let $x_0 \in A_k$, $\varepsilon > 0$ with $K(x_0;\varepsilon) \subset A_k$. Let now $x \in X$ be arbitrary, $x \neq 0$, then

$$z = x_0 + \varepsilon \frac{x}{\|x\|} \in K(x_0;\varepsilon) \subset A_k \,.$$

In addition, let $T \in \mathcal{T}$ be arbitrary. Then

$$\|Tx\| = \left\| T\left(\frac{\|x\|}{\varepsilon}(z - x_0)\right) \right\| = \frac{\|x\|}{\varepsilon}\|Tz - Tx_0\| \leq \frac{\|x\|}{\varepsilon} 2k \,,$$

and therefore

$$\|T\| \leq \frac{2k}{\varepsilon}\,, \quad \text{for all } T \in \mathcal{T}.$$

□

Definition 3.6 (Open Mapping) *Let (X, d_X), (Y, d_Y) be metric spaces. A mapping $T : X \to Y$ is called* ***open*** *if $T(U)$ is open in Y for every open subset U of X.* □

When the mapping T is open, the image of closed subset of X is not necessarily closed in Y; not even when T is linear (Example: exercise.)

Proposition 3.7 (Open mapping theorem)
Let X, Y be Banach spaces, $T : X \to Y$ linear and continuous. Then there are equivalent:

(i) T is open.

(ii) There exists $\delta > 0$ such that $B(0;\delta) \subset T(B(0;1))$.

(iii) T is surjective.

Proof: "(i)⇒(ii)": obvious.
"(ii)⇒(i)": Let U be open in X, let $y \in T(U)$. We choose $x \in U$ such that $Tx = y$, and $\varepsilon > 0$ such that $B(x;\varepsilon) \subset U$. Then

$$T(U) \supset T(B(x;\varepsilon)) = Tx + T(B(0;\varepsilon)) \supset y + B(0;\delta\varepsilon)\,,$$

where $\delta > 0$ is chosen according to (ii). Then $y \in \operatorname{int}(T(U))$.
"(ii)⇒(iii)": As T is linear, it follows from (ii) that $B(0;r\delta) \subset T(B(0;r))$ for all $r > 0$.
"(iii)⇒(ii)": Since T is surjective, we have

$$Y = \bigcup_{n \in \mathbb{N}} \overline{T(B(0;n))}\,.$$

According to Corollary 3.4 we choose a $k \in \mathbb{N}$ such that $\operatorname{int}(\overline{T(B(0;k))}) \neq \emptyset$. We set $V = \overline{T(B(0;k))}$. Let now $y \in V$, $\varepsilon > 0$ such that

$$B(y;\varepsilon) \subset V\,.$$

Since V is symmetric (that is, $z \in V \Rightarrow -z \in V$), we see that $B(-y;\varepsilon) \subset V$. Since V is convex, it follows that

$$B(0;\varepsilon) \subset V = \overline{T(B(0;k))}\,. \tag{3.10}$$

It now suffices to show that

$$B(0;\varepsilon) \subset T(B(0;3k))\,, \tag{3.11}$$

because then (ii) is satisfied for $\delta = \varepsilon/3k$. By (3.10), for every $y \in B(0;\varepsilon)$ we find an $x \in B(0;k)$ such that $\|y - Tx\| < \varepsilon/2$, so

$$2(y - Tx) \in B(0;\varepsilon)\,. \tag{3.12}$$

Let now $y_0 \in B(0;\varepsilon)$ be given. We construct sequences (x_j) in X, (y_j) in Y by choosing $x_0, y_1, x_1, y_2, x_2, \dots$ such that

$$x_j \in B(0;k)\,, \quad y_{j+1} = 2(y_j - Tx_j)\,, \quad y_{j+1} \in B(0;\varepsilon)\,, \quad j \in \mathbb{N}\,. \tag{3.13}$$

It follows that

$$2^{-(j+1)} y_{j+1} = 2^{-j} y_j - T(2^{-j} x_j)\,, \quad j \in \mathbb{N}\,,$$

therefore

$$T\left(\sum_{j=0}^{m} 2^{-j} x_j\right) = y_0 - 2^{-(m+1)} y_{m+1} \to y_0 \tag{3.14}$$

f"ur $m \to \infty$. Since

$$\sum_{j=0}^{m} 2^{-j} \|x_j\| \le 2k\,, \tag{3.15}$$

the series $\sum_j 2^{-j} x_j$ is absolutely convergent, hence convergent (as X is a Banach space). Let

$$x = \sum_{j=0}^{\infty} 2^{-j} x_j\,.$$

Now (3.14) implies that $Tx = y_0$, and (3.15) implies that $\|x\| \le 2k < 3k$. Thus (3.11) is proved. □

Corollary 3.8 *Let X, Y be Banach spaces, $T : X \to Y$ linear, continuous and bijective. Then $T^{-1} : Y \to X$ is linear and continuous.*

Proof: That T^{-1} is linear, is a result of linear algebra. It follows from Proposition 3.7 that T is open. Therefore, for every open $U \subset X$ it follows that $(T^{-1})^{-1}(U) = T(U)$ is open. Thus, T^{-1} is continuous. □

Let us consider the situation when a vector space X is endowed with two different norms $\|\cdot\|_1$ and $\|\cdot\|_2$ so that

$$\|x\|_1 \le C\|x\|_2\,, \quad \text{for all } x \in X, \tag{3.16}$$

holds with a constant C which does not depend on x. This means that $\|\cdot\|_1$ is weaker than $\|\cdot\|_2$ in the sense that convergence w.r.t. $\|\cdot\|_2$ implies convergence w.r.t. $\|\cdot\|_1$, but not the other way round. (3.16) means that

$$id : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$$

is continuous. We now conclude from Corollary 3.8 that

$$id : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$$

is continuous, too, if $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ both are Banach spaces; in this case it follows that both norms are equivalent, and

$$(X, \|\cdot\|_1) \simeq (X, \|\cdot\|_2) .$$

If (3.16) holds, but the two norms are not equivalent, then at least one of the spaces $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ is not complete. Example: $X = C([a, b])$. $C([a, b])$ is complete for $\|\cdot\|_2 = \|\cdot\|_\infty$, and for the L^1-Norm we have

$$\|x\|_1 \le (b - a)\|x\|_\infty ,$$

but these two norms are not equivalent (there exist sequences $(x_n)_{n\in\mathbb{N}}$ with $\|x_n\|_1 \to 0$, but $\|x_n\|_\infty \not\to 0$). Therefore $(C([a, b]); \|\cdot\|_1)$ is not complete.

Corollary 3.9 *Let X, Y be Banach spaces, $T : X \to Y$ linear and continuous, assume that $\overline{T(X)} = Y$ and $T(X) \neq Y$. Then there exists $y \in Y$ such that $\|x_n\| \to \infty$ holds for every sequence $(x_n)_{n\in\mathbb{N}}$ in X which satisfies $\lim_{n\to\infty} Tx_n = y$.*

Proof: Exercise. □

In the situation of Corollary 3.9, problems arise if one wants to solve $Tx = y$ for a given y in a "stable" manner. The corollary implies that there exists a sequence (y_n) in Y of approximations y_n of y, namely $y_n \to y$, for which $\|x\|_n \to \infty$ holds for the exact solutions x_n of the equation $Tx = y_n$. We might want to use (x_n) as a sequence of approximate solutions of $Tx = y$; but they are worthless for large values of n.

The graph of a mapping $T : X \to Y$ is defined as

$$\mathrm{graph}\,(T) = \{(x, Tx) : x \in X\} . \tag{3.17}$$

Proposition 3.10 (Closed graph theorem)
Let X, Y be Banach spaces, let $T : X \to Y$ be linear. Then there are equivalent:

(i) graph (T) *is a closed subset of $X \times Y$.*

(ii) T is continuous.

Proof: "(ii)⇒(i)": Let (x_n, y_n) be a sequence in graph (T) such that $(x_n, y_n) \to (x, y) \in X \times Y$; then $y_n = Tx_n$, $x_n \to x$ and $Tx_n \to Tx$, thus $y = Tx$.
"(i)⇒(ii)": As T is linear, graph (T) is a closed subspace of $X \times Y$ and therefore a Banach space. The projection $P_X|\mathrm{graph}\,(T) : \mathrm{graph}\,(T) \to X$ is linear, continuous and bijective. By Proposition 3.8, it has a linear and continuous inverse $Q : X \to \mathrm{graph}\,(T)$. It follos that $T = P_Y \circ Q$ is continuous. □

This result is useful when one wants to prove that a given T is continuous; it says that it suffices to prove that if $x_n \to x$ and $Tx_n \to y$, then $Tx = y$. (Otherwise one has to prove first that if $x_n \to x$, then Tx_n converges to some y.)

4 Extension, Reflexivity, Separation

Extension of functionals. We know so far that operators which are linear and continuous on a dense subset of a normed space can be extended to a linear continuous operator on the whole space while preserving its norm. The main result of this subsection (the Hahn-Banach theorem) states that such an extension to the whole space is possible even when starting from an arbitrary subspace.

Definition 4.1 (Sublinear functional)
*Let X be a vector space. A mapping $p : X \to \mathbb{R}$ is called **sublinear** if*

$$p(\alpha x) = \alpha p(x)\,, \quad \text{for all } x \in X,\ \alpha \geq 0, \tag{4.1}$$

$$p(x+y) \leq p(x) + p(y)\,, \quad \text{for all } x, y \in X. \tag{4.2}$$

□

Seminorms on X and linear mappings $f : X \to \mathbb{R}$ are sublinear.

Definition 4.2 (Minkowski functional)
Let X be a vector space, let $A \subset X$. Then

$$M_A(x) = \inf\{t : t > 0,\ \frac{1}{t}x \in A\} \tag{4.3}$$

*defines a mapping $M_A : X \to [0, \infty]$, it is called the **Minkowski functional** of A. If $M_A(x) < \infty$ for all $x \in X$, A is called **absorbing**.* □

Let$A$ be a subset of a normed space X, let $0 \in \operatorname{int}(A)$. Then A is absorbing, since $\frac{1}{t}x \in A$ for sufficiently large t. If for A we take the unit ball in X, we have $M_A(x) = \|x\|$.

Lemma 4.3 *Let X be a vector space, $A \subset X$ convex and absorbing. Then the Minkowski functional M_A is sublinear.*

Proof: The property (4.1) immediately follows from (4.3). Let now $x, y \in X$, let $\varepsilon > 0$. We choose $t, s > 0$ such that

$$t \leq M_A(x) + \varepsilon\,, \quad \frac{1}{t}x \in A\,, \quad s \leq M_A(y) + \varepsilon\,, \quad \frac{1}{s}y \in A\,.$$

Then we get, since A is convex,

$$\frac{1}{t+s}(x+y) = \frac{t}{t+s} \cdot \frac{1}{t}x + \frac{s}{t+s} \cdot \frac{1}{s}y \in A\,,$$

so $M_A(x+y) \leq t+s \leq M_A(x) + M_A(y) + 2\varepsilon$. As $\varepsilon > 0$ was arbitrary, the assertion follows. □

Proposition 4.4 *Let X be a vector space over $\mathbb{R}$, $p : X \to \mathbb{R}$ sublinear. Let U be a subspace of X, let $f : U \to \mathbb{R}$ be linear such that $f(x) \leq p(x)$ for all $x \in U$. Then there exists a linear extension $F : X \to \mathbb{R}$ of f on X which satisfies $F(x) \leq p(x)$ for all $x \in X$.*

Proof: We first consider the special case

$$X = \mathrm{span}\,(U \cup \{y\})\,, \quad y \in X \setminus U\,. \tag{4.4}$$

Every $x \in X$ can be decomposed uniquely as

$$x = z + \alpha y\,, \quad z \in U\,,\, \alpha \in \mathbb{R}\,. \tag{4.5}$$

We define $F : X \to \mathbb{R}$ by

$$F(x) = f(z) + \alpha r\,, \quad \text{if } x = z + \alpha y, \tag{4.6}$$

where $r \in \mathbb{R}$ will be fixed later. F is linear, and $F|U = f$. The required inequality

$$F(x) = f(z) + \alpha r \le p(z + \alpha y) = p(x)\,, \quad \text{for all } z \in U,\, \alpha \in \mathbb{R}, \tag{4.7}$$

holds for $\alpha = 0$ by assumption. For $\alpha > 0$ it is equivalent to

$$r \le \frac{p(z + \alpha y) - f(z)}{\alpha} = p\left(\frac{z}{\alpha} + y\right) - f\left(\frac{z}{\alpha}\right)\,, \tag{4.8}$$

and for $\alpha < 0$ it is equivalent to

$$r \ge \frac{p(z + \alpha y) - f(z)}{\alpha} = -p\left(-\frac{z}{\alpha} - y\right) + f\left(-\frac{z}{\alpha}\right)\,. \tag{4.9}$$

Such an r exists if

$$\sup_{z \in U}(f(z) - p(z - y)) \le \inf_{z \in U}(p(z + y) - f(z))\,. \tag{4.10}$$

But for arbitrary $z, \tilde{z} \in U$ we have

$$f(z) + f(\tilde{z}) = f(z + \tilde{z}) \le p(z + \tilde{z}) \le p(z - y) + p(\tilde{z} + y)\,,$$

and therefore

$$f(z) - p(z - y) \le p(\tilde{z} + y) - f(\tilde{z})\,, \quad \text{for all } z, \tilde{z} \in U.$$

From this, (4.10) immediately follows. Thus the proposition is proved for the special case (4.4). To prove the claim in the general case we use Zorn's lemma. We define the set

$$\mathcal{M} = \{(V, g) : V \text{ subspace, } U \subset V \subset X,\, g : V \to \mathbb{R} \text{ linear, } g|U = f,\, g \le p \text{ on } V\}\,, \tag{4.11}$$

and endow $\mathcal{M}$ with the partial ordering

$$(V_1, g_1) \le (V_2, g_2) \qquad \Leftrightarrow \qquad V_1 \subset V_2\,,\; g_2|V_1 = g_1\,.$$

We have $(U, f) \in \mathcal{M}$, so $\mathcal{M} \ne \emptyset$. Let $\mathcal{N}$ be a completely ordered subset of $\mathcal{M}$. We define

$$V_* = \bigcup_{(V,g) \in \mathcal{N}} V \tag{4.12}$$

and $g_* : V_* \to \mathbb{R}$ by

$$g_*(x) = g(x)\,, \quad \text{if } x \in V,\, (V, g) \in \mathcal{N}. \tag{4.13}$$

From the definition of $\mathcal{N}$ it now follows that $g_*(x)$ does not depend on the choice of (V, g), that V_* is a subspace and that g_* is linear (we do not carry out the details). Therefore, (V_*, g_*) is an upper bound of $\mathcal{N}$ in $\mathcal{M}$. By Zorn's lemma, $\mathcal{M}$ possesses a maximal element (V, g). We must have $V = X$; otherwise, by what we have proved for the special case, we could construct a $(\tilde{V}, \tilde{g}) \in \mathcal{M}$ such that $\tilde{V} = \mathrm{span}\,(V \cup \{y\})$, $y \in X \setminus V$, contradicting the maximality of (V, g). $\square$

Proposition 4.5 (Hahn-Banach)
Let X be a normed space over $\mathbb{K}$, let U be a subspace of X, let $u^* \in U^*$. Then there exists an $x^* \in X^*$ satisfying $x^*|U = u^*$ and $\|x^*\| = \|u^*\|$.

Proof: We first consider the case $\mathbb{K} = \mathbb{R}$. We define $p : X \to [0, \infty)$ by

$$p(x) = \|u^*\| \cdot \|x\| . \tag{4.14}$$

p is sublinear, and $u^*(x) \leq p(x)$ for all $x \in U$. By Proposition 4.4 there exists a linear functional $x^* : X \to \mathbb{R}$ such that $x^*(x) \leq p(x)$ for all $x \in X$ and $x^*|U = u^*$. We then have for all $x \in X$

$$x^*(x) \leq p(x) = \|u^*\| \cdot \|x\| , \quad -x^*(x) = x^*(-x) \leq p(-x) = \|u^*\| \cdot \|x\| ,$$

so

$$|x^*(x)| \leq \|u^*\| \cdot \|x\| .$$

Therefore, x^* is continuous and $\|x^*\| \leq \|u^*\|$. Since x^* is an extension of u^*, it follows from the definition of the operator norm that $\|x^*\| \geq \|u^*\|$.
Let now $\mathbb{K} = \mathbb{C}$. With $X_{\mathbb{R}}$ and $U_{\mathbb{R}}$ we denote the normed spaces X and U where we restrict the scalar field to $\mathbb{R}$ (the norm remains the same), and set

$$u_R^* = \operatorname{Re} u^* .$$

Then $u_R^* \in U_{\mathbb{R}}^*$, $\|u_R^*\| \leq \|u^*\|$,

$$u^*(x) = \operatorname{Re} u^*(x) + i \operatorname{Im} u^*(x) = u_R^*(x) - i u_R^*(ix) ,$$

for all $x \in U$, as $\operatorname{Im} u^*(x) = -\operatorname{Re} u^*(ix)$. Let now $x_R^* \in X_{\mathbb{R}}^*$ be an extension of u_R^* such that $\|x_R^*\| = \|u_R^*\|$, as we already have constructed. This extension is $\mathbb{R}$-linear. We set

$$x^*(x) = x_R^*(x) - i x_R^*(ix) . \tag{4.15}$$

Then one may check that $x^* : X \to \mathbb{C}$ is a $\mathbb{C}$-linear extension of u^*. For any given $x \in X$ we now choose $c \in \mathbb{C}$, $|c| = 1$, such that $|x^*(x)| = c x^*(x)$. Then

$$|x^*(x)| = c x^*(x) = x^*(cx) = x_R^*(cx) \leq \|x_R^*\| \cdot \|x\| ,$$

therefore $\|x^*\| \leq \|u^*\|$ and, with the same reasoning as above, $\|x^*\| = \|u^*\|$. $\square$

Corollary 4.6 *Let X be a normed space, $x \in X$ with $x \neq 0$. Then there exists an $x^* \in X^*$ such that $\|x^*\| = 1$ and $x^*(x) = \|x\|$.*

Proof: Let $U = \operatorname{span}(\{x\})$, let $u^* : U \to \mathbb{K}$ be defined by $u^*(\alpha x) = \alpha \|x\|$ where $\alpha \in \mathbb{K}$ is arbitrary. Then $|u^*(\alpha x)| = \|\alpha x\|$, so $\|u^*\| = 1$. We choose $x^* \in X^*$ as an extension of u^* according to Proposition 4.5. $\square$

Corollary 4.7 *Let X be a normed space. Then*

$$\|x\| = \max_{\substack{x^* \in X^* \\ \|x^*\| \leq 1}} |x^*(x)| , \quad \textit{for all } x \in X . \tag{4.16}$$

Proof: We have $|x^*(x)| \le \|x^*\| \cdot \|x\| \le \|x\|$, if $\|x^*\| \le 1$. Therefore, (4.16) follows from Corollary 4.6. $\Box$

Corollary 4.8 *Let X be a normed space, let $U \subset X$ be a closed subspace. Let $x \in X$ such that $x \notin U$. Then there exists an $x^* \in X^*$ which satisfies $x^*|U = 0$ and $x^*(x) \neq 0$.*

Proof: By Proposition 1.14, X/U is a normed space, and we have $[x] \neq 0$. According to Corollary 4.6 we choose a $y^* \in (X/U)^*$ with $y^*([x]) \neq 0$, and define $x^* : X \to \mathbb{K}$ by

$$x^*(z) = y^*([z]) , \quad z \in X .$$

Then $x^* \in X^*$, $x^*|U = 0$ and $x^*(x) \neq 0$. $\Box$.

Reflexivity. Let X be a normed space. The space $X^{**} = (X^*)^*$ is called the **bidual space** of X. The association $x \mapsto x^{**}$,

$$x^{**}(x^*) = x^*(x) , \quad \text{for all } x^* \in X^*, \tag{4.17}$$

yields, due to

$$|x^{**}(x^*)| \le \|x\| \cdot \|x^*\| , \tag{4.18}$$

an embedding

$$J_X : X \to X^{**} . \tag{4.19}$$

(This embedding is called the "canonical" embedding of X into X^{**}, since its definition (4.17) arises from the definition of X and X^{**} in a very natural manner. In the same manner, if $Y \subset X$, one calls the mapping $j : Y \to X$ defined by $j(x) = x$ the canonical embedding of Y into X.)

Lemma 4.9 *The mapping $J_X : X \to X^{**}$ defined by (4.17) is linear and isometric.*

Proof: For all $x \in X$ we obtain, using Corollary 4.7,

$$\|x\|_X = \max_{\substack{x^* \in X^* \\ \|x^*\| \le 1}} |x^*(x)| = \max_{\substack{x^* \in X^* \\ \|x^*\| \le 1}} |(J_X x)(x^*)| = \|J_X x\|_{X^{**}} .$$

$\Box$

The image $J_X(X)$ is a subspace of X^{**}. Since J_X is isometric and X^{**} is a Banach space, we have

$$X \text{ complete} \quad \Leftrightarrow \quad J_X(X) \text{ complete} \quad \Leftrightarrow \quad J_X(X) \text{ closed in } X^{**}.$$

If X is not complete, we may consider the closure $\overline{J_X(X)}$ of $J_X(X)$ in X^{**} as the "completion" of X.

Definition 4.10 (Reflexivity)
*A Banach space X is called **reflexive** if $J_X : X \to X^{**}$ is surjective.* $\Box$

By Lemma 4.9 we then have

$$X \cong X^{**} . \tag{4.20}$$

Conversely, in general it does not follow from (4.20) that X is reflexive. There are examples where $X \cong X^{**}$, but $J_X(X)$ is a closed proper subspace of X^{**}.

Every finite-dimensional Banach space X is reflexive, because in this case

$$\dim(X^{**}) = \dim(X^*) = \dim(X)$$

and therefore J_X is surjective since it is injective.

Proposition 4.11 *Every Hilbert space X is reflexive.*

Proof: Let $J : X \to X^*$, $(Jy)(z) = \langle z, y\rangle$, be the conjugate linear isometric isomorphism coming from the Riesz representation theorem 2.12. Let $x^{**} \in X^{**}$. We define

$$x^*(y) = \overline{x^{**}(Jy)} , \quad y \in X .$$

Then we have $x^* \in X^*$. We set

$$x = J^{-1}x^* .$$

Then for all $y \in X$ we get

$$x^{**}(Jy) = \overline{x^*(y)} = \overline{(Jx)(y)} = \overline{\langle y, x\rangle} = \langle x, y\rangle = (Jy)(x) .$$

Since J is surjective, $x^{**} = J_X(x)$ follows. □

Proposition 4.12 *The spaces $\ell^p(\mathbb{K})$ are reflexive for $1 < p < \infty$.*

Proof: Let $X = \ell^p(\mathbb{K})$, let

$$T_1 : \ell^q(\mathbb{K}) \to \ell^p(\mathbb{K})^* , \quad T_2 : \ell^p(\mathbb{K}) \to \ell^q(\mathbb{K})^* ,$$

be the isometric isomorphisms from Proposition 1.20. Let $x^{**} \in X^{**}$ be given. We set

$$x = T_2^{-1}(x^{**} \circ T_1) .$$

Then we have for all $x^* \in X^*$

$$x^{**}(x^*) = (x^{**} \circ T_1)(T_1^{-1}x^*) = (T_2x)(T_1^{-1}x^*) = \sum_{k=1}^{\infty} x_k (T_1^{-1}x^*)_k = x^*(x) .$$

□

Analogously one proves that the space $L^p(D; \mathbb{K})$ is reflexive for $1 < p < \infty$.

Spaces with sup norms and with L^1 norms are usually not reflexive. As an example we consider $X = L^1(D; \mathbb{K})$. We start from the isometric isomorphism

$$T : L^\infty(D; \mathbb{K}) \to X^* , \quad (Ty)(x) = \int_D x(t)y(t)\, dt .$$

If X were reflexive, for every $x^{**} \in X^{**}$ there would exist an $x \in X$ such that

$$x^{**}(Ty) = (Ty)(x) = \int_D x(t)y(t)\,dt\,, \quad \text{for all } y \in L^\infty(D;\mathbb{K}).$$

An equivalent statement is that for every $y^* \in L^\infty(D;\mathbb{K})^*$ there exists an $x \in L^1(D;\mathbb{K})$ such that

$$y^*(y) = \int_D x(t)y(t)\,dt\,. \tag{4.21}$$

Let now $D = (0,1)$, $E_k = (1/k, 1)$. We set

$$U_k = \{y : y \in L^\infty(D;\mathbb{K}),\, y|(D \setminus E_k) = 0\}\,, \quad U = \overline{\bigcup_{k=2}^\infty U_k}\,.$$

Then U is a closed subspace of $L^\infty(D;\mathbb{K})$, and we have $1_D \notin U$, because $d(1_D, U_k) = 1$ for all $k \geq 2$. Therefore, by Corollary 4.8 there exists a $y^* \in L^\infty(D;\mathbb{K})$ with

$$y^*(1_D) \neq 0 \tag{4.22}$$

and $y^*|U = 0$. In particular,

$$y^*(1_{E_k}) = 0\,, \quad \text{for all } k \geq 2. \tag{4.23}$$

But for every $x \in L^1(D;\mathbb{K})$ we have, according to the Lebesgue convergence theorem,

$$\lim_{k\to\infty} \int_D x(t)1_{E_k}(t)\,dt = \int_D x(t)1_D(t)\,dt\,. \tag{4.24}$$

We see that (4.21) – (4.23) cannot hold at the same time.

The spaces $c_0(\mathbb{K})$, $c(\mathbb{K})$, $\ell^1(\mathbb{K})$, $\ell^\infty(\mathbb{K})$, $C(D;\mathbb{K})$, $L^\infty(D;\mathbb{K})$ are not reflexive (we do not prove this here).

Proposition 4.13 *Let X be a Banach space, let U be subspace of X. If X is reflexive and U is closed, then U is reflexive.*

Proof: Let $u^{**} \in U^{**}$ be arbitrary. The formula

$$x^{**}(x^*) = u^{**}(x^*|U)$$

defines an $x^{**} \in X^{**}$. Let $x = J_X^{-1}(x^{**})$, then

$$x^*(x) = x^{**}(x^*) = u^{**}(x^*|U) \tag{4.25}$$

for all $x^* \in X^*$, and in particular $x^*(x) = 0$ for all x^* satisfying $x^*|U = 0$. It follows from Corollary 4.8 that $x \in U$. Let now $u^* \in U^*$ be arbitrary. We choose according to Proposition 4.5 an $x^* \in X^*$ with $x^*|U = u^*$. Then (4.25) becomes

$$u^*(x) = x^*(x) = u^{**}(u^*)\,,$$

thus $J_U x = u^{**}$. □

Proposition 4.14 *Let X be a Banach space. The X is reflexive if and only if X^* is reflexive.*

Proof: Let X be reflexive. We want to prove that $J_{X^*} : X^* \to X^{***}$ is surjective. Let $x^{***} \in X^{***}$ be arbitrary. We set $x^* = x^{***} \circ J_X$. Let now $x^{**} \in X^{**}$ be arbitrary, let $x = J_X^{-1} x^{**}$. Then

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

so $x^{***} = J_{X^*} x^*$. Conversely, let X^* be reflexive. By what we have just proved, X^{**} is reflexive. By Proposition 4.13, the closed subspace $J_X(X)$ of X^{**} is reflexive and therefore X, too. $\Box$

Proposition 4.15 *Let X normed space, $M \subset X$. The following are equivalent:*

(i) M is bounded.

(ii) $x^*(M)$ bounded in $\mathbb{K}$ for all $x^* \in X^*$.

Proof: We apply Proposition 3.5 (Banach-Steinhaus) to the subset $\mathcal{T} \subset L(X^*; \mathbb{K}) = X^{**}$ defined by

$$\mathcal{T} = J_X(M) .$$

Since J_X is an isometric isomorphism, (i) is equivalent to

$$\sup_{T \in \mathcal{T}} \|T\| < \infty .$$

On the other hand, for every $x^* \in X^*$

$$x^*(M) = \{x^*(x) : x \in M\} = \{T x^* : T \in \mathcal{T}\} .$$

The assertion now immediately follows from Proposition 3.5. $\Box$

Separation of convex sets. Let $x^* : X \to \mathbb{K}$ be a linear continuous functional on a normed space X, $x^* \neq 0$. Its level sets

$$H_\alpha = \{x : x \in X,\, x^*(x) = \alpha\} , \quad \alpha \in \mathbb{K} , \tag{4.26}$$

are called **hyperplanes**. Such hyperplanes separate X into two half-spaces

$$\{x : x \in X,\, x^*(x) < \alpha\} \quad \text{and} \quad \{x : x \in X,\, x^*(x) > \alpha\} .$$

Proposition 4.16 (Separation)
Let X be a normed $\mathbb{R}$-vector space, let $K \subset X$ be open and convex, $K \neq \emptyset$, let $x_0 \in X$ with $x_0 \notin K$. Then there exists an $x^* \in X^*$ which satisfies

$$x^*(x) < x^*(x_0) , \quad \textit{for all } x \in K. \tag{4.27}$$

$\Box$

The inequality (4.27) means that K lies completely on one side of the hyperplane H_α, $\alpha = x^*(x_0)$.

Proof: We first consider the case where $0 \in K$. By Lemma 4.3, its Minkowski functional M_K is sublinear. Let $\varepsilon > 0$ with $K(0;\varepsilon) \subset K$. Then $\varepsilon x/\|x\| \in K$ and therefore

$$M_K(x) \leq \frac{1}{\varepsilon}\|x\|\,, \quad \text{for all } x \in X. \tag{4.28}$$

Moreover we have $x_0/t \notin K$ for all $t < 1$, thus

$$M_K(x_0) \geq 1\,. \tag{4.29}$$

For all $x \in K$ there exists a $t < 1$ with $x/t \in K$, so

$$M_K(x) < 1\,, \quad \text{for all } x \in K. \tag{4.30}$$

We define $u^* : \operatorname{span}(\{x_0\}) \to \mathbb{R}$ by

$$u^*(\alpha x_0) = \alpha M_K(x_0)\,, \quad \alpha \in \mathbb{R}\,.$$

Then

$$\begin{aligned} u^*(\alpha x_0) &= M_K(\alpha x_0) \geq 0\,, \quad \alpha \geq 0\,,\\ u^*(\alpha x_0) &\leq 0 \leq M_K(\alpha x_0)\,, \quad \alpha < 0\,. \end{aligned}$$

Therefore, $u^* \leq M_K$ on $\operatorname{span}(\{x_0\})$. Let now $x^* : X \to \mathbb{R}$, according to Proposition 4.4, be a linear extension of u^* with $x^* \leq M_K$ on X. From (4.28) it follows that for all $x \in X$

$$|x^*(x)| = \max\{x^*(x), x^*(-x)\} \leq \max\{M_K(x), M_K(-x)\} \leq \frac{1}{\varepsilon}\|x\|\,,$$

thus $x^* \in X^*$. From (4.29) and (4.30) it follows that

$$x^*(x) \leq M_K(x) < 1 \leq M_K(x_0) = x^*(x_0)\,, \quad \text{for all } x \in K.$$

So (4.27) is proved. Thus, the proposition is proved for the case $0 \in K$. For the general case, we choose $\tilde{x} \in K$ and set $\tilde{K} = K - \tilde{x}$. Then $0 \in \tilde{K}$ and $x_0 - \tilde{x} \notin \tilde{K}$. Let $x^* \in X^*$ with $x^*(z) < x^*(x_0 - \tilde{x})$ for all $z \in \tilde{K}$. Then we have for all $x \in K$

$$x^*(x) = x^*(x - \tilde{x}) + x^*(\tilde{x}) < x^*(x_0 - \tilde{x}) + x^*(\tilde{x}) = x^*(x_0)\,.$$

□

One may also formulate a separation result in the complex case. Formula (4.27) then becomes dann

$$\operatorname{Re} x^*(x) < \operatorname{Re} x^*(x_0)\,, \quad \text{for all } x \in K.$$

We omit a detailed presentation.

Proposition 4.17 (Separation of two convex sets)

Let X be a normed $\mathbb{R}$-vector space, let $K_1, K_2 \subset X$ be convex and nonempty, let K_1 be open and assume that $K_1 \cap K_2 = \emptyset$. Then there exists an $x^* \in X^*$ and an $\alpha \in \mathbb{R}$ such that

$$x^*(x_1) < \alpha \leq x^*(x_2)\,, \quad \textit{for all } x_1 \in K_1,\ x_2 \in K_2. \tag{4.31}$$

Proof: We set

$$K = K_1 - K_2 = \{x_1 - x_2 : x_1 \in K_1,\, x_2 \in K_2\}\,.$$

Then K is open (it follows from $x \in K$ that $x \in K_1 - x_2 \subset K$ for some $x_2 \in K_2$) and $0 \notin K$. We choose according to Proposition 4.16 an $x^* \in X^*$ such that $x^*(x) < x^*(0) = 0$ for all $x \in K$. Then we get

$$x^*(x_1) - x^*(x_2) = x^*(x_1 - x_2) < 0\,, \quad \text{for all } x_1 \in K_1,\, x_2 \in K_2,$$

and therefore

$$x^*(x_1) \le \alpha \le x^*(x_2)\,, \quad \text{for all } x_1 \in K_1,\, x_2 \in K_2,$$

for every $\alpha \in [\sup x^*(K_1), \inf x^*(K_2)]$. As K_1 is open and $x^* \neq 0$, we get $x^*(x_1) < \alpha$ for all $x_1 \in K_1$; indeed, if $x^*(x_1) = \alpha$ for some $x_1 \in K_1$, then there would exist a $\tilde{x}_1 \in K_1$ such that $x^*(\tilde{x}_1) > \alpha$. □

Proposition 4.18 (Strict Separation)
Let X be a normed $\mathbb{R}$-vector space, let $K \subset X$ be closed and convex, $K \neq \emptyset$, let $x_0 \in X$ such that $x_0 \notin K$. Then there exists an $x^* \in X^*$ and an $\alpha \in \mathbb{R}$ such that

$$x^*(x) \le \alpha < x^*(x_0)\,, \quad \textit{for all } x \in K. \tag{4.32}$$

□

Proof: Exercise. (Hint: For sufficiently small $\varepsilon > 0$ we have $B(x_0;\varepsilon) \cap K = \emptyset$.) □

5 Compact subsets of C and L^p

We want to find out which subsets of the spaces $C(K;\mathbb{K})$ (K compact) and $L^p(\Omega;\mathbb{K})$ (Ω open) are compact. We already know that the closed unit balls in these spaces are not compact, as this is true only for finite-dimensional spaces.

Two notions of compactness.

Proposition 5.1
Let (X,d) be a metric space. Then there are equivalent:

(i) X is sequentially compact, that is, every sequence in X has a subsequence which converges to some element of X.

(ii) Every open covering of X has a finite subcovering, that is: If $(U_i)_{i\in I}$ is a family of open subset of X such that
$$X = \bigcup_{i\in I} U_i\,,$$
then there exists $n\in\mathbb{N}$ and indices $i_1,\dots,i_n\in I$ such that
$$X = \bigcup_{k=1}^{n} U_{i_k}\,.$$

*If (ii) holds, we say that X is **compact**.*

Proof: Omitted. □

A side remark: In general topological spaces, (i) and (ii) are not equivalent.

Corollary 5.2 *Let X be a normed space, let F be a compact subset of X. Then F is closed and bounded.*

Proof: Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in F with $x_n\to a\in X$. By Proposition 5.1(i), for some subsequence and some $b\in F$ we have $x_{n_k}\to b$. As $x_{n_k}\to a$, we must have $a=b$. This proves that F is closed. If F were unbounded, then there would exist a sequence $(x_n)_{n\in\mathbb{N}}$ in F with $\|x_n\|\ge n$ for all $n\in\mathbb{N}$. Such a sequence does not have a convergent subsequence. Consequently, F cannot be compact. □

Compact sets in the space of continuous functions. For a subset of $C(\Omega;\mathbb{K})$ to be compact, in addition to being closed and bounded, the following property must be satisfied.

Definition 5.3 (Equicontinuity)
*Let $\Omega\subset\mathbb{R}^n$. A subset F of $C(\Omega;\mathbb{K})$ is called **equicontinuous**, if for all $x\in\Omega$ and all $\varepsilon>0$ there exists a $\delta>0$ such that for all $y\in\Omega$*
$$\|x-y\|<\delta \qquad\Rightarrow\qquad |f(x)-f(y)|<\varepsilon \quad \text{for all } f\in F.$$
("It is possible to find a common δ for all $f\in F$".) □

Lemma 5.4 *Let $\Omega \subset \mathbb{R}^n$, $F \subset C(\Omega; \mathbb{K})$, assume that there exists an L such that*

$$|f(x) - f(y)| \leq L\|x - y\|\,, \quad \text{for all } x, y \in \Omega,\ f \in F. \tag{5.1}$$

Then F is equicontinuous.

Proof: Setting $\delta = \varepsilon/L$ we see that F has the property required in Definition 5.3. □

When $F \subset C^1(\Omega)$, it follows from the mean value theorem that (5.1) holds, if there exists a $C > 0$ such that

$$|\partial_i f(x)| \leq C\,, \quad \text{for all } x \in \Omega,\ f \in F,\ i = 1, \ldots, n. \tag{5.2}$$

This criterion is often used to prove equicontinuity of a given set F of continuous functions.

Proposition 5.5 (Arzela-Ascoli)
Let $K \subset \mathbb{R}^n$ be compact, let $F \subset (C(K; \mathbb{K}), \|\cdot\|_\infty)$. Then there are equivalent

(i) F is relatively compact in $C(K)$, that is, $\overline{F}$ is compact.

(ii) F is bounded and equicontinuous.

Proof: Omitted. □

Smoothing by convolution. We recall the definition of the convolution of two functions f and g,

$$(f * g)(y) = \int_{\mathbb{R}^n} f(x) g(y - x)\, dx\,. \tag{5.3}$$

It is defined for $f, g \in L^1(\mathbb{R}^n)$ and yields a function $f * g \in L^1(\mathbb{R}^n)$ with the properties

$$f * g = g * f\,, \quad \|f * g\|_1 \leq \|f\|_1 \cdot \|g\|_1\,. \tag{5.4}$$

One can approximate a given function by smooth functions if one convolves it with suitable functions. We define

$$\psi : \mathbb{R} \to \mathbb{R}\,, \quad \psi(t) = \begin{cases} \exp\left(-\frac{1}{t}\right)\,, & t > 0\,, \\ 0\,, & t \leq 0\,, \end{cases} \tag{5.5}$$

$$\tilde{\psi} : \mathbb{R} \to \mathbb{R}\,, \quad \tilde{\psi}(r) = \psi(1 - r^2)\,, \tag{5.6}$$

$$\eta_1 : \mathbb{R}^n \to \mathbb{R}\,, \quad \eta_1(x) = \alpha \tilde{\psi}(\|x\|)\,. \tag{5.7}$$

Here, $\alpha > 0$ is chosen such that

$$\int_{\mathbb{R}^n} \eta_1(x)\, dx = 1\,. \tag{5.8}$$

For given $\varepsilon > 0$ we define the "standard mollifier"

$$\eta_\varepsilon : \mathbb{R}^n \to \mathbb{R}\,, \quad \eta_\varepsilon(x) = \frac{1}{\varepsilon^n} \eta_1\left(\frac{x}{\varepsilon}\right)\,. \tag{5.9}$$

The functions η_ε are radially symmetric (that is, they only depend upon $\|x\|$), and have the properties (as developed in integration theory)

$$\eta_\varepsilon \in C_0^\infty(\mathbb{R}^n)\,, \quad \operatorname{supp}(\eta_\varepsilon) = K(0;\varepsilon)\,, \quad \eta_\varepsilon \geq 0\,, \quad \int_{\mathbb{R}^n} \eta_\varepsilon(x)\,dx = 1\,. \tag{5.10}$$

For functions which are not defined on the whole of $\mathbb{R}^n$ we too want to consider convolution with the standard mollifier.

Given $f \in L^1(\Omega)$, we define $\tilde{f}$ by setting $\tilde{f} = f$ in Ω and $\tilde{f} = 0$ outside of Ω, and define f^ε by

$$f^\varepsilon(y) = (\tilde{f} * \eta_\varepsilon)(y) = \int_{\mathbb{R}^n} \tilde{f}(x)\eta_\varepsilon(y-x)\,dx = \int_\Omega f(x)\eta_\varepsilon(y-x)\,dx\,, \quad \text{for all } y \in \mathbb{R}^n. \tag{5.11}$$

Since $f^\varepsilon = \tilde{f} * \eta_\varepsilon = \eta_\varepsilon * \tilde{f}$, we may represent f^ε as

$$\begin{aligned} f^\varepsilon(y) &= \int_{\mathbb{R}^n} \eta_\varepsilon(x)\tilde{f}(y-x)\,dx = \varepsilon^{-n} \int_{\mathbb{R}^n} \eta_1\big(\frac{x}{\varepsilon}\big)\tilde{f}(y-x)\,dx \\ &= \int_{\mathbb{R}^n} \eta_1(z)\tilde{f}(y-\varepsilon z)\,dz = \int_{K(0;1)} \eta_1(z)\tilde{f}(y-\varepsilon z)\,dz\,. \end{aligned} \tag{5.12}$$

In the following, we no longer distinguish between f and $\tilde{f}$, and just write f for both functions.

Lemma 5.6 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, $f \in L^p(\Omega)$, $1 \leq p < \infty$. Then $f^\varepsilon = f * \eta_\varepsilon$ satisfies*

$$f^\varepsilon \in C^\infty(\mathbb{R}^n)\,, \quad \operatorname{supp}(f^\varepsilon) \subset \overline{\Omega + B(0;\varepsilon)}\,, \quad \|f^\varepsilon\|_{L^p(\Omega)} \leq \|f\|_{L^p(\Omega)}\,. \tag{5.13}$$

Proof: We have $f \in L^1(\Omega)$ since Ω is bounded, and

$$f^\varepsilon(y) = \int_\Omega f(x)\eta_\varepsilon(y-x)\,dx\,, \quad \text{for all } y \in \mathbb{R}^n. \tag{5.14}$$

As $\eta_\varepsilon(y-x) = 0$ holds for $\|y-x\| \geq \varepsilon$, we get $\operatorname{supp}(f^\varepsilon) \subset \overline{\Omega + B(0;\varepsilon)}$. For all multi-indices α, the functions $x \mapsto f(x)\partial^\alpha\eta_\varepsilon(y-x)$ are bounded (uniformly with respect to y) by the integrable functions $\|\partial^\alpha\eta_\varepsilon\|_\infty|f|$. By a result from integration theory, all partial derivatives $\partial^\alpha f^\varepsilon$ exist in $\mathbb{R}^n$ and are continuous; thus we have $f^\varepsilon \in C^\infty(\mathbb{R}^n)$. Let now $y \in \Omega$. In the case $p > 1$, we choose q such that $\frac{1}{p} + \frac{1}{q} = 1$. It follows from Hölder's inequality that

$$\begin{aligned} |f^\varepsilon(y)| &= \Big|\int_\Omega f(x)\eta_\varepsilon(y-x)\,dx\Big| \leq \int_\Omega |f(x)|(\eta_\varepsilon(y-x))^{\frac{1}{p}}(\eta_\varepsilon(y-x))^{\frac{1}{q}}\,dx \\ &\leq \Big(\int_\Omega |f(x)|^p\eta_\varepsilon(y-x)\,dx\Big)^{\frac{1}{p}} \underbrace{\Big(\int_\Omega \eta_\varepsilon(y-x)\,dx\Big)^{\frac{1}{q}}}_{\leq 1}\,. \end{aligned}$$

Next, we see that for $p \geq 1$ (for $p > 1$ we use the estimate from the previous line)

$$\begin{aligned}\int_\Omega |f^\varepsilon(y)|^p \, dy &\leq \int_\Omega \int_\Omega |f(x)|^p \eta_\varepsilon(y-x) \, dx \, dy = \int_\Omega |f(x)|^p \int_\Omega \eta_\varepsilon(y-x) \, dy \, dx \\ &\leq \int_\Omega |f(x)|^p \, dx \, .\end{aligned}$$

□

By $C_0(\mathbb{R}^n)$ we denote the space of functions which are continuous on $\mathbb{R}^n$ and have compact support.

Lemma 5.7 *Let $f \in C_0(\mathbb{R}^n)$. Then $f^\varepsilon \to f$ uniformly on $\mathbb{R}^n$ when $\varepsilon \to 0$.*

Proof: For all $y \in \mathbb{R}^n$ we have

$$\begin{aligned}|f^\varepsilon(y) - f(y)| &= \Big| \int_{\mathbb{R}^n} f(x)\eta_\varepsilon(y-x) \, dx - f(y) \Big| = \Big| \int_{\mathbb{R}^n} (f(x) - f(y))\eta_\varepsilon(y-x) \, dx \Big| \\ &\leq \int_{\mathbb{R}^n} \eta_\varepsilon(y-x) \, dx \cdot \sup_{\substack{x \in \mathbb{R}^n \\ \|x-y\| \leq \varepsilon}} |f(y) - f(x)| \leq \sup_{\substack{x,z \in \mathbb{R}^n \\ \|x-z\| \leq \varepsilon}} |f(z) - f(x)|\end{aligned}$$

Since f has compact support and therefore is uniformly continuous on $\mathbb{R}^n$, the assertion follows. □

Lemma 5.8 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, let $f \in L^p(\Omega)$, $1 \leq p < \infty$. Then $f^\varepsilon \to f$ in $L^p(\Omega)$ for $\varepsilon \to 0$.*

Proof: Let $\delta > 0$. We choose $g \in C_0(\mathbb{R}^n)$ such that

$$\|f - g\|_{L^p(\Omega)} \leq \delta \, .$$

This is possible since $C_0(\mathbb{R}^n)$ is dense in $L^p(\Omega)$ (see the exercises). It follows that

$$\begin{aligned}\|f - f^\varepsilon\|_{L^p(\Omega)} &\leq \|f - g\|_{L^p(\Omega)} + \|g - g^\varepsilon\|_{L^p(\Omega)} + \|g^\varepsilon - f^\varepsilon\|_{L^p(\Omega)} \\ &\leq 2\delta + \|g - g^\varepsilon\|_{L^p(\Omega)} \, ,\end{aligned} \tag{5.15}$$

since

$$g^\varepsilon - f^\varepsilon = g * \eta_\varepsilon - f * \eta_\varepsilon = (g-f)^\varepsilon$$

and by Lemma 5.6

$$\|(g-f)^\varepsilon\|_{L^p(\Omega)} \leq \|g - f\|_{L^p(\Omega)} \leq \delta \, .$$

By Lemma 5.7 we have $g^\varepsilon \to g$ uniformly. Consequently,

$$\limsup_{\varepsilon \to 0} \|f - f^\varepsilon\|_{L^p(\Omega)} \leq 2\delta \, .$$

As $\delta > 0$ was arbitrary, the assertion follows. □

Compactness in L^p. We investigate the behaviour of

$$\int_\Omega |f(x+h) - f(x)|^p \, dx \, , \quad h \to 0 \, .$$

Proposition 5.9 *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, let $f \in L^p(\Omega)$, $1 \le p < \infty$. Then*

$$\lim_{h\to 0} \int_\Omega |f(x+h) - f(x)|^p \, dx = 0\,. \tag{5.16}$$

(Again we extend f by 0 outside of Ω.)

Proof: We define $(\tau_h f)(x) = f(x+h)$. Then

$$\int_\Omega |f(x+h) - f(x)|^p \, dx = \|\tau_h f - f\|^p_{L^p(\Omega)}\,.$$

We estimate from above

$$\|\tau_h f - f\|_{L^p(\Omega)} \le \|\tau_h f - \tau_h f^\varepsilon\|_{L^p(\Omega)} + \|\tau_h f^\varepsilon - f^\varepsilon\|_{L^p(\Omega)} + \|f^\varepsilon - f\|_{L^p(\Omega)}\,.$$

Let $\gamma > 0$. For $\varepsilon > 0$ sufficiently small, the third and also the first term on the right side are smaller than $\gamma/3$ for all h with $\|h\| \le \varepsilon$ because of

$$\|\tau_h f - \tau_h f^\varepsilon\|_{L^p(\Omega)} = \|\tau_h (f - f^\varepsilon)\|_{L^p(\Omega)} \le \|f^\varepsilon - f\|_{L^p(\Omega + B(0;\varepsilon))}\,.$$

As f^ε is uniformly continuous on $\overline{\Omega}$, for h sufficiently small the second term, too, is smaller than $\gamma/3$. This proves the assertion. □

Proposition 5.10 (Fréchet-Riesz-Kolmogorov)
Let $\Omega \subset \mathbb{R}^n$ be open and bounded, $1 \le p < \infty$, let $F \subset (L^p(\Omega; \mathbb{K}), \|\cdot\|_p)$. Then there are equivalent:

(i) F is relatively compact in $L^p(\Omega)$, that is, $\overline{F}$ is compact.

(ii) F is bounded, and

$$\lim_{h\to 0} \sup_{f\in F} \int_\Omega |f(x+h) - f(x)|^p \, dx = 0\,. \tag{5.17}$$

Condition (5.17) means that the passage to the limit $h \to 0$ is uniform with respect to F.

Proof: Omitted. □

6 Weak Convergence

Closed and bounded subsets of a normed space X are guaranteed to be compact only if X is finite-dimensional. When the space has infinite dimension, a closed and bounded subset may or may not be compact, and not every bounded sequence has a subsequence which converges in the norm of X.

In order to keep the result that every bounded sequence has a convergent subsequence, one has to weaken the notion of convergence, so that more sequences are convergent in the weaker sense.

Definition 6.1 (Weak convergence, weak star convergence)
*Let X be a normed space. A sequence $(x_n)_{n\in\mathbb{N}}$ in X is called **weakly convergent** to an $x \in X$, if*

$$\lim_{n\to\infty} x^*(x_n) = x^*(x)\,, \quad \text{for all } x^* \in X^*, \tag{6.1}$$

*we write $x_n \rightharpoonup x$. A sequence $(x_n^*)_{n\in\mathbb{N}}$ in X^* is called **weak star convergent** to an $x^* \in X^*$, if*

$$\lim_{n\to\infty} x_n^*(x) = x^*(x)\,, \quad \text{for all } x \in X, \tag{6.2}$$

we write $x_n^* \overset{*}{\rightharpoonup} x^*$. □

Let X be a normed space. If a sequence in X converges to an x in the sense of the norm, we also say that it converges **strongly** to x.

If X is reflexive, weak and weak star convergence on X^* coincide. If X has finite dimension, both coincide with strong convergence.

It follows immediately from the definition that every strongly convergent sequence is weakly resp. weak star convergent.

For $X = \ell^p(\mathbb{K})$, $1 \le p < \infty$, we have $X^* \cong \ell^q(\mathbb{K})$ where $q = p/(p-1)$. A sequence $(x^n)_{n\in\mathbb{N}}$ in X thus converges weakly to an $x \in X$ if and only if

$$\lim_{n\to\infty} \sum_{k=1}^{\infty} x_k^n y_k = \sum_{k=1}^{\infty} x_k y_k\,, \quad \text{for all } y \in \ell^q(\mathbb{K}). \tag{6.3}$$

If in particular $x^n = e^n$ (n-th unit vector), for $p > 1$ (thus $q < \infty$) we have

$$\lim_{n\to\infty} \sum_{k=1}^{\infty} e_k^n y_k = \lim_{n\to\infty} y_n = 0\,, \quad \text{for all } y \in \ell^q(\mathbb{K}). \tag{6.4}$$

Therefore, $e^n \rightharpoonup 0$ for $p > 1$, but on the other hand $\|e^n\|_p = 1$ and thus e_n does not converge strongly to 0.

For $p = 1$ we also have (6.4) for all $y \in \ell^q(\mathbb{K})$, $q < \infty$, but for $y = (-1, 1, -1, \dots) \in \ell^\infty(\mathbb{K})$ and the corresponding functional $y^* \in X^*$ we get

$$y^*(e^n) = \sum_{k=1}^{\infty} e_k^n y_k = y_n = (-1)^n\,. \tag{6.5}$$

Therefore, $(e^n)_{n\in\mathbb{N}}$ is not weakly convergent in $\ell^1(\mathbb{K})$.

In general it holds (exercise): If X is a Hilbert space, $(x_n)_{n\in\mathbb{N}}$ is sequence in X and $x \in X$, then x_n strongly converges to x if and only if x_n converges weakly to x and $\|x_n\|$ converges to $\|x\|$.

As a further example we consider $X = L^p(D;\mathbb{K})$, $1 \le p < \infty$. We have $X^* \cong L^q(D;\mathbb{K})$, $q = p/(p-1)$. A sequence $(x_n)_{n\in\mathbb{N}}$ in X converges weakly to an $x \in X$ if and only if

$$\lim_{n\to\infty} \int_D x_n(t)y(t)\,dt = \int_D x(t)y(t)\,dt\,, \quad \text{for all } y \in L^q(D;\mathbb{K}).$$

For $X = L^\infty(D;\mathbb{K})$ we have $X \cong L^1(D;\mathbb{K})^*$, and a sequence $(x_n)_{n\in\mathbb{N}}$ in X converges weak star to an $x \in X$ if and only if

$$\lim_{n\to\infty} \int_D x_n(t)y(t)\,dt = \int_D x(t)y(t)\,dt\,, \quad \text{for all } y \in L^1(D;\mathbb{K}).$$

Lemma 6.2 *Weak and weak star limits are uniquely determined.*

Proof: Let $x_n^* \stackrel{*}{\rightharpoonup} x^*$ and $x_n^* \stackrel{*}{\rightharpoonup} y^*$. Then

$$x^*(x) = \lim_{n\to\infty} x_n^*(x) = y^*(x)\,, \quad \text{for all } x \in X,$$

and therefore $x^* = y^*$. If $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$, then

$$x^*(x) = \lim_{n\to\infty} x^*(x_n) = x^*(y)\,, \quad \text{for all } x^* \in X^*,$$

and therefore $x = y$; otherwise, the separation theorem would imply that there exists an $x^* \in X^*$ such that $x^*(x) \ne x^*(y)$. □

Lemma 6.3 *Let X be a normed space, let $(x_n^*)_{n\in\mathbb{N}}$ be a sequence in X^* such that $x_n^* \stackrel{*}{\rightharpoonup} x^*$, $x^* \in X^*$. Then*

$$\|x^*\| \le \liminf_{n\to\infty} \|x_n^*\|\,. \tag{6.6}$$

If $(x_n)_{n\in\mathbb{N}}$ is a sequence in X such that $x_n \rightharpoonup x$, $x \in X$, then

$$\|x\| \le \liminf_{n\to\infty} \|x_n\|\,. \tag{6.7}$$

Proof: For all $x \in X$ we have

$$|x_n^*(x)| \le \|x_n^*\|\,\|x\|\,.$$

Therefore it follows from $x_n^* \stackrel{*}{\rightharpoonup} x^*$ that

$$|x^*(x)| = \lim_{n\to\infty} |x_n^*(x)| \le \liminf_{n\to\infty}(\|x_n^*\|\,\|x\|) = \big(\liminf_{n\to\infty} \|x_n^*\|\big)\,\|x\|\,, \quad \text{for all } x \in X,$$

so (6.6) holds. For the proof of (6.7) we refer to the exercises. □

Let $(e^n)_{n\in\mathbb{N}}$ be the sequence of the unit vectors in $\ell^p(\mathbb{K})$, $1 < p < \infty$. It was shown above in (6.4) that

$$e^n \rightharpoonup 0\,, \quad 0 < \lim_{n\to\infty} \|e^n\|_p = 1\,.$$

Proposition 6.4 *Let X be a normed space. Then every in X weakly convergent sequence is bounded in X (w.r.t. the norm of X). If moreover X is a Banach space, every in X^* weak star convergent sequence is bounded in X^* (w.r.t. the norm of X^*).*

Proof: From $x_n^* \overset{*}{\rightharpoonup} x^*$ it follows that $|x_n^*(x)| \to |x^*(x)|$ for all $x \in X$, so

$$\sup_{n\in\mathbb{N}} |x_n^*(x)| < \infty\,, \quad \text{for all } x \in X.$$

The Banach-Steinhaus theorem (Proposition 3.5) now implies

$$\sup_{n\in\mathbb{N}} \|x_n^*\| < \infty\,.$$

From $x_n \rightharpoonup x$ it follows that $|x^*(x_n)| \to |x^*(x)|$ for all $x^* \in X^*$, so

$$\sup_{n\in\mathbb{N}} |x^*(x_n)| < \infty\,, \quad \text{for all } x^* \in X^*.$$

Proposition 4.15 now implies

$$\sup_{n\in\mathbb{N}} \|x_n\| < \infty\,.$$

□

Lemma 6.5 *Let X be a normed space, let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X such that $x_n \to x$ strongly in X, let $(x_n^*)_{n\in\mathbb{N}}$ be sequence in X^* such that $x_n^* \overset{*}{\rightharpoonup} x^*$. Then*

$$\lim_{n\to\infty} x_n^*(x_n) = x^*(x)\,. \tag{6.8}$$

If $x_n \rightharpoonup x$ and $x_n^* \to x^*$ strongly, then (6.8) also holds.

Proof: Es gilt

$$|x^*(x) - x_n^*(x_n)| \le |(x^* - x_n^*)(x)| + \|x_n^*\|\, \|x - x_n\|\,.$$

Now $x_n^* \overset{*}{\rightharpoonup} x^*$ implies $|(x^* - x_n^*)(x)| \to 0$, and since $\|x_n^*\|$ is bounded by Proposition 6.4, (6.8) follows. The second assertion is proved analogously. □

If $x_n \rightharpoonup x$ and $x_n^* \overset{*}{\rightharpoonup} x^*$, in general it does **not** follow that $x_n^*(x_n) \to x^*(x)$. Here is an example: $X = \ell^2(\mathbb{R})$, $X^* \cong \ell^2(\mathbb{R})$, $x_n = e^n$, $x_n^* = e^n$. We have $e^n \rightharpoonup 0$ and $e^n \overset{*}{\rightharpoonup} 0$, but $\langle e^n, e^n \rangle = 1$.

Proposition 6.6 *Let X be a normed space, $K \subset X$ convex and closed, let $(x_n)_{n\in\mathbb{N}}$ be a sequence in K with $x_n \rightharpoonup x \in X$. Then we have $x \in K$.*

One says that K is **weakly sequentially closed**. Thus, convex closed subsets of a normed space are weakly sequentially closed.

Proof: We consider the case $\mathbb{K} = \mathbb{R}$. Assume that $x \notin K$. Then by the separation theorem there exists an $x^* \in X^*$ such that

$$x^*(x) > \sup_{z\in K} x^*(z) =: c \ge x^*(x_n)\,, \quad \text{for all } n \in \mathbb{N},$$

which contradicts $x^*(x_n) \to x^*(x)$. In the case $\mathbb{K} = \mathbb{C}$ one needs the complex version of the separation theorem (which we did not discuss in Chapter 4). □

Definition 6.7 (Weak sequential compactness)
Let X be a normed space. A subset M of X is called ***weakly sequentially compact*** *if every sequence in M has a weakly convergent subsequence whose limit is an element of M. A subset M of X^* is called* ***weak star sequentially compact****, if every sequence in M has a weak star convergent subsequence whose limit is an element of M.* □

Proposition 6.8 *Let X be a separable normed space. Then the closed unit ball $K(0;1)$ in X^* is weak star sequentially compact.*

Proof: Let $(x_n^*)_{n\in\mathbb{N}}$ be a sequence in $K(0;1)$, so $\|x_n^*\| \le 1$ for all $n \in \mathbb{N}$. Let $\{x_m : m \in \mathbb{N}\}$ be a dense subset of X. Since $|x_n^*(x_m)| \le \|x_m\|$, the sequences $(x_n^*(x_m))_{n\in\mathbb{N}}$ are bounded in $\mathbb{K}$ for all $m \in \mathbb{N}$. Passing to a suitable subsequence (w.r.t. n) we want to achieve that for all $m \in \mathbb{N}$ the corresponding subsequences of $(x_n^*(x_m))_{n\in\mathbb{N}}$ converge. This can be done with a "diagonalization" argument. First, we choose a sequence $(n_{k1})_{k\in\mathbb{N}}$ such that $(x_{n_{k1}}^*(x_1))_{k\in\mathbb{N}}$ converges. Next, we choose a subsequence $(n_{k2})_{k\in\mathbb{N}}$ of $(n_{k1})_{k\in\mathbb{N}}$ such that $(x_{n_{k2}}^*(x_2))_{k\in\mathbb{N}}$ converges. Using induction we correspondingly choose subsequences $(n_{km})_{k\in\mathbb{N}}$ for each $m \in \mathbb{N}$. The subsequence $(n_{kk})_{k\in\mathbb{N}}$ then has the property that $(x_{n_{kk}}^*(x_m))_{k\in\mathbb{N}}$ converges for all $m \in \mathbb{N}$. We set $z_k^* = x_{n_{kk}}^*$ and look for a weak star limit of this subsequence of (x_n^*).

We set $Z = \operatorname{span}\{x_m : m \in \mathbb{N}\}$ and define

$$z^* : Z \to \mathbb{K}, \quad z^*(z) = \lim_{k\to\infty} z_k^*(z)\,.$$

This limit is well-defined, since every $z \in Z$ is a linear combination of the x_m. Moreover, we have $|z_k^*(z)| \le \|z\|$, so $|z^*(z)| \le \|z\|$ for all $z \in Z$, and therefore z^* is continuous on Z and $\|z^*\| \le 1$ holds. By Proposition 1.19, z^* can be extended to an $x^* \in X^*$ with $\|x^*\| \le 1$. Let now $x \in X$ be arbitrary. For all $z \in Z$ we have

$$|x^*(x) - z_k^*(x)| \le |(x^* - z_k^*)(x - z)| + |x^*(z) - z_k^*(z)| \le 2\|x - z\| + |z^*(z) - z_k^*(z)|\,.$$

Let $\varepsilon > 0$ be arbitrary. We choose a $z \in Z$ such that $\|x - z\| \le \varepsilon$ (this is possible since Z is dense in X) and an $N > 0$ such that $|x^*(z) - z_k^*(z)| \le \varepsilon$ for all $k \ge N$. Then

$$|x^*(x) - z_k^*(x)| \le 3\varepsilon$$

for all $k \ge N$. Consequently, $z_k^*(x) \to x^*(x)$ when $k \to \infty$. As x was arbitrary, $z_k^* \xrightharpoonup{*} x^*$ follows. □

Corollary 6.9 *Let X be a separable normed space. Then every bounded sequence in X^* has a weak star convergent subsequence.* □

When $1 < p \le \infty$, the space $L^p(D)$ is isomorphic to the dual space of the separable space $L^q(D)$, $q = p/(p-1) < \infty$. Therefore, if $(x_n)_{n\in\mathbb{N}}$ is a sequence in $L^p(D)$ such that $\|x_n\|_p \le C$ for all $n \in \mathbb{N}$, there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ and an $x \in L^p(D)$ satisfying

$$\lim_{k\to\infty} \int_D x_{n_k}(t)y(t)\,dt = \int_D x(t)y(t)\,dt\,, \quad \text{for all } y \in L^q(D). \tag{6.9}$$

As a further example we consider $X = C([a,b])$. Let $(t_n)_{n\in\mathbb{N}}$ be a sequence in $[a,b]$. Then

$$x_n^* = \delta_{t_n}\,, \quad \delta_{t_n}(x) = x(t_n)\,,$$

defines a sequence $(x_n^*)_{n\in\mathbb{N}}$ in X^*. If $(t_{n_k})_{k\in\mathbb{N}}$ is a convergent subsequence with $t_{n_k} \to t \in [a,b]$, we have

$$x_{n_k}^*(x) = x(t_{n_k}) \to x(t) = \delta_t(x)\,, \quad \text{for all } x \in C[a,b].$$

It follows that $x_{n_k}^* \overset{*}{\rightharpoonup} x^* = \delta_t$. A generalization: If $(\mu_n)_{n\in\mathbb{N}}$ is a sequence of probability measures on $[a,b]$, it follows from Corollary 6.9 that there exists a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ and a probability measure μ such that

$$\lim_{k\to\infty} \int_a^b x(t)\, d\mu_{n_k}(t) = \int_a^b x(t)\, d\mu(t)\,, \quad \text{for all } x \in C[a,b].$$

When X is not separable there may exist bounded sequences in X^* which do not possess a weak star convergent subsequence. As an example we consider $X = L^\infty(0,1)$. Let $(\varepsilon_n)_{n\in\mathbb{N}}$ be a monotone decreasing sequence such that

$$\varepsilon_n \in (0,1)\,, \quad \varepsilon_n \to 0\,, \quad \frac{\varepsilon_{n+1}}{\varepsilon_n} \to 0\,. \tag{6.10}$$

We define

$$x_n^*(x) = \frac{1}{\varepsilon_n} \int_0^{\varepsilon_n} x(t)\, dt\,, \quad x \in L^\infty(0,1)\,. \tag{6.11}$$

We have $x_n^* \in X^*$, $\|x_n^*\| = 1$. Let $x \in L^\infty(0,1)$ be defined by

$$x(t) = (-1)^k\,, \quad \text{if } \varepsilon_{k+1} < t < \varepsilon_k. \tag{6.12}$$

Then

$$\begin{aligned} x_n^*(x) &= \frac{1}{\varepsilon_n}\left(\int_0^{\varepsilon_{n+1}} x(t)\, dt + (\varepsilon_n - \varepsilon_{n+1})(-1)^n \right) \\ &= (-1)^n + \frac{1}{\varepsilon_n} \int_0^{\varepsilon_{n+1}} x(t)\, dt - \frac{\varepsilon_{n+1}}{\varepsilon_n}(-1)^n\,, \end{aligned}$$

so

$$|x_n^*(x) - (-1)^n| \le 2\frac{\varepsilon_{n+1}}{\varepsilon_n} \to 0\,.$$

Therefore, the sequence $(x_n^*(x))_{n\in\mathbb{N}}$ does not converge. Since the same argument can be applied to every subsequence $(x_{n_k}^*)_{n\in\mathbb{N}}$ (the function x has to be chosen correspondingly), the sequence $(x_n^*)_{n\in\mathbb{N}}$ does not possess a weak star convergent subsequence. If, however, we view it as a sequence in X^*, $X = C([0,1])$, we obtain for all $x \in X$

$$|x_n^*(x) - x(0)| = \left| \frac{1}{\varepsilon_n} \int_0^{\varepsilon_n} x(t) - x(0)\, dt \right| \le \sup_{0\le s\le \varepsilon_n} |x(s) - x(0)| \to 0\,,$$

therefore $x_n^* \overset{*}{\rightharpoonup} \delta_0$.

Lemma 6.10 *Let X be a normed space. If X^* is separable, then so is X.*

Proof: Let $\{x_n^* : n \in \mathbb{N}\}$ be a dense subset of X^*. According to the definition of the operator norm we choose for every $n \in \mathbb{N}$ an $x_n \in X$ such that

$$|x_n^*(x_n)| \geq \frac{1}{2}\|x_n^*\|\,, \quad \|x_n\| = 1\,.$$

Let $Y = \text{span}\,\{x_n : n \in \mathbb{N}\}$. Let $x^* \in X^*$ be arbitrary such that $x^*|Y = 0$. It follows that

$$\|x^* - x_n^*\| \geq |x^*(x_n) - x_n^*(x_n)| = |x_n^*(x_n)| \geq \frac{1}{2}\|x_n^*\| \geq \frac{1}{2}(\|x^*\| - \|x_n^* - x^*\|)\,. \qquad (6.13)$$

Passing to the infimum w.r.t. n on both sides we see that $\|x^*\| = 0$ and thus $x^* = 0$. From Corollary 4.8 it follows that $\overline{Y} = X$ (if not, there would exist an $x^* \neq 0$ such that $x^*|\overline{Y} = 0$). Applying Proposition 1.17 we conclude that X is separable. □

Proposition 6.11 *Let X be a reflexive Banach space. Then the closed unit ball $K(0;1)$ in X is weakly sequentially compact.*

Proof: Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X such that $\|x_n\| \leq 1$ for all $n \in \mathbb{N}$. We set

$$Y = \overline{\text{span}\,\{x_n : n \in \mathbb{N}\}}\,.$$

The space Y is separable by Proposition 1.17 and reflexive by Proposition 4.13. Since $Y^{**} \cong Y$, the bidual Y^{**} is separable. Lemma 6.10 now implies that Y^* is separable. The sequence $(J_Y x_n)_{n\in\mathbb{N}}$ in Y^{**} is bounded in Y^{**} and therefore has, according to Corollary 6.9, a weak star convergent subsequence $(J_Y x_{n_k})_{k\in\mathbb{N}}$, assume that $J_Y x_{n_k} \overset{*}{\rightharpoonup} y^{**}$. We set $x = J_Y^{-1} y^{**}$. For arbitrary $x^* \in X^*$ we get setting $y^* = x^*|Y \in Y^*$

$$x^*(x_{n_k}) = y^*(x_{n_k}) = (J_Y x_{n_k})(y^*) \to y^{**}(y^*) = y^*(x) = x^*(x)\,.$$

It follows that $x_{n_k} \rightharpoonup x$. □

The converse of Proposition 6.11 is valid, too. That is: If X is not reflexive, then $K(0;1)$ is not weakly sequentially compact. (We do not give the proof.)

Corollary 6.12 *Let X be a reflexive Banach space. Then every bounded sequence in X has a weakly convergent subsequence.* □

There arises the question: What about weak star compactness in X^* in the case where X is neither separable nor reflexive ?

Proposition 6.13 (Alaoglu)
Let X be a normed space. Then the closed unit ball $K(0;1)$ in X^* is weak star compact.
□

In order to understand this, one needs a notion which is more general than that of a metric space, because it turns out that in general there does not exist a metric on $K(0;1)$ for which the convergent sequences are just the weak star convergent sequences. This more general notion is that of a topological space. In a topological space, one does not axiomatize the metric (that is, the distance between two points), but rather the system of open sets. "Compact" then means the existence of a finite subcovering for every open covering (see Chapter 5 for a precise definition). Here, we do not go further into this subject.

We return to the approximation problem: let X be a normed space, $K \subset X$, $x \in X$. We want to find a $y \in K$ such that

$$\|x - y\| = \inf_{z \in K} \|x - z\| \,. \tag{6.14}$$

Proposition 6.14 *Let X be a reflexive Banach space, $K \subset X$ convex, closed and nonempty. Then for every $x \in X$ there exists a $y \in K$ such that*

$$\|x - y\| = \inf_{z \in K} \|x - z\| \,. \tag{6.15}$$

Proof: Let $(y_n)_{n \in \mathbb{N}}$ be a sequence in K such that $\lim_{n \to \infty} \|x - y_n\| = \inf_{z \in K} \|x - z\| =: d$. Since $\|y_n\| \leq \|x - y_n\| + \|x\|$, the sequence $(y_n)_{n \in \mathbb{N}}$ is bounded and due to Corollary 6.12 therefore has a weakly convergent subsequence, assume that $y_{n_k} \rightharpoonup y$, $y \in X$. By Proposition 6.6, K is weakly sequentially closed, thus $y \in K$. Since moreover $x - y_{n_k} \rightharpoonup x - y$, Lemma 6.3 implies that

$$\|x - y\| \leq \liminf_{k \to \infty} \|x - y_{n_k}\| = d \,,$$

which proves (6.15). $\square$

Definition 6.15 (Strict convexity)
*A normed space X is called **strictly convex** if for all $x_1, x_2 \in X$ with $\|x_1\| = \|x_2\| = 1$ and $x_1 \neq x_2$ we have that*

$$\left\| \frac{1}{2}(x_1 + x_2) \right\| < 1 \,. \tag{6.16}$$

$\square$

Proposition 6.16 *Let X be a normed space, $K \subset X$ convex, $x \in X$. If X is strictly convex, then there exists at most one $y \in K$ such that*

$$\|x - y\| = \inf_{z \in K} \|x - z\| \,. \tag{6.17}$$

Proof: Let $y, \tilde{y} \in K$ be two different solutions (6.17),

$$\|x - y\| = \|x - \tilde{y}\| = d \,.$$

We have $d > 0$ since otherwise $y = \tilde{y} = x$. We set

$$x_1 = \frac{1}{d}(x - y) \,, \quad x_2 = \frac{1}{d}(x - \tilde{y}) \,,$$

in Definition 6.15. Then

$$\frac{1}{d} \|x - \frac{1}{2}(y + \tilde{y})\| < 1 \,, \quad \frac{1}{2}(y + \tilde{y}) \in K \,,$$

which contradicts the minimality of y and $\tilde{y}$. $\square$

Vector spaces endowed with the p norm typically are strictly convex when $1 < p < \infty$ and not strictly convex for $p = 1$ and $p = \infty$ (sup norm). This is true e.g. for $\mathbb{K}^n$, $\ell^p(\mathbb{K})$, $L^p(D;\mathbb{K})$, $C(D;\mathbb{K})$. Since $\mathbb{K}^n$, $\ell^p(\mathbb{K})$ and $L^p(D;\mathbb{K})$ are also reflexive for $1 < p < \infty$, Propositions 6.14 and 6.16 imply that the approximation problem (6.14) has a unique solution in these spaces, for arbitrary closed convex sets K.

Conversely we have (no proof given here): If the approximation problem is uniquely solvable for arbitrary closed convex subsets K of X, then X must be strictly convex and reflexive.

7 Sobolev Spaces

Sobolev spaces arise when one wants to work with complete function spaces endowed with L^p norms which take into account not only the function but also their derivatives. In order to get there, however, one has to generalize the notion of a derivative.

Assume for the moment that $f : C[a,b] \to \mathbb{R}$ is continuously differentiable. Let $\varphi \in C_0^\infty(a,b)$. Then

$$\int_a^b f'(x)\varphi(x)\,dx = (f\varphi)\Big|_a^b - \int_a^b f(x)\varphi'(t)\,dx = -\int_a^b f(x)\varphi'(x)\,dx\,. \tag{7.1}$$

We use (7.1) as the starting point for the definition of the weak derivative. Let $f, g \in L^1(a,b)$. The function g is called a **weak derivative** of f on (a,b) if

$$\int_a^b g(x)\varphi(x)\,dx = -\int_a^b f(x)\varphi'(x)\,dx\,, \quad \text{for all } \varphi \in C_0^\infty(a,b). \tag{7.2}$$

Since g is, if it exists, uniquely determined by (7.2), we call g "the" weak derivative of f and denote it by f', too.

As an example we consider $f : [-1,1] \to \mathbb{R}$, $f(x) = |x|$. We have

$$\begin{aligned}
-\int_{-1}^1 f(x)\varphi'(x)\,dx &= \int_{-1}^0 x\varphi'(x)\,dx - \int_0^1 x\varphi'(x)\,dx \\
&= x\varphi(x)\Big|_{x=-1}^{x=0} - \int_{-1}^0 \varphi(x)\,dx - x\varphi(x)\Big|_{x=0}^{x=1} + \int_0^1 \varphi(x)\,dx \\
&= \int_{-1}^1 \operatorname{sign}(x)\varphi(x)\,dx
\end{aligned}$$

for all $\varphi \in C_0^\infty(a,b)$. Therefore f has the weak derivative $f'(x) = \operatorname{sign}(x)$.

Let us make one further step and compute

$$\begin{aligned}
-\int_{-1}^1 \operatorname{sign}(x)\varphi'(x)\,dx &= \int_{-1}^0 \varphi'(x)\,dx - \int_0^1 \varphi'(x)\,dx = (\varphi(0) - \varphi(-1)) - (\varphi(1) - \varphi(0)) \\
&= 2\varphi(0)\,.
\end{aligned} \tag{7.3}$$

This shows that f' does not possess a weak derivative in the sense of (7.2), because there does not exist an integrable function g such that

$$\int_{-1}^1 g(x)\varphi(x)\,dx = 2\varphi(0)\,, \quad \text{for all } \varphi \in C_0^\infty(a,b).$$

In order to interpret the right side as a derivative of the sign function, one has to further generalize the notion of the derivative; this leads to the notion of a distribution and of a distributional derivative. We will not discuss this.

In multidimensional space, one defines weak partial derivatives. Let $\Omega \subset \mathbb{R}^n$ be open, let f be continuously differentiable on Ω. The rule for partial integration then says that the function $g = \partial_i f$ satisfies

$$\int_\Omega g(x)\varphi(x)\,dx = -\int_\Omega f(x)\partial_i\varphi(x)\,dx\,, \quad \text{for all } \varphi \in C_0^\infty(\Omega). \tag{7.4}$$

Accordingly, g is called the weak i-th partial derivative of f in Ω, if f is integrable and (7.4) holds. For higher partial derivatives we use the notation $\partial^\alpha f$ with the multi-index $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$.

Definition 7.1 (Weak derivative)
*Let $\Omega \subset \mathbb{R}^n$ open, $1 \le p \le \infty$, let $u \in L^p(\Omega)$ and α be a multi-index. A $w \in L^p(\Omega)$ is called **weak α-th derivative** of u, if*

$$\int_\Omega w(x)\varphi(x)\,dx = (-1)^{|\alpha|} \int_\Omega u(x)\partial^\alpha \varphi(x)\,dx\,, \quad \text{for all } \varphi \in C_0^\infty(\Omega). \tag{7.5}$$

We denote it by $\partial^\alpha u$. □

Definition 7.2 (Sobolev space)
Let $\Omega \subset \mathbb{R}^n$ open. For $k \in \mathbb{N}$, $1 \le p \le \infty$ we define

$$W^{k,p}(\Omega) = \{v : v \in L^p(\Omega),\, \partial^\alpha v \in L^p(\Omega) \text{ for all } |\alpha| \le k\}\,. \tag{7.6}$$

□

The phrase "$\partial^\alpha v \in L^p(\Omega)$" means that the α-th weak derivative exists in the sense of Definition 7.1.

For $k = 0$, (7.6) means

$$W^{0,p}(\Omega) = L^p(\Omega)\,. \tag{7.7}$$

Proposition 7.3 *Let $\Omega \subset \mathbb{R}^n$ open, $1 \le p \le \infty$. The space $W^{k,p}(\Omega)$ is a Banach space when endowed with the norm*

$$\|v\|_{W^{k,p}(\Omega)} = \left(\sum_{|\alpha|\le k} \|\partial^\alpha v\|_p^p\right)^{\frac{1}{p}} = \left(\sum_{|\alpha|\le k} \int_\Omega |\partial^\alpha v(x)|^p\,dx\right)^{\frac{1}{p}}, \quad 1 \le p < \infty\,, \tag{7.8}$$

$$\|v\|_{W^{k,\infty}(\Omega)} = \sum_{|\alpha|\le k} \|\partial^\alpha v\|_\infty\,, \quad p = \infty\,. \tag{7.9}$$

Proof: For $p < \infty$, the triangle inequality follows from

$$\begin{aligned}\|u+v\|_{W^{k,p}(\Omega)} &= \left(\sum_{|\alpha|\le k} \|\partial^\alpha u + \partial^\alpha v\|_p^p\right)^{\frac{1}{p}} \le \left(\sum_{|\alpha|\le k} (\|\partial^\alpha u\|_p + \|\partial^\alpha v\|_p)^p\right)^{\frac{1}{p}} \\ &\le \left(\sum_{|\alpha|\le k} \|\partial^\alpha u\|_p^p\right)^{\frac{1}{p}} + \left(\sum_{|\alpha|\le k} \|\partial^\alpha v\|_p^p\right)^{\frac{1}{p}} = \|u\|_{W^{k,p}(\Omega)} + \|v\|_{W^{k,p}(\Omega)}\,.\end{aligned}$$

All other properties of the norm are immediate consequences of the definitions. Let now $(u_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $W^{k,p}(\Omega)$. Because of

$$\|\partial^\alpha u_n - \partial^\alpha u_m\|_p \le \|u_n - u_m\|_{W^{k,p}(\Omega)}\,,$$

the sequences $(\partial^\alpha u_n)_{n\in\mathbb{N}}$ are Cauchy sequences in $L^p(\Omega)$ for all $|\alpha| \le k$. Therefore there exist functions $u \in L^p(\Omega)$, $u_\alpha \in L^p(\Omega)$ such that

$$u_n \to u\,, \quad \partial^\alpha u_n \to u_\alpha \tag{7.10}$$

in $L^p(\Omega)$. Let $p < \infty$, $\varphi \in C_0^\infty(\Omega)$ be arbitrary, then

$$\begin{aligned}\int_\Omega u(x)\partial^\alpha\varphi(x)\,dx &= \lim_{n\to\infty}\int_\Omega u_n(x)\partial^\alpha\varphi(x)\,dx = \lim_{n\to\infty}(-1)^{|\alpha|}\int_\Omega \partial^\alpha u_n(x)\varphi(x)\,dx \\ &= (-1)^{|\alpha|}\int_\Omega u_\alpha(x)\varphi(x)\,dx\,.\end{aligned}$$

Therefore, for all $|\alpha| \le k$ the functions u_α satisfy the requirements in the definition of the weak derivative of u, so $\partial^\alpha u$ exists and $\partial^\alpha u = u_\alpha$ for all $|\alpha| \le k$. Therefore, $u \in W^{k,p}(\Omega)$. The proof in the case $p = \infty$ is analogous. □

The norm in $W^{k,p}(\Omega)$ is defined in such a way that we may easily embed $W^{k,p}(\Omega)$ into a product of L^p spaces. We set

$$X = \prod_{|\alpha|\le k} X_\alpha\,, \quad X_\alpha = L^p(\Omega)\,. \tag{7.11}$$

Thus, an element $v \in X$ has the form $v = (v_\alpha)_{|\alpha|\le k}$ with $v_\alpha \in X_\alpha = L^p(\Omega)$.

We endow X with the p-norm of the product,

$$\|v\|_X^p = \sum_{|\alpha|\le k} \|v_\alpha\|_{L^p(\Omega)}^p\,.$$

According to Proposition 1.7, X is a Banach space. Moreover, X is reflexive for $1 < p < \infty$, because products of reflexive Banach spaces are again reflexive; X is separable for $1 \le p < \infty$, because products of separable metric spaces are separable. We define

$$T : W^{k,p}(\Omega) \to X\,, \quad (Tv)_\alpha = \partial^\alpha v\,. \tag{7.12}$$

Lemma 7.4 *We have*

$$\|Tv\|_X = \|v\|_{W^{k,p}(\Omega)}\,, \quad \textit{for all } v \in W^{k,p}(\Omega). \tag{7.13}$$

$T(W^{k,p}(\Omega))$ is a closed subspace of X. For $1 < p < \infty$, $W^{k,p}(\Omega)$ is reflexive and for $1 \le p < \infty$ it is separable.

Proof: The equality (7.13) immediately follows from the definition of the two norms. Since $W^{k,p}(\Omega)$ is complete by Proposition 7.3, the space $T(W^{k,p}(\Omega))$, too, is complete and therefore closed in X. Let now $p < \infty$. $W^{k,p}(\Omega)$ is reflexive, since closed subspaces of reflexive spaces are reflexive by Proposition 4.13. $W^{k,p}(\Omega)$ is separable, since arbitrary subsets of a separable metric space are separable (exercise). □

Lemma 7.4 provides another way to prove that $W^{k,p}(\Omega)$ is complete; one proves directly that $W^{k,p}(\Omega)$ is a closed subspace of X (and therefore complete, since X is complete).

It then suffices to consider convergent (in $L^p(\Omega)$) sequences $u_n \to u$, $\partial^\alpha u_n \to u_\alpha$ and to prove, as in the final part of the proof of Proposition 7.3, that $u_\alpha = \partial^\alpha u$ holds.

We recall the standard mollifier $\eta_\varepsilon : \mathbb{R}^n \to \mathbb{R}$ and the regularization

$$v^\varepsilon = v * \eta_\varepsilon\,, \quad \varepsilon > 0\,. \tag{7.14}$$

Lemma 7.5 *Let $\Omega, U \subset \mathbb{R}^n$ be open, assume that $U \subset\subset \Omega$. Let $v \in W^{k,p}(\Omega)$, $p < \infty$. Then for all ε satisfying $0 < \varepsilon < \operatorname{dist}(U, \partial\Omega)$ we have that $v^\varepsilon \in C^\infty(U) \cap W^{k,p}(U)$, and $v^\varepsilon \to v$ in $W^{k,p}(U)$ for $\varepsilon \to 0$.*

Proof: Let $\varepsilon < \operatorname{dist}(U, \partial\Omega)$. By Lemma 5.6, $v^\varepsilon \in C^\infty(U)$. Moreover, for all $y \in U$

$$\begin{aligned}\partial^\alpha v^\varepsilon(y) &= \int_\Omega \partial_y^\alpha \eta_\varepsilon(y-x) v(x)\,dx = (-1)^{|\alpha|} \int_\Omega \partial_x^\alpha \eta_\varepsilon(y-x) v(x)\,dx \\ &= \int_\Omega \eta_\varepsilon(y-x) \partial^\alpha v(x)\,dx = (\eta_\varepsilon * \partial^\alpha v)(y)\,.\end{aligned} \tag{7.15}$$

Since $\partial^\alpha v \in L^p(\Omega)$, we obtain that $\partial^\alpha v^\varepsilon \in L^p(U)$ by Lemma 5.6, and that $\partial^\alpha v^\varepsilon \to \partial^\alpha v$ in $L^p(U)$ by Lemma 5.8. □

In order that the first equality in (7.15) holds, one needs that $\varepsilon < \operatorname{dist}(U, \partial\Omega)$; for $U = \Omega$ the assertion of the Lemma in general does not hold.

Proposition 7.6 *Let $\Omega \subset \mathbb{R}^n$ be open, let $v \in W^{k,p}(\Omega)$, $1 \le p < \infty$. Then there exists a sequence $(v_n)_{n\in\mathbb{N}}$ in $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ such that $v_n \to v$ in $W^{k,p}(\Omega)$.*

Proof: We define

$$U_j = \{x : x \in \Omega,\ \operatorname{dist}(x, \partial\Omega) > \frac{1}{j} \text{ and } |x| < j\}\,, \quad V_j = U_{j+3} \setminus \overline{U}_{j+1}\,, \quad j \ge 1\,, \tag{7.16}$$

and set $V_0 = U_3$. Then

$$\Omega = \bigcup_{j=0}^\infty V_j\,.$$

Let $(\beta_j)_{j\ge 0}$ be a partition of unity for Ω such that

$$0 \le \beta_j \le 1\,, \quad \beta_j \in C_0^\infty(V_j)\,, \quad \sum_{j=0}^\infty \beta_j = 1\,. \tag{7.17}$$

Since $v \in W^{k,p}(\Omega)$, we have $\beta_j v \in W^{k,p}(\Omega)$ (see an exercise) and $\operatorname{supp}(\beta_j v) \subset V_j$. Let now $\delta > 0$ be arbitrary. We choose $\varepsilon_j > 0$ sufficiently small such that for

$$w_j = (\beta_j v) * \eta_{\varepsilon_j}$$

we have (by Lemma 7.5 applied to $\beta_j v$)

$$\|w_j - \beta_j v\|_{W^{k,p}(\Omega)} \le 2^{-(j+1)}\delta\,, \tag{7.18}$$

$$\operatorname{supp}(w_j) \subset W_j := U_{j+4} \setminus \overline{U}_j\,, \quad j \ge 1\,, \quad W_0 := U_4\,. \tag{7.19}$$

We set

$$w = \sum_{j=0}^{\infty} w_j \,.$$

By construction, on each W_j only finitely many summands are different from zero. It follows that $w \in C^\infty(\Omega)$, because each $w_j \in C^\infty(\Omega)$ by Lemma 7.5. We then get

$$\|w - v\|_{W^{k,p}(\Omega)} = \left\| \sum_{j=0}^{\infty} w_j - \sum_{j=0}^{\infty} \beta_j v \right\|_{W^{k,p}(\Omega)} \le \sum_{j=0}^{\infty} \|w_j - \beta_j v\|_{W^{k,p}(\Omega)} \le \delta \sum_{j=0}^{\infty} 2^{-(j+1)} = \delta \,.$$

Since $\delta > 0$ was arbitrary, the assertion follows. □

Definition 7.7 *Let $\Omega \subset \mathbb{R}^n$ be open, let $1 \le p < \infty$, $k \in \mathbb{N}$. We define $W_0^{k,p}(\Omega) \subset W^{k,p}(\Omega)$ by*

$$W_0^{k,p}(\Omega) = \overline{C_0^\infty(\Omega)} \,, \tag{7.20}$$

where the closure is taken w.r.t. the norm of $W^{k,p}(\Omega)$. □

$W_0^{k,p}(\Omega)$ is a closed subspace of $W^{k,p}(\Omega)$ and thus a Banach space, too (endowed with the norm of $W^{k,p}(\Omega)$). $W_0^{k,p}(\Omega)$ is a function space whose elements have zero values on $\partial\Omega$ in a certain "weak" sense. (From the outset, it is not clear that for a general element $v \in W^{k,p}(\Omega)$ the assertion "$v = 0$ on $\partial\Omega$" makes sense, because v is actually an equivalence class of functions which may differ on a set of measure zero; but $\partial\Omega$ usually has measure 0 as a subset of $\mathbb{R}^n$.)

Definition 7.8 *Let $\Omega \subset \mathbb{R}^n$ be open. For $k \in \mathbb{N}$ we define*

$$H^k(\Omega) = W^{k,2}(\Omega) \,, \quad H_0^k(\Omega) = W_0^{k,2}(\Omega) \,. \tag{7.21}$$

□

Proposition 7.9 *Let $\Omega \subset \mathbb{R}^n$ be open, $k \in \mathbb{N}$. Then $H^k(\Omega)$ and $H_0^k(\Omega)$ are Hilbert spaces with the scalar product*

$$\langle u, v \rangle_{H^k(\Omega)} = \sum_{|\alpha| \le k} \langle \partial^\alpha u, \partial^\alpha v \rangle_{L^2(\Omega)} = \sum_{|\alpha| \le k} \int_\Omega \partial^\alpha u(x) \cdot \overline{\partial^\alpha v(x)} \, dx \,, \tag{7.22}$$

and

$$\|v\|_{W^{k,2}(\Omega)} = \sqrt{\langle v, v \rangle_{H^k(\Omega)}} \,, \quad v \in H^k(\Omega) \,. \tag{7.23}$$

Proof: The properties of the scalar product immediately follow from the corresponding properties of the scalar product in $L^2(\Omega)$. Obviously (7.23) holds, and $H^k(\Omega)$ is complete by Proposition 7.3. □

Lemma 7.10 *Let $\Omega \subset \mathbb{R}^n$ be open, $1 \le p < \infty$, $1/p + 1/q = 1$. Let $u \in C_0^\infty(\Omega)$. Then for all $h \in \mathbb{R}^n$ we have*

$$\int_\Omega |u(x+h) - u(x)|^p \, dx \le \|\nabla u\|_{L^p(\Omega)}^p \|h\|_q^p \,, \tag{7.24}$$

where again we set $u = 0$ outside of Ω.

Proof: We set $g(t) = u(x + th)$. We have

$$u(x+h) - u(x) = g(1) - g(0) = \int_0^1 g'(t)\, dt = \int_0^1 \langle \nabla u(x+th), h \rangle \, dt$$

and moreover, due to Hölder's inequality in $\mathbb{K}^n$,

$$\begin{aligned} |u(x+h) - u(x)| &\leq \int_0^1 |\langle \nabla u(x+th), h \rangle| \, dt \leq \int_0^1 \|\nabla u(x+th)\|_p \|h\|_q \, dt \\ &\leq \int_0^1 \|\nabla u(x+th)\|_p \, dt \cdot \|h\|_q \, . \end{aligned} \tag{7.25}$$

In the case $p = 1$ we obtain (7.24) by integrating over Ω. In the case $p > 1$ we estimate as

$$\begin{aligned} \int_\Omega |u(x+h) - u(x)|^p \, dx &\leq \int_\Omega \Big(\int_0^1 \|\nabla u(x+th)\|_p \, dt \cdot \|h\|_q \Big)^p \, dx \\ &\leq \int_{\mathbb{R}^n} \int_0^1 \|\nabla u(x+th)\|_p^p \, dt \cdot \underbrace{\Big(\int_0^1 1^q \, dt \Big)^{p/q}}_{=1} \, dx \cdot \|h\|_q^p \\ &= \int_0^1 \int_{\mathbb{R}^n} \|\nabla u(x+th)\|_p^p \, dx \, dt \cdot \|h\|_q^p \\ &= \|\nabla u\|_{L^p(\Omega)}^p \cdot \|h\|_q^p \, , \end{aligned}$$

note that the norm of the gradient of u on $\mathbb{R}^n$ remains unchanged when we translate every vector by the vector th, and that $\nabla u = 0$ outside of Ω. $\Box$

Proposition 7.11 *Let $\Omega \subset \mathbb{R}^n$ be open, $1 \leq p < \infty$, let F be a bounded subset of $W_0^{1,p}(\Omega)$. Then F is relatively compact in $L^p(\Omega)$.*

Proof: Let (u_n) be a sequence in F. According to the definition of $W_0^{1,p}(\Omega)$ we can find a sequence (v_n) in $C_0^\infty(\Omega)$ such that

$$\|v_n - u_n\|_{W^{1,p}(\Omega)} \leq \frac{1}{n} \, . \tag{7.26}$$

Thus, the sequence (v_n), too, is bounded in $W^{1,p}(\Omega)$. Therefore, both (v_n) and (∇v_n) are bounded in $L^p(\Omega)$. Using Lemma 7.10 we obtain

$$\int_\Omega |v_n(x+h) - v_n(x)|^p \, dx \leq \|\nabla v_n\|_{L^p(\Omega)}^p \|h\|_q^p \leq C \|h\|_q^p$$

with a constant C which does not depend on n. It follows that

$$\lim_{h \to 0} \sup_{n \in \mathbb{N}} \int_\Omega |v_n(x+h) - v_n(x)|^p \, dx = 0 \, .$$

By the theorem of Fréchet-Riesz-Kolmogorov, (v_n) is relatively compact in $L^p(\Omega)$. Therefore, there exists a $u \in L^p(\Omega)$ and a subsequence (v_{n_k}) such that $v_{n_k} \to u$ in $L^p(\Omega)$. Due to

(7.26) we also have $u_{n_k} \to u$. As the sequence (u_n) was arbitrary, F is relatively compact in $L^p(\Omega)$. $\square$

The corresponding result for $W^{1,p}(\Omega)$ may be deduced from Proposition 7.11, but this requires some additional constructions. If $\Omega \subset \mathbb{R}^n$ is open and bounded, and if $\partial\Omega$ is sufficiently smooth, one may construct, for an arbitrary given open and bounded set $V \subset \mathbb{R}^n$ with $\Omega \subset\subset V$, a linear and continuous extension operator

$$E : W^{1,p}(\Omega) \to W_0^{1,p}(V)$$

satisfying $(Eu)|\Omega = u$ for all $u \in W^{1,p}(\Omega)$. (See e.g. Evans, Partial Differential Equations.) Let now F be bounded in $W^{1,p}(\Omega)$. Then $E(F)$ is bounded in $W_0^{1,p}(V)$. For every sequence (u_n) in $W^{1,p}(\Omega)$, the sequence (Eu_n) has a subsequence which converges to some $w \in L^p(V)$ by Proposition 7.11. In particular, the corresponding subsequence (u_{n_k}) converges to a $u = w|\Omega \in L^p(\Omega)$. With this reasoning one proves the following proposition.

Proposition 7.12 (Rellich)
Let $\Omega \subset \mathbb{R}^n$ be open and bounded, with sufficiently smooth boundary, let $1 \leq p < \infty$. Assume that F is a bounded subset of $W^{1,p}(\Omega)$. Then F is relatively compact in $L^p(\Omega)$.
$\square$

8 Compact operators

Definition 8.1 (Compact operator)
*Let X, Y be Banach spaces. A linear mapping $T : X \to Y$ is called **compact**, or **compact operator** if the image $T(K(0;1))$ of the closed unit ball $K(0;1)$ is relatively compact in Y.* □

Since a linear mapping is continuous if and only if the image of the unit ball is bounded, every compact operator is continuous. By

$$K(X;Y) = \{T : T \in L(X;Y),\, T \text{ compact operator}\} \tag{8.1}$$

we denote the set of all compact operators from X to Y and set

$$K(X) = K(X;X)\,. \tag{8.2}$$

By what we just said, $K(X;Y) \subset L(X;Y)$.

Lemma 8.2 *Let X, Y be Banach spaces, $T : X \to Y$ linear. Then there are equivalent:*

(i) T is compact.

(ii) $T(B)$ is relatively compact in Y for every bounded subset $B \subset X$.

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

Proof: "(i)⇒(ii)": If $T(K(0;1))$ is relatively compact, then so is $T(K(0;R)) = RT(K(0;1))$ for all $R > 0$. If B is bounded, so is $T(B) \subset T(K(0;R))$ for R sufficiently large. Therefore, $\overline{T(B)}$ is compact as a closed subset of the compact set $\overline{T(K0;R))}$.
"(ii)⇒(iii)": Immediate. "(iii)⇒(i)": Exercise. □

Compact operators map weakly convergent sequences to strongly convergent sequences.

Proposition 8.3 *Let X, Y be Banach spaces, let $T : X \to Y$ be a compact linear operator. let (x_n) be a bounded sequence in X. Then:*
(i) $(Tx_n)_{n\in\mathbb{N}}$ has a subsequence which strongly converges in Y.
(ii) From $x_n \rightharpoonup x \in X$ it follows that $Tx_n \to Tx$.

Proof: Exercise. □

As an example we consider

$$T : C[0,1] \to C[0,1]\,, \quad (Tx)(t) = \int_0^1 k(s,t)x(s)\,ds\,, \tag{8.3}$$

where $k : [0,1] \times [0,1] \to \mathbb{R}$ is continuous. We have shown in an exercise that the image of the unit ball in $C[0,1]$ under T is relatively compact in $C[0,1]$. Therefore, T is compact.

As a second example we again consider the integral operator defined by

$$(Tx)(t) = \int_0^1 k(s,t)x(s)\,ds\,, \tag{8.4}$$

this time as an operator from $L^2(0,1)$ to $L^2(0,1)$.

Proposition 8.4 *Let $k \in L^2(\Omega)$ mit $\Omega = (0,1) \times (0,1)$. Then (8.4) defines a linear continuous operator $T : L^2(0,1) \to L^2(0,1)$ satisfying $\|T\| \le \|k\|_{L^2(\Omega)}$.*

Proof: For $x \in L^2(0,1)$ we obtain, using the Cauchy-Schwarz inequality in $L^2(0,1)$,

$$\begin{aligned}\|Tx\|_2^2 &= \int_0^1 \Big| \int_0^1 k(s,t)x(s)\,ds\Big|^2 dt \le \int_0^1 \Big(\int_0^1 |k(s,t)|^2\,ds \cdot \int_0^1 |x(s)|^2\,ds\Big)\,dt \\ &= \int_0^1 \int_0^1 |k(s,t)|^2\,ds\,dt \cdot \int_0^1 |x(s)|^2\,ds \\ &= \|k\|_{L^2(\Omega)}^2 \cdot \|x\|_2^2\,,\end{aligned}$$

so

$$\|Tx\|_2 \le \|k\|_{L^2(\Omega)} \|x\|_2\,, \quad \text{for all } x \in L^2(\Omega).$$

This yields the assertion. That the integrals are well-defined is a consequence of Fubini's theorem resp. its variants. □

Proposition 8.5 *The operator $T : L^2(0,1) \to L^2(0,1)$ considered in Proposition 8.4 is compact.*

Proof: Let $x \in L^2(0,1)$ and $h \in \mathbb{R}$ be arbitrary. We set $k(s,t) = 0$ and correspondingly $(Tx)(t) = 0$ for $t \notin (0,1)$. Analogously as in the proof of Proposition 8.4 we obtain

$$\begin{aligned}\int_0^1 |(Tx)(t+h) - (Tx)(t)|^2\,dt &= \int_0^1 \Big| \int_0^1 (k(s,t+h) - k(s,t))x(s)\,ds\Big|^2 dt \\ &\le \int_0^1 \int_0^1 |k(s,t+h) - k(s,t)|^2\,ds \cdot \int_0^1 |x(s)|^2\,ds\,dt \qquad (8.5)\\ &= \int_0^1 \int_0^1 |k(s,t+h) - k(s,t)|^2\,ds\,dt \cdot \|x\|_2^2\,.\end{aligned}$$

Using Proposition 5.9 applied to k in $L^2(\Omega)$ we get that

$$\lim_{h\to 0} \int_0^1 \int_0^1 |k(s,t+h) - k(s,t)|^2\,ds\,dt = 0\,.$$

From (8.5) it follows that

$$\lim_{h\to 0} \sup_{\|x\|_2 \le 1} \int_0^1 |(Tx)(t+h) - (Tx)(t)|^2\,dt \le \lim_{h\to 0} \int_0^1 \int_0^1 |k(s,t+h) - k(s,t)|^2\,ds\,dt = 0\,.$$

The theorem of Fréchet-Riesz-Kolmogorov (Proposition 5.10) now implies that the image of the unit ball, namely $\{Tx : \|x\|_2 \le 1\}$, is relatively compact in $L^2(0,1)$. □

As an immediate consequence of Proposition 8.3 we obtain:

Corollary 8.6 *Let T the operator considered in Proposition 8.4, let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in $L^2(0,1)$. Then the sequence $(Tx_n)_{n\in\mathbb{N}}$ has a subsequence which strongly converges in $L^2(0,1)$.* □

As a third example we consider the embedding

$$I : W^{1,p}(\Omega) \to L^p(\Omega) . \tag{8.6}$$

It is compact by the result of Rellich (Proposition 7.12). We obtain, again using Proposition 8.3, that every sequence which is bounded in the norm of $W^{1,p}(\Omega)$ has a subsequence which is strongly convergent in $L^p(\Omega)$.

Let $T \in L(X;Y)$, let

$$\dim(X) < \infty , \quad \text{or} \quad \dim(T(X)) < \infty . \tag{8.7}$$

Then $T(K(0;1))$ is a bounded subset of a finite-dimensional space and thus relatively compact. Therefore, T is compact.

Lemma 8.7 *Let X, Y, Z be Banach spaces, $T \in L(X;Y)$, $S \in L(Y;Z)$. If T or S is compact, then $S \circ T$ is compact.*

Proof: Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in X. If T is compact, the sequence $(Tx_n)_{n\in\mathbb{N}}$ has a convergent subsequence $(Tx_{n_k})_{k\in\mathbb{N}}$, and because S is continuous, the sequence $(S(Tx_{n_k}))_{k\in\mathbb{N}}$ converges, too. If S is compact, the bounded sequence $(Tx_n)_{n\in\mathbb{N}}$ (T is linear and continuous) gives rise to a convergent subsequence $(S(Tx_{n_k}))_{k\in\mathbb{N}}$. □

Proposition 8.8 *Let X, Y be Banach spaces. The space $K(X;Y)$ is a closed subspace of $L(X;Y)$.*

Proof: Let $(T_m)_{m\in\mathbb{N}}$ be a sequence in $K(X;Y)$ with $T_m \to T$, $T \in L(X;Y)$. We want to prove that T is compact. Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in X, let $\|x_n\| \le C$ for all $n \in \mathbb{N}$. We choose a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that $(T_m x_{n_k})_{k\in\mathbb{N}}$ converges in Y for all $m \in \mathbb{N}$. In order to achieve this we utilize a diagonalization argument like the one we already carried out in the proof of Proposition 6.8. Next, we show that $(Tx_{n_k})_{k\in\mathbb{N}}$ is a Cauchy sequence: for a given $\varepsilon > 0$ we choose $m \in \mathbb{N}$ such that $\|T - T_m\| \le \varepsilon$, and $K > 0$ such that

$$\|T_m x_{n_k} - T_m x_{n_j}\| < \varepsilon , \quad \text{for all } k, j \ge K.$$

Then for all $j, k \ge K$ we have

$$\begin{aligned}
\|Tx_{n_k} - Tx_{n_j}\| &\le \|Tx_{n_k} - T_m x_{n_k}\| + \|T_m x_{n_k} - T_m x_{n_j}\| + \|T_m x_{n_j} - Tx_{n_j}\| \\
&\le \|T - T_m\| \, \|x_{n_k}\| + \|T_m x_{n_k} - T_m x_{n_j}\| + \|T_m - T\| \, \|x_{n_j}\| \\
&\le C\varepsilon + \varepsilon + C\varepsilon = (2C+1)\varepsilon .
\end{aligned}$$

Thus, $(Tx_{n_k})_{k\in\mathbb{N}}$ is a Cauchy sequence and therefore convergent, since Y is complete. □

Corollary 8.9 *Let X, Y be Banach spaces, let $T : X \to Y$ be linear and continuous, assume that there exists a sequence $(T_n)_{n\in\mathbb{N}}$ of linear continuous operators $T_n : X \to Y$ such that*

$$\lim_{n\to\infty} \|T_n - T\| = 0 , \quad \dim(T_n(X)) < \infty \quad \textit{for all } n \in \mathbb{N}. \tag{8.8}$$

Then T is compact.

Proof: Every linear continuous operator, whose range is finite dimensional, is compact. The assertion then follows directly from Proposition 8.8. □

There arises the question whether every compact operator $T : X \to Y$ between Banach spaces X and Y can be obtained as the limit of a sequence $(T_n)_{n\in\mathbb{N}}$ of operators $T_n : X \to Y$ with finite dimensional range. This is not the case. In 1973, P. Enflo has published a counterexample.

9 Adjoint Operators

Let X, Y be normed spaces, let $T : X \to Y$ be linear and continuous. For a given $y^* \in Y^*$,

$$T^* y^* = y^* \circ T \tag{9.1}$$

defines a linear continuous mapping $T^* y^* \in X^*$. Obviously $T^* : Y^* \to X^*$ is linear.

Lemma 9.1 *T^* is continuous, and $\|T^*\| = \|T\|$.*

Proof: We have

$$\sup_{\|y^*\| \le 1} \|T^* y^*\| = \sup_{\|y^*\| \le 1} \sup_{\|x\| \le 1} |(T^* y^*)(x)| = \sup_{\|x\| \le 1} \sup_{\|y^*\| \le 1} |y^*(Tx)| = \sup_{\|x\| \le 1} \|Tx\| = \|T\|\,.$$

The next-to-last equality holds because of Corollary 4.7. □

Definition 9.2 (Adjoint Operator)
Let X, Y be normed spaces, let $T \in L(X; Y)$. The operator $T^* \in L(Y^*; X^*)$ defined by (9.1) is called the adjoint operator to T. □

Lemma 9.3 *The association $T \mapsto T^*$ defines an isometric linear mapping from $L(X; Y)$ to $L(Y^*; X^*)$.*

Proof: That T is linear is an immediate consequence of the definition; that it is isometric follows from Lemma 9.1. □

The mapping $T \mapsto T^*$ is not always surjective. (A counterexample can be found in the book of D. Werner.)

As an example we consider the integral operator $T : L^2(0,1) \to L^2(0,1)$,

$$(T\xi)(t) = \int_0^1 k(s,t)\xi(s)\,ds\,, \quad k \in L^2((0,1) \times (0,1))\,. \tag{9.2}$$

The adjoint operator $T^* : L^2(0,1)^* \to L^2(0,1)^*$ satisfies

$$(T^* y^*)(\xi) = y^*(T\xi)\,. \tag{9.3}$$

According to the isometric isomorphism between $L^2(0,1)^*$ and $L^2(0,1)$, see Chapter 1, y^* can be represented by a $y \in L^2(0,1)$,

$$y^*(T\xi) = \int_0^1 (T\xi)(t) y(t)\,dt\,. \tag{9.4}$$

It then follows that

$$(T^* y^*)(\xi) = \int_0^1 (T\xi)(t) y(t)\,dt = \int_0^1 \int_0^1 k(s,t)\xi(s)\,ds\,y(t)\,dt \tag{9.5}$$

$$= \int_0^1 \int_0^1 k(s,t) y(t)\,dt\,\xi(s)\,ds\,, \tag{9.6}$$

that is, $x^* = T^* y^*$ is represented by the function

$$x(s) = \int_0^1 k(s,t) y(t)\, dt\,, \quad \text{oder} \qquad x(t) = \int_0^1 k(t,s) y(s)\, ds\,. \tag{9.7}$$

We summarize: If T is defined by the kernel function "$k = k(s,t)$" then T^* has the kernel function "$k = k(t,s)$".

Applying Definition 9.2 twice, we obtain for a given $T \in L(X;Y)$ the linear and continuous operator

$$T^{**} : X^{**} \to Y^{**}\,.$$

Lemma 9.4 *Let X, Y be normed spaces, let $T \in L(X;Y)$. Then*

$$T^{**} \circ J_X = J_Y \circ T\,. \tag{9.8}$$

Proof: We have

$$\begin{aligned}[(T^{**} \circ J_X)(x)](y^*) &= [T^{**}(J_X x)](y^*) = [(J_X x) \circ T^*](y^*) = (J_X x)(T^* y^*) \\ &= (T^* y^*)(x) = y^*(Tx) = [J_Y(Tx)](y^*) = [(J_Y \circ T)(x)](y^*)\,.\end{aligned}$$

□

Proposition 9.5 *Let X, Y be Banach spaces, let $T : X \to Y$ be linear and continuous. Then T is compact if and only if T^* is compact.*

Proof: Let T be compact. Then, the subset of Y defined by $K = \overline{T(K(0;1))}$ is compact. Let now $(y_n^*)_{n\in\mathbb{N}}$ be a bounded sequence in Y^*, let

$$\|y_n^*\|_{Y^*} \le M\,. \tag{9.9}$$

We consider the sequence $(y_n^*|K)_{n\in\mathbb{N}}$ in $C(K)$. This sequence is bounded in $C(K)$ since

$$\|y_n^*|K\|_\infty = \sup_{y\in K} |y_n^*(y)| \le \|y_n^*\|_{Y^*} \sup_{y\in K} \|y\| \le M \sup_{y\in K} \|y\| < \infty\,. \tag{9.10}$$

Moreover, for arbitrary $y, \tilde{y} \in K$ we have

$$|y_n^*(y) - y_n^*(\tilde{y})| \le \|y_n^*\|_{Y^*}\, \|y - \tilde{y}\|_Y \le M \|y - \tilde{y}\|_Y\,. \tag{9.11}$$

Therefore, $(y_n^*)_{n\in\mathbb{N}}$ is equicontinuous. It follows from the Arzela-Ascoli Theorem that there exists a subsequence $(y_{n_k}^*|K)_{k\in\mathbb{N}}$ which uniformly converges on K. It follows that

$$\|T^* y_{n_k}^* - T^* y_{n_l}^*\| = \sup_{\|x\|\le 1} |y_{n_k}^*(Tx) - y_{n_l}^*(Tx)| \le \sup_{y\in K} |y_{n_k}^*(y) - y_{n_l}^*(y)| = \|(y_{n_k}^* - y_{n_l}^*)|K\|_\infty\,; \tag{9.12}$$

Therefore, $(T^* y_{n_k}^*)_{k\in\mathbb{N}}$ is a Cauchy sequence and thus convergent. This proves that T^* is compact. For the converse, assume that T^* is compact. By what we have just proved, T^{**} is compact. By Lemma 8.7, $T^{**} \circ J_X$ is compact. By Lemma 9.4, $J_Y \circ T$ is compact. Let now $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in X. Then $((J_Y \circ T)x_n)_{n\in\mathbb{N}}$ has a convergent subsequence, let $(J_Y \circ T)x_{n_k} \to y^{**} \in Y^{**}$. Since Y is a Banach space, $J_Y(Y)$ is closed. Thus, there exists a $y \in Y$ such that $y^{**} = J_Y y$. Since J_Y is isometric, we get $Tx_{k_n} \to y$. We have found a convergent subsequence of $(x_n)_{n\in\mathbb{N}}$; by Lemma 8.2, T is compact. □

Proposition 9.6 *Let X, Y be normed spaces, let $T : X \to Y$ be linear and continuous. Then*

$$\overline{T(X)} = (\ker T^*)^o\,, \tag{9.13}$$

where

$$(\ker T^*)^o = \{y : y \in Y,\; y^*(y) = 0 \text{ for all } y^* \in \ker T^*\}\,. \tag{9.14}$$

Proof: "$\subset$": Let $y \in T(X)$, so $y = Tx$ for some $x \in X$. For arbitrary $y^* \in \ker T^*$ we get $y^*(y) = y^*(Tx) = (T^*y^*)(x) = 0$. It follows that $T(X) \subset (\ker T^*)^o$. Since $(\ker T^*)^o$ is closed, as one easily verifies, the assertion follows.
"$\supset$": Let $y \in Y$, $y \notin \overline{T(X)}$. According to Corollary 4.8 we choose a $y^* \in Y^*$ such that $y^* = 0$ on $\overline{T(X)}$ and $y^*(y) \neq 0$. For all $x \in X$ we have $0 = y^*(Tx) = (T^*y^*)(x)$, so $T^*y^* = 0$ and thus $y^* \in \ker T^*$. Since $y^*(y) \neq 0$, we obtain $y \notin (\ker T^*)^o$. $\square$

10 Complements, Factorization

We recall: if X is a normed space and U a closed subspace of X, then

$$\|[x]\| = \inf_{z\in U} \|x - z\| = \inf_{\tilde{x}\in[x]} \|\tilde{x}\| \tag{10.1}$$

defines a norm on X/U, and the quotient mapping

$$Q : X \to X/U\,, \quad Q(x) = [x]\,, \tag{10.2}$$

is linear and continuous and has norm $\|Q\| = 1$ unless $X = \{0\}$.

Proposition 10.1 *Let X, Y be normed spaces, U a closed subspace of X, let $T \in L(X;Y)$ with $T|U = 0$. Then there exists a unique $\tilde{T} \in L(X/U;Y)$ such that $\tilde{T} \circ Q = T$, and we have $\|\tilde{T}\| = \|T\|$. If in particular $U = \ker T$, then $\tilde{T}$ is injective.*

Proof: We define

$$\tilde{T} : X/U \to Y\,, \quad \tilde{T}([x]) = T(x)\,. \tag{10.3}$$

$\tilde{T}$ is well-defined since if $\tilde{x} \in [x]$ we have $\tilde{x} - x \in U$, so

$$T(\tilde{x}) - T(x) = T(\tilde{x} - x) = 0\,.$$

A direct computation shows that $\tilde{T}$ is linear. Moreover,

$$\|T(x)\| = \|T(\tilde{x})\| \le \|T\|\,\|\tilde{x}\|\,, \quad \text{for all } \tilde{x} \in [x],$$

therefore

$$\|\tilde{T}([x])\| \le \inf_{\tilde{x}\in[x]} \|T\|\,\|\tilde{x}\| = \|T\|\,\|[x]\|\,.$$

Thus, $\|\tilde{T}\| \le \|T\|$ and conversely

$$\|T\| = \|\tilde{T} \circ Q\| \le \|\tilde{T}\|\,\|Q\| = \|\tilde{T}\|\,.$$

If $U = \ker T$, it follows from (10.3) that

$$\tilde{T}([x]) = 0 \quad \Leftrightarrow \quad x \in U \quad \Leftrightarrow \quad [x] = 0\,,$$

so $\ker \tilde{T} = 0$. □

Corollary 10.2 *Let X, Y be Banach spaces, $T : X \to Y$ linear and continuous, $T(X)$ closed. Then*

$$X/\ker T \simeq T(X)\,, \tag{10.4}$$

and the mapping $\tilde{T}$ defined by $\tilde{T} \circ Q = T$ is an isomorphism between $X/\ker T$ and $T(X)$.

Proof: By Proposition 10.1, $\tilde{T} : X/\ker T \to T(X)$ is bijective, linear and continuous. Since Y is a Banach space and $T(X)$ is a closed subspace of Y, $T(X)$ also is a Banach space. From Corollary 3.8 we obtain that $\tilde{T}^{-1}$ is continuous, too. □

Corollary 10.3 *Let X, Y be Banach spaces, let $T : X \to Y$ linear, continuous and surjective. Then*
$$X/\ker T \simeq Y\,. \tag{10.5}$$

Proof: This is an immediate consequence of Corollary 10.2. □

Proposition 10.4 *Let X, Y be normed spaces, let $T : X \to Y$ linear and continuous, let $T(X)$ closed in Y. Then*
$$(Y/T(X))^* \cong \ker T^*\,. \tag{10.6}$$

Proof: Let $z^* \in (Y/T(X))^*$. We set $y^* = z^* \circ Q$ where $Q : Y \to Y/T(X)$ is the quotient mapping. We then have $y^* \in Y^*$, $y^* = 0$ on $T(X)$, therefore $T^*y^* = 0$ and thus $y^* \in \ker T^*$. We define

$$I : (Y/T(X))^* \to \ker T^*\,, \quad Iz^* = z^* \circ Q\,. \tag{10.7}$$

The mapping I is linear and continuous and, applying Proposition 10.1 with $(y^*, z^*, T(X))$ in place of $(T, \tilde{T}, U)$, we get that $\|Iz^*\| = \|z^*\|$ for all z^*. Moreover, I is surjective: If $y^* \in \ker T^*$, then $y^* \circ T = 0$, therefore $y^*|T(X) = 0$, and we have $y^* = z^* \circ Q = Iz^*$ for a suitable $z^* \in (Y/T(X))^*$. □

Definition 10.5 (Complement)
*Let X be a vector space, U a subspace of X. A subspace V of X is called an **algebraic complement** of U if*
$$U \cap V = \{0\}\,, \quad U + V = X\,. \tag{10.8}$$
*If moreover X is a Banach space, and U and V are closed, then V is called a **complement** of U. We also say that X is the direct sum of U and V, written as*

$$X = U \oplus V\,. \tag{10.9}$$

□

In linear algebra we have the following situation. Every subspace U of a vector space has an algebraic complement V. Indeed, if $(u_i)_{i\in I}$ is a basis of U, then by enlarging it with suitable vectors $(v_j)_{j\in J}$ we may obtain a basis of X;

$$V = \operatorname{span}\{v_j : j \in J\}$$

then becomes an algebraic complement of U. Algebraic complements are not uniquely determined (unless $U = X$ or $U = \{0\}$), but for every algebraic complement V the quotient mapping $Q : X \to X/U$ yields a bijective linear mapping

$$Q|V : V \to X/U\,. \tag{10.10}$$

Conversely, every subspace V for which $Q|V$ is bijective is an algebraic complement. In this manner we obtain bijective linear mappings

$$U \times X/U \to U \times V \to X\,, \tag{10.11}$$

the latter one defined by $(u, v) \mapsto u + v$.

In functional analysis one is interested in decompositions $X = U \oplus V$ where X, U and V are Banach spaces.

When X is a Hilbert space, every closed subspace has a complement, namely $U^\perp$. If X is only a Banach space, then there may exist closed subspaces which have no complement. For example, $C[0,1]$ does not have a complement in $L^\infty(0,1)$. (We do not prove this here.)

Proposition 10.6 *Let X be a Banach space, U a subspace of X with* $\dim(U) < \infty$. *Then U has a complement.*

Proof: Let $\dim(U) = n$, let $\{x_1, \dots, x_n\}$ be a basis of U, let $\{u_1^*, \dots, u_n^*\}$ be the corresponding dual basis, that is, the basis of U^* defined by

$$u_i^*(x_j) = \delta_{ij} \,.$$

According to Proposition 4.5 (Hahn-Banach) we choose extensions $x_i^* \in X^*$ of u_i^* and define

$$P : X \to U \,, \quad Px = \sum_{i=1}^n x_i^*(x) x_i \,. \tag{10.12}$$

P is linear and continuous, and therefore $V = \ker P$ is a closed subspace of X. We now prove that V is a complement of U. If $z \in U \cap V$, we have $Pz = 0$, thus $0 = x_i^*(z) = u_i^*(z)$ for all i and therefore $z = 0$. Let now $x \in X$ be arbitrary. Then

$$x = Px + (x - Px) \,.$$

As $Px_j = x_j$ for all j, it follows that $P|U = id|U$, so $P \circ P = P$,

$$P(x - Px) = Px - PPx = 0 \,.$$

Therefore we have $x - Px \in V$ and, since x was arbitrary, $X = U + V$. □

Definition 10.7 (Codimension)
Let X be a vector space, U a subspace of X. Then $\dim(X/U)$ *is called the codimension of U in X, written* $\operatorname{codim}(U)$. □

Because of (10.10) we obviously have

$$\operatorname{codim}(U) = \dim(V) \tag{10.13}$$

for every complement V of U.

Proposition 10.8 *Let X be a Banach space, U a closed subspace of X with* $\operatorname{codim}(U) < \infty$. *Then U has a complement.*

Proof: Let V be an algebraic complement of U in X. Then we have $\dim(V) = \operatorname{codim}(U) < \infty$, therefore V is closed. □

Proposition 10.9 *Let X, Y be Banach spaces, let $T : X \to Y$ be linear and continuous, let the codimension of $T(X)$ in Y be finite. Then $T(X)$ is closed in Y.*

Proof: We first consider the case where is T injective. Let $\operatorname{codim}(T(X)) = n$, let $y_1, \dots, y_n \in Y$ such that $\{[y_1], \dots, [y_n]\}$ is a basis of $Y/T(X)$. We define

$$S : \mathbb{K}^n \times X \to Y\,, \quad S(\alpha, x) = Tx + \sum_{i=1}^{n} \alpha_i y_i\,. \tag{10.14}$$

S is linear and continuous. S is injective, since $S(\alpha, x) = 0$ implies that

$$0 = [S(\alpha, x)] = \sum_{i=1}^{n} \alpha_i [y_i]\,,$$

so $\alpha_i = 0$ for all i, thus $Tx = 0$ and therefore $x = 0$. S is surjective: Let $y \in Y$. There exist $\alpha_i \in \mathbb{K}$ such that

$$[y] = \sum_{i=1}^{n} \alpha_i [y_i]\,, \quad \text{so} \quad \left[y - \sum_{i=1}^{n} \alpha_i y_i \right] = 0\,, \quad \text{so} \quad y - \sum_{i=1}^{n} \alpha_i y_i \in T(X)\,.$$

Thus S is bijective. Using Corollary 3.8 we see that S^{-1} is continuous. Therefore $T(X) = S(\{0\} \times X) = (S^{-1})^{-1}(\{0\} \times X)$ is closed. Let now T be arbitrary. We consider the linear continuous mapping $\tilde{T} : X/\ker T \to Y$ defined by $\tilde{T} \circ Q = T$. $\tilde{T}$ is injective due to Proposition 10.1. By what we just have proved, $T(X) = \tilde{T}(X/\ker T)$ is closed. $\Box$

Here are a few remarks concerning the general situation.

- In 1971, Lindenstrauss and Tzafriri have proved: A Banach space X possesses the property that every closed subspace has a complement if and only if X is isomorphic to a Hilbert space. (If moreover X is isometrically isomorphic to a Hilbert space, the parallelogram identity holds and the norm is generated by a suitable scalar product.)
- In 1993, Gowers and Maurey have found an example of a reflexive Banach space X with the property that no closed subspace U has a complement (except those subspaces which have finite dimension or finite codimension, see Proposition 10.6 and Proposition 10.9).

11 Fredholm operators

Definition 11.1 (Fredholm operator)
Let X, Y be Banach spaces. A linear continuous operator $T : X \to Y$ is called a Fredholm operator if $\dim(\ker T) < \infty$ *and* $\operatorname{codim}(T(X)) = \dim(Y/T(X)) < \infty$*. If T is a Fredholm operator, we define the index of T by*

$$\operatorname{ind}(T) = \dim(\ker T) - \operatorname{codim}(T(X)) . \tag{11.1}$$

By $F(X;Y)$ we denote the set of all Fredholm operators from X to Y; if $X = Y$ we also write $F(X)$. □

In this chapter, I denotes the identity mapping in $L(X)$.

When both X and Y have finite dimension, we have $F(X;Y) = L(X;Y)$. Otherwise, $0 \notin F(X;Y)$, and $F(X;Y)$ is not a subspace of $L(X;Y)$. If $T \in F(X;Y)$ and $\alpha \in \mathbb{K}$ with $\alpha \neq 0$, then $\alpha T \in F(X;Y)$. We always have $I \in F(X)$, $\operatorname{ind}(I) = 0$.

If X and Y have finite dimension, say $\dim(X) = n$ and $\dim(Y) = m$, then $n = \dim(\ker T) + \dim(T(X))$ and $m = \dim(T(X)) + \operatorname{codim}(T(X))$, therefore $\operatorname{ind}(T) = n - m$. Thus, the index contains relevant information concerning T only in the infinite dimensional case.

Lemma 11.2 *Let X, Y be Banach spaces, $T : X \to Y$ a Fredholm operator. Then $T(X)$ is closed.*

Proof: This is a direct consequence of Proposition 10.9. □

Proposition 11.3 *Let X be a Banach space, let $S : X \to X$ compact operator. Then $T = I - S$ is a Fredholm operator.*

Proof: First we note that $I|\ker T = S|\ker T$. Thus, $I|\ker T$ is a compact operator and therefore $\dim(\ker T) < \infty$. We want to prove that $T(X)$ is closed. According to Proposition 10.6, let us choose a complement V of $\ker T$ in X and consider $T|V : V \to T(X)$. This mapping is injective, since $V \cap \ker T = \{0\}$. It is surjective, because for every $y = Tx \in T(X)$ we have $x = v + u$ with $v \in V$ and $Tu = 0$, so $Tv = Tx$. Let us assume

$$(T|V)^{-1} : T(X) \to V \quad \text{is not continuous.} \tag{11.2}$$

We will show that this leads to a contradiction. We choose a sequence $(y_n)_{n\in\mathbb{N}}$ in $T(X)$ such that $\|y_n\| = 1$ and $\|T^{-1}y_n\| \geq n$. For

$$v_n = \frac{T^{-1}y_n}{\|T^{-1}y_n\|} \tag{11.3}$$

we have

$$v_n \in V\,, \quad \|v_n\| = 1\,, \quad \|Tv_n\| \leq \frac{1}{n}\,. \tag{11.4}$$

Since S is compact, there exists a convergent subsequence $(Sv_{n_k})_{k\in\mathbb{N}}$. Let $Sv_{n_k} \to v \in X$. Then

$$v_{n_k} = Sv_{n_k} + Tv_{n_k} \to v\,,$$

so $v \in V$, $\|v\| = 1$, $Tv = 0$ in contradiction to $V \cap \ker T = \{0\}$. Therefore, (11.2) does not hold. It follows that $T(X)$ and V are isomorphic. As V is complete, so is $T(X)$, therefore $T(X)$ is closed in X. Proposition 10.4 now implies that

$$(X/T(X))^* \cong \ker T^* = \ker(I - S^*)\,. \tag{11.5}$$

By Proposition 9.5, S^* is compact. By what we just have proved,

$$\infty > \dim(\ker(I - S^*)) = \dim((X/T(X))^*) = \dim(X/T(X)) = \operatorname{codim} T(X)\,.$$

□

Proposition 11.4 *Let X be a Banach space, $T \in L(X)$. If $\|T\| < 1$ then $I - T$ is bijective, $(I - T)^{-1} \in L(X)$ and*

$$(I - T)^{-1} = \sum_{k=0}^{\infty} T^k\,. \tag{11.6}$$

Proof: Exercise. □

Corollary 11.5 *Let X, Y be Banach spaces, $T \in L(X;Y)$, T bijective. Let $\tilde{T} \in L(X;Y)$ such that*

$$\|\tilde{T} - T\| < \frac{1}{\|T^{-1}\|}\,. \tag{11.7}$$

Then $\tilde{T}$ is bijective and $\tilde{T}^{-1} \in L(Y;X)$.

Proof: We have

$$\tilde{T} = T(I - T^{-1}(T - \tilde{T}))\,, \tag{11.8}$$

as one checks by performing the multiplications on the right side. From (11.7) it follows that $\|T^{-1}(T - \tilde{T})\| \le \|T^{-1}\|\,\|T - \tilde{T}\| < 1$. The assertion now follows from Proposition 11.4. □

Proposition 11.6 *Let X, Y be Banach spaces. The set $F(X;Y)$ is an open subset of the Banach space $L(X;Y)$, and the mapping* ind $: F(X;Y) \to \mathbb{Z}$ *is locally constant, that is, for every $T \in F(X;Y)$ there exists an $\varepsilon > 0$ such that*

$$\operatorname{ind}(S) = \operatorname{ind}(T)\,, \quad \text{for all } S \in F(X;Y) \text{ with } \|S - T\| < \varepsilon. \tag{11.9}$$

Proof: Let $T : X \to Y$ be a Fredholm operator. By Lemma 11.2, $T(X)$ is closed. According to Proposition 10.6 and Proposition 10.8 we choose closed subspaces V of X and W of Y such that

$$\ker T \oplus V = X\,, \quad T(X) \oplus W = Y\,. \tag{11.10}$$

We then have

$$\dim(W) = \operatorname{codim}(T(X)) < \infty\,, \quad \operatorname{codim}(V) = \dim(\ker T) < \infty\,. \tag{11.11}$$

For arbitrary given $S \in L(X;Y)$ we define

$$\tilde{S} : V \times W \to Y\,, \quad \tilde{S}(v,w) = Sv + w\,, \tag{11.12}$$

where $V \times W$ is endowed with the maximum norm $\|(v,w)\|_\infty = \max\{\|v\|, \|w\|\}$. Obviously, $\tilde{S}$ is linear and continuous. We have

$$(\tilde{S} - \tilde{T})(v,w) = (S-T)(v)$$

and therefore

$$\|\tilde{S} - \tilde{T}\| = \|S - T\|\,. \tag{11.13}$$

From (11.10) it follows that $T(V) = T(X)$ and that $T|V$ is injective. Therefore, $\tilde{T} : V \times W \to Y$ is linear, bijective and continuous. According to Corollary 11.5 we choose $\varepsilon > 0$ such that $\tilde{S} : V \times W \to Y$ is bijective for all $S \in L(X;Y)$ satisfying $\|S - T\| = \|\tilde{S} - \tilde{T}\| < \varepsilon$. Let now S be an arbitrary operator with this property. It follows from (11.12) that

$$\tilde{S}(V \times \{0\}) = S(V)\,, \quad \tilde{S}(\{0\} \times W) = W\,,$$

therefore

$$Y = S(V \times W) = S(V) \oplus W\,, \quad \operatorname{codim}(S(V)) < \infty\,. \tag{11.14}$$

Since $\tilde{S}$ is bijective, $S|V$ is injective, and thus

$$\ker S \cap V = \{0\}\,.$$

Let now Z be a complement of $\ker S \oplus V$ in X,

$$\ker S \oplus V \oplus Z = X\,, \tag{11.15}$$

then both $\ker S \oplus Z$ and $\ker T$ are complements of V in X, the latter because of (11.10). They therefore have the same dimension,

$$\dim(\ker S) + \dim(Z) = \dim(\ker T) < \infty\,, \quad \dim(Z) < \infty\,. \tag{11.16}$$

From (11.15) it moreover follows that $S|(V \oplus Z)$ is injective and

$$S(X) = S(V \oplus Z) = S(V) \oplus S(Z)\,. \tag{11.17}$$

Adding on both sides a complement of $S(X)$ in Y, we obtain

$$\infty > \operatorname{codim}(S(V)) = \dim(S(Z)) + \operatorname{codim}(S(X))\,. \tag{11.18}$$

Furthermore, by (11.11) and (11.14) we get, since $S|Z$ is injective,

$$\operatorname{codim}(T(X)) = \dim(W) = \operatorname{codim}(S(V)) = \dim(Z) + \operatorname{codim}(S(X))\,. \tag{11.19}$$

From (11.16) and (11.19) we see that S is a Fredholm operator. Adding both equations yields

$$\dim(\ker S) + \operatorname{codim}(T(X)) = \dim(\ker T) + \operatorname{codim}(S(X))\,,$$

thus $\operatorname{ind}(S) = \operatorname{ind}(T)$. □

Corollary 11.7 *Let X, Y be Banach spaces, let $T : [0,1] \to F(X;Y)$ be continuous. Then*

$$t \mapsto \mathrm{ind}\,(T(t)) \tag{11.20}$$

is constant.

Proof: By Proposition 11.6, the mapping $\mathrm{ind} \circ T : [0,1] \to \mathbb{R}$ is continuous. It must be constant because the index has only integer values. □

A locally constant mapping is constant on connected subsets of its domain. Thus one has the more general result that the index is constant on each connected component of $F(X;Y)$.

Corollary 11.8 *Let X be a Banach space, $S : X \to X$ compact. Then*

$$\mathrm{ind}\,(I - S) = 0\,. \tag{11.21}$$

Proof: The operator $T(t) = I - tS$ is a Fredholm operator by Proposition 11.3. We have $\mathrm{ind}\,(T(0)) = \mathrm{ind}\,(I) = 0$, therefore $0 = \mathrm{ind}\,(T(1)) = \mathrm{ind}\,(I - S)$ by Corollary 11.7. □

Let $T : X \to X$ be a Fredholm operator of index 0. Then we have $\dim(\ker T) = 0$ if and only if $\mathrm{codim}\, T(X) = 0$. Consequently,

$$T \text{ injective} \quad \Leftrightarrow \quad T \text{ surjective.} \tag{11.22}$$

This is an important generalization of the corresponding result from linear algebra for $X = \mathbb{K}^n$.

We consider the equation

$$Tx = y\,, \tag{11.23}$$

where $y \in X$ is given and $x \in X$ is to be determined.

The classical formulation of the above properties of Fredholm operators of index 0 is the **Fredholm alternative**:

> **Either** the equation (11.23) has a unique solution $x \in X$ for every given $y \in X$ (this corresponds to (11.22), the case $\dim(\ker T) = 0$),
>
> **or** the homogeneous equation $Tx = 0$ has finitely many linearly independent solutions, and for every given $y \in X$ the equation (11.23) has a solution if and only if $y^*(y) = 0$ for all $y^* \in \ker T^*$. The maximal number of linearly independent solutions of $Tx = 0$ is equal to the maximal number of linearly independent solutions of $T^* y^* = 0$. (This is the case $\dim(\ker T) > 0$.)

The second part of this alternative is based on the formulas

$$\overline{T(X)} = (\ker T^*)^o\,, \quad (Y/T(X))^* \cong \ker T^*$$

from Proposition 9.6 and Proposition 10.4, and on the fact that $\overline{T(X)} = T(X)$.

12 The Spectrum

In linear algebra, a structure theory for linear mappings is developed. It is mainly based on the notions of an eigenvalue and an eigenvector. We recall these notions: let $T : \mathbb{K}^n \to \mathbb{K}^n$ be linear. If

$$Tx = \lambda x\,, \quad \lambda \in \mathbb{K}\,, \quad x \in \mathbb{K}^n\,, \quad x \neq 0\,, \tag{12.1}$$

then λ is called an eigenvalue of T and x is called an eigenvector for λ.

A $\lambda \in \mathbb{K}$ is an eigenvalue of $T : \mathbb{K}^n \to \mathbb{K}^n$ if and only if the mapping $\lambda I - T$ is not bijective. Instead of $\lambda I - T$, in the following we write

$$\lambda - T\,.$$

The space $L(X;X)$ of linear and continuous mappings from X to itself is denoted by

$$L(X)\,.$$

In this and the following chapters we always assume that $X \neq \{0\}$ (otherwise $L(X) = \{0\}$ and $\|I\| = 0$.)

Definition 12.1 (Resolvent)
Let X be a Banach space, $T \in L(X)$. The subset

$$\rho(T) = \{\lambda : \lambda \in \mathbb{K},\, \lambda - T \textit{ is bijective}\} \tag{12.2}$$

of $\mathbb{K}$ is called the resolvent set of T. The mapping

$$R : \rho(T) \to L(X)\,, \quad R_\lambda = (\lambda - T)^{-1}\,, \tag{12.3}$$

is called the resolvent of T. □

The definition of R in (12.3) makes sense since $(\lambda - T)^{-1}$ is linear and continuous if $\lambda - T$ is linear, continuous and bijective, see Corollary 3.8.

Proposition 12.2 *Let X be a Banach space, $T \in L(X)$. The resolvent set $\rho(T)$ is an open subset of $\mathbb{K}$. If $\lambda_0 \in \rho(T)$ then*

$$R_\lambda = (\lambda - T)^{-1} = \sum_{k=0}^{\infty} (\lambda_0 - \lambda)^k [(\lambda_0 - T)^{-1}]^{k+1} = \sum_{k=0}^{\infty} (\lambda_0 - \lambda)^k R_{\lambda_0}^{k+1} \tag{12.4}$$

for all $\lambda \in \mathbb{K}$ with

$$|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|}\,. \tag{12.5}$$

Proof: We obtain

$$(\lambda_0 - T)[1 - (\lambda_0 - T)^{-1}(\lambda_0 - \lambda)] = \lambda - T\,. \tag{12.6}$$

by performing the multiplication on the left side. Because $\|(\lambda - T) - (\lambda_0 - T)\| = |\lambda - \lambda_0|$, it follows from (12.5) that the expression in brackets is invertible. Thus, $\lambda - T$ is invertible. From (12.6) it moreover follows that

$$(\lambda - T)^{-1} = [1 - (\lambda_0 - T)^{-1}(\lambda_0 - \lambda)]^{-1}(\lambda_0 - T)^{-1}$$

We replace the term in brackets with its series expansion according to (11.6). This yields (12.4). □

Definition 12.3 (Spectrum, Point Spectrum)
Let X be a Banach space, $T \in L(X)$. The subset

$$\sigma(T) = \mathbb{K} \setminus \rho(T) = \{\lambda : \lambda \in \mathbb{K},\ \lambda - T \text{ is not bijective}\} \tag{12.7}$$

is called the spectrum of T. Every $\lambda \in \sigma(T)$ is called a spectral value of T. If $\lambda - T$ is not injective, λ is called an eigenvalue of T with the corresponding eigenspace $\ker(\lambda - T)$, *and each $x \in X$ with $\lambda x = Tx$ and $x \neq 0$ is called an eigenvector for the eigenvalue λ. The set*

$$\sigma_p(T) = \{\lambda : \lambda \text{ is an eigenvalue of } T\} \tag{12.8}$$

is called the point spectrum of T. □

In the case $X = \mathbb{K}^n$, the operator $\lambda - T$ is injective if and only if $\lambda - T$ is bijective, we then have $\sigma(T) = \sigma_p(T)$. If X is infinite dimensional, in general we have $\sigma(T) \neq \sigma_p(T)$.

Definition 12.4 (Continuous Spectrum, Residual Spectrum)
Let X be a Banach space, $T \in L(X)$. The set

$$\sigma_c(T) = \{\lambda : \lambda \in \sigma(T),\ \lambda - T \text{ is injective, not surjective, and } \overline{(\lambda - T)(X)} = X\} \tag{12.9}$$

is called the continuous spectrum of T, the subset

$$\sigma_r(T) = \{\lambda : \lambda \in \sigma(T),\ \lambda - T \text{ is injective, not surjective, and } \overline{(\lambda - T)(X)} \neq X\} \tag{12.10}$$

is called the residual spectrum of T. □

By the definitions,

$$\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$$

yields a disjoint partition of the spectrum.

As an example we consider the right shift

$$T : \ell^2(\mathbb{K}) \to \ell^2(\mathbb{K}), \quad (Tx)_k = \begin{cases} 0, & k = 1, \\ x_{k-1}, & k > 1. \end{cases}$$

We have $0 \in \sigma_r(T)$ since T is injective and $T(\ell^2(\mathbb{K})) = \{y : y_1 = 0\}$ is a proper closed subspace of $\ell^2(\mathbb{K})$. For

$$T : \ell^2(\mathbb{K}) \to \ell^2(\mathbb{K}), \quad (Tx)_k = \frac{1}{k} x_k,$$

we have $0 \in \sigma_c(T)$ since T is injective and, because of $(1/k)_{k \in \mathbb{N}} \notin T(\ell^2(\mathbb{K}))$, is not surjective, and has a dense range due to $c_e(\mathbb{K}) \subset T(\ell^2(\mathbb{K}))$.

Proposition 12.5 *Let X be a Banach space, $T \in L(X)$. The spectrum $\sigma(T)$ is compact, and $|\lambda| \leq \|T\|$ for all $\lambda \in \sigma(T)$. In the case $\mathbb{K} = \mathbb{C}$, the spectrum $\sigma(T)$ is not empty.*

Proof: If $T = 0$ then $\sigma(T) = \{0\}$. Let $T \neq 0$. If $|\lambda| > \|T\|$, then it follows from

$$\lambda - T = \lambda \left(1 - \frac{1}{\lambda} T \right) \tag{12.11}$$

and from Proposition 11.4 that $\lambda \in \rho(T)$. Therefore, $\sigma(T)$ is bounded by $\|T\|$ and, because of Proposition 12.2, closed in $\mathbb{K}$ und thus compact. Let now $\mathbb{K} = \mathbb{C}$. For an arbitrary given $\ell \in L(X)^*$ we consider the function

$$f : \rho(T) \to \mathbb{C}, \quad f(\lambda) = \ell(R_\lambda). \tag{12.12}$$

Let $\lambda_0 \in \rho(T)$ be arbitrary. Corollary 12.2 implies that

$$f(\lambda) = \ell(R_\lambda) = \sum_{k=0}^{\infty} (\lambda_0 - \lambda)^k \ell(R_{\lambda_0}^{k+1}), \tag{12.13}$$

if $\lambda \in B(\lambda_0; 1/\|R_{\lambda_0}\|)$. This means that f is represented by a power series which converges in this open disc. Therefore, f is holomorphic on $\rho(T)$.

Let us now assume that $\sigma(T) = \emptyset$, so $\rho(T) = \mathbb{C}$. On the compact set $B(0; 2\|T\|)$, f is bounded. For $|\lambda| > 2\|T\|$ it follows from (12.11) that

$$R_\lambda = (\lambda - T)^{-1} = \frac{1}{\lambda} \sum_{k=0}^{\infty} \left(\frac{T}{\lambda} \right)^k, \tag{12.14}$$

and moreover

$$|f(\lambda)| = |\ell(R_\lambda)| = \left| \frac{1}{\lambda} \sum_{k=0}^{\infty} \frac{\ell(T^k)}{\lambda^k} \right| \leq \frac{1}{|\lambda|} \|\ell\| \sum_{k=0}^{\infty} \frac{\|T\|^k}{|\lambda|^k} \leq 2\|\ell\| \frac{1}{|\lambda|}. \tag{12.15}$$

Therefore, f is bounded on $\mathbb{C}$. The theorem of Liouville (from complex analysis) implies that f is constant on $\mathbb{C}$. From (12.15) it follows that $f = 0$. Since ℓ was arbitrary, we have that $\ell(R_\lambda) = 0$ for all $\ell \in L(X)^*$. By the theorem of Hahn-Banach we see that $R_\lambda = 0$, a contradiction to the fact that R_λ is bijective. $\square$

Let p be a polynomial,

$$p(z) = \sum_{j=0}^{n} c_j z^j, \quad z \in \mathbb{C}.$$

For $T \in L(X)$, X Banach space, we consider

$$p(T) = \sum_{j=0}^{n} c_j T^j. \tag{12.16}$$

We have $p(T) \in L(X)$. It turns out that in the complex case we can compute the spectrum of $p(T)$ from the spectrum of T. First we need a lemma from algebra.

Lemma 12.6 *Let X be a Banach space, let $R, S, T \in L(X)$.*
(i) If $RT = TS = I$, then T has an inverse in $L(X)$ and $T^{-1} = R = S$.
(ii) If ST has an inverse in $L(X)$ and $ST = TS$, then S and T have inverses in $L(X)$.

Proof: We have $S = (RT)S = R(TS) = R$, so (i) follows. To prove (ii), we set $Q = (ST)^{-1} \in L(X)$ and get

$$S(TQ) = (ST)Q = I = Q(ST) = Q(TS) = (QT)S\,.$$

By part (i), S has an inverse in $L(X)$. Since we may interchange S and T, the same applies to T. $\square$

Proposition 12.7 (Spectral mapping theorem for polynomials)
Let X be a Banach space over $\mathbb{C}$, $T \in L(X)$, p a polynomial. Then

$$\sigma(p(T)) = p(\sigma(T)) := \{p(\lambda) : \lambda \in \sigma(T)\} \tag{12.17}$$

Proof: Let n be the degree of p. For $n = 0$ we have $p(z) = c_0$ and $p(T) = c_0 I$, $\sigma(T) = \{c_0\}$, so (12.17) holds. Let $n \geq 1$.
"$\supset$": Let $\lambda \in \sigma(T)$. Since λ is a zero of the polynomial $z \mapsto p(z) - p(\lambda)$, we have

$$p(z) - p(\lambda) = (z - \lambda)q(z) = q(z)(z - \lambda)\,, \quad z \in \mathbb{C}\,,$$

for some polynomial q of degree $n - 1$. Consequently,

$$p(T) - p(\lambda) = (T - \lambda)q(T) = q(T)(T - \lambda)\,.$$

Since $\lambda \in \sigma(T)$, $T - \lambda$ does not have an inverse in $L(X)$. By Lemma 12.6(ii), $p(T) - p(\lambda)$ does not have an inverse in $L(X)$, so $p(\lambda) \in \sigma(p(T))$.
"$\subset$": Let $\lambda \in \sigma(p(T))$. By the fundamental theorem of algebra, $p - \lambda$ has n zeroes $\{\lambda_j\}$, and we can factorize

$$p(z) - \lambda = \gamma \sum_{j=1}^{n} (z - \lambda_j)\,, \quad p(T) - \lambda = \gamma \sum_{j=1}^{n} (T - \lambda_j)\,, \quad \gamma \neq 0\,. \tag{12.18}$$

Since $\lambda \in \sigma(p(T))$, the operator $p(T) - \lambda$ does not have an inverse in $L(X)$. It follows from (12.18) that at least one of the operators $T - \lambda_j$ does not have an inverse in $L(X)$, so $\lambda_j \in \sigma(T)$ for such a j. As $p(\lambda_j) = \lambda$, we obtain $\lambda \in p(\sigma(T))$. $\square$

Definition 12.8 (Spectral radius)
Let X be a Banach space, $T \in L(X)$. The number

$$r(T) = \sup_{\lambda \in \sigma(T)} |\lambda| \tag{12.19}$$

is called the spectral radius of T. $\square$

Because the spectrum is compact by Proposition 12.5, the supremum in (12.19) is actually attained, so

$$r(T) = \max_{\lambda \in \sigma(T)} |\lambda|\,, \tag{12.20}$$

unless the spectrum is empty.

Proposition 12.9 *Let X be a complex Banach space (that is, $\mathbb{K} = \mathbb{C}$), let $T \in L(X)$. Then*

$$r(T) = \lim_{n \in \mathbb{N}} \sqrt[n]{\|T^n\|} = \inf_{n \in \mathbb{N}} \sqrt[n]{\|T^n\|} . \qquad (12.21)$$

Proof: First, we show that $r(T) \le \inf_{n\in\mathbb{N}} \sqrt[n]{\|T^n\|}$. By Proposition 12.5, $\sigma(T)$ is not empty. Let $\lambda \in \sigma(T)$. By the spectral mapping theorem (Proposition 12.7), $\lambda^n \in \sigma(T^n)$ for all $n \in \mathbb{N}$. By Proposition 12.5 applied to T^n we get $|\lambda^n| \le \|T^n\|$, so $|\lambda| \le \sqrt[n]{\|T^n\|}$ for all $n \in \mathbb{N}$. Taking the maximum with respect to $\lambda \in \sigma(T)$, the assertion follows.

Next, we show that $r(T) \ge \limsup_{n\to\infty} \sqrt[n]{\|T^n\|}$. Let $\ell \in L(X)^*$ be arbitrary. We consider

$$f(\lambda) = \ell(R_\lambda) , \quad R_\lambda = (\lambda - T)^{-1} .$$

If $|\lambda| > \|T\|$ then

$$R_\lambda = (\lambda - T)^{-1} = \frac{1}{\lambda} \sum_{k=0}^{\infty} \left(\frac{T}{\lambda}\right)^k , \quad f(\lambda) = \frac{1}{\lambda} \sum_{k=0}^{\infty} \frac{\ell(T^k)}{\lambda^k} . \qquad (12.22)$$

In the proof of Proposition 12.5 we have seen that $f : \rho(T) \to \mathbb{C}$ is a holomorphic function. In particular, f is holomorphic in the domain $G = \{|\lambda| > r(T)\}$. By a result of complex analysis, f has a unique Laurent expansion on G. Since (12.22) gives a Laurent expansion of f in $\{|\lambda| > \|T\|\}$, the two expansions coincide. Therefore,

$$f(\lambda) = \frac{1}{\lambda} \sum_{k=0}^{\infty} \frac{\ell(T^k)}{\lambda^k}$$

is valid on G. The elements of this series converge to 0,

$$0 = \lim_{k\to\infty} \frac{1}{\lambda} \frac{\ell(T^k)}{\lambda^k} = \lambda^{-1} \lim_{k\to\infty} \ell(\lambda^{-k} T^k) .$$

Since ℓ was arbitrary,

$$\lambda^{-k} T^k \rightharpoonup 0 \qquad \text{in } L(X).$$

The sequence $(\lambda^{-k}T^k)_{k\in\mathbb{N}}$ is bounded in the norm of $L(X)$ according to Proposition 6.4, let

$$\|\lambda^{-k} T^k\| \le M .$$

It follows that

$$\|T^k\|^{1/k} \le |\lambda| M^{1/k} , \quad \limsup_{k\to\infty} \|T^k\|^{1/k} \le |\lambda| , \qquad (12.23)$$

for all $\lambda \in \rho(T)$. Passing to the limit $|\lambda| \to r(T)$ yields

$$\limsup_{k\to\infty} \|T^k\|^{1/k} \le r(T) . \qquad (12.24)$$

We thus have proved that

$$r(T) \le \inf_{n\in\mathbb{N}} \sqrt[n]{\|T^n\|} \le \liminf_{n\to\infty} \sqrt[n]{\|T^n\|} \le \limsup_{n\to\infty} \sqrt[n]{\|T^n\|} \le r(T) ,$$

as the two inequalities in the middle are satisfied for arbitrary sequences. From this, (12.21) immediately follows. □

Proposition 12.10 *Let X be a Banach space, $S : X \to X$ compact. If $\lambda \in \mathbb{C}$ is a spectral value of S with $\lambda \neq 0$, then λ is an eigenvalue of S. The corresponding eigenspace* $\ker(\lambda - S)$ *has finite dimension. If* $\dim(X) = \infty$ *then 0 is a spectral value of S.*

Proof: Let $\lambda \neq 0$. By Corollary 11.8, $1 - \lambda^{-1}S$ is a Fredholm operator of index 0, the same holds for $\lambda - S$. If λ is not eigenvalue of S, then $\lambda - S$ is injective and thus bijective, according to (11.22). Therefore, λ is not a spectral value of S. Since $\lambda - S$ is a Fredholm operator, $\ker(\lambda - S)$ has finite dimension.
If $0 \in \rho(S)$, then S is bijective, linear and continuous, therefore S^{-1} is continuous and thus $I = S^{-1}S$ compact. This can only happen if $\dim(X) < \infty$. □

As an example, let S be the integral operator

$$(Sx)(t) = \int_0^1 k(s,t)x(s)\,dx\,. \tag{12.25}$$

The equation $Tx = (\lambda - S)(x) = y$ becomes

$$\lambda x(t) - \int_0^1 k(s,t)x(s)\,ds = y(t)\,, \quad t \in [0,1]\,. \tag{12.26}$$

As we have seen, S is compact on $L^2(0,T)$ if $k \in L^2((0,T)\times(0,T))$. We apply Proposition 12.10 and obtain that, for a given $\lambda \neq 0$, the equation (12.26) is uniquely solvable for every $y \in L^2(0,T)$ if and only if λ is not eigenvalue of S, and that the space of its solutions has finite dimension if λ is an eigenvalue of S.

Proposition 12.11 *Let X be a Banach space, $S : X \to X$ a compact linear operator. Then the set*

$$\{\lambda : \lambda \in \sigma(S),\ |\lambda| \geq \varepsilon\} \tag{12.27}$$

is finite for every $\varepsilon > 0$. (It may be empty.) The spectrum $\sigma(T)$ is a finite or countably infinite set.

Proof: Let us assume that there exists a sequence $(\lambda_n)_{n\in\mathbb{N}}$ of distinct spectral values of S with $|\lambda_n| \geq \varepsilon$. By Proposition 12.10, all those λ_n are eigenvalues of S. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of corresponding eigenvectors of S. As λ_n are distinct, by a result of linear algebra the set $\{x_n : n \in \mathbb{N}\}$ is linearly independent. Setting

$$X_n = \operatorname{span}\{x_1, \ldots, x_n\}$$

we have $X_n \subset X_{n+1}$, $X_n \neq X_{n+1}$ and $S(X_n) \subset X_n$ for all $n \in \mathbb{N}$. We choose $y_n \in X_n$ according to Lemma 2.13 with

$$\|y_n\| = 1\,, \quad \operatorname{dist}(y_n, X_{n-1}) \geq \frac{1}{2}\,.$$

Then $y_n = \alpha x_n + z_{n-1}$ for some $\alpha \in \mathbb{K}$ and $z_{n-1} \in X_{n-1}$. It follows that

$$\lambda_n y_n - Sy_n = \lambda_n \alpha x_n + \lambda_n z_{n-1} - \alpha S x_n - S z_{n-1} = \lambda_n z_{n-1} - S z_{n-1} \in X_{n-1}\,,$$

so for all $m \in \mathbb{N}$ mit $m < n$ we get

$$\|Sy_n - Sy_m\| = \|\lambda_n y_n - (Sy_m + \lambda_n y_n - Sy_n)\| = |\lambda_n| \|y_n - \underbrace{\frac{1}{\lambda_n}(Sy_m + \lambda_n y_n - Sy_n)}_{\in X_{n-1}}\| \geq \frac{\varepsilon}{2}\,.$$

Therefore, $(Sy_n)_{n\in\mathbb{N}}$ does not have a convergent subsequence, a contradiction since S is compact. This proves that the set (12.27) is finite. Since $\sigma(T)$ equals the countable union of those sets for $\varepsilon = 1/n$ with $n \in \mathbb{N}$, it is at most countably infinite. □

13 Spectral decomposition for compact normal operators on Hilbert space

We consider operators on Hilbert spaces.

Proposition 13.1 (Hilbert adjoint)
Let X, Y be Hilbert spaces and $T \in L(X;Y)$. Then

$$\langle x, T^* y \rangle = \langle Tx, y \rangle \ , \quad \text{for all } x \in X \ , \ y \in Y \, , \tag{13.1}$$

defines a linear continuous operator $T^* : Y \to X$, it is called the Hilbert adjoint of T. We have

$$T^{**} = T \, , \quad \|T^*\| = \|T\| \, . \tag{13.2}$$

Moreover,

$$\langle T^* y, x \rangle = \langle y, Tx \rangle \ , \quad \text{for all } x \in X \ , \ y \in Y \, . \tag{13.3}$$

Proof: Let $y \in Y$. The mapping $x \mapsto \langle Tx, y \rangle$ is linear and continuous, because the scalar product is linear in the first argument, and because

$$|\langle Tx, y \rangle| \leq \|T\| \|x\| \|y\| \tag{13.4}$$

holds for all $x \in X$. It follows from the Riesz representation theorem (Proposition 2.12) that (13.1) specifies a unique element $T^* y \in X$, so T^* is a well-defined mapping. Taking the complex conjugate in (13.1), (13.3) follows.

T^* is linear since

$$\begin{aligned} \langle x, T^*(\alpha y + \beta z) \rangle &= \langle Tx, \alpha y + \beta z \rangle = \overline{\alpha} \langle Tx, y \rangle + \overline{\beta} \langle Tx, z \rangle = \overline{\alpha} \langle x, T^* y \rangle + \overline{\beta} \langle x, T^* z \rangle \\ &= \langle x, \alpha T^*(y) + \beta T^*(z) \rangle \ . \end{aligned}$$

Since

$$|\langle x, T^* y \rangle| = |\langle Tx, y \rangle| \leq \|T\| \|x\| \|y\| \, ,$$

for all $x \in H$, we have $\|T^* y\| \leq \|T\| \|y\|$, so T^* is continuous and satisfies $\|T^*\| \leq \|T\|$. We have $T^{**} = T$ because of

$$\langle y, T^{**} x \rangle = \langle T^* y, x \rangle = \langle y, Tx \rangle \ .$$

Finally, we see that $\|T\| = \|T^{**}\| \leq \|T^*\|$ and therefore $\|T^*\| = \|T\|$. □

In Chapter 9, the adjoint of $T \in L(X;Y)$ for Banach spaces X and Y has been defined as an operator from Y^* to X^*. For the moment, let us denote this operator by $T' : Y^* \to X^*$. It is related to the Hilbert adjoint $T^* : Y \to X$ defined by (13.1) through the duality mappings $R_X : X \to X^*$ and $R_Y : Y \to Y^*$ (they were denoted by "J" in the Riesz theorem),

$$T^* = R_X^{-1} \circ T' \circ R_Y \, .$$

Directly from the definitions we obtain the formulas

$$(S + T)^* = S^* + T^* \, , \quad (\alpha T)^* = \overline{\alpha} T^* \, , \quad (S \circ T)^* = T^* \circ S^* \, . \tag{13.5}$$

The middle equation shows that, in the case $\mathbb{K} = \mathbb{C}$, the mapping $T \mapsto T^*$ is not linear, but conjugate linear.

Definition 13.2 *Let X be a Hilbert space. An operator $T \in L(X)$ is called **normal** if $T^*T = TT^*$, and **hermitian** ($\mathbb{K} = \mathbb{C}$) or **self-adjoint** ($\mathbb{K} = \mathbb{R}$), if $T^* = T$. It is called **positive** if $\langle Tx, x\rangle \geq 0$ for all $x \in X$.* □

Obviously, every hermitian or self-adjoint operator is normal.

Lemma 13.3 *Let X, Y be Hilbert spaces, $T \in L(X;Y)$. Then T^*T as well as TT^* are positive and hermitian resp. self-adjoint. Moreover,*

$$\|T^*T\| = \|T\|^2 . \tag{13.6}$$

Proof: As $(T^*T)^* = T^*T^{**} = T^*T$, T^*T is hermitian resp. self-adjoint. Next, we have for all $x \in X$

$$0 \leq \|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle \leq \|T^*T\|\|x\|^2 .$$

This shows that T^*T is positive. Passing on both sides to the supremum with respect to x on the unit sphere $\{\|x\| = 1\}$ we see that $\|T\|^2 \leq \|T^*T\|$. The reverse equality holds, since $\|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2$ according to Proposition 13.1. As $T^{**} = T$, the assertions concerning TT^* follow from those for T^*T. □

Proposition 13.4 *Let X be a Hilbert space, $T \in L(X)$ normal. Then T^n is normal, and*

$$\|T^n\| = \|T\|^n , \quad \text{for all } n \in \mathbb{N}. \tag{13.7}$$

In the case $\mathbb{K} = \mathbb{C}$ we have $\|T\| = r(T)$. If moreover T is compact,

$$\|T\| = r(T) = \max_{\lambda \in \sigma_p(T)} |\lambda| , \tag{13.8}$$

that is, there exists an eigenvalue λ such that $\|T\| = |\lambda|$.

Proof: Using induction one checks that $(T^n)^* = (T^*)^n$; the computation

$$(T^n)^* = (TT^{n-1})^* = (T^{n-1})^*T^* = (T^*)^{n-1}T^* = (T^*)^n$$

yields the induction step. From this, for all $n \in \mathbb{N}$ we obtain

$$(T^n)^*T^n = (T^*)^nT^n = T^n(T^*)^n = T^n(T^n)^* , \tag{13.9}$$

the middle equality follows by successively interchanging T and T^*, which is possible since T is normal. Therefore, T^n is normal for all $n \in \mathbb{N}$. We now prove (13.7). For $n = 2$, this follows from the computation, using Lemma 13.3,

$$\|T^2\|^2 = \|(T^2)^*T^2\| = \|(T^*T)(T^*T)^*\| = \|T^*T\|^2 = \|T\|^4 .$$

In order to perform the induction step $n \to n+1$, we estimate (using that T^n is normal)

$$\|T\|^{2n} = (\|T\|^n)^2 = \|T^n\|^2 = \|(T^n)^2\| \leq \|T^{n+1}\|\|T\|^{n-1} ,$$

so $\|T\|^{n+1} \leq \|T^{n+1}\| \leq \|T\|^{n+1}$ and therefore $\|T\|^{n+1} = \|T^{n+1}\|$. In the case $\mathbb{K} = \mathbb{C}$ it follows from (13.7) and Proposition 12.9 that

$$r(T) = \lim_{n\to\infty} \|T^n\|^{1/n} = \|T\| .$$

That the spectral radius is equal to the maximum of the absolute values of the eigenvalues follows from Proposition 12.10. □

Lemma 13.5 *Let X be a Hilbert space, $T \in L(X)$ normal. Then $\|Tx\| = \|T^*x\|$ for all $x \in X$, and therefore* $\ker T = \ker T^*$. *Moreover,*

$$Tx = \lambda x \qquad \Leftrightarrow \qquad T^*x = \overline{\lambda}x\,. \tag{13.10}$$

Consequently, $\lambda \in \mathbb{C}$ is an eigenvalue of T if and only if $\overline{\lambda}$ is an eigenvalue of T^*.

Proof: Since T is normal, for all $x \in X$ we have

$$0 = \langle T^*Tx - TT^*x, x\rangle = \langle Tx, Tx\rangle - \langle T^*x, T^*x\rangle = \|Tx\|^2 - \|T^*x\|^2\,.$$

Because $\lambda - T$, too, is normal, the second assertion follows from

$$\ker(\lambda - T) = \ker((\lambda - T)^*) = \ker(\overline{\lambda} - T^*)\,.$$

□

Lemma 13.6 *Let X be a Hilbert space, $T \in L(X)$ normal. If $Tx = \lambda x$, $Ty = \mu y$ and $\lambda \neq \mu$, then $\langle x, y\rangle = 0$, that is, eigenvectors to different eigenvalues are orthogonal to each other.*

Proof: Using (13.10), the assertion follows from the computation

$$\lambda \langle x, y\rangle = \langle \lambda x, y\rangle = \langle Tx, y\rangle = \langle x, T^*y\rangle = \langle x, \overline{\mu}y\rangle = \mu \langle x, y\rangle\ .$$

□

We now present the spectral decomposition theorem for compact normal operators in Hilbert space. (Here we restrict ourselves to the case $\mathbb{K} = \mathbb{C}$). This generalizes the result from finite dimensions that for every normal linear mapping $T : \mathbb{C}^n \to \mathbb{C}^n$ there exists an orthonormal basis of $\mathbb{C}^n$ which consists of eigenvectors of T, so the matrix corresponding to T with respect to this basis is diagonal.

Let $T \in L(X)$ be a compact normal operator in a Hilbert space X. We set

$$U_\lambda = \ker(\lambda - T)\,, \quad \lambda \in \mathbb{C}\,. \tag{13.11}$$

By $P_\lambda : X \to X$ we denote the orthogonal projection onto U_λ. Since $(\lambda - T)Tx = T(\lambda - T)x$ we have

$$T(U_\lambda) \subset U_\lambda\,, \quad \text{for all } \lambda \in \mathbb{C}, \tag{13.12}$$

that is, T leaves invariant the subspace U_λ. Let now E_λ be an orthonormal basis of U_λ, if $\lambda \in \mathbb{C}$ is an eigenvalue of T; otherwise we set $E_\lambda = \emptyset$. We define

$$E = \bigcup_{\lambda \neq 0} E_\lambda\,. \tag{13.13}$$

It follows from Propositions 12.10 and 12.11 that the eigenspaces U_λ are finite dimensional for $\lambda \neq 0$, and that T possesses at most countably infinitely many different nonzero eigenvalues. Therefore, the set E is finite or countably infinite. Let

$$E = \{e_1, e_2, \dots\} = \{e_j : j \in J\} \tag{13.14}$$

with $J = \{1, \dots, |J|\}$, resp. $J = \mathbb{N}$ and $|J| = \infty$.

Proposition 13.7 (Spectral decomposition)
Let X be a separable Hilbert space over $\mathbb{K} = \mathbb{C}$, let $T \in L(X)$ be compact and normal. Then the set $E_0 \cup E$ is an orthonormal basis of X, and

$$X = \overline{\text{span}\,(E_0)} \oplus \overline{\text{span}\,(E)}\,, \quad \overline{\text{span}\,(E_0)} = \ker T\,, \quad \overline{\text{span}\,(E)} = \overline{T(X)}\,. \tag{13.15}$$

For all $x \in X$ we have

$$Tx = \sum_{\lambda \in \sigma_p(T)} \lambda P_\lambda x = \sum_{j \in J} \lambda_j \langle x, e_j \rangle e_j\,. \tag{13.16}$$

In addition,

$$\|T\| = \max_{\lambda \in \sigma_p(T)} |\lambda|\,. \tag{13.17}$$

Proof: The second equality in (13.15) holds by definition of E_0. In order to prove that $E_0 \cup E$ is an orthonormal basis (and thus the first equality in (13.15)), we set

$$V = \overline{\text{span}\,(E_0 \cup E)}^{\perp}\,.$$

According to the characterization of orthonormal bases in Proposition 2.20 it suffices to prove that V equals the null space, $V = \{0\}$.
First we show that $T(V) \subset V$. Let $x \in V$. For $y \in E_\lambda$, $\lambda \in \mathbb{C}$ arbitrary, we get according to Lemma 13.5

$$\langle Tx, y \rangle = \langle x, T^* y \rangle = \left\langle x, \overline{\lambda} y \right\rangle = \lambda \langle x, y \rangle = 0$$

by definition of V. It follows that $Tx \in (E_0 \cup E)^\perp$ and therefore $Tx \in V$.
Next we show that $T|V = 0$. If this would not be true, $T|V$ has, according to Proposition 13.4, an eigenvalue μ satisfying $0 \neq \|T|V\| = |\mu|$. For a corresponding eigenvector $x \in V$ we would have $x \in U_\mu \subset V^\perp$, a contradiction to $V \cap V^\perp = \{0\}$. Therefore, $T|V = 0$.
Now, because of $V \subset \ker T = U_0 \subset V^\perp$ we obtain $V = \{0\}$. Therefore $E_0 \cup E$ is an orthonormal basis of X. Consequently,

$$X = \ker T \oplus (\ker T)^\perp\,, \quad (\ker T)^\perp = \overline{\text{span}\,(E)}\,.$$

According to Proposition 2.20(iv), every $x \in X$ can be represented as

$$x = P_0 x + \sum_{j \in J} \langle x, e_j \rangle e_j\,.$$

We have $TP_0 = 0$, since $P_0(X) = \ker T$, and $Te_j = \lambda_j e_j$. This yields (13.16) as well as the third equality in (13.15). Finally, (13.17) was already obtained in Proposition 13.4. □

The preceding proposition remains valid if H is not separable. In this case, every orthonormal basis of the eigenspace $U_0 = \ker T$ is uncountable. For the proof one then uses a more general version of the characterization result 2.20 which is valid for arbitrary (separable or nonseparable) Hilbert spaces.

We define an approximation T_n of T by

$$T_n x = \sum_{j=1}^{n} \lambda_j \langle x, e_j \rangle e_j \tag{13.18}$$

for $n < |J|$; if $|J|$ is finite, we set $T_n = T$ for $n \geq |J|$.

Proposition 13.8 *In the situation of Proposition 13.7 we have*

$$\|T - T_n\| = \max_{j>n} |\lambda_j| \,, \tag{13.19}$$

and $\|T - T_n\| \to 0$ *as* $n \to \infty$.

Proof: We have $T_n e_j = \lambda_j e_j = T e_j$ for $j \le n$ as well as $T_n e_j = 0$ for $j > n$. It follows that

$$T - T_n = T \quad \text{on} \quad \overline{\operatorname{span}\{e_j : j > n\}} \,.$$
$$\ker(T - T_n) = \ker(T) \oplus \operatorname{span}\{e_j : j \le n\}$$

Consequently, $T - T_n$ has the spectral decomposition

$$(T - T_n)x = \sum_{j>n} \lambda_j \langle x, e_j \rangle \, e_j \,, \quad x \in X \,.$$

We thus have $\sigma_p(T - T_n) \subset \sigma_p(T) \cup \{0\}$ and

$$\max_{\lambda \in \sigma_p(T - T_n)} |\lambda| = \max_{j>n} |\lambda| \,. \tag{13.20}$$

This proves (13.19). Since $\{\lambda : \lambda \in \sigma_p(T), \, |\lambda| \ge \varepsilon\}$ is finite for all $\varepsilon > 0$, (13.20) converges to zero for $n \to \infty$. □

We now construct a decomposition for an arbitrary (not necessarily normal) compact operator $T \in L(X)$, where X is a Hilbert space.

From Lemma 13.3 we know that T^*T is normal and positive. It is compact since T is compact. By Proposition 13.7, T^*T has the spectral decomposition

$$(T^*T)(x) = \sum_{k \in K} \nu_k \langle x, e_k \rangle \, e_k \,, \tag{13.21}$$

where ν_k are the nonzero eigenvalues of T^*T and $\{e_k\}$ is a corresponding system of orthonormal eigenvectors. Since T^*T is positive, we have

$$\nu_k = \langle \nu_k e_k, e_k \rangle = \langle T^*T e_k, e_k \rangle > 0$$

for all $k \in K$. The numbers

$$s_k = \sqrt{\nu_k} \,, \quad k \in K \,, \tag{13.22}$$

are called **singular values** of T. We define $f_k \in X$ by

$$f_k = \frac{1}{s_k} T e_k \,, \quad k \in K \,. \tag{13.23}$$

Proposition 13.9 (Singular value decomposition)
Let X be a Hilbert space, $T \in L(X)$ compact. Let $e_k, f_k \in X$ and $\nu_k, s_k > 0$, $k \in K$, be defined by (13.21) – (13.23). Then $\{f_k\}_{k \in K}$ is an orthonormal system in X, and for all $x \in X$ we have

$$Tx = \sum_{k \in K} s_k \langle x, e_k \rangle \, f_k \,. \tag{13.24}$$

Moreover,

$$\|T\| = \max_{k \in K} s_k \,. \tag{13.25}$$

Proof: According to Proposition 13.7,

$$X = \ker(T^*T) \oplus \overline{\operatorname{span}(E)}\,, \quad E = \{e_k : k \in K\}\,.$$

We claim that $\ker(T^*T) = \ker(T)$. Indeed, $\ker(T) \subset \ker(T^*T)$, and if $T^*Tx = 0$ then

$$0 = \langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\|^2\,,$$

so $Tx = 0$ and $x \in \ker(T)$. It follows that

$$X = \ker(T) \oplus \overline{\operatorname{span}(E)}\,.$$

Let again P_0 denote the projection onto $\ker(T)$. Then for all $x \in X$

$$x = P_0 x + \sum_{k\in K} \langle x, e_k\rangle\, e_k$$

by Proposition 2.20(iv). Therefore, since T is linear and continuous,

$$Tx = T\Big(P_0 x + \sum_{k\in K} \langle x, e_k\rangle\, e_k\Big) = \sum_{k\in K} \langle x, e_k\rangle\, Te_k = \sum_{k\in K} s_k \langle x, e_k\rangle\, f_k\,.$$

This proves (13.24). Next, we check that $\{f_k\}_{k\in K}$ is an orthonormal system. This follows from the computation

$$\langle f_k, f_m\rangle = \frac{1}{s_k^2}\langle Te_k, Te_m\rangle = \frac{1}{\nu_k}\langle T^*Te_k, e_m\rangle = \langle e_k, e_m\rangle\,, \quad k, m \in K\,,$$

since $\{e_k\}_{k\in K}$ is an orthonormal system and $T^*Te_k = \nu_k e_k$. Finally, since $\|T^*T\| = \max_k \nu_k$ by Proposition 13.7, (13.25) follows from the computation, using Lemma 13.3,

$$\|T\|^2 = \|T^*T\| = \max_{k\in K} \nu_k = \max_{k\in K} s_k^2 = \big(\max_{k\in K} s_k\big)^2\,.$$

$\square$

Customarily, the nonzero singular values are arranged in decreasing order $s_1 \geq s_2 \geq \dots$, with multiplicity equal to the dimension of the corresponding eigenspace of T^*T, which has finite dimension as we know. In addition, one attaches zero singular values to the vectors of an orthonormal basis of $\ker(T)$ to obtain an orthonormal basis of X. One may also extend $\{f_k\}_{k\in K}$ to an orthonormal basis of X.

For compact and normal operators $T \in L(H)$, it turns out that the singular values s_k and the eigenvalues λ_k are related by $s_k = |\lambda_k|$.

The singular values are closely related to finite-rank approximations of arbitrary compact operators $T \in L(X)$. Namely, $T_n \in L(X)$ defined by

$$T_n x = \sum_{k=1}^{n} s_k \langle x, e_k\rangle\, f_k \tag{13.26}$$

satisfies $\dim(T_n(X)) = n$ if $s_n > 0$. (For $T \in L(X)$, one calls $\dim(T(X))$ the **rank** of T.) We have

$$(T - T_n)(x) = \sum_{k>n} s_k \langle x, e_k\rangle\, f_k\,,$$

and one may check that this yields the singular value decomposition of $T - T_n$. By Proposition 13.9 we therefore get that

$$\|T - T_n\| = s_{n+1}\,, \quad \lim_{n\to\infty} \|T - T_n\| = 0\,. \tag{13.27}$$

This yields:

Corollary 13.10 *Let X be a Hilbert space. Every compact operator $T \in L(X)$ can be obtained as the limit of a sequence of finite-rank operators.* □

We remark that the singular value decomposition of Proposition 13.9 can be extended to the situation where $T \in L(X;Y)$ for Hilbert spaces X and Y. Corollary 13.10 can then be extended, too, to arbitrary compact operators between Hilbert spaces. For general Banach spaces instead of Hilbert spaces, however, it does not hold; see the remark at the end of the chapter on compact operators.

Another approximation property of the singular values is the following. We present it without proof.

Proposition 13.11 *Let H be a Hilbert space, $T \in L(H)$ compact. Then we have*

$$s_n(T) = \inf\{\|T - Q\| : Q \in L(H), \operatorname{rank}(Q) < n\} \tag{13.28}$$

for all $n \in \mathbb{N}$. (Here, $s_n(T)$ denotes the n-th singular value of T.) □

Corollary 13.12 *Let H be a Hilbert space, $T, S \in L(H)$ compact. Then*

$$s_{n+m-1}(T + S) \leq s_n(T) + s_m(S) \tag{13.29}$$

for all $n, m \in \mathbb{N}$.

Proof: Let $Q, R \in L(H)$ with $\operatorname{rank}(Q) < n$ and $\operatorname{rank}(R) < m$ be arbitrary. Then

$$\operatorname{rank}(Q + R) \leq n + m - 2 < n + m - 1\,,$$

so by Proposition 13.11

$$s_{n+m-1}(T + S) \leq \|(T + S) - (Q + R)\| \leq \|T - Q\| + \|S - R\|$$

Passing to the infimum with respect to Q and R, the claim follows. □

Corollary 13.13 *Let H be a Hilbert space, $T, \tilde{T} \in L(H)$ compact. Then*

$$|s_n(T) - s_n(\tilde{T})| \leq \|T - \tilde{T}\| \tag{13.30}$$

for all $n \in \mathbb{N}$.

Proof: We apply (13.29) with $S = \tilde{T} - T$ and $m = 1$. This yields

$$s_n(\tilde{T}) \leq s_n(T) + s_1(\tilde{T} - T) \leq s_n(T) + \|\tilde{T} - T\|\,.$$

Interchanging the role of T and $\tilde{T}$ completes the proof. □

This means that the singular values of compact operators T are continuous functions of T, considered as a mapping from $L(H) \to \mathbb{R}$. In contrast to that, the eigenvalues in general do not have this property. This is one of the reasons why the singular value decomposition is an important computational tool when one deals with linear problems for general (not necessarily normal) operators, in the finite dimensional as well as in the infinite dimensional case.