Functional Analysis 1

LECTURE NOTES

Prof. Dr. Gitta Kutyniok

June 17, 2014

Contents

1 Metric Spaces

In this chapter we recall the basic notions of metric spaces and prove Baire's Theorem and the theorem of Arzelà-Ascoli. Throughout this lecture $\mathbb{K}$ will always denote either $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.1. Let X be a set. Then a map $d\colon X \times X \to [0, \infty)$ is called a *metric on* X, if for all $x, y, z \in X$

(i) $d(x, y) = 0$ if and only if $x = y$,

(ii) $d(x, y) = d(y, x)$ and

(iii) $d(x, z) \leq d(x, y) + d(y, z)$. (Triangle inequality)

(X, d) is then called a *metric space*, and $d(x, y)$ is referred to as the *distance* between x and y. If $Y \subset X$, then $d|_{Y \times Y}$ is the *induced metric on* Y.

Notice that the non-negativity of a metric already follows from

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

Next we will give some important examples of metrics on function spaces, sequence spaces and $\mathbb{K}^n$. Also, we can define a metric on every set as the first example will show.

Beispiel 1.2.

(1) Let X be a set and let $d\colon X \times X \to [0, \infty)$ be defined by

$$d(x, y) := \begin{cases} 1 & x \neq y\,, \\ 0 & \text{else}\,. \end{cases}$$

This is the so-called *discrete metric*. Hence, this always defines a metric.

(2) Let X be a set and define

$$B(X) = \{f\colon X \to \mathbb{K} : f \text{ is bounded}\}\,.$$

Then

$$d(f, g) := \sup_{x \in X} |f(x) - g(x)|$$

is a metric on $B(X)$, the so-called *supremum metric*. Let now $X = [a, b]$ and set

$$C[a, b] = \{f\colon [a, b] \to \mathbb{K} : f \text{ is continuous}\}\,.$$

Then

$$C[a, b] \subset B[a, b]$$

and hence d induces a metric on $C[a, b]$.

(3) For $1 \leq p < \infty$, let $d_p\colon \mathbb{K}^n \times \mathbb{K}^n \to [0, \infty)$ be defined by

$$d_p(x, y) := \left(\sum_{j=1}^{n} |x_j - y_j|^p\right)^{\frac{1}{p}}\,,\ x = (x_j)_{j=1}^n,\ y = (y_j)_{j=1}^n\,,$$

and let $d_\infty\colon \mathbb{K}^n \times \mathbb{K}^n \to [0,\infty)$ be defined by

$$d_\infty(x,y) := \max_{1\le j\le n} |x_j - y_j|\,.$$

For $p = 1, \infty$ this is obviously a metric. Theorem 1.4 will imply that this also is the case for $1 < p < \infty$.

(4) The spaces $(\mathbb{K}^n, d_p)$ can be generalized to "infinite-dimensional sequence spaces". For this, for $1 \le p < \infty$, set

$$\ell_p := \left\{ x = (x_n)_{n\in\mathbb{N}} : x_n \in \mathbb{K},\ \sum_{n=1}^{\infty} |x_n|^p < \infty \right\},$$

and define $d_p\colon \ell_p \times \ell_p \to [0,\infty)$ by

$$d_p(x,y) = \left(\sum_{n=1}^{\infty} |x_n - y_n|^p \right)^{\frac{1}{p}}\,.$$

This is well-defined, since by Theorem 1.4 ℓ_p is a linear space. Let further ℓ_∞ be defined by

$$\ell_\infty := \{x = (x_n)_{n\in\mathbb{N}} : (x_n)_{n\in\mathbb{N}} \text{ is bounded}\}$$

and define $d_\infty\colon \ell_\infty \times \ell_\infty \to [0,\infty)$ by

$$d_\infty(x,y) := \sup_{n\in\mathbb{N}} |x_n - y_n|\,.$$

Then (ℓ_p, d_p), $1 \le p \le \infty$, are metric spaces, again partly proven by Theorem 1.4.

To show the triangle inequality for the ℓ_p-spaces we need another inequality, which is important in its own right. Hölder's inequality gives upper bounds on a series of products in terms of products of series.

Theorem 1.3 (Hölder's inequality)**.** *Let $1 < p < \infty$, and let $1 < q < \infty$ be defined by $q := \frac{p}{p-1}$ (hence $\frac{1}{p} + \frac{1}{q} = 1$). Then, for $x \in \ell_p$ and $y \in \ell_q$, we have*

$$\sum_{n=1}^{\infty} |x_n y_n| \le \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{\frac{1}{p}} \left(\sum_{n=1}^{\infty} |y_n|^q \right)^{\frac{1}{q}}\,.$$

(Case $p = q = 2$: (Cauchy-)Schwarz' inequality)

Proof. Let $c = \frac{1}{p}$ and define $\varphi\colon [0,\infty) \to \mathbb{R}$ by $\varphi(t) = t^c - ct$. Then

$$\varphi'(t) = ct^{c-1} - c \text{ and } \varphi''(t) = c(c-1)t^{c-2}\,.$$

Thus φ has a global maximum value in $t = 1$. This implies

$$1 - c \ge t^c - ct \text{ for all } t > 0\,,$$

hence

$$t^c - 1 \le c(t-1)\,. \tag{1.1}$$

Let now $a, b > 0$, and set $t = \frac{a^p}{b^q}$. Then, by (1.1), we obtain

$$\frac{a}{b^{\frac{q}{p}}} - 1 \leq \frac{1}{p}\left(\frac{a^p}{b^q} - 1\right)$$

and thus

$$\frac{a}{b^{q\left(\frac{1}{p}-1\right)}} - b^q \leq \frac{1}{p}\left(a^p - b^q\right).$$

Since $1 = \frac{1}{p} + \frac{1}{q}$, this implies

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \tag{1.2}$$

We now set

$$A = \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{1}{p}} \quad \text{and}$$

$$B = \left(\sum_{n=1}^{\infty} |y_n|^q\right)^{\frac{1}{q}}$$

as well as $\widetilde{x}_n := \frac{x_n}{A}$ and $\widetilde{y}_n := \frac{y_n}{B}$. Without loss of generality, we assume $A, B > 0$. By (1.2), we obtain

$$|\widetilde{x}_n \widetilde{y}_n| \leq \frac{1}{p}|\widetilde{x}_n|^p + \frac{1}{q}|\widetilde{y}_n|^q.$$

Hence

$$\sum_{n=1}^{\infty} |\widetilde{x}_n \widetilde{y}_n| \leq \frac{1}{p}\sum_{n=1}^{\infty}|\widetilde{x}_n|^p + \frac{1}{q}\sum_{n=1}^{\infty}|\widetilde{y}_n|^q = \frac{1}{p} + \frac{1}{q} = 1.$$

And, finally

$$\sum_{n=1}^{\infty} |x_n y_n| \leq AB,$$

which is the assertion. □

The following Minkowski's inequality sets the ground for the triangle inequality of the metric d_p.

Theorem 1.4 (Minkowski's inequality). *For $1 < p < \infty$ and $x, y \in \ell_p$,*

$$\left(\sum_{n=1}^{\infty} |x_n + y_n|^p\right)^{\frac{1}{p}} \leq \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{1}{p}} + \left(\sum_{n=1}^{\infty} |y_n|^p\right)^{\frac{1}{p}}.$$

Proof. With $z_n := x_n + y_n$, we first obtain

$$|z_n|^p = |x_n + y_n| \cdot |z_n|^{p-1} \leq \left(|x_n| + |y_n|\right)|z_n|^{p-1}.$$

This implies

$$\sum_{n=1}^{m} |z_n|^p \leq \sum_{n=1}^{m} |x_n| \cdot |z_n|^{p-1} + \sum_{n=1}^{m} |y_n| \cdot |z_n|^{p-1} \quad \text{for all } m \in \mathbb{N}.$$

By Theorem 1.3,

$$\sum_{n=1}^{m} |z_n|^p \leq \left(\sum_{n=1}^{m} |x_n|^p\right)^{\frac{1}{p}} \left(\sum_{n=1}^{m} |z_n|^{(p-1)q}\right)^{\frac{1}{q}} + \left(\sum_{n=1}^{m} |y_n|^p\right)^{\frac{1}{p}} \left(\sum_{n=1}^{m} |z_n|^{(p-1)q}\right)^{\frac{1}{q}}.$$

Since $(p-1)q = p$, we conclude that

$$\left(\sum_{n=1}^{m} |z_n|^p\right)^{\frac{1}{p}} = \left(\sum_{n=1}^{m} |z_n|^p\right)^{1-\frac{1}{q}} \leq \left(\sum_{n=1}^{m} |x_n|^p\right)^{\frac{1}{p}} + \left(\sum_{n=1}^{m} |y_n|^p\right)^{\frac{1}{p}} .$$

We now consider $m \to \infty$, which we are allowed to do since the right-hand-side converges. This proves the theorem. □

The triangle inequality $d_p(u,w) \leq d_p(u,v) + d_p(v,w)$ for all $u, v, w \in \ell_p$ for the metric d_p can now directly be concluded from Theorem 1.4 by setting $x_n = u_n - v_n$ and $y_n = v_n - w_n$.

Definition 1.5. Let (X, d) be a metric space.

(1) For $x \in X$ and $r > 0$, the set $U_r(x)$ defined by

$$U_r(x) := \{y \in X : d(x,y) < r\}$$

is called the *open ball* of radius r and center x. $U \subset X$ is called *open*, if for each $x \in U$ there exists some $\varepsilon > 0$ such that $U_\varepsilon(x) \subset U$.

(2) A set $A \subset X$ is *closed*, if $X \setminus A$ is open. The set

$$K_r(x) := \{y \in X : d(x,y) \leq r\}, \ x \in X, \ r > 0,$$

is called the *closed ball* of radius r and center x.

(3) If $E \subset X$, then $x \in E$ is an *interior point of* E, if there exists some open set $U \subset E$ with $x \in U$. E is then called a *neighbourhood of* x. The set of all interior points is referred to as the *interior of* E and is denoted by $\mathring{E}$.

(4) A point $x \in X$ is called *limit point of* E if $U \cap E \neq \emptyset$ for each neighbourhood U of x. The set of all limit points of E is the *closure of* E, which is denoted by $\overline{E}$. E is *dense* in X if $\overline{E} = X$.

The openness of a set and all properties that can be defined with reference only to open sets are called topological. In particular, all terms just defined are topological. The open sets in a metric space form a system of sets called topology. This terminology will be generalized to the notion of a topological space in Chapter 6 on page 44.

Lemma 1.6. *Let (X, d) be a metric space.*

(i) *We have*

(a) *$\emptyset$, X are open.*

(b) *If $U_1, \dots U_r \subset X$ are open, then $\bigcap_{i=1}^{r} U_i$ is open.*

(c) *If $U_i \subset X$, $i \in I$, are open, then $\bigcup_{i \in I} U_i$ is open.*

Hence d defines a topology *on X with $U_\varepsilon(x)$, $x \in X$, $\varepsilon > 0$, as basis.*

(ii) *We have*

(a) *$\emptyset$, X are closed.*

(b) *If* $A_i \subset X$, $i \in I$, *are closed, then* $\bigcap_{i\in I} A_i$ *is closed.*

(c) *If* $A_1, \ldots, A_r \subset X$ *are closed, then* $\bigcup_{i=1}^{r} A_i$ *is closed.*

(iii) *For each* $x \in X$, $r > 0$, *the set* $K_r(x)$ *is closed.*

(iv) *For* $E \subset X$, $\overline{E}$ *is the smallest closed set containing* E.

(v) *For* $E \subset X$, $\mathring{E}$ *is the biggest open set contained in* E.

Proof. Tutorials □

The next definition generalizes the notion of convergence from $\mathbb{K}^n$ (with the Euclidian metric) to general metric spaces.

Definition 1.7. Let (X, d) be a metric space.

(i) A sequence $(x_n)_{n\in\mathbb{N}} \subset X$ *converges* to $x \in X$ if for each $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbb{N}$ with $d(x_n, x) < \varepsilon$ for all $n \geq N_\varepsilon$. We then write $x_n \to x$, as $n \to \infty$ or $x = \lim_{n\to\infty} x_n$. The point x is called the *limit* of $(x_n)_{n\in\mathbb{N}}$.

(ii) A sequence $(x_n)_{n\in\mathbb{N}} \subset X$ is a *Cauchy-sequence*, if for each $\varepsilon > 0$ there exists some $N_\varepsilon \in \mathbb{N}$ with

$$d(x_n, x_m) < \varepsilon \quad \text{for all } n, m \geq N_\varepsilon\,.$$

(iii) The space (X, d) is *complete*, if each Cauchy-sequence in X converges.

Convergence of a sequence is a topological property. The sequence $(x_n)_{n\in\mathbb{N}}$ converges to x, if and only if every neighborhood of x contains all but finitely many elements of the sequence. In particular, the limit of a sequence is independent of the ordering of the sequence's terms. Which sequences are Cauchy-sequences does not only depend on the open sets but also on the chosen metric (see Remark 1.11).

In general, topological properties in metric spaces can be tested by sequences. We note that there is a characterization of closedness of a set by convergent sequences.

Lemma 1.8. *Let* (X, d) *be a metric space.*

(i) *A sequence can have at most one limit.*

(ii) *Let* $E \subset X$. *Then* $x \in \overline{E}$ *if and only if there exists a sequence* $(x_n)_{n\in\mathbb{N}} \subset E$ *with* $x_n \to x$ *as* $n \to \infty$.

(iii) *If* $(x_n)_{n\in\mathbb{N}} \subset X$ *is convergent, then* $(x_n)_{n\in\mathbb{N}} \subset X$ *is a Cauchy-sequence. The converse is not always true*[1]. *A Cauchy-sequence is convergent, if it contains a convergent subsequence.*

(iv) *If* X *is complete and* $E \subset X$ *closed, then* E *is complete. If* $E \subset X$ *is complete, then* E *is closed in* X.

Proof. Tutorials □

The following example provides the reader with some complete metric spaces.

[1] For example, consider $X = (0, 1]$, $x_n = \frac{1}{n}$.

Beispiel 1.9.

(1) The space $B(X)$ is complete.

Proof. Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy-sequence in $B(X)$, and for $\varepsilon > 0$ let $N_\varepsilon \in \mathbb{N}$ be such that

$$d(f_n, f_m) < \varepsilon \text{ for all } n, m \geq N_\varepsilon\,.$$

This implies $|f_n(x) - f_m(x)| < \varepsilon$ for all $x \in X$, $n, m \geq N_\varepsilon$. Hence, for all $x \in X$, $(f_n(x))_{n\in\mathbb{N}}$ is a Cauchy-sequence in $\mathbb{K}$. Setting $f(x) := \lim_{n\to\infty} f_n(x)$, we obtain

$$|f(x) - f_m(x)| = \lim_{n\to\infty} |f_n(x) - f_m(x)| < \varepsilon$$

for all $m \geq N_\varepsilon$. Hence $|f(x)| \leq |f_n(x)| + \varepsilon$, which implies $f \in B(X)$. Further, for $m \geq N_\varepsilon$,

$$d(f, f_m) = \sup_{x\in X} |f(x) - f_m(x)| < \varepsilon\,,$$

and thus $f = \lim_{n\to\infty} f_n$. □

(2) The space $C[a, b]$ is complete, since it is closed in $B[a, b]$ (see lemma 1.8), the reason being that a uniform limit of continuous functions is again continuous.

(3) The spaces $(\mathbb{K}^n, d_p)$, $n \in \mathbb{N}$, $1 \leq p \leq \infty$, are complete, since convergence in $\mathbb{K}^n$ w.r.t. d_p is the same as convergence in $\mathbb{K}^n$ w.r.t. the component sequences.

(4) The spaces ℓ_p, $1 \leq p \leq \infty$, are complete.

Proof. Let $(x_k)_{k\in\mathbb{N}}$ be a Cauchy-sequence in ℓ_p, $x_k = (x_{k,n})_{n\in\mathbb{N}}$, and for $\varepsilon > 0$ let $N_\varepsilon \in \mathbb{N}$ be with

$$\begin{aligned} d_p(x_k, x_l) &= \left(\sum_{n=1}^{\infty} |x_{k,n} - x_{l,n}|^p\right)^{\frac{1}{p}} < \varepsilon && \text{and} \\ d_\infty(x_k, x_l) &= \sup_{n\in\mathbb{N}} |x_{k,n} - x_{l,n}| < \varepsilon && \text{for all } k, l > N_\varepsilon\,. \end{aligned} \tag{1.3}$$

Thus, for fixed $n \in \mathbb{N}$, $(x_{k,n})_{k\in\mathbb{N}}$ is a Cauchy-sequence in $\mathbb{K}$. Now set $y_n := \lim_{k\to\infty} x_{k,n}$ and $y := (y_n)_{n\in\mathbb{N}}$. Then $y \in \ell_p$ and $y = \lim_{n\to\infty} x_n$.
Indeed, consider $l \to \infty$, which implies (by 1.3)

$$\sum_{n=1}^{m} |x_{k,n} - y_n|^p < \varepsilon^p$$

for all $m \in \mathbb{N}$ and thus

$$\sum_{n=1}^{\infty} |x_{k,n} - y_n|^p < \varepsilon^p$$

for all $k \geq N_\varepsilon$ and $|x_{k,n} - y_n| < \varepsilon$ for all $k \geq N_\varepsilon$, $n \in \mathbb{N}$. Hence $x_n - y \in \ell_p$, and thus $y \in \ell_p$ and $y = \lim_{k\to\infty} x_k$. □

The following theorem of Baire only holds in complete metric spaces. It is a key ingredient in the proofs of the fundamental theorems of functional analysis. Thus, they will only hold under some completeness assumption.

Theorem 1.10 (Baire's Theorem)**.** *Let (X, d) be a complete metric space, and let D_n, $n \in \mathbb{N}$, be open, dense subsets of X. Then also $\bigcap_{n\in\mathbb{N}} D_n$ is dense in X.*

Proof. We need to prove that for all $x \in X$ and $r > 0$ we have

$$U_r(x) \cap \bigcap_{n=1}^{\infty} D_n \neq \emptyset\,.$$

For this, let $x \in X$ and $r > 0$ be arbitrary, but fixed. By induction define a sequence $(x_n)_{n\in\mathbb{N}} \subset X$ and $(r_n)_{n\in\mathbb{N}} \subset \mathbb{R}^+$ by

(a) $K_{r_{n+1}}(x_{n+1}) \subset D_n \cap U_{r_n}(x_n)$ and

(b) $r_n \leq \frac{1}{n}$.

This can be done as follows: First, set $x_1 = x$ and $r_1 = \min\{1, r\}$. Second, assume that $x_1, \dots, x_n, r_1, \dots, r_n$ be already chosen $(n \geq 1)$. Since D_n is open and dense, also $D_n \cap U_{r_n}(x_n) \neq \emptyset$ is open. Hence there exists $x_{n+1} \in X$ and $r_{n+1} > 0$ with

$$U_{2r_{n+1}}(x_{n+1}) \subset D_n \cap U_{r_n}(x_n) \quad \text{and} \quad r_{n+1} \leq \frac{1}{n+1}\,.$$

This implies (a) and (b), since $K_{r_{n+1}}(x_{n+1}) \subset U_{2r_{n+1}}(x_{n+1})$.
Having constructed sequences $(x_n)_{n\in\mathbb{N}}$ and $(r_n)_{n\in\mathbb{N}}$ satisfying a) and b), we obtain

$$x_n \in K_{r_n} \subset D_{n-1} \cap U_{r_{n-1}}(x_{n-1}) \subset U_{r_{n-1}}(x_{n-1}) \subset \cdots \subset U_{r_m}(x_m)$$

for all $n > m$. Thus $d(x_n, x_m) < r_m \leq \frac{1}{m}$ for all $n > m$. This implies that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy-sequence in X.
Now set $x_0 := \lim_{n\to\infty} x_n$ (remember that X is complete). Since $d(x_n, x_m) \leq r_m$ for all $n > m$, we obtain $d(x_0, x_m) \leq r_m$ for all $m \in \mathbb{N}$. Thus, finally,

$$x_0 \in \bigcap_{m=1}^{\infty} K_{r_{m+1}}(x_{m+1}) \subset \bigcap_{m=1}^{\infty} D_m \cap U_{r_m}(x_m) \subset U_{r_1}(x_1) \cap \bigcap_{m=1}^{\infty} D_m \subset U_r(x) \cap \bigcap_{m=1}^{\infty} D_m\,,$$

and the theorem is proved. □

Bemerkung 1.11.

(a) Theorem 1.10 is in general false if X is not complete. As an example choose $X = \mathbb{Q} = \{q_1, q_2, \dots\}$ and $D_n = X \setminus \{q_n\}$, $n \in \mathbb{N}$, which are open and dense. We immediately see that however

$$\bigcap_{n=1}^{\infty} D_n = \emptyset\,.$$

(b) Let (X, d) be a complete, non-empty metric space and $A_n \subset X$, $n \in \mathbb{N}$, closed with $X = \bigcup_{n=1}^{\infty} A_n$. Then there exists at least one $n \in \mathbb{N}$ with

$$\mathring{A}_n \neq \emptyset\,.$$

Proof. Towards a contradiction, assume that

$$\mathring{A}_n = \emptyset \text{ for all } n \in \mathbb{N}\,.$$

Then $X \setminus A_n$ are open and dense for all $n \in \mathbb{N}$. By Baire's Theorem 1.10, $\bigcap\limits_{n=1}^{\infty} (X \setminus A_n)$ dense in X. But $\bigcap\limits_{n=1}^{\infty} (X \setminus A_n) = X \setminus \bigcup\limits_{n=1}^{\infty} A_n = \emptyset$. ↯ □

(c) Completeness is a property of the particular metric and not the convergence in X. For example, consider $X = (0, 1]$, $d_1(x, y) := \left|\frac{1}{x} - \frac{1}{y}\right|$ and $d_2(x, y) = |x - y|$. Then we have

$$x_n \to x \text{ in } (X, d_1) \Leftrightarrow x_n \to x \text{ in } (X, d_2)\,,$$

but (X, d_1) is complete and (X, d_2) is not (see tutorials).

Definition 1.12. Let (X, d) be a metric space.

(i) Let $\varepsilon > 0$. Then $M \subset X$ is called *ε-net*, if $X = \bigcup\limits_{x \in M} U_\varepsilon(x)$. The space X is called *totally bounded*, if for each $\varepsilon > 0$ there exists a finite ε-net. A subset $A \subset X$ is *totally bounded*, if $(A, d|_{A \times A})$ is totally bounded.

(ii) The space X is *compact*, if every open cover of X (that is, a family of open sets U_i, $i \in I$, such that $X = \bigcup_{i \in I} U_i$) has a finite subcover. The metric space $(A, d|_{A\times A})$ is *compact* if and only if every open cover of A (of open sets in X) has a finite subcover.

Compactness and total boundedness are intrinsic properties, that is a subset $A \subset (X, d)$ of some metric space is compact (totally bounded) if the metric space $(A, d|_{A\times A})$ is compact (totally bounded).

It is easy to see that every compact metric space is totally bounded. The following theorem shows that the two notions coincide for complete metric spaces. Note that this does not imply that these two properties coincide for all subsets of a complete metric space (see Corollary 1.15).

Theorem 1.13. *Let (X, d) be a metric space. Then the following are equivalent:*

(i) *The space (X, d) is complete and totally bounded.*

(ii) *The space (X, d) is compact.*

(iii) *Each sequence in X has a convergent subsequence.*

Proof. (i)⇒(ii). Towards a contradiction, assume that X is not compact. Let $\mathfrak{U}$ be an open cover of X which does not contain a finite subcover. By induction, we now define a sequence $(x_n)_{n\in\mathbb{N}} \subset X$ such that

(a) each neighborhood $U_{2^{-n}}(x_n)$, $n \in \mathbb{N}$, is not covered by finitely many $U \subset \mathfrak{U}$ and

(b) any neighborhoods $U_{2^{-n}}(x_n)$ and $U_{2^{-(n+1)}}(x_{n+1})$, $n \in \mathbb{N}$, do intersect, i. e. $U_{2^{-n}}(x_n) \cap U_{2^{-(n+1)}}(x_{n+1}) \neq \emptyset$.

First, for $n = 1$, notice that X is totally bounded. Hence $X = \bigcup_{y \in M} U_{\frac{1}{2}}(y)$, $|M| < \infty$, which implies that there exists $y_{i_0} =: x_1 \in X$ such that $U_{\frac{1}{2}}(x_1)$ is not covered by finitely many $U \subset \mathfrak{U}$. Second ($n \mapsto n+1$), again by totally boundedness, there exists a finite M such that $X = \bigcup_{y \in M} U_{2^{-(n+1)}}(y)$. Assume $x_1, \ldots, x_n$ are chosen such that (a) and (b) are satisfied. Towards a contradiction assume that for each $y \in M$ with $U_{2^{-(n+1)}}(y) \cap U_{2^{-n}}(x_n) \neq \emptyset$, the set $U_{2^{-(n+1)}}(y)$ is covered by finitely many $U \in \mathfrak{U}$. Then this is also true for $U_{2^{-n}}(x_n)$. $\lightning$
Hence there exists $x_{n+1} \in X$ with $U_{2^{-(n+1)}}(x_{n+1})$ is not covered by finitely many $U \in \mathfrak{U}$ and $U_{2^{-n}}(x_n) \cap U_{2^{-(n+1)}}(x_{n+1}) \neq \emptyset$.
For each $n \in \mathbb{N}$, let $z_n \in U_{2^{-n}}(x_n) \cap U_{2^{-(n+1)}}(x_{n+1})$. Then, for $m > n$,

$$\begin{aligned} d(x_m, x_n) &\leq \sum_{\nu=n}^{m-1} d(x_{\nu+1}, x_\nu) \leq \sum_{\nu=n}^{m-1} \left(d(x_{\nu+1}, z_\nu) + d(z_\nu, x_\nu) \right) \\ &\leq \sum_{\nu=n}^{m-1} \left(2^{-(\nu+1)} + 2^{-\nu} \right) \leq 2 \sum_{\nu=n}^{m-1} 2^{-\nu} \leq \frac{1}{2^{n-2}} . \end{aligned}$$

This implies that $(x_n)_{n \in \mathbb{N}}$ is a Cauchy-sequence. Since X is complete, there exists $x = \lim_{n \to \infty} x_n$.
Now choose $U \in \mathfrak{U}$ with $x \in U$ and choose $\varepsilon > 0$ such that $U_\varepsilon(x) \subset U$. Then $x_n \in U_{\frac{\varepsilon}{2}}(x)$ for all $n \geq N$, hence $U_{2^{-n}}(x_n) \subset U$ for all $n \geq N$ with $2^{-n} < \frac{\varepsilon}{2}$. $\lightning$ to choice of $U_{2^{-n}}(x_n)$.
(ii)$\Rightarrow$(iii). Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in X and set $A_n := \overline{\{x_\nu : \nu > n\}} \subset X$. Towards a contradiction assume that

$$\bigcap_{n \in \mathbb{N}} A_n = \emptyset .$$

This implies $\bigcup_{n \in \mathbb{N}} (X \setminus A_n) = X$. Since X is compact, the open cover $\{X \setminus A_n\}_{n \in \mathbb{N}}$ contains an open subcover $\{X \setminus A_{n_j} : 1 \leq j \leq r\}$. Since $A_{n+1} \subset A_n$, hence $X \setminus A_n \subset X \setminus A_{n+1}$, for $N := \max\{n_j : 1 \leq j \leq r\}$ we have

$$X = \bigcup_{j=1}^{r} X \setminus A_{n_j} = X \setminus A_N .$$

Thus $A_N = \emptyset$ $\lightning$, which is impossible. This proves that $\bigcap_{n \in \mathbb{N}} A_n \neq \emptyset$. Choosing $x \in \bigcap_{n \in \mathbb{N}} A_n$, there exists a sequence $(n_k)_{k \in \mathbb{N}} \subset \mathbb{N}$ with $n_{k+1} > n_k$ and $d(x_{n_k}, x) \leq \frac{1}{k}$ [if n_k is chosen, then $x \in A_{n_{k+1}}$]. This shows (iii), since $(x_{n_k})_{k \in \mathbb{N}}$ is a convergent subsequence of $(x_n)_{n \in \mathbb{N}}$ in X.
(iii)$\Rightarrow$(i). Each Cauchy-sequence in X contains by hypothesis a convergent susequence, is hence itself convergent. This implies that X is complete.
Towards a contradiction, we now assume that X is not totally bounded. Then there exists $\varepsilon > 0$ such that X is not covered by finitely many $U_\varepsilon(x)$, $x \in X$. By induction, we define a sequence $(x_n)_{n \in \mathbb{N}} \subset X$ with

$$x_n \notin U_\varepsilon(x_j),\ 1 \leq j \leq n-1 .$$

This can be achieved in the following way: Let $x_1 \in X$ be arbitrary. Then assume $x_1, \ldots, x_n$ are already constructed. Since

$$X \setminus \bigcup_{j=1}^{n} U_\varepsilon(x_j) \neq \emptyset ,$$

choose $x_{n+1} \in X \setminus \bigcup_{j=1}^{n} U_\varepsilon(x_j)$. Then, for $n \neq m$, we have

$$d(x_n, x_m) \geq \varepsilon\,.$$

By (iii), $(x_n)_{n\in\mathbb{N}}$ contains a convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$. Let $x := \lim_{k\to\infty} x_{n_k}$. Then $d(x_{n_k}, x) < \frac{\varepsilon}{2}$ for all $k > k_0$, hence $d(x_{n_k}, x_{n_l}) < \varepsilon$ for all $k, l > k_0$. $\lightning$ □

Lemma 1.14. *Let (X, d) be a metric space, and let $A \subset X$ be a non-empty subset.*

(i) *If X is totally bounded, then also A is totally bounded.*

(ii) *If A is totally bounded, then also $\overline{A}$ is totally bounded.*

Proof. (i). Let $\varepsilon > 0$. By hypothesis, there exists an $\frac{\varepsilon}{2}$-net $\{x_1, \ldots x_n\}$ of X. Without loss of generality, let $A \cap U_{\frac{\varepsilon}{2}}(x_j) \neq \emptyset$ if and only if $1 \leq j \leq m$, $m \leq n$. For each $1 \leq j \leq m$, choose $y_j \in A \cap U_{\frac{\varepsilon}{2}}(x_j)$. Let $y \in A$. Then there exists $1 \leq j \leq m$ with $y \in U_{\frac{\varepsilon}{2}}(x_j)$, and hence

$$d(y, y_j) \leq d(y, x_j) + d(x_j, y_j) < \varepsilon\,.$$

This implies that $\{y_1, \ldots, y_m\}$ is an ε-net for A.

(ii). Let $\varepsilon > 0$. By hypothesis, there exists an $\frac{\varepsilon}{2}$-net $\{y_1, \ldots, y_n\}$ for A. Let $x \in \overline{A}$. Then there exists $y \in A$ with $d(x, y) < \frac{\varepsilon}{2}$. Let y_j be such that $d(y, y_j) < \frac{\varepsilon}{2}$. This yields

$$d(x, y_j) \leq d(x, y) + d(y, y_j) < \varepsilon\,,$$

hence $\{y_1, \ldots, y_n\}$ is an ε-net for $\overline{A}$. □

Korollar 1.15. *Let (X, d) be a complete metric space, and let $A \subset X$. Then the following are equivalent.*

(i) *$\overline{A}$ is compact.*

(ii) *A is totally bounded.*

Proof. (i)⇒(ii). Since $\overline{A}$ is compact, by Theorem 1.13, $\overline{A}$ is totally bounded. By Lemma 1.14, A is totally bounded.

(ii)⇒(i). Since A is totally bounded, by Lemma 1.14, $\overline{A}$ is totally bounded. Since X is complete, $\overline{A}$ is also complete. Hence Theorem 1.13 implies that $\overline{A}$ is compact. □

Definition 1.16. Let (X, d) and (X', d') be metric spaces, and let $f\colon X \to X'$.

(1) f is *continuous in* $x \in X$, if for each $\varepsilon > 0$ there exists $\delta > 0$ such that $d(x, y) < \delta$ implies $d'(f(x), f(y)) < \varepsilon$ for all $y \in X$. If f is continuous in each $x \in X$, f is called *continuous*.

(2) f is a *homeomorphism*, if f is bijective and f and f^{-1} are both continuous. f is an *isometry*, if $d(x, y) = d'(f(x), f(y))$ for all $x, y \in X$. If further f is bijective, f is an isometric isomorphism. X and X' are then called *homeomorphic* resp. *isometric* resp. *isometrically isomorphic*.

(3) f is *uniformly continuous*, if for each $\varepsilon > 0$ there exists $\delta > 0$ such that $d(x, y) < \delta$ implies $d'(f(x), f(y)) < \varepsilon$ for all $x, y \in X$.

Two metric spaces being isometric is a strong notion of equivalence for metric spaces, being homeomorphic is the properly weaker topological equivalence of metric spaces.

Lemma 1.17. *Let (X, d) and (X', d') be metric spaces, and let $f\colon X \to X'$.*

(i) *f is continous $\Leftrightarrow$ $f^{-1}(U)$ is open in X for all $U \subset X'$ open $\Leftrightarrow$ $f(x_n) \to f(x)$ for all $x_n \to x$ in X.*

(ii) *Let X be compact and f continuous. Then f is automatically uniformly continuous.*

Proof. Exercises. □

We want to relate the relative compactness, that is the compactness of the closure, of a set of continuous real functions to the pointwise relative compactness of these functions. The relatively compact sets in $\mathbb{R}$ are the bounded sets by Heine-Borel theorem. If a set of continuous real functions is relatively compact, we obtain pointwise relative compactness, by continuity of $C(X) \to \mathbb{R}, f \mapsto f(x)$ for every $x \in X$. However, to prove the converse a second condition is needed: the equicontinuity of the functions.

Definition 1.18. $F \subset C(X)$ is *equicontinuous in $x \in X$*, if for each $\varepsilon > 0$ there exists a neighbourhood U of x with $|f(x) - f(y)| < \varepsilon$ for all $y \in U$ and $f \in F$. F is called *equicontinuous*, if it is equicontinuous in each $x \in X$.

Theorem 1.19 (Arzelà-Ascoli)**.** *Let X be a compact metric space and $F \subset C(X)$. Then the following are equivalent.*

(i) *$\overline{F}$ is compact.*

(ii) *F is equicontinuous and pointwise bounded.*

Proof. (i)⇒(ii). Exercise.
(ii)⇒(i). For $x \in X$ we write $F(x) := \{f(x) : f \in F\}$. Let F be equicontinuous and $F(x) \in \mathbb{K}$ bounded for all $x \in X$. Since $C(X)$ is complete, by Corollary 1.15 it remains to prove that F is totally bounded. For this, let $\varepsilon > 0$, and, for each $x \in X$, let U_x be an open neighbourhood of x with

$$|f(y) - f(x)| < \frac{\varepsilon}{4} \text{ for all } f \in F \text{ and } y \in U_x\,.$$

Let now $x_1, \ldots, x_n \in X$ be chosen such that $X = \bigcup\limits_{i=1}^{n} U_{x_i}$ and set

$$K := \bigcup_{i=1}^{n} F(x_i) \subset \mathbb{K}\,.$$

Since K is bounded, there exist $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$ with

$$K \subset \bigcup_{j=1}^{m} U_{\frac{\varepsilon}{4}}(\lambda_j)\,.$$

Define Φ to be the set of maps $\varphi\colon \{1, \ldots, n\} \to \{1, \ldots, m\}$. For $\varphi \in \Phi$, set

$$F_\varphi := \{f \in F : |f(x_i) - \lambda_{\varphi(i)}| < \frac{\varepsilon}{4} \text{ for } 1 \le i \le n\}\,.$$

Then

$$F = \bigcup_{\varphi \in \Phi} F_\varphi \,.$$

To see this, let $f \in F$. Then for each $1 \leq i \leq n$, there exists $j \in \{1, \ldots, m\}$ such that $f(x_i) \in U_{\frac{\varepsilon}{4}}(\lambda_j)$. Hence, there exists $\varphi \in \Phi$ with $f(x_i) \in U_{\frac{\varepsilon}{4}}(\lambda_{\varphi(i)})$. Hence $f \in F_\varphi$.
For $f, g \in F_\varphi$ and $y \in U_{x_i}$, $i \in \{1, \ldots, n\}$, we then obtain

$$|f(y) - g(y)| \leq |f(y) - f(x_i)| + |f(x_i) - \lambda_{\varphi(i)}| + |\lambda_{\varphi(i)} - g(x_i)| + |g(x_i) - g(y)| < \varepsilon \,.$$

Thus $d(f, g) < \varepsilon$ for all $f, g \in F_\varphi$, and hence a finite ε-net for F_φ (and thus also for F) exists. □

2 Normed Spaces

We define the fundamental spaces of our study, namely the normed spaces.

Definition 2.1. Let E be a linear space over $\mathbb{K}$.

1. A map $\|\cdot\| \colon E \to [0, \infty)$ is called a *norm an* E, and $(E, \|\cdot\|)$ a *normed space*, if for all $x, y \in E$, $\lambda \in \mathbb{K}$ the following are satisfsfied:
   (a) $\|x\| = 0$ if and only if $x = 0$,
   (b) $\|\lambda x\| = |\lambda| \cdot \|x\|$ and
   (c) $\|x + y\| \leq \|x\| + \|y\|$.

   The space E is called a *Banach space*, if $\left(E, d_{\|\cdot\|}\right)$ is complete (for the definition of $d_{\|\cdot\|}$ see Remark 2.2).

2. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on E are *equivalent* if there exist $\alpha, \beta > 0$ such that

$$\alpha\|x\|_1 \leq \|x\|_2 \leq \beta\|x\|_1$$

   for all $x \in E$.

We remark that a normed space is a special case of a metric space and that the topology induced by a norm is in some sense compatible with the vector space structure.

Remark 2.2. Let $(E, \|\cdot\|)$ be a normed space.

1. Letting $d_{\|\cdot\|}(x, y) := \|x - y\|$ defines a metric on E.

2. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be equivalent norms on E. Then $(E, \|\cdot\|_1)$ is complete if and only if $(E, \|\cdot\|_2)$ is complete.

3. For all $x, y \in E$ there holds $\big|\|x\| - \|y\|\big| \leq \|x - y\|$. In particular $\|\cdot\| \colon E \to \mathbb{R}$ is Lipschitz-continuous.

4. The algebraic operations

$$\begin{aligned} +\colon E \times E \mapsto E\,,\ (x, y) \mapsto x + y \quad \text{and} \\ \cdot\colon \mathbb{K} \times E \mapsto E\,,\ (\lambda, y) \mapsto \lambda \cdot y \end{aligned}$$

   are continuous since

   - $\|(x + y) - (x_0 + y_0)\| \leq \|x - x_0\| + \|y - y_0\|$ and
   - $\|\lambda x - \lambda_0 x_0\| \leq |\lambda|\|x - x_0\| + |\lambda - \lambda_0|\|x_0\|$.

5. If $F \subset E$ is a subspace, so is its closure $\overline{F}$.

We recall the concept of a quotient vector space.

Definition 2.3 (Quotient space)**.** Let E be a linear space, $F \subset E$ a subspace. Then defining $x \sim y$ by $x - y \in F$ for $x, y \in E$ is an equivalence relation. The equivalence classes are given by

$$[x]_\sim = \{y \in E : y - x \in F\} = \{y \in E : y \in x + F\} = x + F\,.$$

Thus $[x]_\sim$ is an *affine subspace*. The *quotient space* E/F is defined by $E/F := \{x + F : x \in E\}$. Via

$$[x]_\sim + [y]_\sim := [x+y]_\sim \qquad (x+F)+(y+F) = (x+y)+F \quad \text{and}$$
$$\lambda[x]_\sim := [\lambda x]_\sim \qquad \lambda(x+F) = (\lambda x) + F$$

the space E/F becomes a linear space.

Now we combine the concepts of a normed space and quotient spaces by endowing reasonable quotients of normed spaces with a natural norm.

Lemma 2.4. *Let $(E, \|\cdot\|)$ be a normed space and let $F \subset E$ be a* closed *subspace. Then*

$$\|x + F\| := \inf\{\|x+y\| : y \in F\}$$

defines a norm on E/F. Moreover, if E is a Banach space so is E/F.

Proof. First we check the norm properties.
(i). If $\|x+F\| = 0$, there exists a sequence $(y_n)_n \subset F$, such that

$$\|x - y_n\| \to 0\,,$$

for $n \to \infty$. Since $y_n \in F$ and the subspace F is closed, we conclude $x \in F$, i. e.

$$x + F = F = F = 0 + F = 0\,.$$

(ii) There holds $\|\lambda(x+F)\| = \|(\lambda x) + F\| = \inf\{\|\lambda x + y\| : y \in F\}$. For $\lambda = 0$ we have:

$$|\lambda(x+F)\| = 0 = |\lambda|\|x+F\|\,.$$

For $\lambda \neq 0$ we have

$$\begin{aligned}\|\lambda(x+F)\| &= \inf\{\|\lambda x + y\| : y \in F\}\\ &= |\lambda| \inf\{\|x+y\| : y \in F\}\\ &= |\lambda|\|x+F\|\,.\end{aligned}$$

(iii). For $x, y \in E$ and $\varepsilon > 0$ choose $z_1, z_2 \in F$ such that

$$\|x+F\| \geq \|x + z_1\| - \frac{\varepsilon}{2} \quad \text{and} \quad \|y+F\| \geq \|y+z_2\| - \frac{\varepsilon}{2}\,.$$

This yields

$$\|(x+F)+(y+F)\| = \|(x+y)+F\| \leq \|x+z_1+y+z_2\| \leq \|x+F\| + \|y+F\| + \varepsilon\,.$$

Now, let E be a Banach space. To show that also E/F is complete, let $(x_n + F)_{n\in\mathbb{N}}$ be a Cauchy-sequence in E/F, i.e. for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $\|(x_n - x_m) + F\| \leq \varepsilon$ for all $n, m \geq N$. So for all $i \in \mathbb{N}$ we can find n_i, such that $\|x_{n_{i+i}} - x_{n_i} + F\| < 2^{-i}$. In particular there exists $y_i \in F$ such that $\|x_{n_{i+1}} - x_{n_i} + y_i\| < 2^{-i}$. We may assume $n_i < n_{i+1}$. Now define

$$\begin{aligned} z_1 &:= 0\,,\\ z_{i+1} &:= y_i + z_i \qquad i \geq 1\,. \end{aligned}$$

Then we have

$$\|(x_{n_{i+1}} + z_{i+1}) - (x_{n_i} + z_i)\| < 2^{-i}\,.$$

Now we define $\eta_i := x_{n_i} + z_i$, which gives us $\|\eta_{i+1} - \eta_i\| < 2^{-i}$. By this,

$$\|\eta_{m+k} - \eta_m\| \le \sum_{i=0}^{k-1} \|\eta_{m+i+1} - \eta_{m+i}\| < \sum_{i=0}^{k-1} 2^{-m-i} \le 2^{1-m}\,,$$

which means that (η_n) is a Cauchy-sequence in E and thus converges. Now we set $\lim_{n\to\infty} \eta_n =: x$. We obtain:

$$\begin{aligned}\|(x_n + F) - (x + F)\| &= \|(x_n - x) + F\| \\ &\overset{z_i \in F}{\le} \|x_{n_i} + z_i - x\| = \|\eta_i - x\| \to 0\,.\end{aligned}$$

This gives us a convergent subsequence, so the Cauchy-sequence is covergent itself. □

Indeed, completeness is not only inherited by closed subspaces and quotients by closed subspaces. As a converse, we have the following lemma.

Lemma 2.5. *Let E be a normed space, $F \subset E$ a closed subspace. If F and E/F are Banach spaces, then also E is a Banach space.*

Proof. Let $(x_n) \subset E$ be a Cauchy-sequence in E. Hence

$$\|(x_n + F) - (x_m + F)\| = \|(x_n - x_m) + F\| \le \|x_n - x_m\|\,.$$

So $(x_n + F)_n \subset E/F$ is a Cauchy-sequence in E/F. With $x + F := \lim_{n\to\infty} x_n + F$ we obtain:

$$\inf\{\|x_n - x + y\| : y \in F\} = \|(x_n - x) + F\| \to 0\,.$$

This means that there exists $(y_n) \subset F$ with $\|x_n - x + y_n\| \to 0$. This (y_n) is in fact a Cauchy sequence:

$$\begin{aligned}\|y_n - y_m\| &= \|y_n + x_n - x - x_n + x_m - y_m - x_m + x\| \\ &\le \|y_n + x_n - x\| + \|x_n - x_m\| + \|y_m + x_m - x\| \xrightarrow{m,n\to\infty} 0\,.\end{aligned}$$

Hence, it converges. Put $y := \lim_{n\to\infty} y_n \in F$. Now

$$\|x_n - x + y\| \le \|x_n + y_n - x\| + \|y - y_n\| \xrightarrow{n\to\infty} 0\,.$$

Thus (x_n) converges to $x - y$. □

Using this, we show completeness of finite-dimensional spaces.

Korollar 2.6. *A finite-dimensional normed space E is always a Banach space.*

Proof. We prove the statement by induction over $n = \dim E$. Let $\dim E = 1$ and choose $x \in E$ such that $\|x\| = 1$. Then $q\colon \mathbb{R} \mapsto E, q(\lambda) := \lambda x$, is isometric. So $E \cong \mathbb{R}$ and E is a Banach space. Assume that the claim holds for n and let $\dim E = n + 1$. Choose $x \in E \setminus \{0\}$ and set $F := \operatorname{span}\{x\}$. Because $\dim F = 1$ we know that F is complete, so closed. There holds

$$\dim(E/F) = \dim E - \dim F = n\,.$$

By assumption we get that E/F is complete, and by Lemma 2.5 we see that E is a Banach space. □

Lemma 2.7. *Let F be a closed subspace of a normed space E. Then for each $x \in E \setminus F$ there exist $M, M' > 0$ such that for all $y \in F$ and all $\lambda \in \mathbb{K}$ we have*

$$|\lambda| \le M\|\lambda x + y\| \quad \text{and} \quad \|y\| \le M'\|\lambda x + y\|\,.$$

Proof. Because of $x \notin F$ we have $\|x + F\| \neq 0$. We set

$$M := \|x + F\|^{-1} \quad \text{and} \quad M' := 1 + M\|x\|\,.$$

Then for $y \in F$ and $\lambda \in \mathbb{K}$ we obtain

$$|\lambda| = M|\lambda|\|x + F\| = M\|\lambda x + F\| \le M\|\lambda x + y\|\,,$$

and therefore also

$$\|y\| \le \|y + \lambda x\| + |\lambda|\|x\| \le \|y + \lambda x\| + M\|\lambda x + y\|\|x\| = \|y + \lambda x\|(1 + M\|x\|)\,.$$

This proves the lemma. □

In the finite-dimensional case, also the boundedness of linear operators is always granted.

Lemma 2.8. *Let $T\colon E \to X$ be a linear mapping between normed spaces E and X, where $\dim E < \infty$. Then there exists $c > 0$, such that for all $x \in E$ we have*

$$\|Tx\|_X \le c\|x\|_E\,.$$

Proof. We set $n := \dim E$. Let $n = 1$ and choose $x_0 \in E$, $\|x_0\| = 1$. Then

$$E = \operatorname{span}\{x_0\}$$

and for $x = \lambda x_0 \in E$ we have:

$$\|Tx\|_X = \|T(\lambda x_0)\|_X = |\lambda|\|Tx_0\|_X \overset{\|x_0\|=1}{=} \underbrace{\|Tx_0\|_X}_{=:c}\underbrace{\|\lambda x_0\|_E}_{=\|x\|}\,.$$

Now assume that the statement holds for $\dim E = n$ and let $\dim E = n + 1$. Choose an n-dimensional subspace $F \subset E$. Let $x_0 \in E \setminus F$. Then for each $x \in E$ there are $\lambda \in \mathbb{K}$, $y \in F$, such that

$$x = \lambda x_0 + y\,,$$

i.e. $E = F \dotplus \operatorname{span}\{x_0\}$. By assumption, we find $c' > 0$, such that $\|Ty\|_X \le c'\|y\|_E$ for all $y \in F$. Now, by Corollary 2.6, F is closed and with Lemma 2.7 we obtain:

$$\|T(\lambda x_0 + y)\|_X \le |\lambda|\|Tx_0\|_X + c'\|y\|_E \le (\|Tx_0\|_X M + c'M')\|\lambda x_0 + y\|\,.$$

The lemma is proven. □

We arrive at the following important corollary

Corollary 2.9. *The following statements hold true:*

(a) Each two norms on a finite-dimensional space are equivalent.

(b) Let E, F be normed spaces and $\dim E < \infty$. Then for every linear and bijective mapping $T\colon E \to F$ there exist $c, d > 0$, such that

$$\|Tx\| \le c\|x\| \quad \text{and} \quad \|T^{-1}y\| \le d\|y\|\,.$$

(c) Let E be a normed space with $\dim E < \infty$ and $A \subset E$ a subset. Then A is compact if and only if A is bounded and closed.

Proof. (a). Consider the linear map

$$T\colon (E, \|\cdot\|_1) \to (E, \|\cdot\|_2)\,,\ Tx := x\,,$$

$x \in E$, where $\|\cdot\|_i$ are norms on E for $i = 1, 2$. Then Lemma 2.8 implies

$$\|Tx\|_1 \le c\|x\|_2 \quad \text{and} \quad \|Tx\|_2 \le d\|x\|_1$$

for some $c, d > 0$. Since $Tx = x$, this proves the claim.
(b). The inverse operator T^{-1} is linear, thus Lemma 2.8 implies the claim.
(c). It is well known that the statement holds for $\mathbb{K}^n$ with the Euclidean norm. Let $\dim E = n$. Then there exists a bijective linear map $T\colon \mathbb{K}^n \to E$. By (b) there exist $c, d > 0$ with $\|Tx\| \le c\|x\|$ and $\|T^{-1}y\| \le d\|y\|$. This implies that A is bounded/closed if and only if $T^{-1}(A)$ is bounded/closed in $\mathbb{K}^n$ and A is compactif and only if $T^{-1}A$ is compact, which finishes the proof. □

Finally we show that finite-dimensional normed spaces are the only locally compact ones.

Theorem 2.10. *For a normed space E, the following are equivalent.*

(i) We have $\dim E < \infty$.

(ii) The space E is locally compact, i.e. each $x \in E$ has a compact neighborhood.

(iii) The closed ball $K_1(0)$ is compact.

Proof. (i)⇒(ii). This follows from Corollary 2.9(c).
(ii)⇒(iii). Let K be a compact neighborhood of 0. Then there exists $\delta > 0$ with $K_\delta(0) \subset K$. But $K_\delta(0)$ is a closed subset of K, thus compact. Since the mapping $x \mapsto \frac{x}{q}$ maps compact sets to compact sets, (iii) follows.
(iii)⇒(i). The compactness of $K_1(0)$ implies that $K_1(0)$ is totally bounded. Hence there exists a finite $\frac{1}{2}$-net $y_1, \dots, y_n$ of $K_1(0)$. Let $F := \operatorname{span}(y_1, \dots, y_n)$. If we prove $E = F$, we are done. Assume $F \ne E$. Then there exists $x_0 \in E \backslash F$, i.e. $\operatorname{dist}(x_0, F) = \inf_{y \in F} \|x_0 - y\| > 0$ (since F is finite dimensional, thus closed). Let $y_0 \in F$ such that

$$\|x_0 - y_0\| \le 2 \operatorname{dist}(x_0, F)\,.$$

Define

$$x = \frac{x_0 - y_0}{\|x_0 - y_0\|} \in K_1(0)\,.$$

For all $y \in F$ we then have

$$\|x - y\| = \|x_0 - y_0\|^{-1} \Big\| x_0 - \big(y_0 + \|x_0 - y_0\| y\big) \Big\|\,.$$

Since $y_0 + \|x_0 - y_0\| y \in F$, this implies

$$\Big\| x_0 - (y_0 + \|x_0 - y_0\| y) \Big\| \geq \operatorname{dist}(x_0, F) \,.$$

Hence

$$\|x - y\| \geq \|x_0 - y_0\|^{-1} \operatorname{dist}(x_0, F) \geq \frac{1}{2 \operatorname{dist}(x_0, F)} \operatorname{dist}(x_0, F) = \frac{1}{2} \,.$$

Thus x is not in $\bigcup_{i=1}^n U_{\frac{1}{2}}(y_i)$. This contradiction shows $F = E$. □

Example 2.11.

(1) The metric $d_\infty(f, g) = \sup_{x \in X} |f(x) - g(x)|$ on $B(X)$ is induced by the norm $\|f\|_\infty := \sup_{x \in X} |f(x)|$. Hence $(B(X), \|\cdot\|_\infty)$ is a Banach space.

(2) The spaces ℓ_p and L_p are Banach spaces.

(3) The space $(\mathbb{K}^n, \|\cdot\|_p)$ is a closed subspace of ℓ_p, thus $\mathbb{K}^n$ is a Banach space.

(4) Let E, F be normed spaces over $\mathbb{K}$. Then $E \times F$ is a normed space (the so-called *product space*) if we define the operations

$$\lambda(x, y) := (\lambda x, \lambda y) \,,$$
$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2)$$

and the norm $\|(x, y)\| = \max(\|x\|_E, \|y\|_F)$. If E and F are Banach spaces, so is $E \times F$.

Proof. Let $(x_n, y_n))_{n \in \mathbb{N}}$ be a Cauchy sequence in $E \times F$. Then (x_n) and (y_n) are Cauchy sequences in E and F respectively, thus they converge to, say, $x \in E$, $y \in F$. But then also $(x_n, y_n) \to (x, y)$. □

More generally, let $E_1, \dots, E_r$ be normed spaces over $\mathbb{K}$ with norms $\|\cdot\|_1, \dots, \|\cdot\|_r$. Then $E_1 \times \cdots \times E_r$ is also a normed space with norm

$$\|(x_1, \dots, x_r)\|_\infty = \max\left(\|x_1\|_1, \dots, \|x_r\|_r\right) \quad \text{or}$$
$$\|(x_1, \dots, x_r)\|_p = \left(\sum_{i=1}^r \|x_i\|_i^p\right)^{\frac{1}{p}} \,.$$

If $E_1, \dots, E_r$ are complete, then $E_1 \times \cdots \times E_r$ as well, and vice versa. Furthermore, convergence in $E_1 \times \cdots \times E_r$ is equivalent to convergence of each component $x_i \in E_i$.

3 Linear Operators, Dual Space

The next main objects are (continuous) linear operators between normed spaces. The following characterisation of continuity of linear operators will be of grave importance (it would, however, be false in the mor general setting of topological vector spaces).

Lemma 3.1. *Let E, F be normed spaces over $\mathbb{K}$ and $T\colon E \to F$ be a linear operator (a linear map). Then the following are equivalent:*

(i) *The operator T is continuous on E.*

(ii) *The operator T is continuous in one point $x_0 \in E$.*

(iii) *The operator T is bounded, i.e. $\|Tx\| \leq c\|x\|$ for all $x \in E$ and some $c > 0$.*

Proof. (i)⇒(ii). This is trivial.
(ii)⇒(iii). Let T be continuous in x_0. Then there exists $\delta > 0$ such that for all $x \in E$ with $\|x - x_0\| \leq \delta$ we have $\|Tx - Tx_0\| \leq 1$. Let $y = \delta^{-1} \cdot (x_0 - x)$, then $\left\| \frac{\delta \cdot y}{\|y\|} \right\| \leq \delta$. We obtain

$$\left\| T\left(\frac{\delta \cdot y}{\|y\|}\right) \right\| \leq 1$$

and thus

$$\|Ty\| \leq \frac{\|y\|}{\delta}\,.$$

(iii)⇒(i). Let $\varepsilon > 0$. Then $\|T(x - x_0)\| \leq c\|x - x_0\| < \varepsilon$ if and only if $\|x - x_0\| \leq \frac{\varepsilon}{c}$. This implies $T(K_{\frac{\varepsilon}{c}}(x_0)) \subset K_{\frac{\varepsilon}{c}}(Tx_0)$. Thus, T is continous. □

Now, we turn the set of continuous or, equivalently, bounded operators between two normed spaces into another normed space.

Definition 3.2. Let E, F be normed spaces over $\mathbb{K}$.

(i) We denote the set $L(E, F)$ of bounded linear operators from E to F by

$$L(E, F) := \{T\colon E \to F \;\;:\;\; T \text{ is linear and bounded}\}\,.$$

If $E = F$ we write $L(E)$ instead of $L(E, F)$.

(ii) We define the *operator norm* on $L(E, F)$ by

$$\|T\| := \sup_{\|x\|_E \leq 1} \|Tx\|_F\,.$$

Lemma 3.3. *Let E, F, G be normed spaces over $\mathbb{K}$.*

(i) *The space $(L(E, F), \|\cdot\|)$ is a normed linear space. Also for $T \in L(E, F)$ we have*

$$\|T\| = \sup_{\|x\|=1} \|Tx\| = \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} = \inf\{c \geq 0 : \|Tx\| \leq c\|x\| \textit{ for all } x \in E\}\,.$$

(ii) *If $T \in L(E, F)$ and $S \in L(F, G)$ then $S \circ T \in L(E, G)$ and $\|S \circ T\| \leq \|S\|\|T\|$.*

Proof. (i). It is well-known that the space of linear maps from E to F is a linear space. Now, for $T_1, T_2, T \in L(E,F)$, $\lambda \in \mathbb{K}$, and arbitrary $x \in E$ with $\|x\| = 1$ we have

$$\|(T_1 + T_2)x\| \leq \|T_1 x\| + \|T_2 x\| \leq \|T_1\| + \|T_2\|$$

and

$$\|\lambda T x\| = |\lambda| \|Tx\| \leq |\lambda| \|T\| .$$

Because of x being arbitrary, we have proven that $(T_1 + T_2), \lambda T \in L(E,F)$, thus $L(E,F)$ is a linear space. It is also clear that $\|T\| = 0$ implies $T = 0$. Thus $\|\cdot\|$ is a norm on it. Further,

$$\begin{aligned} \sup_{\|x\| \leq 1} \|Tx\| \geq \sup_{\|x\|=1} \|Tx\| &= \sup_{y \neq 0} \left\| T\left(\frac{y}{\|y\|}\right) \right\| = \sup_{y \neq 0} \frac{\|Ty\|}{\|y\|} \\ &\geq \sup_{\substack{y \neq 0 \\ \|y\| \leq 1}} \frac{\|Ty\|}{\|y\|} \geq \sup_{\substack{y \neq 0 \\ \|y\| \leq 1}} \|Ty\| \geq \sup_{\|y\| \leq 1} \|Ty\| . \end{aligned}$$

This chain proves

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

Finally, if $\|Tx\| \leq c\|x\|$ for all $x \in E$ then for all $x \in E$ we have $\frac{\|Tx\|}{\|x\|} \leq c$ and thus $\|T\| \leq c$. By this, $\|T\|$ is a lower bound for $\{c \geq 0 : \|Tx\| \leq c\|x\| \text{ for all } x \in E\}$. On the other hand, we have for each $x \neq 0$

$$\|Tx\| = \frac{\|Tx\|}{\|x\|} \|x\| \leq \|T\| \|x\| .$$

Thus $\|T\| \in \{c : \|Tx\| \leq c\|x\| \text{ for all } x \in E\}$ and $\|T\|$ is actually the greatest lower bound for that set, i. e.

$$\|T\| = \inf\{c : \|Tx\| \leq c\|x\| \text{ for all } x \in E\} .$$

(ii). Exercise. □

The case $F = \mathbb{K}$ will be of particular interest.

Definition 3.4. Let E be a normed space over $\mathbb{K}$. Then a map $\ell \colon E \to \mathbb{K}$ is called a *functional* on E. The space $L(E, \mathbb{K})$ of all bounded linear functionals is called the *dual space* of E and is denoted by E^*.

Theorem 3.5. *Let E be a normed space over $\mathbb{K}$ and let F be a Banach space over $\mathbb{K}$. Then $L(E,F)$ is a Banach space. In particular, E^* is a Banach space.*

Proof. Let $(T_n)_n$ be a Cauchy sequence in $L(E,F)$. Then since

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

also $(T_n x)_n$ is a Cauchy sequence in F for each $x \in E$. Since F is complete, we can find $Tx = \lim_{n\to\infty} T_n x$. This defines a mapping $T \colon E \to F$.

We show that T is linear: Let $x, y \in E$ and $\lambda, \mu \in \mathbb{K}$. Then

$$T(\lambda x + \mu y) = \lim_{n\to\infty} T_n(\lambda x + \mu y) = \lambda \cdot \lim_{n\to\infty} T_n x + \mu \cdot \lim_{n\to\infty} T_n y = \lambda \cdot Tx + \mu \cdot Ty .$$

We show that T is bounded: Let $\varepsilon > 0$, $N_\varepsilon \in \mathbb{N}$ such that $\|T_n - T_m\| < \varepsilon$ for all $m, n \geq N_\varepsilon$. This implies

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

for all $m, n \geq N_\varepsilon$ and $x \in E$. Letting $n \to \infty$ we obtain $\|Tx - T_m x\| \leq \varepsilon \|x\|$ for all $m \geq N_\varepsilon$. Hence $\|Tx - T_{N_\varepsilon} x\| \leq \varepsilon \|x\|$. This implies

$$\|Tx\| \leq \varepsilon \|x\| + \|T_{N_\varepsilon} x\| \leq (\varepsilon + \|T_{N_\varepsilon}\|) \cdot \|x\|$$

for all $x \in E$. Thus $T \in L(E, F)$.

We show $T_n \to T$: With $\|Tx - T_m x\| \leq \varepsilon \|x\|$ for all $m \geq N_\varepsilon$ we obtain $\|T - T_m\| \leq \varepsilon$. Hence $\lim_{m \to \infty} T_m = T$. □

Now we show that a continuous linear operator, defined on a subspace can be continuously extended to the closure of this subspace leaving its norm fixed.

Lemma 3.6. *Let E be a normed space over $\mathbb{K}$, $L \subset E$ a subspace, F a Banach space and $T\colon L \to F$ a continous linear operator. Then there exists a unique $S \in L(\overline{L}, F)$ with $S|_L = T$. We have*

$$\|S\| = \|T\| \,.$$

Proof. Let $x \in \overline{L}$ and let $(x_n) \subset L$ be a sequence which converges to x. We observe

$$\|Tx_n - Tx_m\| \leq \|T\| \|x_n - x_m\| \,.$$

This implies that (Tx_n) is a Cauchy sequence in F. The space F is complete – hence (Tx_n) converges. For any other sequence (y_n) which also converges to x we have

$$\|Tx_n - Ty_n\| \leq \|T\| \|x_n - y_n\| \,.$$

Thus $\lim_{n \to \infty} Tx_n = \lim_{n \to \infty} Ty_n$. We can now define

$$Sx = \lim_{n \to \infty} Tx_n$$

and have no concerns about well-defining issues. It follows immediately that $S|_L = T$, because for $x \in L$ we can just choose the constant sequence $x_n := x$ for all n as "defining sequence". The linearity of S is also easily proven:

$$S(x + y) = \lim_{n \to \infty} T(x_n + y_n) = \lim n \to \infty T x_n + T y_n = Sx + Sy$$

and analogously $S(\lambda x) = \lambda S x$. Moreover, $\|S\| \geq \|S|_L\| = \|T\|$. To prove the reverse inequality, let $0 \neq x \in \overline{L}$ and let (x_n) be a sequence in L converging to x. Then

$$\|Sx\| = \lim_{n \to \infty} \|Tx_n\| \leq \|T\| \lim_{n \to \infty} \|x_n\| = \|T\| \|x\| \,,$$

which implies that also $\|S\| \leq \|T\|$. Hence, $\|S\| = \|T\|$, and S is bounded. It only remains to prove the uniqueness of S. Consider another continous linear operator R with $R|_L = T$, and $x \in \overline{L}$. For any sequence converging to x we have

$$Rx = \lim_{n \to \infty} Rx_n = \lim_{n \to \infty} Tx_n = Sx \,.$$

We used the continuity of R. □

We draw a corollary. It follows from the uniqueness part of the lemma.

Korollar 3.7. *If two bounded linear operators $S, T \in L(E, F)$, where F is a Banach space and E a normed space, coincide on a dense subspace of E, then they coincide on E.*

Now we consider the inverse of a continuous linear operator.

Lemma 3.8. *Let E, F be normed spaces over $\mathbb{K}$ and $T\colon E \to F$ linear. Then the following are equivalent:*

(i) *There exists a linear, continous inverse operator*

$$T^{-1}\colon T(E) \to E\,.$$

(ii) *There exists $c > 0$ such that $c\|x\| \leq \|Tx\|$.*

Proof. (i)⇒(ii). Assume T^{-1} exists. Its continuity gives us the existence of a $\gamma > 0$ such that

$$\|T^{-1}y\| \leq \gamma\|y\|$$

for all $y \in T(E)$. For an arbitrary $x \in E$ we put $y = Tx$ to obtain

$$\|x\| \leq \gamma\|Tx\|\,.$$

Putting $c := \frac{1}{\gamma}$ we have proven (ii).
(ii)⇒(i). We observe that (ii) secures the injectivity of T (if $x \in \ker(T)$, then $\|x\| = 0$). Thus $T^{-1}\colon \operatorname{ran} E \to E$ exists. Now letting $y = Tx$ in (ii) assures the existence of a $c > 0$ with $c\|T^{-1}y\| \leq \|y\|$ for all $y \in T(E)$. Thus the inverse operator is continous. □

Note that, unlike in the finite-dimensional case, this inverse does not have to be bounded automatically.

Example 3.9. Consider $E = C[0, 1]$ with $\|\cdot\|_\infty$-norm and let

$$T\colon E \to F\,,\ (Tf)(t) = \int_0^t f(s)\,\mathrm{d}s\,,$$

with $F = \{g \in C^1[0, 1] : g(0) = 0\}$ having $\|\cdot\|_\infty$ as a norm.

- The operator T is linear.
- The operator T is bounded, since $\|Tf\|_\infty \leq \|f\|_\infty$.
- The operator T is injective, since $Tf = 0$ implies $f = 0$.
- The operator T is surjective, since for $g \in F$ we have $T(g') = g$.

But: T^{-1} is not continuous! For this, choose $f_n(t) = t^n$. Then $(Tf_n)(t) = \frac{1}{n+1}t^{n+1}$, but

$$\|T^{-1}(Tf_n)\|_\infty = \|f_n\|_\infty = 1$$

and $\|Tf_n\|_\infty = \frac{1}{n+1}$.

Next we define closed operators.

Definition 3.10 (Graph of T)**.** Let E, F be normed spaces, $L \subset E$ a subspace and $T\colon L \to F$ a linear operator.

(i) We define the *graph of* T by

$$G_T = \{(x, Tx) : x \in L\} \subset L \times F \subset E \times F\,.$$

(ii) The operator T is called *closed* if its graph G_T is closed in $E \times F$.

Remark 3.11. Note that we require the graph to be closed in $E \times F$, not in $L \times F$. For example, let $L \subset E$ not be closed and consider the identity $\mathrm{Id}\colon L \to E$. Its graph is closed as a subset of $L \times E$ but not as a subset of $E \times F$, so this operator is not closed. In particular, closedness of $T\colon L \to F$ also depends on the superspace E of L.

Luckily, we have quite an easy method to check whether a given operator is closed.

Lemma 3.12. *Let E, F, T and L be as above. Then the following are equivalent:*

(i) *The operator T is closed.*

(ii) *If $(x_n) \subset L$ converges to $x \in E$ and (Tx_n) converges to $y \in F$, then $x \in L$ and $y = Tx$.*

Proof. Since

$$\|(x_n, Tx_n) - (x, y)\| = \max(\|x_n - x\|, \|Tx_n - y\|)$$

we have that if $x_n \to x$ and $Tx_n \to y$, then

$$\lim_{n\to\infty} (x_n, Tx_n) = (x, y)\,.$$

Because of G_T being closed, $(x, y) \in G_T$. Thus $x \in L$, $y = Tx$. Now consider a convergent sequence $(x_n, y_n) \to (x, y)$ in G_T. Because of the convergence of $y_n = Tx_n$ and (ii), it follows that $x \in L$ and $y = Tx$, thus $(x, y) \in G_T$. □

Remark 3.13. If L is closed, and T is continous, then T is closed. In particular, each $T \in L(E, F)$ is closed.

Proof. If $(x_n, Tx_n) \to (x, y)$, then (L closed) $x \in L$. Continuity of T now implies $Tx_n \to Tx$, thus $(x, y) \in G_T$. □

We will now give a concrete computation of a dual space.

Theorem 3.14. *Let $1 \le p < \infty$. Define q such that $\frac{1}{p} + \frac{1}{q} = 1$, i.e.*

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

Moreover, for $y \in \ell_q$ define

$$f_y\colon \ell_p \to \mathbb{K}\,, \quad x = (x_n) \mapsto \sum_{n=1}^{\infty} x_n y_n\,.$$

Then $f_y \in \ell_p^*$ and $y \mapsto f_y$ is an isometric isomorphism. In particular $\ell_q \cong \ell_p^*$.

Proof. First of all, f_y is well-defined (i.e. the series converges) because of the Hölder-inequality:

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \|x\|_p \|y\|_q \,.$$

It is evident that f_y is linear. Furthermore, by the above, we have

$$\|f_y(x)\| \leq \|x\|_p \|y\|_q \,.$$

Thus, f_y is bounded. We conclude that $f_y \in \ell_p^*$. We now claim that $\|f_y\| = \|y\|_q$. We already proved $\|f_y\| \leq \|y\|_q$. To prove the inverse inequality, we consider the cases $p = 1$, $p > 1$ separately.

First let $p = 1$: Let $\varepsilon > 0$. There exists an $n \in \mathbb{N}$ such that

$$|y_n| \geq \|y\|_q - \varepsilon \,.$$

Set $x = e_n \in \ell_1$. We obtain $f_y(x) = y_n$. Because of $\|x\| = 1$ it follows that $\|f_y\| \geq \|y\|_\infty - \varepsilon$. Because of ε being arbitrary, we conclude $\|f_y\| \geq \|y\|_q$.

Now let $p > 1$: Define x by

$$x_n = \begin{cases} 0 & y_n = 0, \\ \frac{|y_n|^q}{y_n} & \text{otherwise} \,. \end{cases}$$

Then

$$\sum_{n=1}^{\infty} |x_n|^p = \sum_{n=1}^{\infty} |y_n|^{p(q-1)} = \sum_{n=1}^{\infty} |y_n|^q < \infty \,,$$

thus $x \in \ell_p$.

We now compute $f_y(x)$:

$$f_y(x) = \sum_{n=1}^{\infty} x_n y_n = \sum_{n=1}^{\infty} |y_n|^q = \|y\|_q^q \,.$$

Thus $\frac{|f_y(x)|}{\|x\|} = \|y\|^{\frac{q(p-1)}{p}} = \|y\|_q$, and we conclude $\|f_y\| \geq \|y\|_q$.

This proves that $y \mapsto f_y$ is isometric. For the surjectivity let $f \in \ell_p^*$ and put

$$y_n := f(e_n) \,.$$

To prove $y = (y_n) \in \ell_q$, we again treat the two cases $p = 1$, $p > 1$ separately .

Let $p = 1$. We have for all n

$$|y_n| = |f(e_n)| \leq \|f\| \|e_n\| = \|f\|.$$

Thus, $\|y\|_\infty \leq \|f\|$, $y \in \ell_\infty$.

Let $p > 1$. For all $m \in \mathbb{N}$ there holds

$$\sum_{n=1}^{m} |y_n|^q = \sum_{\substack{n=1 \\ y_n \neq 0}}^{m} \frac{|y_n|^q}{y_n} f(e_n) = f\left(\sum_{\substack{n=1 \\ y_n \neq 0}}^{m} \frac{|y_n|^q}{y_n} e_n \right) \leq \|f\| \left\| \sum_{\substack{n=1 \\ y_n \neq 0}}^{m} \frac{|y_n|^q}{y_n} e_n \right\|_p \,.$$

We have

$$\left\| \sum_{\substack{n=1 \\ y_n \neq 0}}^{m} \frac{|y_n|^q}{y_n} e_n \right\|_p = \left(\sum_{\substack{n=1 \\ y_n \neq 0}}^{n} |y_n|^{p(q-1)} \right)^{\frac{1}{p}} \,.$$

And thus

$$\sum_{n=1}^{m} |y_n|^q \leq \|f\| \left(\sum_{n=1}^{m} |y_n|^q \right)^{\frac{1}{p}},$$

which implies

$$\left(\sum_{n=1}^{m} |y_n|^q \right)^{\frac{1}{q}} \leq \|f\| .$$

Letting $m \to \infty$, we get $\|y\|_q < \infty$, implying $y \in \ell_q$.
Finally, $f = f_y$ because for $x = \sum_{n=1}^{m} x_n e_n$, there holds

$$f(x) = \sum_{n=1}^{m} x_n f(e_n) = \sum_{n=1}^{m} x_n y_n = f_y(x) .$$

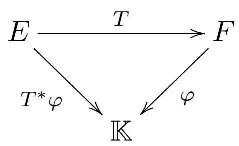
Thus f coincides with f_n on the dense linear subspace $\operatorname{span}(e_n)_{n \in \mathbb{N}}$, hence (Lemma 3.6, $\mathbb{K}$ is a Banach space) they are equal. □

Finally, we will introduce the important concept of the dual operator.

Lemma 3.15. *Let E, F be normed spaces, and let $T \in L(E, F)$. Then the operator $T^* \colon F^* \to E^*$ defined by*

$$T^* \varphi = \varphi \circ T, \quad \varphi \in F^* ,$$

satisfies $T^* \in L(F^*, E^*)$.



Proof. Obviously, T^* is linear. Further

$$\|(T^* \varphi) x\| = \|\varphi(Tx)\| \leq \|\varphi\| \|T\| \|x\| .$$

This proves that T^* is bounded and $\|T^*\| \leq \|T\|$. □

Definition 3.16. Let E, F be normed spaces and $T \in L(E, F)$. Then the operator $T^* \colon F^* \to E^*$, $\varphi \mapsto \varphi \circ T =: T^* \varphi$, is called the *dual operator* of T.

4 Hahn-Banach Theorem and Corollaries

In this chapter we prove one of the most important theorems of this course; the Hahn-Banach theorem. This will ensure us that we can extend continuous linear functionals from subspaces to the entire space, actually guaranteeing the existence of nontrivial continuous linear functionals on normed spaces. This simple statement will lead to a few important corollaries and several forms of separation theorems.

First, we need some terminology.

Definition 4.1.

a) Let E be a linear space. Then the *algebraic dual space* of E is the space of linear maps $E \to \mathbb{K}$. It is denoted E'.

b) Let E be an $\mathbb{R}$-vector space. Then $\rho\colon E \to \mathbb{R}$ is a *sublinear functional* on E if all $x, y \in E$ and $\lambda > 0$ satisfy $\rho(x+y) \le \rho(x) + \rho(y)$ and $\rho(\lambda x) = \lambda \rho(x)$.

On our way to the Hahn-Banach theorem, we need some lemmas. As we mentioned, our goal is to extend a continuous linear functional from a subspace of a normed space to the entire space, keeping some kind of bound. This will be done similarly to induction: First we prove that we can extend functionals to a space with "one dimension more".

Lemma 4.2. *Let E be an $\mathbb{R}$-vector space, F a linear subspace and $x_0 \in E \setminus F$. Let further L be the space generated by x_0 and F, i.e. $L = F + \mathbb{R}x_0$, $f \in F'$ and ρ be a sublinear functional on E such that*

$$f(x) \le \rho(x) \quad \text{for all } x \in F\,.$$

Then there exists an $\ell \in L'$ such that $\ell|_F = f$ and $\ell(x) \le \rho(x)$ for all $x \in L$.

Proof. Note that each element $y \in L$ has a unique representation $y = x + \lambda x_0$, $x \in F$, $\lambda \in \mathbb{R}$. This follows from the fact that x_0 is not an element of F. To define $\ell \in L'$, observe that

$$\ell(y) = \ell(x) + \lambda \ell(x_0) = f(x) + \lambda \ell(x_0)\,.$$

Thus, it is sufficient to choose $\ell(x_0)$. In other words, it suffices to show the existence of a $\gamma \in \mathbb{R}$ with

$$f(x) + \lambda\gamma \le \rho(x + \lambda x_0) \tag{4.1}$$

for all $x \in F$, $\lambda \in \mathbb{R}$. First, we have for $x, y \in F$

$$\begin{aligned} f(x) + f(y) = f(x+y) \le \rho(x+y) &= \rho\big((x+x_0) + (y-x_0)\big) \\ &\le \rho(x+x_0) + \rho(y-x_0)\,. \end{aligned}$$

This implies

$$f(y) - \rho(y - x_0) \le \rho(x + x_0) - f(x) \tag{4.2}$$

for all $x, y \in F$. Next, define

$$\begin{aligned} A &:= \{f(x) - \rho(x - x_0) : x \in F\} \quad \text{and} \\ B &:= \{\rho(x + x_0) - f(x) : x \in F\}\,. \end{aligned}$$

By (4.2), we get $\sup A \le \inf B$. Choose $\gamma \in [\sup A, \inf B]$. Then we have

$$f(x) - \gamma \le \rho(x - x_0) \quad \text{and} \quad f(x) + \gamma \le \rho(x + x_0)$$

for all $x \in F$. Now for $\lambda > 0$ we have

$$f(x) - \lambda\gamma = \lambda\left(f(\lambda^{-1}x) - \gamma\right) \leq \lambda\rho(\lambda^{-1}x - x_0) = \rho(x - \lambda x_0)\,,$$

and

$$f(x) + \lambda\gamma = \lambda\left(f(\lambda^{-1}x) + \gamma\right) \leq \lambda\rho(\lambda^{-1}x + x_0) = \rho(x + \lambda x_0)\,.$$

Thus, (4.1) is proven. □

Now, will use Zorn's Lemma for some kind of "transfinite induction". We will show the existence of a maximal extension. To show that such a maximal extension is defined everywhere, we apply the previous lemma, which says: If it were not, we could add yet another dimension.

Lemma 4.3. *Let E be an $\mathbb{R}$-vector space, $F \subseteq E$ a linear subspace and ρ a sublinear functional on E. Further let $f \in F'$ with $f(x) \leq \rho(x)$ for all $x \in F$. Then there exists $\ell \in E'$ with $\ell|_F = f$ and $\ell(x) \leq \rho(x)$ for all $x \in E$.*

Proof. Set

$$\mathcal{L} = \left\{(L,\ell) : \begin{array}{c} L \text{ linear subspace of } E \text{ with } L \supset F \text{ and} \\ \ell \in L' \text{ with } \ell|_F = f \text{ and } \ell(x) \leq \rho(x) \text{ for all } x \in L\,. \end{array}\right\}\,.$$

This is the space of all extensions of f satisfying the desired inequality. We have to prove the existence of a pair of the form $(E,\ell) \in \mathcal{L}$. As mentioned above, we will show that there exists an extension which cannot be extended any further, a maximal extension. To do this, we define a partial ordering on $\mathcal{L}$:

$$(L_1,\ell_1) \leq (L_2,\ell_2) :\Leftrightarrow L_1 \subset L_2 \ \wedge\ \ell_2|_{L_1} = \ell_1\,.$$

We know that $\mathcal{L} \neq \varnothing$ since $(F,f) \in \mathcal{L}$.

Claim: Let $\mathscr{K}$ be a chain in $\mathcal{L}$. Then $\mathscr{K}$ has an upper bound in $\mathcal{L}$.

This is proved with a standard argument for Zorn's Lemma. We take all our domains of definition, unite them and use this as the domain of definition for our upper bound. So we set

$$\widetilde{L} = \bigcup\{L : \text{ there is } \ell \in L' \text{ with } (L,\ell) \in \mathscr{K}\}$$

and let $\tilde{\ell}\colon \widetilde{L} \to \mathbb{R}$ be defined by

$$\tilde{\ell}(x) := \ell(x) \quad \text{if } x \in L \text{ and } \ell \in L' \text{ with } (L,\ell) \in \mathscr{K}\,.$$

First, $\widetilde{L}$ is a linear subspace since $\mathscr{K}$ is linearly ordered: if $x, y \in \widetilde{L}$, there exist L_x, L_y with $x \in L_x$, $y \in L_y$ such that there exist ℓ_x, ℓ_y so that $(L_x,\ell_x), (L_y,\ell_y) \in \mathscr{K}$. Now because of the fact that $\mathscr{K}$ is a chain we can without loss of generality assume that $L_x \subset L_y$. But then $x + y \in L_y$, hence $x + y \in \widetilde{L}$. Checking $\lambda x \in L$ is trivial.

We show that $\tilde{\ell}$ is well-defined: Let $x \in \widetilde{L}$. We know by the definition of $\widetilde{L}$ that an $(L_1,\ell_1) \in \mathscr{K}$ such that $x \in L_1$ and $\ell_1 \in L_1'$ always exists. If $(L_2,\ell_2) \in \mathscr{K}$ is another pair with $x \in L_2$, then one of the pairs is bigger with respect to the ordering ($\mathscr{K}$ is a chain). Without loss of generality, $(L_1,\ell_1) \leq (L_2,\ell_2)$. Then, because of $x \in L_1 \subset L_2$ we have $\ell_2(x) = \ell_1(x)$.

Now we show $\tilde{\ell} \in \tilde{L}'$: Let $x_i \in L_i$, $\alpha_i \in \mathbb{R}$, $(L_i, \ell_i) \in \mathscr{K}$, $i = 1, 2$. Without loss of generality, $(L_1, \ell_1) \leq (L_2, \ell_2)$. Then $x_1 \in L_2$ and thus $\alpha_1 x_1 + \alpha_2 x_2 \in L_2$. Hence, we have

$$\begin{aligned}\tilde{\ell}(\alpha_1 x_1 + \alpha_2 x_2) = \ell_2(\alpha_2 x_1 + \alpha_2 x_2) &= \alpha_1 \underbrace{\ell_2(x_1)}_{=\ell_1(x_1)} + \alpha_2 \ell_2(x_2) \\ &= \alpha_1 \tilde{\ell}(x_1) + \alpha_2 \tilde{\ell}(x_2).\end{aligned}$$

Finally, $\tilde{\ell}(x) = \ell(x) \leq \rho(x)$ for some ℓ for every x.

Zorn's Lemma now provides the existence of a maximal element $(L, \ell) \in \mathcal{L}$ of $\mathcal{L}$. We have to prove that $L = E$. Suppose the opposite, then there exists $x_0 \in E \setminus L$. By Lemma 4.2, there exists $g \in (L + \mathbb{R}x_0)'$ with $g|_L = \ell$ and $g(x) \leq \rho(x)$ for all $x \in (L + \mathbb{R}x_0)$. But then we have $(L, \ell) < (L + \mathbb{R}x_0, g)$. This contradiction shows $L = E$, which proves the lemma. □

The following is also called the analytic version of the Hahn-Banach theorem.

Theorem 4.4. *Let E be a vector space over $\mathbb{K}$, F a linear subspace and $f \in F'$. Let $\rho\colon E \to \mathbb{R}$ be a seminorm on E, i. e. for all $x, y \in E$, $\lambda \in \mathbb{K}$*

$$\begin{aligned}\rho(x+y) &\leq \rho(x) + \rho(y) \quad \text{and} \\ \rho(\lambda x) &= |\lambda| \rho(x).\end{aligned}$$

Suppose that $|f(x)| \leq \rho(x)$ for all $x \in F$. Then there exists an $\ell \in E'$ with $\ell|_F = f$ and $|\ell(x)| \leq \rho(x)$ for all $x \in E$.

Proof. First consider $\mathbb{K} = \mathbb{R}$. Then $f(x) \leq \rho(x)$ for all $x \in F$ and $\rho(\alpha x) = \alpha \rho(x)$ for all $x \in E$, $\alpha \geq 0$. By Lemma 4.3, there exists some $\ell \in E'$ with

$$\ell|_F = f \text{ and } \ell(x) \leq \rho(x)$$

for all $x \in E$. Since also

$$-\ell(x) = \ell(-x) \leq \rho(-x) = \rho(x)\,,$$

we have $|\ell(x)| \leq \rho(x)$. The case $\mathbb{K} = \mathbb{C}$ will be discussed in the exercises. □

Finally we arrive at what the Hahn-Banach Theorem for normed spaces. It says essentially that we may extend continuous linear functionals from (arbitrary) subspaces to all of our space without enlarging the norm.

Theorem 4.5 (Hahn-Banach Theorem)**.** *Let E be a normed space, F a linear subspace of E. Then for each $f \in F^*$ there exists some $\ell \in E^*$ with*

$$\ell|_F = f \quad \text{and} \quad \|\ell\| = \|f\|\,.$$

Proof. Let ρ be defined by

$$\rho(x) = \|f\| \|x\|\,.$$

Then ρ is a seminorm - the properties are inherited from the norm properties of $\|\cdot\|$. Furthermore $|f(x)| \leq \rho(x)$ for all $x \in F$. By Theorem 4.4, there exists $\ell \in E'$ with

$$\ell|_F = f \quad \text{and} \quad |\ell(x)| \leq \rho(x) = \|f\| \|x\|\,.$$

This proves in particular that $\ell \in E^*$ and $\|\ell\| \leq \|f\|$. Because of $\ell|_F = f$, the reverse inequality holds. Thus $\|\ell\| = \|f\|$. □

Remark 4.6. If G is another normed space, we remark that it is in general not possible to extend an operator $T \in L(F,G)$ to an operator $S \in L(E,G)$ having the same norm. One can show, however, that this can be done if $G = \ell_\infty$ (by applying the Hahn-Banach Therem to each component). In the case $G = c_0$, by a little more effort, it can be shown that we can extend $S \in L(F, c_0)$ to $T \in L(E, c_0)$, where at least $\|S\| \leq 2\|T\mathbb{R}$. On the other hand, let E be a Banach space, $F \subset E$ a non-closed subspace and $T \in L(F,F)$ the identity. Then there is no $S \in L(E,F)$ extending T.

Now we turn to some important corollaries.

Korollar 4.7. *Let E be a normed space and F a linear subspace of E and $x \in E$ such that*

$$\delta := \inf_{y \in F} \|x - y\| > 0\,.$$

Then there exists an $\ell \in E^*$ with

$$\ell|_F = 0\,,\ \|\ell\| = 1 \quad \textit{and} \quad \ell(x) = \delta\,.$$

In particular, for any $x \neq 0$ there exists an $\ell \in E^*$ with $\|\ell\| = 1$ and $\ell(x) = \|x\|$.

Proof. Let $G = F + \mathbb{K}x$ and $g\colon G \to \mathbb{K}$ be defined by $g(y + \lambda x) = \lambda\delta$ for $y \in F$, $\lambda \in \mathbb{K}$. g is well defined, since $x \notin F$ implies $G = F \dotplus \mathbb{K}x$. Further g is linear, $g|_F = 0$ and $g(x) = \delta$. We now claim $\|g\| = 1$. There holds

$$\begin{aligned}|g(y + \lambda x)| &= |\lambda|\delta = |\lambda| \inf_{z \in F} \|z - x\| \\ &= \inf_{z \in F} \|\lambda z - \lambda x\| = \inf_{z \in F} \|z + \lambda x\| \leq \|y + \lambda x\|\,.\end{aligned}$$

Thus $\|g\| \leq 1$. Secondly, for every $\varepsilon > 0$ there exists a $z_\varepsilon \in F$ with $\delta \leq \|z_\varepsilon + x\| \leq \delta + \varepsilon$. It follows that

$$g(x + z_\varepsilon) = \delta \geq \|x + z_\varepsilon\| - \varepsilon$$

and thus

$$g\left(\|x + z_\varepsilon\|^{-1}(x + z_\varepsilon)\right) = \frac{\delta}{\|x + z_\varepsilon\|} \geq 1 - \frac{\varepsilon}{\|x + z_\varepsilon\|} \geq 1 - \frac{\varepsilon}{\delta}\,.$$

Now apply Theorem 4.5 to lift g up to E^*. For the in-particular part, choose $F = \{0\}$. $\square$

As a particular case we obtain a way to represent a norm which will be of grave importance.

Korollar 4.8. *Let E be a normed space. Then for each $x \in E$ we have*

$$\|x\| = \sup\{|\ell(x)| : \ell \in E^*,\ \|\ell\| \leq 1\}\,.$$

Moreover, this supremum is attained.

Proof. Let $x \in E$. Then $S := \sup\{|\ell(x)| : \ell \in E^*\,,\ \|\ell\| \leq 1\} \leq \|x\|$. By Corollary 4.7 there exists some $\ell \in E^*$ with $\|\ell\| = 1$ and $|\ell(x)| = \|x\|$, hence $S = \|x\|$ and the supremum is attained. $\square$

We now turn to a geometric separation problem (the solution of which will lay on the Hahn-Banach Theorem): Given two subsets U, V of a normed space E, under which conditions is it possible to separate them by a closed hyperplane (the kernel of a continuous linear functional). More explicit, we ask whether it is possible to find some $f \in E^*$ such that

$$\sup_{x \in U} \operatorname{Re} f(x) \leq \inf_{x \in V} \operatorname{Re} f(x)\,, \tag{4.3}$$

or – in the real case – just $\sup_U f(x) \leq \inf_V f(x)$. Thinking of subsets of two- or three-dimensional space and separating by a line or a plane, it is intuitive to restrict on convex subsets U, V (also, they should not overlap "too much"). Thus, we adress the following question: Let E be a normed space and $U, V \in E$ convex subsets. Under which additional assumptions does there exist some $f \in E^*$ satisfying (4.3)?

To investigate this we need the help of Minkowski functionals.

Definition 4.9. Let E be an arbitrary vector space over $\mathbb{K}$ and let $C \subset E$ be some subset. Then the *Minkowski functional* $\rho_C \colon E \to [0, \infty]$ is defined by

$$\rho_C(x) := \inf\{\alpha > 0 : x \in \alpha C\}\,.$$

The set C is called *absorbing* if $\rho_C(x) < \infty$ for all $x \in E$.

Lemma 4.10. *Let E be a normed space and $C \subset E$ be an open, convex subset containing 0. Then*

(i) *The Minkowski functional ρ_C is sublinear.*

(ii) *The set C is absorbing. In fact, there exists $M \geq 0$ such that $\rho_C(x) \leq M\|x\|$ for all $x \in E$.*

(iii) *The set C can be described as*

$$C = \{x \in E : \rho_C(x) < 1\}\,.$$

Proof. Homework. □

We are ready to state and prove our first separation theorem.

Theorem 4.11 (Hahn-Banach separation theorem)**.** *Let E be a normed space and $U, V \subset E$ disjoint convex subsets. If further U is open, then there exists $f \in E^*$ such that*

$$\operatorname{Re} f(u) < \operatorname{Re} f(v)$$

for all $u \in U$ and $v \in V$.

Proof. Step 1: We assume that V consists of a single point, i. e. $V = \{x_0\}$ for some $x_0 \in E$. Further, we restrict to the case $\mathbb{K} = \mathbb{R}$ – the proof for the complex case is done similarly as you did for theorem 4.4 in the exercises.

We have to prove the existence of some $f \in E^*$ sich that $f(x) < f(x_0)$ for all $x \in U$. W. l. o. g. assume $0 \in U$. Otherwise fix $u_0 \in U$ and translate by u_0: the shifted spaces $U - u_0$ and $V - u_0$ still satisfy all of our assumptions and $f(x) < f(x_0)$ for all $x \in U$ is equivalent to $f(x - u_0) < f(x_0 - u_0)$ for all $x \in U$ by linearity of f. Now that we justified the assumption $0 \in U$ let ρ_U be the Minkowski functional of U. We shall prove

$$\rho_U(x_0) = \inf\{\alpha > 0 : x_0 \in \alpha U\} \geq 1\,. \tag{4.4}$$

Clearly, $x_0 \notin 1 \cdot U = U$ as $U \cap V = \emptyset$. Thus for $\alpha \in (0, 1)$ we have $x_0 \notin \alpha U$ since $\alpha U \subset \alpha U + (1 - \alpha)U \subset 1 \cdot U$ by convexity of U. This shows (4.4).

Now let $F := \operatorname{span}\{x_0\}$ and define $\varphi \colon F \to \mathbb{R}$, $\lambda x_0 \mapsto \lambda$. Then φ is obviously linear and φ is dominated by ρ_U, i. e. $\varphi(y) \leq \rho_U(y)$ for all $y \in F$. To see this, note first that for $\lambda \geq 0$ we have

$$\varphi(\lambda x_0) = \lambda \leq \lambda \rho_U(x_0) = \rho_U(\lambda x_0)$$

by (4.4) while for $\lambda < 0$ there holds

$$\varphi(\lambda x_0) = \lambda < 0 \leq \rho_U(\lambda x_0)\,.$$

Now, lemma 4.3 ensures the existence of $f \in E'$ such that $f|_F = \varphi$ and still $f(x) \leq \rho_U(x)$ for all $x \in E$. In particular, $f(x_0) = \varphi(x_0) = 1$. By lemma 4.10 there exists $M \geq 0$ such that $\rho_U(x) \leq M\|x\|$ for all $x \in E$, implying continuity of f by

$$\pm f(x) = f(\pm x) \leq \rho_U(\pm x) \leq M\|x\|\,,$$

thus $f \in E^*$. The claim now follows from lemma 4.10 as

$$f(x) \leq \rho_U(x) < 1 = f(x_0)$$

for all $x \in U$.
Step 2: Now, let V be an arbitrary convex subset disjoint to U. Define

$$W := U - V = \bigcup_{v \in V} (U - v)$$

which is nonempty and convex (since U and V are nonempty and convex) and open as a union of open sets. Applying step 1 to W and $\{0\}$ yields $f \in E^*$ such that $f(u - v) < f(0) = 0$ for all $u \in U$, $v \in V$, i. e. $f(u) < f(v)$ for all $u \in U$, $v \in V$. □

In a second separation theorem we will separate closed sets from compact sets.

Lemma 4.12. *Let (X, d) be a metric space and let $A, B \subset X$ be two nonempty subsets where A is compact and B is closed. Then*

$$\operatorname{dist}(A, B) := \inf\{d(a, b) : a \in A, b \in B\} > 0\,.$$

Proof. Tutorials. □

Theorem 4.13 (Hahn-Banach strict separation theorem)**.** *Let E be a normed space, $U, V \subset E$ nonempty, disjoint and convex subsets. Further, let U be closed and V be compact. Then there exists $f \in E^*$ and $\alpha_1 < \alpha_2 \in \mathbb{R}$ such that*

$$\operatorname{Re} f(u) \leq \alpha_1 < \alpha_2 \leq \operatorname{Re} f(v)$$

for all $u \in U$ and $v \in V$.

Proof. By the previous lemma, $d := \operatorname{dist}(U, V) > 0$. Let $0 < r < d$ and let

$$U_r := U + U_r(0) = \bigcup_{u \in U} (u + U_r(0))$$

which is open and convex (a neighborhood of U. Furthermore, $U_r \cap V = \emptyset$ since for all $u + x \in U_r$, $v \in V$ we have

$$\|(u + x) - v\| \geq \|u - v\| - \|x\| \geq d - r > 0\,.$$

By our first separation theorem 4.11 there exists $f \in E^*$ such that

$$\operatorname{Re} f(u + x) < \operatorname{Re} f(v)$$

for all $u + x \in U_r$, $v \in V$. Now, let $y \in U_1(0)$, hence also $-y \in U_1(0)$ and $\pm\frac{r}{2}y \in U_r(0)$. Therefore,

$$\operatorname{Re} f\left(u \pm \frac{r}{2}y\right) = \operatorname{Re} f(u) \pm \frac{r}{2} \operatorname{Re} f(y) < \operatorname{Re} f(v)$$

for all $u \in U$, $v \in V$ and thus

$$\operatorname{Re} f(u) + \frac{r}{2} \sup_{x \in U_1(0)} |\operatorname{Re} f(x)| \leq \operatorname{Re} f(v) .$$

As we take the supremum of the interior of the unit ball, we may rewrite this as

$$\operatorname{Re} f(u) + \frac{r}{2}\|f\| \leq \operatorname{Re} f(v) .$$

Letting $\varepsilon := \frac{r}{2}\|f\|$, we have

$$\operatorname{Re} f(u) < \operatorname{Re} f(u) + \frac{\varepsilon}{2} < \operatorname{Re} f(u) + \varepsilon \leq \operatorname{Re} f(v) .$$

Now, we have proved the claim by setting

$$\alpha_1 := \sup_{u \in U} \operatorname{Re} f(u) \quad \text{and} \quad \alpha_2 := \inf_{v \in V} \operatorname{Re} f(v) .$$

□

In the rest of this chapter we will use the Hahn-Banach Theorem to study annihilators and reflexivity.

Definition 4.14. Let E be a normed space, $M \subset E$ an arbitrary subset of E and $L \subset E^*$. Then the *annihilator of M in E^** is defined by

$$M^\perp := \{\ell \in E^* : \ell(x) = 0 \text{ for all } x \in M\}$$

and the *annihilator of L in E* is given by

$$L_\perp := \{x \in E : \ell(x) = 0 \text{ for all } \ell \in L\} .$$

Remark 4.15. The annihilators are closed linear subspaces of E^* and E, respectively. This follows from the continuity of $\ell \mapsto \ell(x)$, $x \mapsto \ell(x)$.

Our first result concerns annihilators of annihilators.

Lemma 4.16. *Let E be a normed space and $\varnothing \neq M \subseteq E$. Then $(M^\perp)_\perp$ is the* closed linear hull of M*, i.e. the smallest closed linear subspace of E which contains M.*

Proof. If $x \in M$, then $\ell(x) = 0$ for all $\ell \in M^\perp$, thus $x \in (M^\perp)_\perp$. Hence, $M \subset (M^\perp)_\perp$. Now let F be the closed linear hull of M. By Remark 4.15, $F \subseteq (M^\perp)_\perp$. Assume there exists $x \in (M^\perp)_\perp \setminus F$. Corollary 4.7 secures the existence of an $\ell \in (M^\perp)_\perp^*$ with $\ell|_F = 0$ and $\ell(x) \neq 0$. Theorem 4.5 now implies the existence of an $f \in E^*$ with $f|_{(M^\perp)_\perp} = \ell$. The functional f is in $M^\perp$ because of $f|_F = \ell|_F = 0$ and $M \subseteq F$. But $f(x) \neq 0$. A contradiction! □

Also, we are able now to characterize the dual space of a subspace and of quotient space.

Theorem 4.17. *Let E be a normed space over $\mathbb{K}$, and $F \subset E$ a linear subspace.*

(i) The linear operator

$$\Phi\colon E^*/F^\perp \to F^*\,,\ \Phi(f+F^\perp) = f|_F\,,$$

$f \in E^*$, is an isometric isomorphism.

(ii) If F is closed, then the linear operator

$$\Phi\colon (E/F)^* \to F^\perp\,,\ (\Phi f)(x) = f(x+F)\,,$$

$x \in E$, $f \in (E/F)^*$, is an isometric isomorphism.

Proof. (i). Consider the map $T\colon E^* \to F^*$, $Tf := f|_F$, $f \in E^*$. We have $\ker T = F^\perp$. Hence, Φ is well-defined, linear, and injective. By Theorem 4.5, for each $\ell \in F^*$ there exists some $f \in E^*$ with $f|_F = \ell$. Hence, Φ is surjective. Finally, let $f \in E^*$, and choose $g \in E^*$ such that

$$\|g\| = \|f|_F\| \quad \text{and} \quad g|_F = f|_F\,.$$

Then, we obtain

$$\|f + F^\perp\| \le \|f + (g-f)\| = \|g\| = \|f|_F\| = \|\Phi(f+F^\perp)\|\,.$$

On the other hand, for all $g \in F^\perp$,

$$\|\Phi(f+F^\perp)\| = \|f|_F\| = \|(f+g)|_F\| \le \|f+g\|\,.$$

Hence, $\|\Phi(f+F^\perp)\| \le \|f+F^\perp\|$.

(ii). First, $\Phi f\colon E \to \mathbb{K}$, $x \mapsto f(x+F)$, is linear. Since

$$|(\Phi f)(x)| = |f(x+F)| \le \|f\| \cdot \|x+F\| \le \|f\| \cdot \|x\|\,,$$

we have $\Phi f \in E^*$. If $x \in F$, then $(\Phi f)(x) = f(F) = 0$, hence $\Phi f \in F^\perp$. Therefore, indeed $\Phi\colon (E/F)^* \to F^\perp$. It is obvious that Φ is linear and injective. To prove surjectivity, let $g \in F^\perp$, and let $f\colon E/F \to \mathbb{K}$ be defined by

$$f(x+F) = g(x), \quad x \in E\,.$$

The functional f is is well-defined since $F \subset \ker g$. Moreover, f is linear, and for all $x \in E$, $y \in F$ we have

$$|f(x+F)| = |g(x)| = |g(x+y)| \le \|g\| \cdot \|x+y\|\,,$$

which implies $f \in (E/F)^*$. In addition, $\Phi f = g$, and surjectivity is proved.

It remains to show that Φ is isometric. For this, note that $|(\Phi f)(x)| \le \|f\| \cdot \|x\|$, $x \in E$, implies $\|\Phi f\| \le \|f\|$ for all $f \in (E/F)^*$. On the other hand, for each $\varepsilon > 0$ there exists $x \in E$ with

$$\|x+F\| = 1 \quad \text{and} \quad |f(x+F)| \ge \|f\| - \varepsilon\,.$$

Since $1 = \|x+F\| = \inf_{y\in F} \|x+y\|$, there exists $y \in F$ with $\|x+y\| \le 1+\varepsilon$. This implies $\left\|\frac{x+y}{1+\varepsilon}\right\| \le 1$ and hence

$$\left|(\Phi f)\left(\frac{x+y}{1+\varepsilon}\right)\right| = \frac{|f(x+F)|}{1+\varepsilon} \ge \frac{\|f\|-\varepsilon}{1+\varepsilon}\,.$$

Thus $\|\Phi f\| \ge \frac{\|f\|-\varepsilon}{1+\varepsilon}$, which yields $\|\Phi f\| \ge \|f\|$. This proves that Φ is isometric. □

We observe that one may naturally identify elements of a normed space with elements of its second dual space.

Lemma 4.18. *Let E be a normed space over $\mathbb{K}$. For $x \in E$ let $\hat{x}\colon E^* \to \mathbb{K}$ be defined by $\hat{x}(\ell) = \ell(x)$, $\ell \in E^*$. Then $\Lambda_E\colon E \to (E^*)^*$, $\Lambda_E x = \hat{x}$, is an isometric linear operator.*

Proof. The operator Λ_E is linear, and for $x \in E$ we have

$$\sup\{|\hat{x}(\ell)| : \ell \in E^*, \ \|\ell\| = 1\} = \sup\{|\ell(x)| : \ell \in E^*, \ \|\ell\| = 1\} = \|x\| ,$$

where the last equality follows from Corollary 4.7. This proves that indeed $\hat{x} \in (E^*)^*$ and that $\|\Lambda_E x\| = \|\hat{x}\| = \|x\|$ for all $x \in E$. □

If we can identify a normed space with all of its bidual in this way (!), the space will be called reflexive.

Definition 4.19. Let E be a normed space over $\mathbb{K}$, and let Λ_E be defined as above. Then Λ_E is called *canonical map* or *canonical embedding of E in* $E^{**} := (E^*)^*$. The space E is called *reflexive*, if Λ_E is surjective. The space E^{**} is the *bi-dual* of E, and $\overline{\Lambda_E(E)}$ is the *completion* of E.

Remark 4.20. Only Banach spaces can be reflexive. The class of reflexive spaces is a highly important class of Banach spaces. Intriguingly, there exist non-reflexive Banach spaces, which are isometrically isomorphic to their bi-dual (like the *James space*). Moreover, notice that finite-dimensional spaces are always reflexive because of $\dim E^{**} = \dim E$.

We relate reflexivity of a space to reflexivity of subspaces, quotient spaces and dual spaces.

Theorem 4.21. *Let E be a normed space over $\mathbb{K}$.*

(i) *If E is reflexive and $F \subset E$ a closed linear subspace, then F is also reflexive.*

(ii) *If E is a Banach space, then E is reflexive if and only if E^* is reflexive.*

Proof. (i). We have to show that for each $\varphi \in F^{**}$ there exists some $y \in F$ with $\varphi(f) = f(y)$ for all $f \in F^*$. For this, let $\varphi \in F^{**}$ and let $\psi\colon E^* \to \mathbb{K}$ be defined by $\psi(\ell) := \varphi(\ell|_F)$, $\ell \in E^*$. Since

$$|\psi(\ell)| \le \|\varphi\| \cdot \|\ell|_F\| \le \|\varphi\| \cdot \|\ell\| ,$$

we have $\psi \in E^{**}$. The space E being reflexive then implies that there exists $y \in E$ with $\psi(\ell) = \ell(y)$ for all $\ell \in E^*$. Next, towards a contradiction, assume that $y \notin F$. Then there exists some $\ell \in E^*$ with $\ell(y) \neq 0$ and $\ell|_F = 0$. Hence, $0 \neq \ell(y) = \psi(\ell) = \varphi(\ell|_F) = 0$, which is a contradiction. Finally, for $f \in F^*$ there exists some $\ell \in E^*$ with $\ell|_F = f$, hence

$$\varphi(f) = \varphi(\ell|_F) = \psi(\ell) = \ell(y) = f(y) .$$

This shows that F is reflexive.

(ii). Let E be reflexive. We need to show that for each $u \in E^{***}$ there exists some $f \in E^*$ with $u(\varphi) = \varphi(f)$ for all $\varphi \in E^{**}$. For this, let $u \in E^{***}$, and set $f(x) := u(\hat{x})$, $x \in E$. Then $f \in E^*$. Next, let $\varphi \in E^{**}$. Since there hence exists some $x \in E$ with $\hat{x} = \varphi$, we obtain

$$u(\varphi) = u(\hat{x}) = f(x) = \hat{x}(f) = \varphi(f) .$$

Therefore, E^* is reflexive.

For the converse, let E^* be reflexive. Then, by the above, E^{**} is reflexive. By (i), also $\Lambda_E(E) = \overline{\Lambda_E(E)}$ is reflexive. The claim now follows from the fact that E and $\Lambda_E(E)$ are (isometrically) isomorphic (see Exercise Sheet 6, Exercise 1(ii)). □

Theorem 4.22. *Let E be a Banach space and $F \subset E$ a closed linear subspace. Then the following are equivalent:*

(i) *E is reflexive.*

(ii) *F and E/F are reflexive.*

Proof. (i)⇒(ii). By (i) and Theorem 4.21(i), also F is reflexive. By Theorem 4.21(ii), E^* is reflexive, hence $F^\perp$ is reflexive. By Theorem 4.17(ii), $F^\perp$ is isometrically isomorphic to $(E/F)^*$ which is therefore also reflexive. By Theorem 4.21(ii), this finally implies that E/F is reflexive.

(ii)⇒(i). Let $\varphi \in E^{**}$. We will again use the isometric isomorphism

$$\Phi\colon (E/F)^* \to F^\perp \subset E^*\,,\ (\Phi u)(x) = u(x+F)\,,$$

$u \in (E/F)^*$, $x \in E$ from Theorem 4.17. Then we can define $\psi \in (E/F)^{**}$ by $\psi(u) := \varphi(\Phi u)$ for $u \in (E/F)^*$. Since E/F is reflexive, there exists some $x \in E$ with $\widehat{x+F} = \psi$, i. e.

$$\varphi(\Phi u) = \psi(u) = (\widehat{x+F})(u) = u(x+F) = (\Phi u)(x) = \hat{x}(\Phi u)\,,$$

$u \in (E/F)^*$. Hence, $(\varphi - \hat{x})|_{F^\perp} = 0$.

To utilize the reflexivity of F, we next define a suitable $\rho \in F^{**}$. For each $f \in F^*$, choose some $g \in E^*$ with $g|_F = f$ and $\|g\| = \|f\|$. Then define $\rho(f) := (\varphi - \hat{x})(g)$. This is a proper definition since for two extensions $g, h \in E^*$ of f we have $(g-h)|_F = 0$ and thus $g - h \in F^\perp$. A similar argument shows that ρ is linear. Moreover,

$$|\rho(f)| \le \|\varphi - \hat{x}\| \cdot \|g\| = \|\varphi - \hat{x}\| \cdot \|f\|\,.$$

Thus, $\rho \in F^{**}$. As F is reflexive, there exists some $y \in F$ with $\rho(f) = f(y)$ for all $f \in F^*$. Now, we conclude that for all $g \in E^*$ we have

$$\hat{y}(g) = g(y) = (g|_F)(y) = \rho(g|_F) = (\varphi - \hat{x})(h)$$

with some $h \in E^*$ satisfying $h|_F = g|_F$ and $\|h\| = \|g|_F\|$. Hence, $h - g \in F^\perp$ and thus $\hat{y}(g) = (\varphi - \hat{x})(g)$ for all $g \in E^*$. Equivalently, $\varphi = \hat{x} + \hat{y} = \widehat{x+y} \in \Lambda_E(E)$, which shows that E is reflexive. □

The next result is just another nice consequence of the Hahn-Banach Theorem.

Theorem 4.23. *Let E and F be normed spaces, $E \neq \{0\}$. If $L(E,F)$ is complete, then so is F.*

Proof. First, choose $x_0 \in E$ with $\|x_0\| = 1$. Then there exists some $f \in E^*$ with $f(x_0) = \|x_0\| = 1 = \|f\|$. Next, let $(y_n)_{n\in\mathbb{N}} \subset F$ be a Cauchy sequence, and define $T_n\colon E \to F$ by $T_n x := f(x) y_n$, $x \in E$. Since $\|T_n x\| \le \|f\|\|x\|\|y_n\| = \|y_n\|\|x\|$ for $x \in E$, we have $T_n \in L(E,F)$. Further, $\|T_n x - T_m x\| = |f(x)|\|y_n - y_m\| \le \|y_n - y_m\|\|x\|$ implies $\|T_n - T_m\| \le \|y_n - y_m\|$. Hence $(T_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L(E,F)$ and thus converges to some $T \in L(E,F)$. This implies $y_n = T_n x_0 \to T x_0$ as $n \to \infty$, i. e. $(y_n)_{n\in\mathbb{N}}$ converges in F. □

5 The Open Mapping, Closed Graph and Banach-Steinhaus Theorem

In this section we will prove three theorems which are of fundamental importance in operator theory – the field in functional analysis which deals with linear operators on and between Banach spaces. These are the open mapping theorem, the closed graph theorem and the Banach-Steinhaus theorem.

Starting with the open mapping theorem, we will need some lemmas.

Lemma 5.1. *Let E be a normed space, F a Banach space and $T \in L(E,F)$ surjective. Then $K_1(0_F) \subset \overline{T(K_r(0_E))}$ for some $r > 0$.*

Proof. Since T is surjective, $F = \bigcup_{n=1}^{\infty} T(K_n(0_E))$. As F is complete, there exists some $m \in \mathbb{N}$ with

$$\left(\overline{T(K_m(0_E))}\right)^{\circ} \neq \emptyset$$

by Remark 1.11 on page 7. Therefore, there exist $y_0 \in F$ and $s > 0$ with

$$K_s(y_0) \subset \overline{T(K_m(0_E))} =: A\,.$$

Now take $y \in K_s(0_F)$. Then $\|(y + y_0) - y_0\| \leq s$ and $\|(y_0 - y) - y_0\| \leq s$. Hence,

$$y \pm y \in K_s(y_0) \subset A\,.$$

If $z_1, z_2, z \in A$, then $-z, \frac{1}{2}(z_1 + z_2) \in A$. This implies $y = \frac{1}{2}((y + y_0) - (y_0 - y)) \in A$. Thus,

$$K_s(0_F) \subset \overline{T(K_m(0_E))}\,.$$

Hence, we conclude

$$K_1(0_F) = \frac{1}{s} K_s(0_F) \subset \overline{\tfrac{1}{s} T(K_m(0_E))} = \overline{T(K_{\frac{m}{s}}(0_E))}\,.$$

The lemma is proved with $r = m/s$. □

Lemma 5.2. *Let E be a Banach space, F a normed space and $T \in L(E,F)$. Suppose further that there exists $r > 0$ with*

$$K_1(0_F) \subset \overline{T(K_r(0_E))}\,.$$

Then

$$K_1(0_F) \subset T(K_{2r}(0_E))$$

and T is surjective.

Proof. Let $y \in K_1(0_F)$. By induction, we define a sequence $(y_n)_{n \in \mathbb{N}} \subset T(K_r(0_E))$ with

$$\left\| y - \sum_{k=1}^{n} \frac{y_k}{2^{k-1}} \right\| \leq \frac{1}{2^n}$$

for all $n \in \mathbb{N}$. By hypothesis, there exists $y_1 \in T(K_r(0_E))$ with $\|y - y_1\| \leq \frac{1}{2}$. Assume that $y_1, \ldots, y_n$ are constructed, i.e.

$$2^n \left(y - \sum_{k=1}^{n} \frac{y_k}{2^{k-1}} \right) \in K_1(0_F) \subset \overline{T(K_r(0_E))}\,.$$

Then there exists $y_{n+1} \in T(K_r(0_E))$ with

$$\left\|2^n\left(y - \sum_{k=1}^{n+1} \frac{y_k}{2^{k-1}}\right)\right\| = \left\|2^n\left(y - \sum_{k=1}^{n} \frac{y_k}{2^{k-1}}\right) - y_{n+1}\right\| \leq \frac{1}{2}\,.$$

This shows that such a sequence exists.
For each $n \in \mathbb{N}$, let $x_n \in K_r(0_E)$ be such that $T(x_n) = y_n$. Since $\|2^{-k+1}x_n\| \leq 2^{-k+1}r$, the sequence

$$\left(\sum_{k=1}^{n} \frac{x_k}{2^{k-1}}\right)_{n\in\mathbb{N}}$$

is a Cauchy-sequence in E. And as E is complete, we can define

$$x = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{x_k}{2^{k-1}}\,.$$

We have

$$\|x\| = \lim_{n\to\infty} \left\|\sum_{k=1}^{n} \frac{x_k}{2^{k-1}}\right\| \leq \lim_{n\to\infty} \sum_{k=1}^{n} \frac{\|x_k\|}{2^{k-1}} \leq \sum_{k=1}^{\infty} \frac{r}{2^{k-1}} = 2r\,.$$

Thus $x \in K_{2r}(0_E)$. As the operator T is continuous,

$$Tx = \lim_{n\to\infty} T\left(\sum_{k=1}^{n} \frac{x_k}{2^{k-1}}\right) = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{y_k}{2^{k-1}} = y\,,$$

and therefore $y \in T(K_{2r}(0_E))$. This shows $K_1(0_F) \subset T(K_{2r}(0_E))$.
It remains to show that T is surjective. For this, let $y \in F$, $y \neq 0$. We have $\frac{y}{\|y\|} \in K_1(0_F) \subset T(E)$. Therefore there exists $x \in E$ with $Tx = \frac{y}{\|y\|}$. Hence $T(\|y\|x) = y$. □

We arrive at the Open Mapping Theorem. First, we should say what we mean by an open map.

Definition 5.3. Let E, F be normed spaces. An operator $T \in L(E, F)$ is called *open*, if $T(U)$ is open in F for each open $U \subset E$.

As we have seen in the exercises, any open operator between normed spaces is surjective. The Open Mapping Theorem is the converse of this statement for operators between Banach spaces.

Theorem 5.4 (Open Mapping Theorem)**.** *For Banach spaces E, F, any surjective operator $T \in L(E, F)$ is open.*

Proof. For an open set $U \subset E$ and $x \in U$ there exists $t > 0$ with $K_t(x) \subset U$. Then

$$K_t(0) = K_t(x) - x = \{y - x : y \in K_t(x)\} \subset U - x\,.$$

From Lemma 5.1 it follows that there exists some $r > 0$ with

$$K_1(0_F) \subset \overline{T(K_r(0_E))}\,.$$

Lemma 5.2 implies

$$K_1(0_F) \subset T(K_{2r}(0_E))\,.$$

Moreover, we have

$$K_{\frac{t}{2r}}(0_F) = \tfrac{t}{2r}K_1(0_F) \subset \tfrac{t}{2r}T(K_{2r}(0_E)) = T(K_t(0_E)) \subset T(U) - Tx$$

and hence

$$K_{\frac{t}{2r}}(Tx) = K_{\frac{t}{2r}}(0_F) + Tx \subset T(U)\,.$$

Thus, $T(U)$ ist open. □

The importance of this theorem may become clearer when we emphasize that it implies the continuity of inverse operators between Banach spaces.

Corollary 5.5.

(i) *Let E, F be Banach spaces and $T \in L(E,F)$ bijective. Then T^{-1} is continuous.*

(ii) *If $\|\cdot\|_1$ and $\|\cdot\|_2$ are Banach space norms on E and if $\|x\|_1 \le c\|x\|_2$ for all $x \in E$ and some $c > 0$, then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.*

Proof. (i). By Theorem 5.4, T is open. Equivalently, $T(U) = \left(T^{-1}\right)^{-1}(U)$ is open for each open $U \subset E$, which is the continuity of T^{-1}.

(ii). Set $T = \mathrm{Id}$, consider $T\colon (E, \|\cdot\|_1) \to (E, \|\cdot\|_2)$ and apply the first statement. □

Recall Example 3.9, where we saw a bijective operator with unbounded inverse. There, the space F was not complete, so if E is a Banach space, F only a normed space, and $T\colon E \to F$ bijective, $T \in L(E,F)$ does *not* imply that T^{-1} is continuous.

Our next fundamental theorem will be the closed graph theorem, giving a criterion for continuity.

Theorem 5.6 (Closed Graph Theorem)**.** *Let E, F be Banach spaces and $T\colon E \to F$ be a closed linear operator. Then T is bounded.*

Proof. The product space $E \times F$ is a Banach space and the operator T is closed, hence the graph G_T is closed in $E \times F$ and thus a Banach space. Now define $S\colon G_T \to E$ by $S(x, Tx) = x$. This operator S is linear, bijective and continuous:

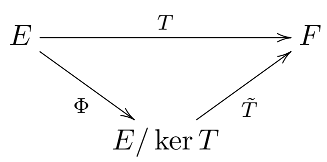
$$\|S(x,Tx)\| = \|x\| \le \max\{\|x\|, \|Tx\|\} = \|(x,Tx)\|\,.$$

By Corollary 5.5, $S^{-1}\colon E \to G_T$ is continuous, meaning

$$\|Tx\| \le \|(x,Tx)\| = \|S^{-1}(x)\| \le \|S^{-1}\|\,\|x\|$$

for all $x \in E$. □

Remark 5.7. While we used the Open Mapping Theorem to prove the Closed Graph Theorem, we could also use the latter to prove the first. Indeed, if E and F are Banach spaces and $T \in L(E,F)$ is surjective, let us first assume that T is also injective. Then, T^{-1} is a closed operator, since T is. But the Closed Graph Theorem now implies continuity of T^{-1}, i. e. openness of T. If T is not injective, we may factorize T as $T = \tilde{T} \cdot \Phi$, where $\Phi\colon E \to E/\ker T$ is the natural projection and $\tilde{T}\colon E/\ker T \to F$ is now injective.



By the first step, $\tilde{T}$ is open and since Φ is open (Exercises), also $T = \tilde{T}\Phi$ is open.

The following theorem gives a characterization of operators with closed range and is another one of the central theorems of functional analysis.

Theorem 5.8 (Closed Range Theorem)**.** *Let X and Y be Banach spaces and let $T \in L(X,Y)$. Then the following are equivalent:*

(i) *The space* $\operatorname{ran} T$ *is closed.*

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

(iii) *The space* $\operatorname{ran} T^*$ *is closed.*

(iv) *We have* $\operatorname{ran} T^* = (\ker T)^\perp$.

Proof. (i)⇔(ii). There holds $\overline{\operatorname{ran} T} = (\ker T^*)_\perp$ (c.f. Exercise).
(i)⇒(iv). Let $\operatorname{ran} T$ be closed. Let $x^* \in \operatorname{ran} T^*$. Then there exists $y^* \in Y^*$ with $T^* y^* = x^*$. Hence, for $x \in \ker T$ we have

$$x^*(x) = (T^*y^*)(x) = y^*(Tx) = y^*(0) = 0\,,$$

in explicit $\ker T \subset \ker x^*$ and thus $x^*|_{\ker T} = 0$ and therefore $x^* \in (\ker T)^\perp$, i.e. $\operatorname{ran} T^* \subset (\ker T)^\perp$. To show the converse inclusion, let $x^* \in (\ker T)^\perp$. Define

$$f\colon X/\ker T \to \mathbb{K},\ f[x] := x^*(x) \qquad \hat{T}\colon X/\ker T \to \operatorname{ran} T,\ \hat{T}[x] = Tx\,.$$

f and $\hat{T}$ are well-defined, bounded and $\hat{T}$ is bijective. Corollary 5.5 implies that $\hat{T}^{-1}$ is bounded. Define $z^* := f \circ \hat{T}^{-1} \in (\operatorname{ran} T)^*$. Let $y^* \in Y^*$ be a Hahn Banach extension of z^*. For $x \in X$ we have

$$(T^*y^*)(x) = y^*(Tx) = z^*(Tx) = f(\hat{T}^{-1}Tx) = f[x] = x^*(x)$$

and thus, $x^* = T^*y^* \in \operatorname{ran} T^*$.
(iv)⇒(iii). Obvious.
To show that (iii) implies (i) we define the sets

$$R_N := \Big\{ y \in Y :\ \text{there is } x \in X \text{ with } \|y - Tx\| \le \tfrac{1}{2}\|y\| \text{ and } \|x\| \le N\|y\| \Big\}\,.$$

<u>First claim:</u> Let $M \subset Y$ be a subspace such that $\operatorname{ran} T \subset M \subset R_N$ for some $N \in \mathbb{N}$. Then $\operatorname{ran} T = M$.
Proof: Let $y \in M$. Then there exists $x_1 \in X$ such that $\|y - Tx_1\| \le \frac{1}{2}\|y\|$ and $\|x_1\| \le N\|y\|$. Now, $y - Tx_1 \in M + \operatorname{ran} T = M \subset R_N$. Thus, there exists $x_2 \in X$ such that

$$\|y - Tx_1 - Tx_2\|\| \le \tfrac{1}{2}\|y - Tx_1\| \le \tfrac{1}{4}\|y\|$$

and $\|x_2\| \le N\|y - Tx_1\| \le \frac{N}{2}\|y\|$. Proceeding like this, we obtain a sequence $(x_n) \subset X$ with

$$\|y - Tx_1 - \cdots - Tx_n\| \le \tfrac{1}{2^n}\|y\| \text{ and } \|x_n\| \le \tfrac{N}{2^{n-1}}\|y\|\,.$$

Setting $u_n = \sum_{k=1}^n x_n$ we have that (u_n) is a Cauchy sequence and thus there exists $u \in X$ with $u_n \to u$. Therefore $Tu_n \to Tu$ but also $Tu_n = Tx_1 + \cdots + Tx_n \to y$. Hence, $y = Tu \in \operatorname{ran} T$.
This proves the first claim.

Second claim: If $M \subset Y$ is a subspace such that $M \not\subset R_N$ for all $N \in \mathbb{N}$. Then for each $n \in \mathbb{N}$ there exists $y_n \in M$ with $\|y_n\| = 1$ such that for all $x \in X$ we have $\|y_n - Tx\| > \frac{1}{2}$ or $\|x\| > n$.

Proof: For all $n \in \mathbb{N}$ there exists $v_n \in M$, $v_n \notin R_n$ (thus $v_n \neq 0$). We put $y_n := \frac{v_n}{\|v_n\|}$. Let $x \in X$. Then $\|v_n - T(\|v_n\|x)\| > \frac{1}{2}\|v_n\|$ or $\|\|v_n\|x\| > n\|v_n\|$. Hence $\|y_n - Tx\| > \frac{1}{2}$ oder $\|x\| > n$.

This proves the second claim.

Third claim: If $\overline{\operatorname{ran} T} \not\subset R_N$ for each $N \in \mathbb{N}$, then $\operatorname{ran} T^*$ is not closed.

Proof: Consider $X \times Y$ with the norm

$$\|(x, y)\|_1 := \|x\| + \|y\| .$$

Then $(X \times Y, \|\cdot\|_1)$ is a Banach space and $\operatorname{graph} T$ is a closed subspace of $X \times Y$. For $n \in \mathbb{N}$ define

$$V_n := \left\{ \left(\tfrac{1}{n}x, Tx \right) : x \in X \right\} = \operatorname{graph}(nT) .$$

V_n is a closed subspace of $X \times Y$. By the second claim, there exists a sequence $(y_n) \subset \overline{\operatorname{ran} T}$ with $\|y_n\| = 1$ such that for all $x \in X$ there holds $\|y_n - Tx\| > \frac{1}{2}$ or $\|x\| > n$. Hence

$$\|(0, y_n) - (\tfrac{1}{n}x, Tx)\|_1 = \|(-\tfrac{1}{n}x, y_n - Tx)\|_1 = \tfrac{1}{n}\|x\| + \|y_n - Tx\| > \tfrac{1}{2}$$

and thus $(0, y_n)$ has a positive distance to V_n. By Hahn-Banach there exists $z_n^* \in (X \times Y)^*$ with $\|z_n^*\| = 1$ such that

$$z_n^*(0, y_n) = 1 \text{ and } z_n^*|_{V_n} = 0 .$$

Now, for $n \in \mathbb{N}$ and $y \in Y$ we set $y_n^*(y) := z_n^*(0, y)$. Then $y_n^* \in Y^*$ and $\operatorname{dist}(y_n^*, \ker T^*) \geq 1$, because

$$(y_n^* - y^*)(y_n) = y_n^*(y_n) = z_n^*(0, y_n) = 1$$

for all $y^* \in \ker T^*$, since $\overline{\operatorname{ran} T} = (\ker T^*)_\perp$. For $x \in X$ and $n \in \mathbb{N}$ we now have

$$\begin{aligned} |(T^* y_n^*)(x)| = |y_n^*(Tx)| = |z_n^*(0, Tx)| &= \left| z_n^*\big((\tfrac{1}{n}x, Tx)_{\in V_n} - (\tfrac{1}{n}x, 0)\big) \right| \\ &= |z_n^*(\tfrac{1}{n}x, 0)| \leq \tfrac{1}{n} \underbrace{\|z_n^*\|}_{=1} \|x\| . \end{aligned}$$

Hence, $T^* y_n^* \to 0$ in X^*. Now define the operator

$$A \colon Y^* / \ker T^* \to \operatorname{ran} T^*, \ A[y^*] := T^* y^* .$$

A is bounded and bijective and if further $\operatorname{ran} T^*$ is closed, A^{-1} is bounded by the Open Mapping Theorem 5.4. We have $A[y_n^*] = T^* y_n^* \to 0$ and $\|[y_n^*]\| = \operatorname{dist}(y_n^*, \ker T^*) \geq 1$. Applying A^{-1} yields $[y_n^*] \to 0$, thus $\operatorname{ran} T^*$ cannot be closed.

This proves the third claim.

Finally, we can show that (iii) implies (i): Assume, $\operatorname{ran} T^*$ is closed. Then by the third claim, we have $\overline{\operatorname{ran} T} \subset R_n$ for some $n \in \mathbb{N}$. This implies $\operatorname{ran} T \subset \overline{ranT} \subset R_n$ (for some $n \in \mathbb{N}$). By the first claim, we have $\operatorname{ran} T = \overline{\operatorname{ran} T}$. ☐

We now tackle the Banach-Steinhaus theorem. This name is frequently given to the following Uniform Boundedness Principle but will here be reserved for another theorem.

Theorem 5.9 (Uniform Boundedness Principle)**.** *Let E be a Banach space, F a normed space and let $\mathcal{T} \subset L(E,F)$. Let $\mathcal{T}$ be pointwise bounded, i.e., for each $x \in E$ there exists $M_x < \infty$ such that $\|Tx\| \leq M_x$ for all $T \in \mathcal{T}$. Then $\mathcal{T}$ is bounded, i.e., there exists some $M < \infty$ such that $\|T\| \leq M$ for all $T \in \mathcal{T}$.*

Proof. For $n \in \mathbb{N}$ let

$$E_n := \{x \in E : \|Tx\| \leq n \text{ for all } T \in \mathcal{T}\}.$$

By hypothesis, $E = \bigcup_{n=1}^{\infty} E_n$. Let $x = \lim_{j\to\infty} x_j$ with $x_j \in E_n$ for fixed n. Since $\|Tx_j\| \leq n$ for all j we have

$$\|Tx\| = \lim_{j\to\infty} \|Tx_j\| \leq n.$$

Thus $x \in E_n$, and hence E_n is closed. By Remark 1.11 on page 7 on Baire's Theorem, there exists some $n_0 \in \mathbb{N}$ with

$$\mathring{E}_{n_0} \neq \emptyset.$$

Hence, $K_r(x) \subset E_{n_0}$ for some $x \in E_{n_0}$, $r > 0$. Let $y \in E$ with $\|y\| \leq r$. Then $y + x \in K_r(x) = x + K_r(0)$. This implies

$$\|Ty\| = \|T(y+x) - Tx\| \leq \|T(y+x)\| + \|Tx\| \leq 2n_0 \tag{5.1}$$

for all $T \in \mathcal{T}$. Now let $y \in E$, $y \neq 0$ arbitrary. Then

$$\frac{r}{\|y\|}\|Ty\| = \left\|T\left(\frac{ry}{\|y\|}\right)\right\| \leq 2n_0,$$

where the inequality follows from (5.1). This implies $\|Ty\| \leq \frac{2n_0}{r}\|y\|$ and thus $\|T\| \leq \frac{2n_0}{r}$. □

Remark 5.10. It is possible to prove the Uniform Boundedness Principle without using Baire's Theorem (instead using a "gliding hump argument"). One may find the outline of such a proof as an Exercise in *An Introduction to Banach Space Theory* by Megginson.

Example 5.11. In general, Theorem 5.9 does not hold, if E is *not* a Banach space. Consider

$$E = c_{00} := \{x = (x_n)_{n\in\mathbb{N}} \in \ell_\infty : x_n = 0 \text{ for almost all } n \in \mathbb{N}\} \subset \ell_\infty,$$

the space F being the field $\mathbb{K}$ and $\mathcal{T}$ being the sequence (f_n), where $f_n(x) = nx_n$ for $x \in E$. We see that $\mathcal{T}$ is pointwise bounded, since $x_n = 0$ from some $n \geq N$ on, but $\|f_n\| = n$ for $n \in \mathbb{N}$.

We obtain a first consequence of the above theorem.

Corollary 5.12. *Let E be a Banach space, F a normed space and $T_n \in L(E,F)$. Suppose that $(T_n x)_{n\in\mathbb{N}}$ is convergent in F for every $x \in E$. Then define $T\colon E \to F$ by*

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

Then $T \in L(E,F)$, $(\|T_n\|)_{n\in\mathbb{N}}$ is bounded, and

$$\|T\| \leq \liminf_{n\to\infty} \|T_n\|.$$

Proof. By definition, T obviously is linear and $(\|T_n x\|)_{n\in\mathbb{N}}$ is bounded. By Theorem 5.9, $\|T_n\| \le M$ for all $n \in \mathbb{N}$. Hence, for all $x \in E$

$$\|Tx\| = \lim_{n\to\infty} \|T_n x\| \le M\|x\| .$$

This shows that $T \in L(E, F)$. Let $(\|T_{n_k}\|)_{k\in\mathbb{N}}$ be a convergent subsequence of $(\|T_n\|)_{n\in\mathbb{N}}$. Then

$$\|Tx\| = \lim_{k\to\infty} \|T_{n_k} x\| \le \|x\| \lim_{k\to\infty} \|T_{n_k}\| .$$

Thus

$$\|T\| \le \lim_{k\to\infty} \|T_{n_k}\| ,$$

and hence

$$\|T\| \le \liminf_{n\to\infty} \|T_n\| .$$

□

The next lemma is a reformulation of Lemma 3.6.

Lemma 5.13. *Let E be a normed space and F a Banach space. Let E_0 be a dense linear subspace of E, and $T_0 \in L(E_0, F)$. Then there exists a unique $T \in L(E, F)$ with*

$$T|_{E_0} = T_0 \quad \text{and} \quad \|T\| = \|T_0\| .$$

We finally arrive at the Banach-Steinhaus theorem.

Theorem 5.14 (Banach-Steinhaus Theorem)**.**

(i) *Let E be a Banach space and F a normed space. Further, let $T_n \in L(E, F)$, $n \in \mathbb{N}$. If $(T_n)_{n\in\mathbb{N}}$ is pointwise convergent to some $T : E \to F$ which is linear, then*

$$\sup_{n\in\mathbb{N}} \|T_n\| < \infty .$$

(ii) *Let E be a normed space and F a Banach space. Further, let $T_n \in L(E, F)$, $n \in \mathbb{N}$. If*

  (a) $\sup_{n\in\mathbb{N}} \|T_n\| < \infty$ *and*

  (b) *there exists a dense linear subspace E_0 of E such that $(T_n x)_{n\in\mathbb{N}}$ is convergent in F for each $x \in E_0$,*

  then there exists some $T \in L(E, F)$ with $Tx = \lim_{n\to\infty} T_n x$ for all $x \in E$.

Proof. (i). This is Corollary 5.12
(ii). For each $y \in E_0$ set $T_0 y := \lim_{n\to\infty} T_n y$, which exists by (iib). The operator T_0 is linear and, by (iia),

$$\|T_0 y\| = \lim_{n\to\infty} \|T_n y\| \le \sup_{n\in\mathbb{N}} \|T_n\| \|y\| < \infty .$$

Hence $T_0 \in L(E_0, F)$. By Lemma 5.13, there exists some $T \in L(E,F)$ with $T|_{E_0} = T_0$. Let $x \in E$, and $\varepsilon > 0$. Then let $y \in E_0$ with $\|x - y\| \le \varepsilon$ and choose $N \in \mathbb{N}$ such that $\|T_n y - T_0 y\| \le \varepsilon$ for all $n \ge N$, which is possible by (iib). Then for all $n \ge N$,

$$\begin{aligned}\|T_n x - Tx\| &\le \underbrace{\|T_n x - T_n y\|}_{\le \|T_n\|\,\|x-y\|} + \underbrace{\|T_n y - T_0 y\|}_{\le \varepsilon} + \underbrace{\|T_0 y - Tx\|}_{\le \|T\|\,\|y-x\|} \\ &\le \|T_n\|\|x-y\| + \varepsilon + \|T\|\|y-x\| \\ &\le \varepsilon(\|T_n\| + 1 + \|T\|) \\ &\le \varepsilon\Big(\underbrace{\sup_{n\in\mathbb{N}} \|T_n\|}_{<\infty \text{ by (iia)}} + 1 + \|T\|\Big) .\end{aligned}$$

This implies $T_n x \to Tx$ as $n \to \infty$. □

Also, the Uniform Boundedness Principle gives rise to a criterion for a subset of a normed space to be bounded.

Theorem 5.15. *Let E be a normed space and $M \subset E$. Then the following conditions are equivalent:*

(i) *The set M is bounded.*

(ii) *For each $f \in E^*$ the set $f(M) \subset \mathbb{K}$ is bounded*[2].

Remark 5.16 (Geometric interpretation of Theorem 5.15)*.* Suppose that for every closed hyperplane H in E (kernel of some $f \in E^*$) there exists some c with M lying between $H + c$ and $H - c$. Then M is already contained in a ball.

Proof. (i)⇒(ii). This follows from $\|x\| < c$ implying $\|f(x)\| \le c\|f\|$ for all $x \in M$.
(ii)⇒(i). Consider the set

$$\hat{M} = \{\hat{x} \colon f \mapsto f(x) : x \in M\} \subset L(E^*, \mathbb{K}) = E^{**} .$$

Since $\hat{M}(f) = f(M)$, $\hat{M}$ is pointwise bounded by (ii). By Theorem 5.9, $\hat{M}$ is bounded. Since the embedding is isometric, also M is bounded. □

This gives rise to a criterion for continuity of operators.

Corollary 5.17. *Let E, F be normed spaces and $T \colon E \to F$ be linear. Then T is continuous if and only if $f \circ T \in E^*$ for all $f \in F^*$.*

Proof. The operator T is bounded if and only if $T\big(K_1(0)\big)$ is bounded in F. By Theorem 5.15 this is equivalent to $f\big(T(K_1(0))\big)$ being bounded for all $f \in F^*$, which means $f \circ T \in E^*$ for all $f \in F^*$. □

[2] One also says that M is *weakly bounded*

6 Weak Convergence and Weak Topology

We have already seen that the closed unit ball of an infinite-dimensional normed space is not compact. Wishing to extract convergent subsequences under certain conditions, we weaken the notion of continuity. This will lead us to weak convergence and the weak topology.

Let us outline the concept of weak convergence. The key idea is to reduce the question of convergence in a normed space to a scalar problem by applying any continuous linear functional. As a motivation, remember convergence in $\mathbb{K}^n$ was characterized by componentwise convergence. From our point of view, sending a vector to one of its components defines a continuous linear functional and these functionals in fact span the whole dual space of $\mathbb{K}^n$. Now, weak convergence will be a natural generalization.

Definition 6.1. Let E be a normed space. Then $(x_n)_{n\in\mathbb{N}} \subset E$ is *weakly convergent* to $x \in E$, if

$$f(x_n) \to f(x) \text{ for all } f \in E^* .$$

Then x is called the *weak limit* of (x_n), and we write $x_n \xrightarrow{\text{w}} x$.

Remark 6.2.

(a) Convergence implies weak convergence, i.e. if $(x_n)_{n\in\mathbb{N}} \subset E$ converges to $x \in E$ (w.r.t. the norm), then (x_n) is also weakly convergent to x. To see this, observe $|f(x_n) - f(x)| = |f(x_n - x)| \leq \|f\|\|x_n - x\| \to 0$ as $n \to \infty$.

(b) The weak limit of a weak convergent sequence is unique, i.e. if $(x_n)_{n\in\mathbb{N}} \subset E$ is weakly convergent to both $x \in E$ and $y \in E$, then $x = y$: Since $(x_n)_{n\in\mathbb{N}}$ converges weakly to x and y, it follows that $f(x) = f(y)$ and thus $f(x - y) = 0$ for all $f \in E^*$. Hence $x = y$ (see Corollary 4.8).

(c) If $(x_n)_{n\in\mathbb{N}} \subset E$ and $(y_n)_{n\in\mathbb{N}} \subset E$ are weakly convergent sequences with $x_n \xrightarrow{\text{w}} x$ and $y_n \xrightarrow{\text{w}} y$, then $(x_n + \lambda y_n)_{n\in\mathbb{N}}$, $\lambda \in \mathbb{K}$, is weakly convergent to $x + \lambda y$. Thus, the linear operations on E are compatible with weak convergence. For this, note that $f(x_n + \lambda y_n) = f(x_n) + \lambda f(y_n) \to f(x) + \lambda f(y) = f(x + \lambda y)$.

(d) The converse of (a) is in general not true:

Example: Let $E = C[0,1]$, endowed with the L_2-norm $\|f\|_2 := \left(\int_0^1 |f(t)|^2 \,\mathrm{d}t\right)^{\frac{1}{2}}$. For $f_n(t) = \sin(nt)$ we have

$$\|f_n\|_2^2 = \int_0^1 |\sin(nt)|^2 \,\mathrm{d}t = \frac{1}{2}\left(1 - \frac{1}{2n}\sin(n \cdot 2)\right) \xrightarrow{n\to\infty} \frac{1}{2},$$

so (f_n) is not convergent to 0. But $f_n \xrightarrow{\text{w}} 0$ as $n \to \infty$.

(e) It can be shown that for sequences in ℓ_1 convergence is the same as weak convergence which is sometimes called *Schur's property*. (Proof uses Remark 3.13/Lemma 3.15.)

Another point of view is that weak convergence is just "pointwise convergence on any element of the dual space". This becomes more natural in the following definition.

Definition 6.3. Let E be a normed space. Then $(f_n)_{n\in\mathbb{N}} \subset E^*$ is *weak*-convergent* to $f \in E^*$, if

$$f_n(x) \to f(x) \text{ for all } x \in E.$$

We write $f_n \xrightarrow{\text{w}^*} f$.

Remark 6.4. Remark 6.2 (a), (b), (c) holds similarly.

We will now apply the Banach-Steinhaus Theorem to these new concepts.

Theorem 6.5. *The following statements hold true:*

(i) *Let E be a normed space, $(x_n)_{n\in\mathbb{N}} \subset E$, $x \in E$. Then the following conditions are equivalent:*

  (a) $x_n \xrightarrow{\text{w}} x$.

  (b) *We have:*

    (I) $\sup_{n\in\mathbb{N}} \|x_n\| < \infty$

    (II) *There exists a dense subspace $D \subset E^*$ such that*

$$f(x_n) \to f(x)$$

    for all $f \in D$.

(ii) *Let E be a Banach space, $(f_n)_{n\in\mathbb{N}} \subset E^*$, $f \in E^*$. Then the following conditions are equivalent:*

  (a) $f_n \xrightarrow{\text{w}^*} f$.

  (b) *We have:*

    (I) $\sup_{n\in\mathbb{N}} \|f_n\| < \infty$

    (II) There exists a dense subspace $D \subset E$ such that

$$f_n(x) \to f(x)$$

    for all $x \in D$.

Proof. (i).(a)⇒(b). Since (x_n) converges weakly to x, we have $f(x_n) \to f(x)$ as $n \to \infty$ for all $f \in E^*$. This implies (II). Moreover, $\{f(x_n) : n \in \mathbb{N}\} \subset \mathbb{K}$ is bounded for each $f \in E^*$. By Theorem 5.15, also $\{x_n : n \in \mathbb{N}\}$ is bounded, i.e. (I) is true.

(b)⇒(a). For this, consider the canonical embedding of E in E^{**}. Then, by (II), $\hat{x}_n(f) \to \hat{x}(f)$ for all $f \in D$ and $\|\hat{x}_n\| = \|x_n\|$. By the Banach-Steinhaus Theorem (applied to $(\hat{x}_n)_{n\in\mathbb{N}}$, $\hat{x} \in L(E^*, \mathbb{K})$ and using (I)) we obtain $\hat{x}_n(f) \to \hat{x}(f)$ as $n \to \infty$ for all $f \in E^*$ and thus $f(x_n) \to f(x)$ for all $f \in E^*$.

(ii). This is Banach-Steinhaus Theorem (Theorem 5.14). □

We arrive at a first theorem ensuring the existence of (weak*-)convergent subsequences under certain conditions.

Theorem 6.6. *Let E be a normed space, which is separable (i.e. there exists a countable dense subset). Then every bounded sequence in E^* contains a weak*-convergent subsequence.*

Proof. Let $\{x_1, x_2, x_3, \dots\}$ be a dense countable subset of E and let $(f_n)_{n\in\mathbb{N}} \subset E^*$ be a sequence in E^* with $\|f_n\| \le C$ for all $n \in \mathbb{N}$, $C \in \mathbb{K}$ fixed. First, we have that the sequence

$$(f_n(x_1))_{n\in\mathbb{N}} \subset \mathbb{K}$$

is bounded, hence there exists a convergent subsequence

$$(f_{n,1}(x_1))_{n\in\mathbb{N}} \subset \mathbb{K}\,.$$

Also $(f_{n,1}(x_2))_{n\in\mathbb{N}} \subset \mathbb{K}$ is bounded, hence has a convergent subsequence $(f_{n,2}(x_2))_{n\in\mathbb{N}} \subset \mathbb{K}$. We continue this way and select the diagonal sqeuence

$$g_n := f_{n,n}\,,\ n \in \mathbb{N}\,,$$

of $(f_n)_{n\in\mathbb{N}}$. By construction, $(g_n(x_m))_{n\in\mathbb{N}}$ is convergent for any $m \in \mathbb{N}$. By the Banach-Steinhaus theorem, we know that $(g_n)_{n\in\mathbb{N}}$ is pointwise convergent to some $g \in E^*$. Thus $g_n \xrightarrow{\text{w}^*} g$. □

Example 6.7. The separability of E is neccessary condition in Theorem 6.6. In general, without separability, Theorem 6.6 is not true. For this, consider $E = \ell_\infty$, $x = (x_n)_{n\in\mathbb{N}}$ and define the sequence (f_n) by $f_n(x) = x_n$, $n \in \mathbb{N}$. Then $\|f_n\| = 1$ for all $n \in \mathbb{N}$, which implies boundedness. Let $(f_{n_k})_{k\in\mathbb{N}}$ be a subsequence of $(f_n)_{n\in\mathbb{N}}$. Then, for $x = (x_n)_n \in E$ defined by

$$x_n = \begin{cases} 1 & n = n_k \text{ and } k \text{ is even} \\ 0 & \text{otherwise} \end{cases}$$

we have that

$$f_{n_k}(x) = \begin{cases} 1 & k \text{ is even} \\ 0 & \text{otherwise}\,. \end{cases}$$

Hence $f_{n_k}(x)$ is not convergent and (f_n) has no weak*-convergent subsequence.

After having discussed weak and weak* convergence, we will introduce corresponding topologies, the weak and weak* topologies. The weak topology of an infinite dimensional space and the weak* topology on the dual of an infinite-dimensional Banach space will not be metrizable, i. e. we cannot define them by giving a metric. Instead, we will need the notions of general topology, using the language of open sets to define a topology. Let us mention the most basic definitions.

Definition 6.8 (An excursion to topology)**.** Let X be a nonempty set

(i) A *topology* $\mathcal{T}$ on X is a family of subsets of X with the following properties:

  (T_1) The empty set and X are contained in $\mathcal{T}$, i. e. $\emptyset, X \in \mathcal{T}$.

  (T_2) If $\gamma \subset \mathcal{T}$, then $\bigcup_{S\in\gamma} S \in \mathcal{T}$.

  (T_3) If $S_1, \ldots, S_r \in \mathcal{T}$, then $\bigcap_{i=1}^{r} S \in \mathcal{T}$.

  The pair $(X, \mathcal{T})$ is then called a *topological space*. The sets in $\mathcal{T}$ are called *open* and the sets $X \setminus U$, $U \in \mathcal{T}$, are called *closed*.

  A subset $\mathcal{B} \subset \mathcal{T}$ is called a *basis* for $\mathcal{T}$, if each $U \in \mathcal{T}$ can be written as the union of elements of $\mathcal{B}$.

  Remark: If (X, d) is a metric space, then the set of open subsets (w.r.t. d) is a topology on X, the *topology induced* by d. The set of open balls $U_\varepsilon(x)$ is a basis for $\mathcal{T}$.

(ii) A family $\mathfrak{U} \subset \mathcal{T}$ is an *open covering* of X, if $X \subset \bigcup_{U\in\mathfrak{U}}$. The space X is *compact*, if every open covering contains a finite subcover.

(iii) Let X, Y be topological spaces. Then $f\colon X \to Y$ is *continuous*, if $f^{-1}(U)$ is open in X for every open set $U \subset Y$. If f is bijective and f, f^{-1} are continuous, f is called a *homeomorphism*.

Remark: A continuous function $f\colon X \to Y$ maps compact sets to compact sets.

Proof. Let $K \subset X$ be compact and $\mathfrak{U}$ an open covering of $f(K)$. Then, by continuity of f,

$$f^{-1}(\mathfrak{U}) = \{f^{-1}(U) : U \in \mathfrak{U}\}$$

is an open covering of the compact set K, hence there exists a finite subcover $f^{-1}(U_1), \dots, f^{-1}(U_n)$. Thus, $U_1, \dots, U_n$ is a finite subcover of $\mathfrak{U}$. □

Lemma 6.9. *Let $(X, \mathcal{T})$ be a topological space, $\gamma \subset \mathcal{T}$, and $\mathcal{B}$ the set of all finite intersections of sets in γ. Assume furthermore that $\mathcal{B}$ is a basis for $\mathcal{T}$. Then, if each open cover $\mathfrak{U} \subset \gamma$ of X contains a finite subcover, X is compact.*

Proof. First, towards a contradiction, assume that there exists a cover $\mathfrak{U}_0 \subset \mathcal{T}$ of X without a finite subcover. Define

$$\Omega := \{\mathfrak{U} \subset \mathcal{T} : \mathfrak{U} \supset \mathfrak{U}_0 \text{ and } \mathfrak{U} \text{ does not contain a finite subcover}\} .$$

Then $\Omega \neq \emptyset$, since $\mathfrak{U}_0 \in \Omega$. Also, Ω is partially ordered by inclusion "$\subset$".

Claim 1. Ω satisfies the hypothesis of Zorn's Lemma.

Proof. Let K be a chain in Ω and set

$$\tilde{\mathfrak{U}} := \bigcup \{\mathfrak{U} : \mathfrak{U} \in K\} .$$

If $\tilde{\mathfrak{U}}$ contains a finite subcover, then there exist $\mathfrak{U}_i \in K$, $U_i \in \mathfrak{U}_i$, $i = 1, \dots, r$, with

$$X = \bigcup_{i=1}^{r} U_i .$$

Since K is a chain, there exists $\mathfrak{U}_{i_0} \in K$ with $\mathfrak{U}_i \subset \mathfrak{U}_{i_0}$, $i = 1, \dots, r$. Thus $\mathfrak{U}_{i_0}$ contains the finite subcover $\{U_i\}_{i=1}^{r}$, which is a contradiction. Therefore, $\tilde{\mathfrak{U}}$ does not contain a finite subcover, which implies $\tilde{\mathfrak{U}} \in \Omega$. Claim 1 is proved.

By Zorn's Lemma there exists a maximal element $\mathcal{M} \in \Omega$ (i.e. no element in Ω is larger than $\mathcal{M}$ (contains $\mathcal{M}$)).

Claim 2. If $A, B \in \mathcal{T}$, $A \notin \mathcal{M}$, $B \notin \mathcal{M}$, then $A \cap B \notin \mathcal{M}$.

Proof. Since $A \notin \mathcal{M}$ and $\mathcal{M}$ is maximal, it follows that $\mathcal{M} \cup \{A\}$ contains a finite subcover $\{A, M_1, \dots, M_n\}$. Similarly $\mathcal{M} \cup \{B\}$ contains a finite subcover $\{B, M_1', \dots, M_m'\}$. This implies $\{A \cap B, M_1, \dots, M_n, M_1', \dots, M_m'\}$ is a finite subcover of $\mathcal{M} \cup \{A \cap B\}$ of X, therefore $A \cap B \notin \mathcal{M}$. By induction, $A_1, \dots, A_n \notin \mathcal{M}$ and thus $\bigcap_{i=1}^{n} A_i \notin \mathcal{M}$.

Claim 3. We have

$$\bigcup_{M \in \mathcal{M}} M = \bigcup_{M \in \gamma \cap \mathcal{M}} M .$$

Proof. Let $x \in M \in \mathcal{M}$. Then there exist $S_1, \dots, S_r \in \gamma$, such that

$$x \in \bigcap_{i=1}^{r} S_i \subset M .$$

Since the existence of a finite subcover of $\mathcal{M} \cup \{S_1 \cap \ldots \cap S_r\}$ implies the existence of a finite subcover of $\mathcal{M} \cup \{M\} = \mathcal{M}$, we have that $\mathcal{M} \cup \{S_1 \cap \ldots \cap S_r\} \in \Omega$. Because of $\mathcal{M}$ beeing the maximal element of Ω and claim 2, there exists some $i \in \{1, \ldots, r\}$ such that $S_i \in \mathcal{M}$, hence $x \in S_i \in \mathcal{M} \cap \gamma$. This proves Claim 3.

Since $\mathcal{M}$ covers X, claim 3 implies that $\mathcal{M} \cap \gamma$ is another cover of X which has, by hypothesis of the lemma, a finite subcover. Hence $\mathcal{M}$ has a finite subcover, which contradicts $\mathcal{M} \in \Omega$. □

Lemma 6.10. *Let $X \neq \emptyset$ and let γ be a set of subsets of X, i.e. $\gamma \subset \mathcal{P}(X)$. By $\mathcal{L}$ denote the set consisting of all finite intersections of sets in γ, the empty set and X itself. Let $\mathcal{T}$ be the set which consists of all unions of sets in $\mathcal{L}$. Then $\mathcal{T}$ is a topology on X, and $\mathcal{L}$ is a basis for $\mathcal{T}$.*

Proof. Tutorials. □

Definition 6.11. Let $X \neq \emptyset$ and $\gamma \subset \mathcal{P}(X)$. Then the topology defined in Lemma 6.10 is denoted by $\mathcal{T}(\gamma)$. If $(X, \mathcal{T})$ is a topological space, a set $\gamma \subset \mathcal{P}(X)$ is called a *subbasis* for $\mathcal{T}$, if $\mathcal{T} = \mathcal{T}(\gamma)$.

Lemma 6.12. *The intersection of arbitrarily many topologies on a set X is again a topology on X. For a set $\gamma \subset \mathcal{P}(X)$ we have*

$$\mathcal{T}(\gamma) = \bigcap \{\mathcal{T} : \mathcal{T} \text{ is a topology on } X \text{ with } \gamma \subset \mathcal{T}\} .$$

That is, $\mathcal{T}(\gamma)$ is the smallest topology containing γ.

Proof. Exercise. □

Definition 6.13. Let X be a set, I an index set, let be $(X_i, \mathcal{T}_i)$ topological spaces and $f_i \colon X \to X_i$, $i \in I$. The *weak topology with respect to the mappings f_i* is defined as $\mathcal{T}(\gamma)$, where

$$\gamma := \left\{ f_i^{-1}(V) : V \in \mathcal{T}_i\,,\ i \in I \right\} .$$

This is just the "coarsest" (smallest) topology on X under which all f_i are continuous.

Lemma 6.14. *Let X, I, X_i, $\mathcal{T}_i$ and f_i be as above. For $i \in I$ let γ_i be a subbasis of $\mathcal{T}_i$. Then the weak topology $\mathcal{T}$ with respect to the f_i is given by $\mathcal{T}(\gamma)$, where*

$$\gamma = \left\{ f_i^{-1}(V) : V \in \gamma_i, i \in I \right\} .$$

Moreover, $\mathcal{T}$ is the smallest topology on X with respect to which all f_i are continuous. In particular, each f_i is $(X, \mathcal{T})$-$(X_i, \mathcal{T}_i)$-continuous.

Proof. The second claim is almost immediate from Lemma 6.12. The first is exercise. □

Definition 6.15. Let E be a normed space over $\mathbb{K}$. The *weak topology on E* is defined as the weak topology on E with respect to the mappings $f \colon E \to \mathbb{K}$, $f \in E^*$. It is denoted by $\sigma(E, E^*)$. The *weak* topology on E^** is the weak topology on E^* with respect to the mappings $\hat{x} \colon E^* \to \mathbb{K}$, where $\hat{x}(f) = f(x)$, $x \in E$, $f \in E^*$. It is denoted by $\sigma(E^*, E)$.

Remark 6.16. We have

$$\sigma(E^*, E) \subset \sigma(E^*, E^{**}) .$$

This follows from the fact that each $\hat{x}$ is an element of E^{**} and hence $\sigma(E^*, E^{**})$-continuous.

Lemma 6.17. *Let E be a normed space. Then a subbasis for of the weak topology $\sigma(E, E^*)$ on E is given by the sets $\{x \in E : |f(x) - a| < \varepsilon\}$, where $f \in E^*$, $a \in \mathbb{K}$ and $\varepsilon > 0$. For the weak*-topology $\sigma(E^*, E)$ on E^*, a subbasis is given by the sets $\{f \in E^* : |f(x) - a| < \varepsilon\}$, where $x \in E$, $a \in \mathbb{K}$ and $\varepsilon > 0$.*

Proof. This is a consequence of Lemma 6.14. According to this Lemma, a subbasis is given by sets of the form $f^{-1}(U_\varepsilon(a))$, where $f \in E^*$, $a \in \mathbb{K}$ and $\varepsilon > 0$, since $U_\varepsilon(a)$, $a \in \mathbb{K}$, $\varepsilon > 0$ form a basis of the standard topology on $\mathbb{K}$. Now

$$f^{-1}(U_\varepsilon(a)) = \{x \in E : |f(x) - a| < \varepsilon\}.$$

The second claim is proven by a similar argument. $\square$

Definition 6.18. Let I be an index set and let $(X_i, \mathcal{T}_i), i \in I$, be topological spaces. Denote by $X = \prod_{i \in I} X_i$ the cartesian product of the X_i. Let further p_i be the canonical projection from X onto X_i, i.e. $p_j((x_i)_{i \in I}) = x_j$. Then the weak topology with respect to the p_i is called the *product topology* on X.

The next theorem actually is, again, a statement from topology, but is often counted as a functional analytic one.

Theorem 6.19 (Tychonoff)**.** *Let I be an index set and let $(X_i, \mathcal{T}_i)$ be topological spaces. Denote the product topology on $X = \prod_{i \in I} X_i$ by $\mathcal{T}$. Then $(X, \mathcal{T})$ is compact if and only if each $(X_i, \mathcal{T}_i)$ is.*

Proof. First, let $(X, \mathcal{T})$ be compact. We know that each p_i is continous. Therefore $X_i = p_i(X)$ is compact for all $i \in I$.

For the converse, assume that each $(X_i, \mathcal{T}_i)$ is compact. By definition, $\gamma := \{p_i^{-1}(V_i) : V_i \in \mathcal{T}_i, i \in I\}$ is a subbasis for $\mathcal{T}$. Suppose that $(X, \mathcal{T})$ is not compact. Then, by Lemma 6.9 there exists $W \subset \gamma$ which is a cover of X but does not contain a finite subcover. For $i \in I$ put

$$W_i := \{V_i \in \mathcal{T}_i : p_i^{-1}(V_i) \in W\}.$$

The family W_i is then not a cover of X_i, since otherwise there would exist $V_{i,1}, \dots, V_{i,n} \in W_i$ with $X_i = \bigcup_{k=1}^n V_{i,k}$ (since $(X_i, \mathcal{T}_i)$ is compact). This would imply $X = p_i^{-1}(X_i) = \bigcup_{k=1}^n \underbrace{p_i^{-1}(V_{i,k})}_{\in W}$, which would be a contradiction. For $i \in I$, pick $x_i \in X_i \setminus \bigcup_{V_i \in W_i} V_i$ and set $x := (x_i)_{i \in I} \in X$. Since W is a cover of X, there exists $V \in W$ such that $x \in V$. By the inclusion $W \subset \gamma$ there exist $i \in I$ and $V_i \in \mathcal{T}_i$ such that $V = p_i^{-1}(V_i)$. Now, $V_i \in W_i$ by the definition of W_i. But from $p_i(x) = x_i \notin V_i$, we conclude $x \notin V$, which is a contradiction. $\square$

In the following, B_X for a normed space X will denote the closed unit ball $\overline{K_1(0_X)} \subset X$.

Theorem 6.20 (Alaoglu)**.** *Let E be a normed space and let $M \subset E^*$ be bounded as well as closed in $\sigma(E^*, E)$. Then M is $\sigma(E^*, E)$-compact. In particular, B_{E^*} is $\sigma(E^*, E)$-compact.*

Proof. Put $c := \sup\{\|f\| : f \in M\}$ and for $x \in E$ let $A_x := \{z \in \mathbb{K} : |z| \le c\|x\|\}$. By Theorem 6.19, $A = \prod_{x \in E} A_x$, endowed with the product topology $\mathcal{T}_A$, is compact ($A_x \subset \mathbb{K}$ is bounded and closed, thus compact). Define the mapping

$$\varphi \colon M \to A, \ \varphi(f) = (f(x))_{x \in E}.$$

Indeed, φ is well-defined because of $|f(x)| \leq \|f\|\|x\| \leq c\|x\|$ implying $f(x) \in A_x$. Evidently, f is injective as $\varphi(f) = \varphi(g)$ means $f(x) = g(x)$ for all $x \in E$, thus $f = g$. In fact, φ is a homeomorphism between M and $\varphi(M)$. The surjectivity is clear. To prove that φ is continous, we first conclude from Lemma 6.14 and 6.17 that the collection of sets of the form

$$V(a, x, \varepsilon) = \{g \in E^* : |g(x) - a| < \varepsilon\},$$

$a \in \mathbb{K}$, $x \in E$, $\varepsilon > 0$, is a subbasis for $\sigma(E^*, E)$ and the sets

$$W(a, x, \varepsilon) = \{(z_y)_{y\in E} : |z_x - a| < \varepsilon\},$$

$a \in \mathbb{K}$, $x \in E$, $\varepsilon > 0$, form a subbasis for $\mathcal{T}_A$. We have

$$\varphi(V(a, x, \varepsilon) \cap M) = W(a, x, \varepsilon) \cap \varphi(M). \tag{6.1}$$

Together with the injectivity of φ, this implies that both φ and φ^{-1} are continuous[3].
We will now prove that $\varphi(M)$ is closed in A. We will use that $x \in \overline{A}$ if and only if $U \cap A \neq \emptyset$ for all open neighborhoods U of x (this is a tutorial exercise). Let $a = (a_x)_{x\in E} \in \overline{\varphi(M)}$ and define the functional

$$f : E \to \mathbb{K}, \;\; x \mapsto a_x .$$

To show that f is linear, let $\varepsilon > 0$ be arbitrary and let $x, y \in E, \lambda, \mu \in \mathbb{K}$. Then the set

$$W := W(a_x, x, \varepsilon) \cap W(a_y, y, \varepsilon) \cap W(a_{\lambda x+\mu y}, \lambda x + \mu y, \varepsilon)$$

is a $\mathcal{T}_A$-open neighbourhood of a (a is in all the sets by definition and finite intersections of open sets are open). Therefore, there exists $g \in M$ such that $\varphi(g) \in W$. Hence, we have

$$\begin{aligned}|a_{\lambda x+\mu y} - (\lambda a_x + \mu a_y)| &\leq |a_{\lambda x+\mu y} - g(\lambda x + \mu y)| + |\lambda g(x) - \lambda a_x| + |\mu g(y) - \mu a_y)| \\ &\leq \varepsilon + |\lambda|\varepsilon + |\mu|\varepsilon .\end{aligned}$$

Since ε was chosen arbitrarily, we conclude that $a_{\lambda x+\mu y} = \lambda a_x + \mu a_y$, which is linearity. Moreover, we have $|f(x)| = |a_x| \leq c\|x\|$ for all $x \in E$, thus f is bounded. Now let $U \in \sigma(E^*, E)$ be a neighborhood of f. Then there exist $a_1, \dots, a_n \in \mathbb{K}$ and $x_1, \dots, x_n \in E$ and $\varepsilon > 0$ such that $f \in \bigcap_{k=1}^n V(a_k, x_k, \varepsilon) \subset U$. This is because the sets of that form form a basis of $\sigma(E^*, E)$. This implies $|a_{x_k} - a_k| = |f(x_k) - a_k| < \varepsilon$ for all k and thus

$$a \in \bigcap_{k=1}^{n} W(a_k, x_k, \varepsilon) .$$

Since $a \in \overline{\varphi(M)}$, it follows that

$$\varphi(M) \cap \bigcap_{k=1}^{n} W(a_k, x_k, \varepsilon) \neq \emptyset .$$

From (6.1) it is seen that also $M \cap \bigcap_{k=1}^n V(a_k, x_k, \varepsilon)$ is nonempty. As M is $\sigma(E^*, E)$-closed, we conclude that $f \in M$ (any neighborhood U of f has non-empty intersection with M, thus $f \in \overline{M} = M$.) This means $a = \varphi(f) \in \varphi(M)$ and thus $\varphi(M)$ is closed. Since A is compact, $\varphi(M)$ as a closed subset is also. As φ is a homeomorphism from M to $\varphi(M)$, the set M is compact in $\sigma(E^*, E)$.

[3] One easily proves that $f : X \to Y$ is continous if for a subbasis γ, $f^{-1}(V)$ is open in X for all $V \in \gamma$

Now we prove that B_{E^*} is compact. It is sufficient to prove that it is weak*-closed, or equivalently, that

$$\{f \in E^* : \|f\| > 1\}$$

is open. Let $f \in E^*$ have norm larger than 1. Then there exists $x \in E$ with $\|x\| = 1$ with $|f(x)| > 1$. And since

$$V(f(x), x, |f(x)| - 1) \cap B_{E^*} = \emptyset$$

the claim is proven. □

Lemma 6.21. *Let E be a normed space, $\varphi \in B_{E^{**}}$, $f_1, \dots, f_n \in E^*$, and define $h\colon E \to \mathbb{R}$ by*

$$h(x) = \sum_{i=1}^{n} |\varphi(f_i) - f_i(x)|^2 .$$

Then

$$\inf_{\|x\| \le 1} h(x) = 0 .$$

Proof. (missing) □

Corollary 6.22. *Let E be a normed space. Then $\widehat{B_E}$ is $\sigma(E^{**}, E^*)$-dense in $B_{E^{**}}$.*

Proof. For $\varphi \in B_{E^{**}}$ define (for $\varepsilon > 0$, $f_1, \dots, f_n \in E^*$)

$$\mathcal{U}(\varphi, f_1, \dots, f_n, \varepsilon) := \{\psi \in E^{**} : |\psi(f_i) - \varphi(f_i)| < \varepsilon \text{ for } i = 1, \dots, n\} .$$

Then the sets $\mathcal{U}(\varphi, f_1, \dots, f_n, \varepsilon)$ are a basis for the neighborhoods of φ in $\sigma(E^{**}, E^*)$. By Lemma 6.21,there exists an $x \in B_E$, such that

$$|\varphi(f_i) - \hat{x}(f_i)| \le \sqrt{h(x)} < \varepsilon$$

for $i = 1, \dots, n$. Hence $\hat{x} \in \mathcal{U}(\varphi, f_1, \dots, f_n, \varepsilon)$. (and $\hat{x} \in \widehat{B_E}$) □

Theorem 6.23. *Let E be a normed space. Then the following are equivalent:*

(i) The space E is reflexive.

(ii) The unit ball B_E is $\sigma(E, E^*)$-compact.

Proof. (i)⇒(ii). First, we have

$$\Lambda_E(\mathcal{U}(x, f_1, \dots, f_n, \varepsilon)) = \hat{E} \cap \mathcal{U}(\hat{x}, f_1, \dots, f_n, \varepsilon).$$

We have:

$$\Lambda_E\colon (E, \sigma(E, E^*)) \to (E^{**}, \sigma(E^{**}, E^*))$$

is continuous and a homeomorphism. By Theorem 6.20, $B_{E^{**}}$ is $\sigma(E^{**}, E^*)$-compact. By (i) $B_E = \Lambda_E^{-1}(B_{E^{**}})$ is $\sigma(E, E^*)$-compact.

(ii)⇒(i). (ii) implies that $\widehat{B_E}$ is $\sigma(E^{**}, E^*)$-compact, hence $\widehat{B_E}$ is closed. By Corollary 6.22, $\widehat{B_E}$ is $\sigma(E^{**}, E^*)$-dense in $B_{E^{**}}$. This implies that

$$\widehat{B_E} = B_{E^{**}}.$$

and thus $\hat{E} = E^{**}$, which is (i). □

7 Hilbert Spaces and Riesz Representation Theorem

Definition 7.1. Let E be a linear space over $\mathbb{K}$. A *Hermitian form* on E is a map

$$\langle \cdot\,,\cdot\rangle \colon E \times E \to \mathbb{K}$$

satisfying

(a) the homogeneity relation $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$ in its first argument,

(b) the additivity relation $\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$ in its first argument and

(c) the hermiticity relation $\langle x, y\rangle = \overline{\langle y, x\rangle}$

for all $x, x', y \in E$, $\lambda \in \mathbb{K}$. The Hermitian form $\langle \cdot\,,\cdot\rangle$ is called *positive semidefinite*, if

$$\langle x, x\rangle \geq 0$$

for all $x \in E$. If further $\langle x, x\rangle = 0 \Leftrightarrow x = 0$, $\langle \cdot\,,\cdot\rangle$ is called *positive definite*. A *scalar product* (inner product) is a positive definite Hermitian form. If $\langle \cdot\,,\cdot\rangle$ is positive definite, then $(E, \langle \cdot\,,\cdot\rangle)$ is called an *inner product space* (space with a scalar product).

Lemma 7.2 (Cauchy-Schwarz Inequality)**.** *Let E be a $\mathbb{K}$-vector space and let $\langle \cdot\,,\cdot\rangle$ be a positive semidefinite Hermitian form on E. Then*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle$$

for all $x, y \in E$. If $\langle \cdot\,,\cdot\rangle$ is positive definite, then equality holds if and only if x and y are linearly dependent.

Proof. For the sake of brevity, we write $\|x\| := \sqrt{\langle x, x\rangle}$ for $x \in E$. As will be seen in the next lemma, this is a norm if $\langle \cdot\,,\cdot\rangle$ is positive definite. By sesqui-linearity and hermiticity of $\langle \cdot\,,\cdot\rangle$ we have for $x, y \in E$:

$$\begin{aligned}\Big\|\|y\|^2 x - \langle x, y\rangle y\Big\|^2 &= \|y\|^4\|x\|^2 - 2\operatorname{Re}\Big\langle \|y\|^2 x, \langle x, y\rangle y\Big\rangle + |\langle x, y\rangle|^2\|y\|^2 \\ &= \|y\|^4\|x\|^2 - \|y\|^2|\langle x, y\rangle|^2 = \|y\|^2\left(\|x\|^2\|y\|^2 - |\langle x, y\rangle|^2\right) .\end{aligned}$$

This proves the Cauchy-Schwarz inequality and also that equality in it, in the positive definite case, implies that x and y are linearly dependent. Finally, it is clear that $|\langle x, y\rangle|^2 = \|x\|^2\|y\|^2$ if x and y are linearly dependent. □

Lemma 7.3. *Let $\langle \cdot\,,\cdot\rangle$ be a scalar product on E. Then*

$$\|x\| := \sqrt{\langle x, x\rangle}$$

defines a norm on E, and the map

$$(x, y) \mapsto \langle x, y\rangle\,,\ E \times E \to \mathbb{K}$$

is continuous.

Proof. By positive definiteness $\|x\| = 0$ is equivalent to $x = 0$. Also

$$\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x\rangle} = \sqrt{|\lambda|^2 \langle x, x\rangle} = |\lambda|\|x\| .$$

By Lemma 7.2,

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

Thus, $\|\cdot\|$ is a norm. For continuity we observe that

$$\begin{aligned}|\langle x,y\rangle - \langle x_0,y_0\rangle| &\le |\langle x-x_0, y\rangle| + |\langle x_0, y-y_0\rangle| \le \|x-x_0\|\|y\| + \|x_0\|\|y-y_0\| \\ &\le \|x-x_0\|\|y-y_0\| + \|x-x_0\|\|y_0\| + \|x_0\|\|y-y_0\| .\end{aligned}$$

This is small, if both $\|x-x_0\|$ and $\|y-y_0\|$ are small. □

Definition 7.4. Let $(\mathcal{H}, \langle\cdot\,,\cdot\rangle)$ be an inner product space. If the normed space $(\mathcal{H}, \|\cdot\|)$ with $\|x\| = \sqrt{\langle x,x\rangle}$ is complete, then $\mathcal{H}$ is called *Hilbert space*.

Lemma 7.5. *Let $(E, \|\cdot\|)$ be a normed space. Then the following are equivalent:*

(i) There exists a scalar product $\langle\cdot\,,\cdot\rangle$ on E such that

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

for all $x \in E$.

(ii) For $\|\cdot\|$, the parallelogram identity

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

holds for all $x, y \in E$.

Proof. (i) $\Rightarrow$ (ii). This is an easy calculation.
(ii) $\Rightarrow$ (i). First, assume $\mathbb{K} = \mathbb{C}$. Let us show that

$$\langle x,y\rangle := \tfrac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right) ,$$

$x, y \in E$, defines an inner product. First, $\langle x,y\rangle = \overline{\langle y,x\rangle}$ is immediate. Second, to prove that $\langle x+x', y\rangle = \langle x,y\rangle + \langle x',y\rangle$, we have to show that

$$\|x+x'+y\|^2 - \|x+x'-y\|^2 + i\|x+x'+iy\|^2 - i\|x+x'-iy\|^2$$

equals

$$\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2 + \|x'+y\|^2 - \|x'-y\|^2 + i\|x'+iy\|^2 - i\|x'-iy\|^2 .$$

By (ii), we have for $z \in E$

$$\begin{aligned}\|x+x'+z\|^2 &= \tfrac{1}{2}\|x+x'+z\|^2 + \tfrac{1}{2}\|x+x'+z\|^2 \\ &= \|x+z\|^2 + \|x'\|^2 - \tfrac{1}{2}\|x+z-x'\|^2 + \|x'+z\|^2 + \|x\|^2 - \tfrac{1}{2}\|x'+z-x\|^2 .\end{aligned}$$

Choose $z = y$ and $z = -y$ and substract. Then

$$\|x+x'+y\|^2 - \|x+x'-y\|^2 = \|x+y\|^2 + \|x'+y\|^2 - (\|x-y\|^2 + \|x'-y\|^2) .$$

Next, choose $z = iy$ and $z = -iy$ and substract. This proves

$$\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle \,.$$

Third, show that

$$\langle \lambda x, y\rangle = \lambda\langle x, y\rangle \tag{7.1}$$

for all $\lambda \in \mathbb{C}$, $x, y \in E$. We already proved that $\langle mx, y\rangle = m\langle x, y\rangle$ for all $m \in \mathbb{N}$, $x, y \in E$. Since obviously $\langle -x, y\rangle = -\langle x, y\rangle$, this also holds for $m \in \mathbb{Z}$. For $m, n \in \mathbb{Z}$, $n \neq 0$, this implies

$$n\left\langle \frac{m}{n}x, y\right\rangle = \langle mx, y\rangle = m\langle x, y\rangle$$

and thus

$$\left\langle \frac{m}{n}x, y\right\rangle = \frac{m}{n}\langle x, y\rangle \,.$$

We know that

$$\lambda \mapsto \langle \lambda x, y\rangle - \lambda\langle x, y\rangle$$

is continuous (since $\|\cdot\|$ is continuous) and zero on $\mathbb{Q}$, hence on all $\mathbb{R}$. Also $\langle ix, y\rangle = i\langle x, y\rangle$ follows by definition of $\langle\cdot\,,\cdot\rangle$, so that (7.1) is proved.

For $\mathbb{K} = \mathbb{R}$ we define

$$\langle x, y\rangle := \frac{1}{2}\left(\|x+y\|^2 - \|x-y\|^2\right) \,.$$

Then, similar arguments as above show that $\langle\cdot\,,\cdot\rangle$ is an inner product inducing the norm $\|\cdot\|$. □

Remark 7.6. Let $(E, \langle\cdot\,,\cdot\rangle)$ be an inner product space and $\Lambda_E\colon E \to E^{**}$ the canonical embedding of E into E^{**}. Endow E with $\|\cdot\| = \sqrt{\langle\cdot\,,\cdot\rangle}$. The completion of $(E, \|\cdot\|)$ is given by

$$\mathcal{H} := \overline{\Lambda_E(E)} \subset E^{**} \,.$$

As Λ_E is isometric, the parallelogram identity also holds for $\Lambda_E(E)$. Since the norm on E^{**} is continuous, this identity also holds for $\mathcal{H}$. On $\Lambda_E(E)$ we set

$$\langle \hat{x}, \hat{y}\rangle := \langle x, y\rangle \,,$$

$x, y \in E$. By continuity, this inner product on $\Lambda_E(E)$ can be extended to $\mathcal{H}$, so that $(\mathcal{H}, \langle\cdot\,,\cdot\rangle)$ is a Hilbert space, the so-called *Hilbert space completion of E*.

Lemma 7.7. *Let $\mathcal{H}$ be a Hilbert space, $K \subset \mathcal{H}$, $K \neq \emptyset$, convex and closed in $\mathcal{H}$. Then there exists a* unique *$x \in K$ with*

$$\inf_{y\in K} \|y\| = \|x\| \,.$$

Proof. Set $d := \inf_{y\in K}\|y\|$ and let $(x_n) \subset K$ with $\|x_n\| \xrightarrow{n\to\infty} d$. We show that (x_n) is s Cauchy-sequence: Since K is convex, we have $\frac{1}{2}(x_n + x_m) \in K$ for all $m, n \in \mathbb{N}$. Thus,

$$\|x_n + x_m\| \geq 2d \,. \tag{7.2}$$

By the parallelogram identity we have

$$0 \leq \|x_n - x_m\|^2 = 2(\|x_n\|^2 + \|x_m\|^2) - \|x_n + x_m\|^2 \leq 2(\|x_n\|^2 + \|x_m\|^2) - 4d^2 \,. \tag{7.3}$$

For each $\varepsilon > 0$, there exists N_ε with $\|x_n\|^2 \leq d^2 + \frac{\varepsilon}{4}$ for all $n \geq N_\varepsilon$. By (7.3) we have

$$0 \leq \|x_n - x_m\|^2 \leq 2(2d^2 + \tfrac{\varepsilon}{2}) - 4d^2 = \varepsilon$$

for all $n, m \geq N_\varepsilon$. Hence (x_n) is a Cauchy sequence. Let $x := \lim_{n\to\infty} x_n$. Then $x \in K$ since K is closed, and

$$\|x\| = \lim_{n\to\infty} \|x_n\| = d\,.$$

It remains to prove uniqueness: For this let $x, y \in K$ with $\|x\| = \|y\| = d$. Since $\frac{1}{2}(x+y) \in K$ by convexity, it follows that

$$d^2 \leq \left\|\frac{1}{2}(x+y)\right\|^2 = 2\left(\left\|\frac{x}{2}\right\|^2 + \left\|\frac{y}{2}\right\|^2\right) - \left\|\frac{1}{2}(x-y)\right\|^2 = \tfrac{d^2}{2} + \tfrac{d^2}{2} - \left\|\frac{x-y}{2}\right\|^2 .$$

This implies $x = y$. □

Definition 7.8. Let $\mathcal{H}$ be an inner product space. Then $x, y \in \mathcal{H}$ are called *orthogonal* $(x \perp y)$ if

$$\langle x, y\rangle = 0\,.$$

For $M \subset \mathcal{H}$ we define by

$$M^\perp := \{y \in \mathcal{H} : \langle y, x\rangle = 0 \text{ for all } x \in M\}$$

the *orthogonal complement of M.*

Lemma 7.9. *Let $\mathcal{H}$ be an inner product space, $M \subset \mathcal{H}$. Then*

$$M \cap M^\perp = \{0\}\,,$$

and $M^\perp$ is a closed linear subspace of $\mathcal{H}$.

Proof. The first part follows directly from definition 7.8. The second part follows from the properties of the inner product $\langle\cdot\,,\cdot\rangle$. □

Let M and N be subspaces of an inner product space with $M \cap N = \{0\}$. Then the sum $M + N$ is direct which we express by writing $M \dot{+} N$. If, in addition, $M \perp N$, then we write $M \oplus N$ for the *orthogonal direct sum.* By Lemma 7.9, $M + M^\perp = M \oplus M^\perp$.

Lemma 7.10. *Let M be a closed subspace of the Hilbert space $\mathcal{H}$. Then*

$$\mathcal{H} = M \oplus M^\perp\,.$$

Proof. Let $x \in \mathcal{H}$ and set

$$K := \{x - y : y \in M\} = x - M\,.$$

In particular K is closed. Since for $y_1, y_2 \in M$

$$\lambda(x - y_1) + (1-\lambda)(x - y_2) = x - (\lambda y_1 + (1-\lambda)y_2) \in K\,,$$

the set K is convex. By 7.7 there exists a unique $x_2 \in K$ with

$$\|x_2\| = \inf_{y\in K} \|y\|\,.$$

By definition of K, $x_2 = x - x_1$ for some $x_1 \in M$. Hence, $x = x_1 + x_2$, and we are done if we can show that $x_2 \in M^\perp$. For this, let $y \in M \setminus \{0\}$ be arbitrary. Then

$$\|x_2\|^2 \le \| \underbrace{x_2 - \lambda y}_{\in K - M = K} \|^2 \tag{7.4}$$

for all $\lambda \in \mathbb{K}$. Choose $\lambda = \frac{\langle x_2, y\rangle}{\langle y, y\rangle}$. By (7.4)

$$0 \le \left\| x_2 - \tfrac{\langle x_2, y\rangle}{\langle y, y\rangle} y \right\|^2 - \|x_2\|^2 = -\tfrac{|\langle x_2, y\rangle|^2}{\langle y, y\rangle} \le 0\,.$$

This implies $\langle x_2, y\rangle = 0$, so $x_2 \in M^\perp$. ☐

Theorem 7.11 (Riesz Representation Theorem)**.** *Let $\mathcal{H}$ be a Hilbert space. For $y \in \mathcal{H}$ define*

$$f_y \colon \mathcal{H} \to K\,,\ \ f_y(x) := \langle x, y\rangle\,.$$

Then $f_y \in \mathcal{H}^*$ and $\|f_y\| = \|y\|$.
Conversely, for each $f \in \mathcal{H}^*$ there exists a unique $y \in \mathcal{H}$ such that

$$f = f_y\,.$$

Finally, the Riesz map

$$y \mapsto f_y\,,\ \ \mathcal{H} \to \mathcal{H}^*\,,$$

is conjugate linear, that is, $f_{y_1} + f_{y_2} = f_{y_1 + y_2}$ and $f_{\lambda y} = \overline{\lambda} f_y$.

Proof. By definition of an inner product, f_y is linear and $y \mapsto f_y$ is conjugate linear. By Cauchy-Schwarz (Lemma 7.2), we have

$$|f_y(x)| = |\langle x, y\rangle| \le \|x\|\|y\|$$

and thus $\|f_y\| \le \|y\|$. Further,

$$|f_y(y)| = \langle y, y\rangle = \|y\|^2 = \|y\|\|y\|$$

and hence $\|f_y\| = \|y\|$.
Now, let $f \in \mathcal{H}^*$, $f \ne 0$, be arbitrary. We have to find $y \in \mathcal{H}$ such that $f(x) = \langle x, y\rangle$ for all $x \in \mathcal{H}$. For this, denote by N the kernel of f and find $z' \in \mathcal{H}$ with $f(z') = 1$. Since N is closed and $\mathcal{H}$ is complete, we can decompose z' as $x + z$ with $x \in N$ and $z \in N^\perp$ (Lemma 7.10). Then

$$1 = f(z') = f(x) + f(z) = f(z).$$

For arbitrary $x \in \mathcal{H}$ we now have $x - f(x)z \in N$ and thus

$$f(x) = \left\langle f(x)z, \|z\|^{-2}z \right\rangle = \left\langle (x - f(x)z) + f(x)z, \|z\|^{-2}z \right\rangle = \left\langle x, \|z\|^{-2}z \right\rangle = \langle x, y\rangle,$$

where $y := \|z\|^{-2}z$. ☐

Remark 7.12. Note that the first claim in Theorem 7.11 also holds if $\mathcal{H}$ is merely an inner product space. If the Riesz map $\mathcal{H} \to \mathcal{H}^*$, $y \mapsto f_y$, is surjective, then $\mathcal{H}$ is a Hilbert space.

Proof. By assumption, $y \mapsto f_y$ is a surjective isometry from $\mathcal{H}$ to $\mathcal{H}^*$. The space $\mathcal{H}^*$ is complete, thus $\mathcal{H}$ is as well. ☐

Theorem 7.13. *Let $\mathcal{H}$ be a Hilbert space and for $f \in \mathcal{H}^*$ let $y_f \in \mathcal{H}$ be defined by $f(x) = \langle x, y_f \rangle$ for all $x \in \mathcal{H}$. Then $\mathcal{H}^*$ is a Hilbert space with inner product*

$$\langle f, g \rangle_{\mathcal{H}^*} := \langle y_g, y_f \rangle.$$

Proof. For $f \neq 0$ we have $y_f \neq 0$ and thus

$$\langle f, f \rangle = \langle y_f, y_f \rangle > 0.$$

Further,

$$\langle f, g \rangle = \langle y_g, y_f \rangle = \overline{\langle y_f, y_g \rangle} = \overline{\langle g, f \rangle}$$

and

$$\begin{aligned}\langle f_1 + f_2, g \rangle &= \langle y_g, y_{f_1+f_2} \rangle = \langle y_g, y_{f_1} + y_{f_2} \rangle \\ &= \langle y_g, y_{f_1} \rangle + \langle y_g, y_{f_2} \rangle = \langle f_1, g \rangle + \langle f_2, g \rangle\end{aligned}$$

and

$$\begin{aligned}\langle \lambda f, g \rangle &= \langle y_g, y_{\lambda f} \rangle = \langle y_g, \overline{\lambda} y_f \rangle \\ &= \lambda \langle y_g, y_f \rangle = \lambda \langle f, g \rangle.\end{aligned}$$

For $f \in \mathcal{H}^*$ we now have

$$\langle f, f \rangle = \langle y_f, y_f \rangle = \|y_f\|^2 = \|f\|^2 \quad \text{(Riesz map is an isometry)}.$$

Thus $\langle \cdot , \cdot \rangle_{\mathcal{H}^*}$ really induces the dual norm. Also, $\mathcal{H}^*$ is complete. □

Corollary 7.14. *Hilbert spaces are reflexive.*

Proof. Let $\varphi \in \mathcal{H}^{**}$ be arbitrary. There is a unique $f_\varphi \in \mathcal{H}^*$ with

$$\varphi(f) = \langle f, f_\varphi \rangle$$

for all $f \in \mathcal{H}^*$. Using the same notation as above,

$$\hat{y}_{f_\varphi}(f) = f(y_{f_\varphi}) = \langle y_{f_\varphi}, y_f \rangle = \langle f, f_\varphi \rangle = \varphi(f).$$

Thus $\hat{y}_{f_\varphi} = \varphi$. □

Corollary 7.15. *Let $\mathcal{H}$ be a Hilbert space and $L \subset \mathcal{H}$ a linear subspace, $g \in L^*$. Then there is a* unique *$f \in \mathcal{H}^*$ with*

$$f|_L = g \text{ and } \|f\| = \|g\|.$$

Proof. Apart from uniqueness this follows from Hahn-Banach, Theorem 4.5 on page 28. But there is a direct proof: Define $\overline{g} \colon \overline{L} \to \mathbb{K}$ by continuously extending g. Then $\|\overline{g}\| = \|g\|$. So without loss of generality, L is closed, hence a Hilbert space.

By Riesz, there exists a $y \in L$ with $g(x) = \langle x, y \rangle$ for all $x \in L$, and $\|g\| = \|y\|$. Define $f \colon \mathcal{H} \to \mathbb{K}, x \mapsto \langle x, y \rangle$. Then $f|_L = g$ and $\|f\| = \|y\| = \|g\|$.

To prove uniqueness, let $f' \in \mathcal{H}^*$ with $f'|_L = g$ and $\|f'\| = \|g\|$. By Riesz $f'(x) = \langle x, y' \rangle$ for some $y' \in \mathcal{H}$ with $\|y'\| = \|f'\|$. Since $f'|_L = g$, $0 = \langle x, y' - y \rangle$ for all $x \in L$. This means $y' - y \in L^\perp$ and thus $y' = y + z$ with $z \in L^\perp$. Then we have

$$\|f'\|^2 = \|y'\|^2 = \|y + z\|^2 = \|y\|^2 + \|z\|^2 = \|g\|^2 + \|z\|^2,$$

and hence $z = 0 \Rightarrow y = y' \Rightarrow f = f'$. □

Theorem 7.16. *Let $\mathcal{H}$ be a Hilbert space. If $A \in L(\mathcal{H})$ and $\varphi(x,y) := \langle Ax, y\rangle$ for $x, y \in \mathcal{H}$, then φ is a sesquilinear form on $\mathcal{H}$, i.e. linear in x and conjugate linear in y. Further,*

$$\sup_{x,y\neq 0} \frac{|\varphi(x,y)|}{\|x\|\|y\|} = \|A\|. \tag{7.5}$$

Conversely for a sesquilinear form φ with

$$\sup_{x,y\neq 0} \frac{|\varphi(x,y)|}{\|x\|\|y\|} < \infty$$

there exists a unique $A \in L(\mathcal{H})$ with $\varphi(x,y) = \langle Ax, y\rangle$ and (7.5).

Proof. Given some operator $A \in L(\mathcal{H})$, φ defined by $\varphi(x,y) := \langle Ax, y\rangle$ is sesquilinear. Now,

$$\begin{aligned}\|A\| &= \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{x\neq 0} \frac{\|f_{Ax}\|}{\|x\|} = \sup_{x\neq 0} \|x\|^{-1} \sup_{y\neq 0} \frac{|\langle Ax, y\rangle|}{\|y\|} \\ &= \sup_{x,y\neq 0} \frac{|\langle Ax, y\rangle|}{\|x\|\|y\|} = \sup_{x,y\neq 0} \frac{|\varphi(x,y)|}{\|x\|\|y\|}.\end{aligned}$$

Conversely, let φ by a sesquilinear form with $M := \sup_{x,y\neq 0} \frac{|\varphi(x,y)|}{\|x\|\|y\|} < \infty$. Then define

$$f_x(y) := \overline{\varphi(x,y)}.$$

f_x is linear and bounded:

$$|f_x(y)| = |\varphi(x,y)| \leq M\|x\|\|y\|.$$

Thus, $f_x \in \mathcal{H}^*$. By Riesz, there is a unique $z_x \in \mathcal{H}$ with $f_x(y) = \overline{\varphi(x,y)} = \langle y, z_x\rangle$ and $\|f_x\| = \|z_x\|$. Define

$$A\colon \mathcal{H} \to \mathcal{H}, \quad x \mapsto z_x.$$

The map $x \mapsto f_x$ is conjugate linear as well as $f_x \mapsto z_x$, therefore their composition $x \mapsto z_x$ is linear. Also,

$$\varphi(x,y) = \overline{f_x(y)} = \overline{\langle y, z_x\rangle} = \langle z_x, y\rangle = \langle Ax, y\rangle,$$

which in particular implies

$$\|Ax\|^2 = \varphi(x, Ax) \leq M\|x\|\|Ax\|,$$

and thus $\|Ax\| \leq M\|x\|$ for all $x \in \mathcal{H}$. Hence, $A \in L(\mathcal{H})$, and $\|A\| = M$ follows from the first part.

For the uniqueness part, let $B \in L(\mathcal{H})$ with $\varphi(x,y) = \langle Bx, y\rangle$. Then $\langle (A-B)x, y\rangle = 0$ for all $x, y \in \mathcal{H}$. In particular, with $y := (A-B)x$ we conclude that $(A-B)x = 0$ for all $x \in \mathcal{H}$ and therefore $A = B$. □

Remark 7.17. Let I be an index set and $\ell_2(I)$ the set of maps $x\colon I \to \mathbb{K}$ with $x(i) \neq 0$ for only countably many i and

$$\sum_{i\in I} |x(i)|^2 < \infty.$$

Then

$$\langle x, y\rangle := \sum_{i\in I} x(i)\overline{y(i)}, \quad x, y \in \ell_2(I),$$

defines an inner product on $\ell_2(I)$, and $(\ell_2(I), \langle\cdot\,,\cdot\rangle)$ is a Hilbert space. For $I = \mathbb{N}$ this is the usual ℓ_2 defined in chapter 1. In this case, the Riesz Representation Theorem coincides with Theorem 3.14 on page 23 (stating that $\ell_p \cong \ell_q^*$) for $p = q = 2$.

8 Orthogonality and Bases

Definition 8.1. Let E be a normed space, $\{x_i : i \in I\} \subset E$ and $x \in E$. Then the set $\{x_i : i \in I\}$ is called *summable to x* if for every $\varepsilon > 0$ there exists some finite $I_\varepsilon \subset I$ such that for all finite $J \subset I$ with $I_\varepsilon \subset J$ we have

$$\Big\|\sum_{i\in J} x_i - x\Big\| \leq \varepsilon .$$

Obviously, $\{x_i : i \in I\}$ can only be summable to at most one $x \in E$. We write

$$x = \sum_{i\in I} x_i .$$

We call $\{x_i : i \in I\}$ *summable*, if it is summable to some $x \in E$.

Remark 8.2. If $\{x_i\}_{i\in I}$ and $\{y_i\}_{i\in I}$ are summable, so are $\{\lambda x_i\}_{i\in I}$ and $\{x_i + y_i : i \in I\}$. There holds

$$\sum_{i\in I} \lambda x_i = \lambda \sum_{i\in I} x_i$$

as well as

$$\sum_{i\in I} x_i + y_i = \sum_{i\in I} x_i + \sum_{i\in I} y_i .$$

Lemma 8.3. *Let E be a Banach space and $\{x_i\}_{i\in I} \subset E$. Then we have*

(i) *The set $\{x_i\}_{i\in I}$ is summable to some $x \in E$ if and only if for every $\varepsilon > 0$ there exists some finite $I_\varepsilon \subset I$ such that for all finite $J \subset I$ with $I_\varepsilon \cap J = \emptyset$ we have*

$$\Big\|\sum_{i\in J} x_i\Big\| < \varepsilon .$$

(ii) *The set $\{x_i\}_{i\in I}$ is summable to x if and only if there exists a countable set $J \subset I$ with $x_i = 0$ $\forall i \in I \setminus J$ and for any bijection $\mathbb{N} \to J, k \mapsto i_k$ we have*

$$\lim_{n\to\infty} \sum_{k=1}^{n} x_{i_k} = x$$

or, if J is finite, $x = \sum_{j\in J} x_j$.

Proof. (i). Let $\{x_i\}_{i\in I}$ be summable to $x \in E$. For every $\varepsilon > 0$, let $I_\varepsilon \subset I$ be finite such that

$$\Big\|\sum_{i\in L} x_i\Big\| \leq \frac{\varepsilon}{2} \quad \text{for all finite } L \supset I_\varepsilon .$$

Let $J \subset I$ be finite with $J \cap I_\varepsilon = \emptyset$. Then we obtain

$$\Big\|\sum_{i\in J} x_i\Big\| \underset{J\cap I_\varepsilon=\emptyset}{=} \Big\|\sum_{i\in J\cup I_\varepsilon} x_i - \sum_{i\in I_\varepsilon} x_i\Big\| \leq \Big\|\sum_{i\in J\cup I_\varepsilon} x_i - x\Big\| + \Big\|\sum_{i\in I_\varepsilon} x_i - x\Big\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} ,$$

where in the last step we used the fact that $J \cup I_\varepsilon$ and I_ε are both finite supersets of I_ε.

Conversely, let for $n \in \mathbb{N}$ the set $J_n \subset I$ be finite such that for each finite $J \subset I$ with $J \cap J_n = \emptyset$ we have $\|\sum_{i\in J} x_i\| < \frac{1}{n}$. Then

$$(y_n) = \Big(\sum_{i\in J_1\cup\cdots\cup J_n} x_i\Big)_{n\in\mathbb{N}}$$

is a Cauchy sequence since for $m \geq n$ we have

$$\|y_m - y_n\| = \left\| \sum_{i \in \bigcup_{k=1}^m J_k \setminus \bigcup_{k=1}^n J_k} x_i \right\| \leq \frac{1}{n}$$

because $\bigcup_{k=1}^m J_k \setminus \bigcup_{k=1}^n J_k$ is finite and disjoint from J_n. Let $y = \lim_{n\to\infty} y_n$. For $\varepsilon > 0$, let $n \in \mathbb{N}$ be such that $\|y - y_n\| < \frac{\varepsilon}{2}$ and $\|\sum_{i \in L} x_i\| < \frac{\varepsilon}{2}$ for all finite sets $L \subset I$ with $L \cap J_n = \emptyset$. We then have with $I_\varepsilon := \bigcup_{k=1}^n J_k$ for $L \supset I_\varepsilon$

$$\left\| y - \sum_{i \in L} x_i \right\| \leq \left\| y - \sum_{i \in I_\varepsilon} x_i \right\| + \left\| \sum_{i \in L \setminus I_\varepsilon} x_i \right\| < \|y - y_n\| + \frac{\varepsilon}{2},$$

since $(L \setminus I_\varepsilon) \cap J_n = \emptyset$. This means that $y = \sum_{i \in I} x_i$.

(ii). Let $\{x_i\}_{i \in I}$ be summable to x. By (i), for $n \in \mathbb{N}$ we can choose a finite $J_n \subset I$ with $\|\sum_{i \in J} x_i\| \leq \frac{1}{n}$ for all finite $J \subset I$ with $J \cap J_n = \emptyset$.

The set $\{i : x_i \neq 0\}$ is contained in $\bigcup_{n \in \mathbb{N}} J_n$. To prove this, let $i \notin \bigcup_{n \in \mathbb{N}} J_n$ and take $J = \{i\}$. Then $J \cap J_n = \emptyset$ for all $n \in \mathbb{N}$. This implies

$$\left\| \sum_{j \in J} x_j \right\| = \|x_i\| \leq \tfrac{1}{n} \quad \forall n \in \mathbb{N}.$$

Thus, $x_i = 0$ and $\{x_i : x_i \neq 0\}$ is contained in a countable set, and hence is countable.

Let $\mathbb{N} \to \bigcup_{n \in \mathbb{N}} J_n$, $k \mapsto i_k$ be a bijection (if $\bigcup_{n \in \mathbb{N}} J_n$ is infinite). Then

$$x = \lim_{n \to \infty} \sum_{k=1}^n x_{i_k}.$$

This can be seen as follows: For $\varepsilon > 0$, let $I_\varepsilon \subset I$ be finite with $\left\| x - \sum_{j \in J} x_j \right\| < \varepsilon$ for all finite supersets J of I_ε. Then choose $N \in \mathbb{N}$ with

$$I_\varepsilon \cap \{i \in I : x_i \neq 0\} \subset \{x_{i_k} : k = 1, \ldots, N\}.$$

This is possible due to the bijectivity of $k \mapsto i_k$. The injectivity and the above now implies that for every $m \leq N$, we have

$$\left\| x - \sum_{k=1}^m x_{i_k} \right\| < \varepsilon.$$

Conversely, assume that $\{x_i : i \in I\}$ is not summable to x. Then there exists an $\varepsilon > 0$ such that for every finite $I' \subset I$, there is some finite $I'' \supset I'$ with

$$\left\| x - \sum_{i \in I''} x_i \right\| \geq \varepsilon.$$

If $J = \{i \in I, x_i \neq 0\}$ is not countable, we are done. Furthermore, J cannot be finite, otherwise $\{x_i\}_{i \in I}$ would be summable to $\sum_{j \in J} x_j$.

Let $\mathbb{N} \to J$, $l \to j_l$ be a bijection. Set $I_1' = \{j_1\}$ and I_1'' some superset of I_1', say $I_1'' = \{i_1, \ldots i_{n_1}\}$, with

$$\left\| x - \sum_{k=1}^{n_1} x_{i_k} \right\| \geq \varepsilon.$$

Then set $I_2' = I_1'' \cup \{j_r\}$ where r is minimally chosen with $j_r \notin I_1''$. Furthermore let $I_2'' = I_1'' \cup \{i_{n_1+1}, \dots, i_{n_2}\}$ with

$$\left\| x - \sum_{k=1}^{n_2} x_{i_k} \right\| \geq \varepsilon \,.$$

Continue this process inductively.

Then $k \mapsto i_k$ is a bijection (we have never chosen some i twice and $k \mapsto i_k$ is surjective due to $k \in I_k''$) and for all $m \in \mathbb{N}$

$$\left\| x - \sum_{k=1}^{n_m} x_{i_k} \right\| \geq \varepsilon$$

thus $\sum_{k=1}^{n} x_{i_k} \not\to x$. □

Korollar 8.4. *Let $a_i \geq 0$, $i \in I$. Then $\{a_i : i \in I\}$ is summable if and only if*

$$S := \sup \left\{ \sum_{i \in J} a_i : J \subset I \text{ finite} \right\} < \infty \,.$$

In this case, $\{a_i : i \in I\}$ is summable to S.

Proof. If $\{a_i : i \in I\}$ is summable, then $S < \infty$ follows directly from Lemma 8.3(ii). Conversely, assume that $S < \infty$ and let $\varepsilon > 0$. Then there exists a finite $I_\varepsilon \subset I$ such that

$$S - \sum_{i \in I_\varepsilon} a_i < \varepsilon \,.$$

Hence, the same holds with I_ε replaced by any finite $J \supset I_\varepsilon$. This shows that $\{a_i : i \in I\}$ is summable to S. □

Remark 8.5. Let $\mathcal{H}$ be an inner product space and let $\{x_i\}_{i \in I}$ be summable in $\mathcal{H}$. Then for all $y \in \mathcal{H}$

$$\langle \cdot\,, \cdot \rangle \sum_{i \in I} x_i, y = \sum_{i \in I} \langle \cdot\,, \cdot \rangle x_i, y \,.$$

Proof. Let $J \subset I$ be finite and $\|x - \sum_{j \in J} x_j\| < \varepsilon$. Then for every $y \in \mathcal{H}$ we have

$$\left| \langle \cdot\,, \cdot \rangle x, y - \sum_{j \in J} \langle \cdot\,, \cdot \rangle x_j, y \right| \leq \|y\| \left\| x - \sum_{j \in J} x_j \right\| \leq \|y\| \varepsilon \,.$$

This proves the claim. □

Lemma 8.6. *Let $\mathcal{H}$ be a Hilbert space and $\{x_i\}_{i \in I}$ a family of pairwise orthogonal elements in $\mathcal{H}$. Then the following are equivalent:*

(i) *The set $\{x_i\}_{i \in I}$ is summable in $\mathcal{H}$.*

(ii) *The set $\{\|x_i\|^2\}_{i \in I}$ is summable in $\mathbb{R}$.*

Moreover $\|\sum_{i \in I} x_i\|^2 = \sum_{i \in I} \|x_i\|^2$. In particular, summability is not equivalent to absolute convergence in infinite dimensions.

Proof. (i)⇒(ii). For $\varepsilon > 0$, let I_ε be finite satisfying

$$\left\|\sum_{j\in J} x_j\right\| < \varepsilon \quad \text{for all finite } J \text{ with } J \cap I_\varepsilon = \emptyset\,.$$

For $x, y \in \mathcal{H}$ with $x \perp y$ we have ("Pythagoras")

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

Hence $\sum_{j\in J}\|x_j\|^2 = \left\|\sum_{j\in J} x_j\right\|^2 < \varepsilon^2$ for all finite J with $J \cap I_\varepsilon = \emptyset$ - this is (ii).

(ii)⇒(i). If $\{\|x_i\|^2\}_{i\in I}$ is summable in $\mathbb{R}$, then for all $\varepsilon > 0$ there exists a finite $I_\varepsilon \subset I$, such that for every finite J with $J \cap I_\varepsilon = \emptyset$,

$$\left\|\sum_{j\in J} x_j\right\|^2 = \sum_{j\in J}\|x_j\|^2 < \varepsilon\,.$$

Thus, $\{x_i\}_{i\in I}$ is summable in $\mathcal{H}$. □

Definition 8.7. Let $\mathcal{H}$ be a an inner product space and $M \subset \mathcal{H}$. The set M is called an *orthonormal system* (ONS) if

$$x \perp y \quad \forall x \neq y \in M \quad \text{and} \quad \|x\| = 1 \quad \forall x \in M\,.$$

An orthonormal system M in $\mathcal{H}$ is called *complete*, *maximal* or an *orthonormal basis* (ONB) if for all orthonormal systems N with $M \subset N$ we have $M = N$.

Remark 8.8. An orthonormal system $M \subset \mathcal{H}$ is complete if and only if $M^\perp = \{0\}$.

Proof. Let M be complete and let $y \in M^\perp$. If $y \neq 0$, then $N := M \cup \{y/\|y\|\}$ is an ONS with $M \subsetneq N$. Therefore, $y = 0$ follows.

Conversely, assume that $M^\perp = \{0\}$, and let $N \supset M$ be an ONS. Suppose there exists $y \in N \setminus M$. Then $\langle y, x\rangle = 0$ for all $x \in M$. Hence, $y = 0$ follows. But $\|y\| = 1$ as N is an ONS. This implies $M = N$. □

Theorem 8.9. *Let $\mathcal{H}$ be an inner product space and $\{x_i\}_{i\in I}$ be an orthonormal system. Then there hold*

(i) *the* Bessel inequality

$$\sum_{i\in I}|\langle x, x_i\rangle|^2 \leq \|x\|^2$$

for all $x \in \mathcal{H}$.

(ii) *and the* Parseval identity, *i. e.*

$$\sum_{i\in I}|\langle x, x_i\rangle|^2 = \|x\|^2$$

holds if and only if

$$x = \sum_{i\in I}\langle x, x_i\rangle x_i\,.$$

Proof. (i). For all finite $J \subset I$, we have

$$\begin{aligned} 0 &\le \Big\| x - \sum_{i\in J} \langle x, x_i\rangle x_i \Big\|^2 \\ &= \|x\|^2 - \sum_{i\in J} \overline{\langle x, x_i\rangle}\langle x, x_i\rangle - \sum_{i\in J}\langle x, x_i\rangle \overline{\langle x, x_i\rangle} + \sum_{i,j\in J} \langle x, x_i\rangle \overline{\langle x, x_j\rangle}\langle x_i, x_j\rangle \\ &= \|x\|^2 - \sum_{i\in J} |\langle x, x_i\rangle|^2 . \end{aligned}$$

This implies that $\sum_{i\in I} |\langle x, x_i\rangle|^2$ exists and satisfies the Bessel inequality for all $x \in \mathcal{H}$.
(ii). From

$$\Big\| x - \sum_{i\in J} \langle x, x_i\rangle x_i \Big\|^2 = \|x\|^2 - \sum_{i\in J} |\langle x, x_i\rangle|^2$$

for all finite $J \subset I$, we obtain that

$$\|x\|^2 = \sum_{i\in I} |\langle x, x_i\rangle|^2 ,$$

is equivalent to: For all $\varepsilon > 0$, finite $I_\varepsilon \subset I$ and finite $J \supset I_\varepsilon$ we have

$$\|x\|^2 - \sum_{i\in J} |\langle x, x_i\rangle|^2 \le \varepsilon .$$

This is again equivalent to: For all $\varepsilon > 0$, finite $I_\varepsilon \subset I$ and finite $J \supset I_\varepsilon$ we have

$$\Big\| x - \sum_{i\in J} \langle x, x_i\rangle x_i \Big\|^2 \le \varepsilon .$$

And this is equivalent to the summability to x of $\{\langle x, x_i\rangle x_i : i \in I\}$. □

Definition 8.10. The *orthogonal sum* $\mathcal{H}$ of Hilbert spaces $\mathcal{H}_i$, $i \in I$, denoted by

$$\mathcal{H} = \bigoplus_{i\in I} H_i,$$

is defined as the set of all

$$x = (x_i)_{i\in I} \in \prod_{i\in I} \mathcal{H}_i$$

with the property that $\sum_{i\in I} \|x_i\|^2$ exists (i.e., $\{\|x_i\|^2 : i \in I\}$ is summable).

Remark 8.11. The orthogonal sum $\mathcal{H} = \bigoplus_{i\in I} H_i$ is a linear space, since for any elements $(x_i)_{i\in I}, (y_i)_{i\in I} \in \mathcal{H}$ we have

$$\begin{aligned} \sum_{i\in I} \|x_i + y_i\|^2 &\le \sum_{i\in I} \|x_i\|^2 + 2\sum_{i\in I} |\langle x_i, y_i\rangle| + \sum_{i\in I} \|y_i\|^2 \\ &\le \sum_{i\in I} \|x_i\|^2 + 2\sum_{i\in I} \|x_i\|\|y_i\| + \sum_{i\in I} \|y_i\|^2 \\ &\le \sum_{i\in I} \|x_i\|^2 + 2\left(\sum_{i\in I} \|x_i\|^2\right)^{1/2} \left(\sum_{i\in I} \|y_i\|^2\right)^{1/2} + \sum_{i\in I} \|y_i\|^2. \end{aligned}$$

Also $\langle \cdot, \cdot\rangle$, defined by

$$\langle (x_i)_i, (y_i)_i\rangle = \sum_{i\in I} \langle x_i, y_i\rangle,$$

is an inner product on $\mathcal{H}$.

Lemma 8.12. *Let $\mathcal{H}_i$, $i \in I$, be Hilbert spaces.*

(i) *The space $\mathcal{H} = \bigoplus_{i\in I} \mathcal{H}_i$ is a Hilbert space.*

(ii) *We can interpret each $\mathcal{H}_i$ as a subspace of $\mathcal{H}$, i.e. there is an isometric isomorphism from $\mathcal{H}_i$ to some subspace of $\mathcal{H}$.*

Proof. (i). Let $(x_n)_n \subset \mathcal{H}$ be a Cauchy-sequence in $\mathcal{H}$, $x_n = (x_{n,i})_i$, and let $\varepsilon > 0$. Then there exists an $N_\varepsilon \in \mathbb{N}$ with

$$\|x_n - x_m\|^2 = \|(x_{n,i})_i - (x_{m,i})_i\|^2 \leq \varepsilon \quad \forall n \geq N_\varepsilon.$$

Then we obtain for all $i \in I$

$$\|x_{n,i} - x_{m,i}\|^2 \leq \sum_{j\in I} \|x_{n,j} - x_{m,j}\|^2 = \|(x_{n,j})_j - (x_{m,j})_j\|^2 \leq \varepsilon.$$

Hence $(x_{n,i})_n$ is a Cauchy-sequence in $\mathcal{H}_i$. Now, let

$$x_i := \lim_{n\to\infty} x_{n,i}.$$

For each finite $J \subset I$ and $n \geq N_\varepsilon$ we have

$$\sum_{i\in J} \|x_{n,i} - x_i\|^2 = \lim_{m\to\infty} \sum_{i\in J} \|x_{n,i} - x_{m,i}\|^2 \leq \varepsilon$$

and thus

$$\left(\sum_{i\in J} \|x_i\|^2\right)^{1/2} \leq \left(\sum_{i\in J} \|x_{n,i} - x_i\|^2\right)^{1/2} + \left(\sum_{i\in J} \|x_{n,i}\|^2\right)^{1/2} \leq \varepsilon^{1/2} + \|x_n\|.$$

This shows that both $\{\|x_{n,i} - x_i\|^2 : i \in I\}$ and $\{\|x_i\|^2 : i \in I\}$ are summable (see Corollary 8.4) and

$$\sum_{i\in I} \|x_{n,i} - x_i\|^2 \leq \varepsilon \quad \forall n, m \geq N_\varepsilon.$$

Hence, $x := (x_i)_i \in \mathcal{H}$ and $\|x_n - x\| \to 0$ as $n \to \infty$.
(ii). Let $i \in I$ and consider the map

$$\mathcal{H}_i \to \mathcal{H},\ y \mapsto (x_j)_{j\in I} \text{ with } x_j = \delta_{ij} y.$$

This map is linear and isometric. □

Theorem 8.13. *For an orthonormal system $\{x_i : i \in I\}$ in a Hilbert space $\mathcal{H}$, the following conditions are equivalent:*

(i) *The system $\{x_i : i \in I\}$ is complete.*

(ii) *The space $\operatorname{span}\{x_i : i \in I\}$ is dense in $\mathcal{H}$.*

(iii) *If $\mathcal{H}_i := \operatorname{span} x_i$, $i \in I$, then $\bigoplus_{i\in I} \mathcal{H}_i$ is isometrically isomorphic to $\mathcal{H}$ by*

$$(\lambda_i x_i)_{i\in I} \mapsto \sum_{i\in I} \lambda_i x_i.$$

(iv) *For all* $x \in \mathcal{H}$*,*

$$x = \sum_{i\in I} \langle x, x_i\rangle x_i \,.$$

(v) *For all* $x, y \in \mathcal{H}$*,*

$$\langle x, y\rangle = \sum_{i\in I} \langle x, x_i\rangle\langle x_i, y\rangle \,.$$

Proof. (i)⇒(ii). Define $L := \operatorname{span}\{x_i : i \in I\}$. Then, by Lemma 7.10,

$$\mathcal{H} = \overline{L} \oplus \overline{L}^{\perp} \,.$$

But $\overline{L}^{\perp} \subset \{x_i : i \in I\}^{\perp} = \{0\}$ by Remark 8.8. Hence $\overline{L} = \mathcal{H}$.
(ii)⇒(iii). Let F be the linear subspace of all $(\lambda_i x_i)_{i\in I} \in \bigoplus_{i\in I} \mathcal{H}_i$ with $\lambda_i \neq 0$ for an at most finite number of indices $i \in I$. Since

$$\Big\|\sum_{i\in I} \lambda_i x_i\Big\|^2 = \sum_{i,j\in I} \lambda_i \bar{\lambda}_j \langle x_i, x_j\rangle = \sum_{i\in I} |\lambda_i|^2 = \|(\lambda_i x_i)_{i\in I}\|^2 \,,$$

the linear map

$$\ell\colon (\lambda_i x_i)_i \mapsto \sum_{i\in I} \lambda_i x_i \quad F \to \ell(F) =: L$$

is isometric. Since F is dense in $\bigoplus_{i\in I} \mathcal{H}_i$ (by definition of F), there exists a unique extension of ℓ to

$$\varphi\colon \bigoplus_{i\in I} \mathcal{H}_i \to \mathcal{H} \,,$$

which is linear and isometric, too. By (ii), also L is dense in $\mathcal{H}$ and $\varphi(\bigoplus_{i\in I} \mathcal{H}_i) \subset \mathcal{H}$ is complete, hence closed. Thus, $\varphi(\bigoplus_{i\in I} \mathcal{H}_i) = \mathcal{H}$. It remains to prove that for all $x = (\lambda_i x_i)_i \in \bigoplus_{i\in I} \mathcal{H}_i$, we have

$$\varphi(x) = \sum_{i\in I} \lambda_i x_i \,,$$

i.e. $\{\lambda_i x_i : i \in I\}$ is summable to $\varphi(x)$. For this, let $\varepsilon > 0$. Then there exists a finite $I_\varepsilon \subset I$ with

$$\sum_{i\notin I_\varepsilon} |\lambda_i|^2 = \sum_{i\notin I_\varepsilon} \|\lambda_i x_i\|^2 < \varepsilon \,.$$

For a finite subset $J \supset I_\varepsilon$, $J \subset I$, define

$$y_i = \begin{cases} \lambda_i x_i & i \in J \,, \\ 0 & \text{otherwise.} \end{cases}$$

We have

$$\|x - y\|^2 = \sum_{i\notin J} |\lambda_i|^2 < \varepsilon \,.$$

This implies that

$$\varepsilon \geq \|\varphi(x) - \varphi(y)\|^2 = \Big\|\varphi(x) - \sum_{i\in J} \lambda_i x_i\Big\|^2 \,.$$

(iii)⇒(iv). By (iii), each element $x \in \mathcal{H}$ can be uniquely written as

$$x = \sum_{i\in I} \lambda_i x_i.$$

Thus

$$\langle x, x_j \rangle = \left\langle \sum_{i \in I} \lambda_i x_i, x_j \right\rangle = \lambda_j .$$

(iv)⇒(v). By (iv),

$$\langle x, y \rangle = \left\langle \sum_{i \in I} \langle x, x_i \rangle x_i, \sum_{j \in I} \langle y, x_j \rangle x_j \right\rangle = \sum_{i \in I} \langle x, x_i \rangle \langle x_i, y \rangle \,.$$

(v)⇒(i). Let $x \in \{x_i : i \in I\}^{\perp}$. Then, by (v),

$$\|x\|^2 = \sum_{i \in I} |\langle x, x_i \rangle|^2 = 0\,.$$

Now, (i) follows from Remark 8.8. □

Theorem 8.14. *Every Hilbert space $\mathcal{H} \neq \{0\}$ possesses an orthogonal basis, and all orthogonal bases have the same cardinalitity.*

Proof. Let γ be the family of all orthonormal systems in $\mathcal{H}$. First, observe that $\gamma \neq \emptyset$, since there exists some $x \in \mathcal{H}$, $x \neq 0$, with $\|x\| = 1$. Now we order γ by inclusion. Let $\mathcal{K}$ be a chain in γ and set

$$T = \bigcup_{S \in \mathcal{K}} S.$$

Let $x_1, x_2 \in T$, $x_1 \neq x_2$ (if this is not possible, we have $T \in \gamma$). Then $x_1 \in S_1$, $x_2 \in S_2$ for some $S_1, S_2 \in \mathcal{K}$. Without loss of generality $S_1 \subset S_2$, hence $x_1, x_2 \in S_2$. Hence

$$\langle x_1, x_2 \rangle = 0.$$

This shows that also $T \in \gamma$. By Zorn's lemma, there exists a maximal ONS in γ, which is, by definition, an orthonormal basis.

For the second claim, let B and C be orthonormal bases of $\mathcal{H}$. For each $x \in B$ define

$$C_x = \{y \in C \colon \langle x, y \rangle \neq 0\}.$$

By Theorem 8.13,

$$x = \sum_{y \in C} \langle x, y \rangle y,$$

hence, C_x is at most countable. By Theorem 8.13, for each $y \in C$ there exists an $x_y \in B$ with

$$\langle x_y, y \rangle \neq 0.$$

Now set

$$M := \{(x, y) \colon x \in B, y \in C_x\},$$

and consider the map

$$y \mapsto (x_y, y),\ C \to M.$$

This map is injective, hence

$$|C| \leq |M|.$$

If $|B| = \infty$, then

$$|C| \leq |M| \leq |B| \cdot \aleph_0 = |B|.$$ [4]

By symmetry, both cardinalities are the same. □

[4] $\aleph_0$ is the cardinality of $\mathbb{N}$.

Definition 8.15. Let $\mathcal{H}$ be a Hilbert space over $\mathbb{K}$. Then the *dimension* of $\mathcal{H}$ ($\dim \mathcal{H}$) is the cardinality of some (and hence of each) orthonormal basis of $\mathcal{H}$.

Theorem 8.16. *Let $\mathcal{H}_1, \mathcal{H}_2 \neq \{0\}$ be Hilbert spaces over $\mathbb{K}$. Then the following conditions are equivalent.*

(i) *The equation* $\dim \mathcal{H}_1 = \dim \mathcal{H}_2$ *holds.*

(ii) *The spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ are isometrically isomorphic.*

Proof. (i)⇒(ii). Let $\{x_i : i \in I\}$ and $\{y_i : i \in I\}$ be orthonormal bases of $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. Then the map

$$\varphi : \mathcal{H}_1 \to \mathcal{H}_2, \quad \sum_{i \in I} \lambda_i x_i \mapsto \sum_{i \in I} \lambda_i y_i, \ (\lambda_i)_i \in \ell_2(I),$$

fulfills the desired properties.
(ii)⇒(i). Let B be an orthonormal basis of $\mathcal{H}_1$ and $\varphi : \mathcal{H}_1 \to \mathcal{H}_2$ a surjective isometry. Then $\varphi(B) = \{\varphi(x) : x \in B\}$ is an ONS in $\mathcal{H}_2$. If $y \in \varphi(B)^\perp$, then $\varphi^{-1}(y) \in B^\perp$ which implies $\varphi^{-1}(y) = 0$ and thus $y = 0$. Hence, $\varphi(B)$ is an orthonormal basis of $\mathcal{H}_2$ which has the same cardinality as B. □

Corollary 8.17. *Let $\mathcal{H}$ be a Hilbert space, and I an index set of an orthonormal basis of $\mathcal{H}$. Then $\mathcal{H}$ is isometrically isomorphic to $\ell^2(I)$.*

Proof. Note that the unit vectors $e_j = (\delta_{ij})_{i \in I}$ form an orthonormal basis for $\ell^2(I)$. The claim now follows from Theorem 8.16. □

Theorem 8.18 (Schmidt Orthogonalization Method). *Let $\mathcal{H}$ be an inner product space, and let $\{x_1, x_2, \dots\}$ be a linearly independent set in $\mathcal{H}$. Then define $\{y_1, y_2, \dots\}$ by*

$$y_1 = \|x_1\|^{-1} x_1$$
$$n \geq 1 : \quad y_{n+1} = \left\| x_{n+1} - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle y_i \right\|^{-1} \cdot \left(x_{n+1} - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle y_i \right)$$

Then $\{y_1, y_2, \dots\}$ is an ONS and

$$\operatorname{span}\{y_1, \dots, y_k\} = \operatorname{span}\{x_1, \dots, x_k\} \quad \forall k \in \mathbb{N}.$$

Proof. For each $n \in \mathbb{N}$, we prove that $\{y_1, \dots, y_n\}$ is an ONS and

$$\operatorname{span}\{y_1, \dots, y_n\} = \operatorname{span}\{x_1, \dots, x_n\}.$$

Nothing is to prove for $n = 1$. Assume that the claim holds for n. First of all, we have $x_{n+1} - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle y_i \neq 0$, since $\operatorname{span}\{y_1, \dots, y_n\} = \operatorname{span}\{x_1, \dots, x_n\}$ and the set $\{x_1, \dots, x_{n+1}\}$ is linearly independent. $\|y_{n+1}\| = 1$ and for all $k = 1, \dots, n$ we have

$$\begin{aligned}
\langle y_{n+1}, y_k \rangle &= \left\| x_{n+1} - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle y_i \right\|^{-1} \cdot \left(\langle x_{n+1} - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle y_i, y_k \rangle \right) \\
&= \left\| x_{n+1} - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle y_i \right\|^{-1} \cdot \left(\langle x_{n+1}, y_k \rangle - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle \underbrace{\langle y_i, y_k \rangle}_{=\delta_{ik}} \right) \\
&= 0.
\end{aligned}$$

We know that $\operatorname{span}\{x_1,\dots,x_n\} = \operatorname{span}\{y_1,\dots,y_n\}$. By definition of y_{n+1}, it then follows that also $\operatorname{span}\{x_1,\dots,x_{n+1}\} = \operatorname{span}\{y_1,\dots,y_{n+1}\}$. □

Theorem 8.19. *Let $\mathcal{H}$ be a Hilbert space. Then the following conditions are equivalent.*

(i) *The space $\mathcal{H}$ possesses an at most countable orthonormal basis.*

(ii) *The space $\mathcal{H}$ is separable (i. e. there exists a dense countable subset of $\mathcal{H}$).*

Proof. (ii)⇒(i). Let $D = \{x_1, x_2, \dots\}$ be a dense countable subset of $\mathcal{H}$. By induction, we delete from $(x_n)_{n\in\mathbb{N}}$ each x_n, which is contained in $\operatorname{span}\{x_1,\dots,x_{n-1}\}$. This creates a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ which is now linearly independent.
Also by construction

$$\operatorname{span}\{x_1, x_2, \dots\} = \operatorname{span}\{x_{n_1}, x_{n_2}, \dots\} =: L.$$

Apply Schmidt (Theorem 8.18) to generate an ONS $\{y_1, y_2 \dots\}$ with

$$\operatorname{span}\{y_1, y_2, \dots\} = L.$$

Since $D \subset L$, also L is dense in $\mathcal{H}$. By Theorem 8.13, $\{y_1, y_2, \dots\}$ is an orthonormal basis of $\mathcal{H}$.
(i)⇒(ii). Let $\{y_1, y_2, \dots\}$ be a countable orthonormal basis of $\mathcal{H}$. Set $M = \mathbb{Q}$, if $\mathbb{K} = \mathbb{R}$, and $M = \mathbb{Q} + i\cdot\mathbb{Q}$, if $\mathbb{K} = \mathbb{C}$. Then define $D := \{\sum_{i=1}^n \lambda_i y_i \colon \lambda_i \in M, n \in \mathbb{N}\}$. Since M is countable, also D is countable. Let $x \in \mathcal{H}$ and $\varepsilon > 0$. By Theorem 8.13, there exist $n \in \mathbb{N}$ and $\mu_i \in \mathbb{K}$ with

$$\left\| x - \sum_{i=1}^n \mu_i y_i \right\| \leq \tfrac{\varepsilon}{2}.$$

For each μ_i there exists $\lambda_i \in M$ with $|\lambda_i - \mu_i| \leq \frac{\varepsilon}{2n}$. This gives

$$\left\| x - \sum_{i=1}^n \lambda_i y_i \right\| \leq \left\| x - \sum_{i=1}^n \mu_i y_i \right\| + \left\| \sum_{i=1}^n (\mu_i - \lambda_i) y_i \right\| \leq \varepsilon.$$

Hence, D is dense in $\mathcal{H}$. □

9 Compact Operators

Definition 9.1. Let E and F be normed spaces. A linear operator $T\colon E \to F$ is called *compact*, if for all bounded sets $B \subset E$ the set $\overline{T(B)}$ is compact. The set of all compact operators from E to F is denoted by $\mathcal{K}(E,F)$. If $E = F$, then usually the notion $\mathcal{K}(E)$ is used instead of $\mathcal{K}(E,E)$.

Lemma 9.2. *Let E and F be normed spaces. Then:*

(i) $\mathcal{K}(E,F) \subset L(E,F)$,

(ii) *For an operator $T\colon E \to F$ the following are equivalent:*

  (a) *T is compact.*

  (b) *$\overline{T(K_1(0))}$ is compact.*

  (c) *If $(x_n) \subset E$ is bounded, $(Tx_n) \subset F$ contains a convergent subsequence.*

Proof. (i). Let $T \in \mathcal{K}(E,F)$. Then $\overline{T(K_1(0))}$ is compact. Thus, $T(K_1(0)) = \{Tx : \|x\| \leq 1\}$ is bounded. Therefore, $\sup\{\|Tx\| : \|x\| \leq 1\} < \infty$ which shows that $T \in L(E,F)$.

(ii). The implication (a)⇒(b) is obvious.

(b)⇒(a). Let $B \subset E$ be bounded. Then there exists $r > 0$ such that $B \subseteq rK_1(0)$. Now,

$$\overline{T(B)} \subseteq \overline{T(rK_1(0))} = r\overline{T(K_1(0))}\,.$$

Hence, the closed set $\overline{T(B)}$ is contained in the compact set $r\overline{T(K_1(0))}$ and is therefore itself compact.

(a)⇔(c). It is well-known, that a metric space X is compact if and only if each sequence in X contains a convergent subsequence, see Theorem 1.13. This implies (a)⇔(c). □

Lemma 9.3. *Let E, F and G be normed spaces. Then:*

(i) $\mathcal{K}(E,F)$ is a linear subspace of $L(E,F)$.

(ii) For $S \in L(F,G)$ and $T \in L(E,F)$ we have $ST \in \mathcal{K}(E,G)$, if $S \in \mathcal{K}(F,G)$ or $T \in \mathcal{K}(E,F)$.

Proof. (i). Let $S,T \in \mathcal{K}(E,F)$, $\alpha,\beta \in \mathbb{K}$ and let $(x_n) \subseteq E$ be bounded. Then (Sx_n) contains a convergent subsequence (Sx_{n_k}). Since (x_{n_k}) is bounded, (Tx_{n_k}) contains a convergent subsequence $(Tx_{n_{k_j}})$. Thus $(\alpha Sx_{n_{k_j}} + \beta Tx_{n_{k_j}})$ converges.

(ii). Let $(x_n) \in E$ be bounded. If T is compact, then (Tx_n) contains a convergent subsequence (Tx_{n_k}). Thus (STx_{n_k}) converges. If S is compact, then (STx_n) contains a convergent subsequence, since (Tx_n) is bounded. □

Lemma 9.4. *Let E be a normed space and let F be a Banach space. Then $\mathcal{K}(E,F)$ is closed in $L(E,F)$.*

Proof. Let $T \in \overline{\mathcal{K}(E,F)}$. For any $\varepsilon > 0$ there exists an operator $S \in \mathcal{K}(E,F)$ such that $\|S - T\| < \frac{\varepsilon}{3}$. Now, we use that $\overline{S(K_1(0))}$ is compact. Hence, the set $\overline{S(K_1(0))}$ is totally bounded. By Lemma 1.14 on page 10, $S(K_1(0))$ is totally bounded. Therefore there exist $x_1, \ldots, x_r \in K_1(0)$ such that $S(K_1(0)) \subseteq \bigcup_{j=1}^r K_{\varepsilon/3}(Sx_j)$ holds. If $x \in K_1(0)$, then $Sx \in K_{\varepsilon/3}(Sx_j)$ for some j. Hence

$$\|Tx - Tx_j\| \leq \|T - S\|\|x\| + \|Sx - Sx_j\| + \|S - T\|\|x_j\| < \varepsilon\,.$$

Therefore, $Tx \in K_\varepsilon(Tx_j)$ and $T(K_1(0)) \subseteq \bigcup_{j=1}^r K_\varepsilon(Tx_j)$. Thus $T(K_1(0))$ is totally bounded, and by Corollary 1.15 on page 10 its closure is compact. □

Example 9.5. (1) The identity operator $\mathrm{Id}\colon E \to E$, $x \mapsto x$ is compact if and only if E is finite-dimensional (by Theorem 2.10 on page 17).

(2) The zero-operator is compact.

(3) An operator $T \in L(E, F)$ is called *finite-dimensional* if $\dim T(E) < \infty$. A finite-dimensional operator is compact.

Proof. Let $(x_n)_n \subseteq E$ be bounded. Then $(Tx_n)_n$ is a bounded sequence in the finite-dimensional space $T(E)$ and hence contains a convergent subsequence. □

(4) By Lemma 9.4, also each limit of finite-dimensional operators $T_n \in L(E, F)$ is compact, if F is a Banach space. The converse does not hold in Banach spaces, i.e. a compact operator is not always the limit of finite-dimensional operators (Enflo 1973), but in Theorem 9.7 we will show that it holds in Hilbert spaces.

(5) Let $E = (C[a, b], \|\cdot\|_\infty)$, let $k\colon [a, b] \times [a, b] \to \mathbb{K}$ be continuous, and define an operator $K\colon E \to E$ by

$$(Kf)(s) := \int_a^b k(s, t)f(t)\mathrm{d}t, \quad f \in E\,.$$

Then K is compact.

Proof. By Arzelà-Ascoli we need to prove that for each bounded $B \subseteq E$, $K(B)$ is equicontinuous and pointwise bounded. For this, observe that

$$|(Kf)(s)| \leq (b - a)\|f\|_\infty\|k\|_\infty\,.$$

This shows that $K(B)$ is pointwise bounded. Since k is uniformly continuous, for $\varepsilon > 0$ there exists a $\delta > 0$ such that for $t \in [a, b]$ and $|s_1 - s_2| < \delta$ we have $|k(s_1, t) - k(s_2, t)| < \varepsilon$. Thus,

$$|(Kf)(s_1) - (Kf)(s_2)| \leq \int_a^b |k(s_1, t) - k(s_2, t)||f(t)|\mathrm{d}t \leq (b - a)\varepsilon\|f\|_\infty\,,$$

if $|s_1 - s_2| < \delta$. Therefore, and since B is bounded, $K(B)$ is equicontinuous. □

Definition 9.6. Let $\mathcal{H}$ be a Hilbert space and $\mathcal{L} \subset \mathcal{H}$ a closed subspace. Then $\mathcal{H} = \mathcal{L} \oplus \mathcal{L}^\perp$. If $x = u + v$ with $u \in \mathcal{L}$, $v \in \mathcal{L}^\perp$, then set $P_\mathcal{L} x := u$. The operator $P_\mathcal{L}\colon \mathcal{H} \to \mathcal{H}$ is then well-defined, linear and bounded. It is called the *orthogonal projection onto* $\mathcal{L}$. Moreover, we have that $\ker P_\mathcal{L} = \mathcal{L}^\perp$, $P_\mathcal{L}^2 = P_\mathcal{L}$ and $\|P_\mathcal{L}\| = 1$ if $\mathcal{L} \neq \{0\}$. More generally, an operator $P\colon \mathcal{H} \to \mathcal{H}$ with $P^2 = P$ is called a *projection*.

Theorem 9.7. *Let E be a normed space, $\mathcal{H}$ a Hilbert space and $T \in \mathcal{K}(E, \mathcal{H})$. Then there exist finite-dimensional operators $T_n \in L(E, \mathcal{H})$ with $\|T - T_n\| \to 0$.*

Proof. Compactness of $\overline{T(K_1(0))}$ implies that $T(K_1(0))$ is totally bounded. Thus, for each $n \in \mathbb{N}$ there exist $x_{n,1}, \dots, x_{n,r_n}$ in $K_1(0)$ such that $T(K_1(0)) \subset \bigcup_{i=1}^{r_n} K_{1/n}(Tx_{n,i})$. Put $\mathcal{L}_n := \operatorname{span}\{Tx_{n,i} : i = 1, \dots, r_n\}$ and $P_n := P_{\mathcal{L}_n}$. Then $\dim \mathcal{L}_n \leq r_n$. In particular, $\mathcal{L}_n$ is closed in $\mathcal{H}$. Now, let $n \in \mathbb{N}$. Then for each $x \in K_1(0)$, $Tx \in K_{1/n}(Tx_{n,i})$ for some i. Thus for $x \in K_1(0)$

$$\begin{aligned} \|Tx - P_nTx\| &\leq \|Tx - Tx_{n,i}\| + \|Tx_{n,i} - P_nTx_{n,i}\| + \|P_nTx_{n,i} - P_nTx\| \\ &\leq \|Tx - Tx_{n,i}\| + \|Tx_{n,i} - Tx\| \leq \tfrac{2}{n}\,, \end{aligned}$$

hence $\|T - P_nT\| \to 0$ as $n \to \infty$, i.e. the finite-dimensional operators P_nT converge to T. □

Theorem 9.8. *Let E, F be normed spaces, $T \in L(E,F)$. Also let $T^*\colon F^* \to E^*$, $(T^*f)(x) = f(Tx)$ be the dual operator. Then the following holds:*

(i) If T is compact, then also T^* is compact.

(ii) If T^* is compact and F is a Banach space, then T is compact.

Proof. (i). Let $(f_n)_n \subseteq F^*$ be bounded. By Lemma 9.2 it is sufficient to prove that $(T^*f_n)_n$ contains a convergent subsequence. $Y := \overline{T(K_1(0))}$ is compact in F. Now, consider

$$\mathcal{F} := \{f_n|_Y : n \in \mathbb{N}\} \subseteq C(Y)\,.$$

Setting $C := \sup_n \|f_n\|$, we first observe that

- $|f_n(y)| \leq C\|y\|$, hence $\mathcal{F}$ is pointwise bounded.
- $|f_n(y_1) - f_n(y_2)| \leq C\|y_1 - y_2\|$, i.e $\mathcal{F}$ is equicontinous.

Hence, by Arzelà-Ascoli $\overline{\mathcal{F}} \subseteq (C(Y), \|\cdot\|_\infty)$ is compact. This implies that there exists a convergent subsequence $(f_{n_k}|_Y)_k$. Hence for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ with

$$\|f_{n_k}|_Y - f_{n_l}|_Y\|_\infty < \varepsilon \quad \forall k,l \geq N\,. \tag{9.1}$$

This yields

$$\|T^*f_{n_k} - T^*f_{n_l}\| = \sup_{\|x\|=1} |f_{n_k}(Tx) - f_{n_l}(Tx)| < \varepsilon$$

for all $k,l \geq N$. Thus $(T^*fn_k)_k$ is a Cauchy-sequence in E^*, and hence converges.

(ii). Since T^* is compact, by (i) also T^{**} is compact. It is easily seen that $\Lambda_F T = T^{**}\Lambda_E$. Hence,

$$\overline{\Lambda_F(T(K_1(0_E)))} = \overline{T^{**}(\Lambda_E(K_1(0_E)))} \subseteq \overline{T^{**}(K_1(0_{E^{**}}))}\,,$$

which is a compact subset of F^{**}. Therefore, $\overline{\Lambda_F(T(K_1(0_E)))}$ is compact. Since F is a Banach space, also $\overline{T(K_1(0_E))}$ is compact. By Lemma 9.2, T is compact. □

Definition 9.9. Let E be a linear space, F, G linear subspaces, such that $E = F \dot{+} G$ (direct sum). Then G is called a *complementary subspace* to F in E.

Bemerkung 9.10. Let E be a Banachspace and let $F, G \subset E$ be closed linear subspaces such that $E = F \dot{+} G$. Then the mapping $P\colon E \to F$, $x + y \mapsto x$, where $x \in F$, $y \in G$, is a continuous projection.

Proof. It is clear that P is a projection, i.e. $P^2 = P$. To see that it is continuous, by the Closed Graph Theorem we only need to prove that P is closed. For this, let $(x_n)_n \subseteq F$ and $(y_n)_n \subseteq G$ such that $x_n + y_n \to u$ and $x_n = P(x_n + y_n) \to x$ as $n \to \infty$. Then $x \in F$ since F is closed, and $y_n \to u - x =: y$. Since also G is closed, we have $y \in G$ and thus $u = x + y \in F \dotplus G$. Therefore, $Pu = x$, which we had to prove. □

Lemma 9.11. *Let E be a normed space, F a closed linear subspace, such that* $\dim F < \infty$ *or* $\dim E/F < \infty$*, then there exists a closed complementary subspace to F in E.*

Proof. Assume that $\dim E/F < \infty$. Let $x_1, \dots, x_n \in E$ be such that $\{x_1 + F, \dots, x_n + F\}$ is a basis of E/F. Then $G := \operatorname{span}\{x_1, \dots, x_n\} \subseteq E$ is closed, since $\dim G < \infty$. Let us show that $E = F + G$ and $F \cap G = \{0\}$. For $x \in E$ we have $x + F = \sum_{i=1}^n \lambda_i (x_i + F)$ with some $\lambda_1, \dots, \lambda_n \in \mathbb{K}$. Put $g := \sum_{i=1}^n \lambda_i x_i \in G$ and $f := x - g$. Then $x = f + g$ and $f + F = (x + F) - (g + F) = 0$, and hence $f \in F$. This shows $E = F + G$. If $x \in F \cap G$, then $x = \sum_{i=1}^n \lambda_i x_i$ since $x \in G$ and $0 = x + F = \sum_{i=1}^n \lambda_i (x_i + F)$ since $x \in F$. As the $x_i + F$ are linearly independent, we conclude that $\lambda_1 = \ldots = \lambda_n = 0$ and thus $x = 0$.

Assume now that $\dim F < \infty$. Let $\{x_1, \dots, x_n\}$ be a basis of F. Now let $f_1, \dots, f_n \in F^*$ such that

$$f_i(x_j) = \delta_{ij} \quad \forall i, j \in \{1, \dots, n\}\,.$$

This defines a basis of F^* (the so-called dual basis). By Hahn-Banach, there exist $\ell_1, \dots, \ell_n \in E^*$ such that $\ell_i|_F = f_i$. Now let $P\colon E \to E$ be defined by

$$Px = \sum_{j=1}^n \ell_j(x) x_j\,.$$

In particular, P is linear and continous and $P|_F = \operatorname{Id}|_F$, since $P(x_k) = x_k$. Further, $P^2 = P$ since

$$P^2 x = \sum_{j=1}^n \ell_j(x) P x_j = \sum_{j=1}^n \ell_j(x) x_j = Px\,.$$

Define $G := \ker P$. Then G is closed, $G \cap F = \{0\}$, and $E = F + G$, since $x = Px + (x - Px)$. □

Theorem 9.12. *Let E be a Banach space, and $K\colon E \to E$ a compact operator. Set $T := \operatorname{Id} - K \in L(E)$, Then*

(i) $\dim(\ker T) < \infty$,

(ii) $T(E)$ *is closed in E and*

(iii) $\dim\big(E/T(E)\big) < \infty$.

Proof. (i). Since

$$\operatorname{Id}|_{\ker T} = K|_{\ker T}\,,$$

$\operatorname{Id}|_{\ker T}$ is compact. Hence, $\ker T$ must be finite-dimensional.

(ii). Set $F = \ker T$. Since $\dim F < \infty$, there exists a closed complementary subspace G to F. Now we consider

$$S\colon G \to T(E), \quad Sx := Tx,\ x \in G\,.$$

S is continuous and bijective. In particular, $S(G) = T(E)$. Let $(y_n)_n \subseteq T(E)$, $y_n \to y$ as $n \to \infty$. Then $y_n = x_n - Kx_n$, where $x_n = S^{-1}y_n \in G$, $n \in \mathbb{N}$. We have to prove that $y \in T(E)$.

Suppose that $(x_n)_n$ has no bounded subsequence. Then $\|x_n\| \to \infty$. Put $u_n := \frac{x_n}{\|x_n\|} \in G$, $n \in \mathbb{N}$. Then $u_n - Ku_n = \frac{y_n}{\|x_n\|} \to 0$ as $n \to \infty$. As $(u_n)_n$ is bounded, there exists a subsequence $(u_{n_j})_j$ such that $Ku_{n_j} \to v$ as $j \to \infty$ for some $v \in E$. But then also $u_{n_j} \to v$, implying $v \in G$ and also $Ku_{n_j} \to Kv$. Consequently, $Sv = v - Kv = 0$ and $\|v\| = \lim_j \|u_{n_j}\| = 1$, contradicting the fact that S is injective.

Hence, $(x_n)_n$ has a bounded subsequence $(x_{n_j})_j$ such that $Kx_{n_j} \to v$ as $j \to \infty$ for some $v \in E$. This implies $x_{n_j} = Sx_{n_j} + Kx_{n_j} \to y + v$ as $j \to \infty$ and thus $y = \lim_j Sx_{n_j} = S(y+v) \in S(G) = T(E)$.

(iii). Since $T(E)$ is closed, by Theorem 4.17 we know that

$$(E/T(E))^* \text{ is isometrically isomorphic to } T(E)^\perp .$$

Note that $T(E)^\perp = \ker T^*$:

$$\begin{aligned} T(E)^\perp &= \{f \in E^* : f(Tx) = 0 \ \forall x \in E\} \\ &= \{f \in E^* : (T^* f)x = 0 \ \forall x \in E\} \\ &= \{f \in E^* : T^* f = 0\} . \end{aligned}$$

Since K^* is compact by Theorem 9.8, and $T^* = \mathrm{Id} - K^*$, by (i) we obtain

$$\dim(\ker T^*) < \infty .$$

Hence, we can conclude that

$$\dim(E/T(E))^* < \infty ,$$

which implies (iii). □

Definition 9.13. An operator $T \in L(E)$ which satisfies (i) - (iii) in Theorem 9.12 is called *Fredholm operator*. Further, the integer

$$\mathrm{ind}(T) := \dim(\ker T) - \dim(E/T(E))$$

is called the *index* of the Fredholm operator T. $\mathcal{F}(E)$ shall denote the set of all Fredholm operators on E, hence $\mathrm{Id} - \mathcal{K}(E) \subseteq \mathcal{F}(E)$.

Lemma 9.14. *Let E and F be Banach spaces and let $T \in L(E,F)$ be bijective. Let $T^{-1} \in L(F,E)$ be the inverse of T and $S \in L(E,F)$ be such that*

$$\|S - T\| < \|T^{-1}\|^{-1} ,$$

then S is also invertible.

Proof. We first write

$$S = T(\mathrm{Id} - T^{-1}(T - S)) ,$$

and set

$$Q = T^{-1}(T - S) .$$

For $c := \|T - S\| \|T^{-1}\| < 1$ we obtain

$$\left\| \sum_{k=m+1}^{n} Q^k \right\| = \left\| \sum_{k=m+1}^{n} (T^{-1}(T - S))^k \right\| \le \sum_{k=m+1}^{n} c^k .$$

Thus the geometric series $\sum_{k=0}^{\infty} Q^k$ is convergent in $L(E)$. Finally,

$$\sum_{k=0}^{\infty} Q^k T^{-1} S = \lim_{n\to\infty} \sum_{k=0}^{n} Q^k (\mathrm{Id} - Q) = \lim_{n\to\infty} (\mathrm{Id} - Q^{n+1}) = \mathrm{Id}\,,$$

and also

$$S \sum_{k=0}^{\infty} Q^k T^{-1} = \sum_{k=0}^{\infty} T(\mathrm{Id} - Q) Q^k T^{-1} = T\left(\sum_{k=0}^{\infty} (\mathrm{Id} - Q) Q^k\right) T^{-1} = \mathrm{Id}\ .$$

This shows that S is invertible with $S^{-1} = \sum_{k=0}^{\infty} Q^k T^{-1}$. □

Bemerkung 9.15. If E is a Banach space and $T \in L(E)$ with $\|T\| < 1$, then Lemma 9.14 implies that $\mathrm{Id} - T$ is invertible, and it follows from the above proof that the inverse of $\mathrm{Id} - T$ is given by

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

This series is called the *Neumann series.*

Theorem 9.16. *Let E be a Banach space. Then $\mathcal{F}(E)$ is open in $L(E)$ and the map*

$$T \mapsto \operatorname{ind} T\,,\ \mathcal{F}(E) \to \mathbb{Z}$$

is continuous.

Proof. to be added □

Bemerkung 9.17. Note that Theorem 9.16 implies that each set

$$\mathcal{F}_k(E) := \{T \in \mathcal{F}(E) : \operatorname{ind} T = k\}$$

is open in $L(E)$. Since these sets are mutually disjoint, it follows that the index is constant on each connected component of $\mathcal{F}(E)$. The set of invertible operators is a (proper) subset of the open set $\mathcal{F}_0(E)$ and is itself open. This follows from Lemma 9.14. Note furthermore that $\mathcal{F}(E) = \mathcal{F}_0(E)$ if E is finite-dimensional.

Corollary 9.18. *Let E be a Banach space and $K \in \mathcal{K}(E)$. Then* $\operatorname{ind}(\mathrm{Id} - K) = 0$.

Proof. By Theorem 9.16 the map

$$\mathbb{R} \to \mathcal{F}(E) \to \mathbb{Z}\,, \quad t \mapsto \mathrm{Id} - tK \mapsto \operatorname{ind}(\mathrm{Id} - tK)\,.$$

is continuous which implies

$$\operatorname{ind}(\mathrm{Id} - K) = \operatorname{ind}(\mathrm{Id}) = 0\,.$$

□

Corollary 9.19. *Let E be a Banach space and $K \in \mathcal{K}(E)$. Define $T := \mathrm{Id} - K$. Then*

$$\dim(\ker T) = \dim(\ker T^*)\,.$$

Proof. By Theorem 9.8, K^* is compact. Also,

$$\ker T^* = T(E)^{\perp} \quad \text{and} \quad T(E)^{\perp} \cong (E/T(E))^*\,.$$

By Corollary 9.18, $\operatorname{ind}(T) = 0$, and we have

$$\dim(\ker(T^*)) = \dim(T(E)^{\perp}) = \dim((E/T(E))^*) = \dim(E/T(E)) = \dim(\ker T)\,.$$

□

10 The Spectrum of an Operator

Definition 10.1. Let E be a normed space over $\mathbb{K}$ and $T \in L(E)$.

(1) $\lambda \in \mathbb{K}$ is an *eigenvalue* of T, if $\ker(\lambda \operatorname{Id} - T) \neq \{0\}$. The elements of $\ker(\lambda \operatorname{Id} - T) \setminus \{0\}$ are called *eigenvectors* of T associated with λ. Then

$$E(\lambda) := \{x \in E : Tx = \lambda x\}$$

is called the *eigenspace* of T with respect to λ.

(2) $\lambda \in \mathbb{K}$ is called a *spectral value* of T, if $\lambda \operatorname{Id} - T$ does not possess an inverse in $L(E)$. The set of all spectral values is called the *spectrum* of T and is denoted by $\sigma(T)$. The complement $\rho(T) := \mathbb{K} \setminus \sigma(T)$ is called the *resolvent set* of T. The elements of $\rho(T)$ are called *regular values* of T.

Remark 10.2. (i) If λ is an eigenvalue of T, then it is also a spectral value of T. If $\dim E < \infty$, then also each spectral value is an eigenvalue (since, if $\lambda \in \sigma(T)$, then $\lambda \operatorname{Id} - T$ is not bijective, hence not injective, thus λ is an eigenvalue.).

(ii) Let E be a Banach space. Then $(\lambda \operatorname{Id} - T)^{-1}$ exists if and only if $\lambda \operatorname{Id} - T$ is bijective. This is a direct consequence of the open mapping theorem.

(iii) If $\dim E = \infty$ and $T \in \mathcal{K}(E)$, then $0 \in \sigma(T)$ since $0 \in \rho(T)$ implies that T is invertible, hence open. Then $\overline{T(K_1(0))}$ is a compact neighborhood of 0 in E.

Lemma 10.3. *$\sigma(T)$ is closed, and for every $\lambda \in \sigma(T)$ we have*

$$|\lambda| \leq \|T\| .$$

Proof. Let $\lambda_0 \in \rho T$ and let $\lambda \in \mathbb{K}$ be such that

$$\|(\lambda \operatorname{Id} - T) - (\lambda_0 \operatorname{Id} - T)\| = |\lambda - \lambda_0| < \|(\lambda_0 \operatorname{Id} - T)^{-1}\|^{-1} .$$

Lemma 9.14 implies that $\lambda \operatorname{Id} - T$ is (boundedly) invertible, thus $\lambda \in \rho T$. This argument shows that ρT is open, thus $\sigma(T)$ is closed.
Now let $\lambda \in \mathbb{K}$ be such that $|\lambda| > \|T\|$. Then

$$\|(\lambda \operatorname{Id} - T) - \lambda \operatorname{Id}\| = \|T\| < |\lambda| = \|(\lambda \operatorname{Id})^{-1}\|^{-1} .$$

Again, Lemma 9.14 implies that $\lambda \operatorname{Id} - T$ is invertible, hence $\lambda \in \rho T$. □

Lemma 10.4. *Let E be a Banach space and $T \in L(E)$.*

(i) *If $|\lambda| > \|T\|$, then*

$$(\lambda \operatorname{Id} - T)^{-1} = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}} \qquad \textit{and} \qquad \|(\lambda \operatorname{Id} - T)^{-1}\| \leq (|\lambda| - \|T\|)^{-1} .$$

(ii) *If T is invertible and $S \in L(E)$ with*

$$\|T - S\| \leq \varepsilon \|T^{-1}\|^{-1}$$

for some $\varepsilon \in (0,1)$, then $\|T^{-1} - S^{-1}\| \leq \frac{1}{1-\varepsilon} \|T^{-1}\|^2 \|S - T\|$.

Proof. (i). It suffices to prove that for $A = \frac{1}{\lambda}T$ we have

$$(\operatorname{Id} - A)^{-1} = \sum_{n=0}^{\infty} A^n \,.$$

We know that $\|A\| < 1$. Hence

$$\sum_{n=0}^{\infty} \|A\|^n < \infty \,.$$

Therefore, $B_m := \sum_{n=0}^{m} A^n$ converges to some $B \in L(E)$ as $m \to \infty$. We have

$$(\operatorname{Id} - A)B_m = \operatorname{Id} - A^{m+1} = B_m(\operatorname{Id} - A) \,.$$

Since $\|A^{m+1}\| \to 0$ as $m \to \infty$, we can let m tend to infinity and obtain

$$(\operatorname{Id} - A)B = \operatorname{Id} = B(\operatorname{Id} - A) \,.$$

This proves the first part. We also have

$$\|(\operatorname{Id} - A)^{-1}\| \leq \sum_{n=0}^{\infty} \|A\|^n = \frac{1}{1 - \|A\|} \,.$$

Multiplication with $|\lambda|$ yields the second claim.
(ii). This is proved similarly as Lemma 9.14. □

Theorem 10.5 (Gelfand-Mazur, 1941)**.** *Let E be a Banach space over $\mathbb{C}$ and $T \in L(E)$. Then $\sigma(T) \neq \emptyset$.*

Proof. Let $f \in L(E)^*$ and define

$$\varphi : \rho T \to \mathbb{C}, \quad \varphi(\lambda) := f\left((\lambda \operatorname{Id} - T)^{-1}\right), \, \lambda \in \rho(T).$$

For $\lambda, \lambda_0 \in \rho T$ we have

$$\begin{aligned}\varphi(\lambda) - \varphi(\lambda_0) &= f\left((\lambda \operatorname{Id} - T)^{-1} - (\lambda_0 \operatorname{Id} - T)^{-1}\right) \\ &= (\lambda_0 - \lambda) f\left((\lambda \operatorname{Id} - T)^{-1}(\lambda_0 \operatorname{Id} - T)^{-1}\right).\end{aligned}$$

By Lemma 10.4, the inverse is continous. Therefore

$$\lim_{\lambda \to \lambda_0} \frac{\varphi(\lambda) - \varphi(\lambda_0)}{\lambda - \lambda_0} = -f\left((\lambda_0 \operatorname{Id} - T)^{-2}\right).$$

Hence φ is holomorphic on $\rho(T)$.
Towards a contradiction, assume $\sigma(T) = \emptyset$. Then φ is holomorphic on $\mathbb{C}$ – in other words, it is an entire function. Moreover, $\lim_{|\lambda| \to \infty} \varphi(\lambda) = 0$ by Lemma 10.4(i). Liouville's Theorem (from complex analysis) implies that φ is constant, and hence zero. Since this is true for any $f \in L(E)^*$, Corollary 4.8 implies $(\lambda \operatorname{Id} - T)^{-1} = 0$ for each $\lambda \in \rho(T)$. A contradiction. □

Theorem 10.6 (Formula for the spectral radius, Gelfand 1941)**.** *Let E be a Banach space over $\mathbb{C}$ and $T \in L(E)$. Then*

$$\sup_{\lambda \in \sigma(T)} |\lambda| = \lim_{n \to \infty} \|T^n\|^{\frac{1}{n}} \,.$$

The left hand side of the above equation is called the *spectral radius* of T.

Proof. Let $\lambda \in \sigma(T)$. From

$$T^n - \lambda^n \operatorname{Id} = (T - \lambda \operatorname{Id}) \sum_{k=1}^{n} \lambda^{k-1} T^{n-k} = \left(\sum_{k=1}^{n} \lambda^{k-1} T^{n-k} \right) (T - \lambda \operatorname{Id})$$

it is seen that $\lambda^n \in \sigma(T^n)$. Lemma 10.3 now implies $|\lambda|^n \leq \|T^n\|$. Thus, if $r(T)$ denotes the spectral radius,

$$r(T) \leq \liminf_{n\to\infty} \|T^n\|^{\frac{1}{n}} .$$

Next we consider

$$\varphi(\lambda) := f\left((\lambda \operatorname{Id} - T)^{-1} \right)$$

for some $f \in L(E)^*$ and $\lambda \in \rho(T)$. We know by the proof of Theorem 10.5 that φ is holomorphic on $\{\lambda \in \mathbb{C} : |\lambda| > r(T)\}$. As seen from methods in complex analysis, φ is given by the Laurent series

$$\varphi(\lambda) = \sum_{n=0}^{\infty} \tfrac{1}{\lambda^{n+1}} f(T^n).$$

We conclude that this series converges for all $|\lambda| > r(T)$. In particular,

$$\sup_n \left| \tfrac{f(T^n)}{\lambda^n} \right| < \infty \quad \forall |\lambda| > r(T), f \in L(E)^*.$$

By Theorem 5.15, for each $\lambda \in \mathbb{C}$ with $|\lambda| > r(T)$ there exists some $M_\lambda > 0$ such that

$$\left\| \tfrac{T^n}{\lambda^n} \right\| \leq M_\lambda \quad \forall n \in \mathbb{N}.$$

This implies that $\limsup_{n\to\infty} \|T^n\|^{\frac{1}{n}} \leq |\lambda|$ for all $\lambda \in \mathbb{C}$ with $|\lambda| > r(T)$. Hence

$$\limsup_{n\to\infty} \|T^n\|^{\frac{1}{n}} \leq r(T).$$

Finally, $r(T) = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}$. □

Lemma 10.7. *Let E be a normed space and $T \in L(E)$. Further, let F and G be closed subspaces of E with $F \subseteq G$, $F \neq G$, and*

$$(\operatorname{Id} - T)G \subseteq F.$$

Then there exists some $a \in G$ with $\|a\| = 1$ and

$$\|Ta - Tx\| \geq \tfrac{1}{2} \quad \forall x \in F.$$

Proof. Choose $b \in G \backslash F$. Consider

$$\alpha = \operatorname{dist}(b, F) = \inf_{x \in F} \|x - b\|.$$

$\alpha > 0$, since F is closed. This implies that there exists a $y \in F$ such that $\|b - y\| < 2\alpha$. Define

$$a := \frac{b - y}{\|b - y\|} \in G.$$

Then a has norm 1 and for arbitrary $z \in F$, we have

$$\|z - a\| = \tfrac{1}{\|b-y\|} \| \underbrace{z\|b-y\| + y}_{\in F} - b\| \geq \frac{\alpha}{2\alpha} = \tfrac{1}{2}.$$

Finally, for each $x \in F$

$$\|Tx - Ta\| = \|\underbrace{x - (\operatorname{Id} - T)x + (\operatorname{Id} - T)a}_{\in F} - a\| \geq \tfrac{1}{2}.$$

□

Lemma 10.8. *Let E be a linear space, $T \colon E \to E$ a linear operator and M be a set of eigenvectors of T such that any two elements in M are eigenvectors to different eigenvalues. Then M is linearly independent, i.e. each finite subset of M is linearly independent.*

Proof. Let $\mathcal{M}_n$ be the collection of all subsets of M with n elements. We show the claim via induction over n. Clearly, each set in $\mathcal{M}_1$ is linearly independent. Assume now that the claim is proven for $n-1$ and let $x_1, \ldots x_n \in M$, $Tx_i = \lambda_i x_i$. Then $\sum_{i=1}^n \alpha_i x_i = 0$ implies

$$\sum_{i=1}^n \alpha_i \lambda_i x_i = \sum_{i=1}^n \alpha_i T x_i = T\left(\sum_{i=1}^n \alpha_i x_i\right) = 0\,.$$

We also have $\sum_{i=1}^n \alpha_i \lambda_n x_i = 0$. This implies

$$0 = \alpha_1(\lambda_1 - \lambda_n)x_1 + \cdots + \alpha_{n-1}(\lambda_{n-1} - \lambda_n)x_{n-1}.$$

By induction hyphothesis the vectors $x_1, \ldots, x_{n-1}$ are linearly independent. Hence

$$\alpha_i(\lambda_i - \lambda_n) = 0 \quad \forall i = 1, \ldots, n-1.$$

Since the λ_i are mutually distinct, this yields $\alpha_i = 0$ for $i = 1, \ldots, n-1$. Finally, $\alpha_n = 0$. □

Theorem 10.9. *Let E be a normed space and $K \colon E \to E$ a compact operator. Then the set M of eigenvalues of K is at most countable and can only accumulate to 0.*

Proof. It suffices to prove that for each $\delta > 0$, the set

$$M_\delta := \{\mu \in M : |\mu| > \delta\}$$

is finite. Towards a contradiction, assume that M_δ is infinite for some $\delta > 0$, i.e. there exist $\mu_n \in M_\delta$, $n \in \mathbb{N}$, with $\mu_n \neq \mu_m$, $m \neq n$. Now, let $0 \neq x_n \in E$ with

$$Kx_n = \mu_n x_n$$

for all $n \in \mathbb{N}$. Next define

$$F_n = \operatorname{span}\{x_1, \ldots, x_n\}.$$

By Lemma 10.8, the x_n are linearly independent, hence

$$F_n \subsetneq F_{n+1}.$$

If $y = \sum_{i=1}^n \alpha_i x_i \in F_n$, then

$$\begin{aligned}(K - \mu_n \operatorname{Id})y &= \sum_{i=1}^n \alpha_i K x_i - \mu_n \sum_{i=1}^n \alpha_i x_i = \sum_{i=1}^n \alpha_i \mu_i x_i - \mu_n \sum_{i=1}^n \alpha_i x_i \\ &= \sum_{i=1}^n \alpha_i (\mu_i - \mu_n) x_i \in F_{n-1},\end{aligned}$$

thus

$$(\operatorname{Id} - \tfrac{1}{\mu_n} K) F_n \subseteq F_{n-1}.$$

For each $n \in \mathbb{N}$, F_n is closed, hence by Theorem 10.7 (inductively) there exists a sequence $(y_n)_{n\in\mathbb{N}} \subseteq E$ with

$$y_n \in F_n, \ \ \|y_n\| = 1 \text{ and } \|Ky_n - Ky_m\| \geq \tfrac{1}{2}|\mu_n| \ \forall m > n, n \in \mathbb{N}.$$

This implies

$$\|Ky_n - Ky_m\| \geq \tfrac{1}{2}\delta \quad \forall m > n, n \in \mathbb{N}.$$

Thus $(Ky_n)_{n\in\mathbb{N}}$ cannot contain a convergent subsequence although $(y_n)_{n\in\mathbb{N}}$ is bounded which contradicts the compactness of K. □

Theorem 10.10. *Let E be a Banach space and let $K\colon E \to E$ be compact.*

(i) If $0 \neq \lambda \in \sigma(K)$, then λ is an eigenvalue.

(ii) The eigenspace $E(\lambda)$, $\lambda \neq 0$, is finite-dimensional.

(iii) $\sigma(K)$ is at most countable and can only accumulate to 0.

Proof. (i). By Corollary 9.18,

$$\operatorname{ind}(\operatorname{Id} - \tfrac{1}{\lambda} K) = 0.$$

Therefore, $\operatorname{Id} - \frac{1}{\lambda}K$ (and thus also $\lambda \operatorname{Id} - K$) is injective if and only if it is surjective.
(ii). $\operatorname{Id} - \frac{1}{\lambda}K$ is a Fredholm operator, hence

$$\dim(\ker(\operatorname{Id} - \tfrac{1}{\lambda} K)) < \infty.$$

Since $E(\lambda) = \ker(\operatorname{Id} - \frac{1}{\lambda}K)$, we conclude $\dim(E(\lambda)) < \infty$.
(iii). This follows from Theorem 10.9 and (i). □

11 Spectral Theory for Compact Operators

Let $\mathcal{H}$ be a Hilbert space, $T \in L(\mathcal{H})$ and $\varphi\colon \mathcal{H} \times \mathcal{H} \to \mathbb{K}$ defined by

$$\varphi(x, y) = \langle x, Ty \rangle .$$

Then φ is a sesquilinear form, and we have

$$\sup_{x,y \neq 0} \frac{|\varphi(x, y)|}{\|x\|\|y\|} \leq \|T\| < \infty .$$

By Theorem 7.16 on page 58 there exists a unique operator $T^* \in L(\mathcal{H})$ with

$$\langle T^* x, y \rangle = \varphi(x, y) .$$

This implies $\langle x, Ty \rangle = \langle T^* x, y \rangle$ and $\langle Tx, y \rangle = \langle x, T^* y \rangle$ for all $x, y \in \mathcal{H}$.

Definition 11.1. Let $\mathcal{H}$ be a Hilbert space and $T \in L(\mathcal{H})$. Then T^* from the above discussion is called the *adjoint operator* to T. If $T = T^*$, T is called *self-adjoint.*

Remark 11.2. Let $j\colon \mathcal{H} \to \mathcal{H}^*$ be the conjugate linear map $y \mapsto f_y$ (Riesz map). Further let $T \in L(\mathcal{H})$, $\tilde{T}^*$ be the associated dual operator ("Banach space adjoint") and let T^* be the (Hilbert space) adjoint. Then

$$T^* = j^{-1} \tilde{T}^* j .$$

Remark 11.3. Let $\mathcal{H}$ be a finite-dimensional inner product space and $T\colon \mathcal{H} \to \mathcal{H}$ be self-adjoint. Linear algebra gives us the existence of an orthonormal basis $\{u_1, \ldots, u_n\}$ of $\mathcal{H}$ which consists of eigenvectors of T ($Tu_i = \lambda_i u_i$, $i = 1, \ldots, n$). For $x = \sum_{i=1}^n \langle x, u_i \rangle u_i \in \mathcal{H}$, we have

$$Tx = \sum_{i=1}^n \lambda_i \langle x, u_i \rangle u_i .$$

We now aim for a corresponding result for compact self-adjoint operators on a Hilbert space.

Lemma 11.4. *If $T \in L(\mathcal{H})$ is self-adjoint, then the eigenvalues of T are real. Also, the eigenspaces corresponding to two different eigenvalues are orthogonal.*

Proof. Let λ be an eigenvalue of T. Then for an eigenvector x to λ we have

$$\lambda \|x\|^2 = \lambda \langle x, x \rangle = \langle Tx, x \rangle = \langle x, Tx \rangle = \bar{\lambda} \langle x, x \rangle = \bar{\lambda} \|x\|^2$$

and hence $\lambda = \bar{\lambda}$.
Now let $\lambda \neq \mu$ be two eigenvalues of T. For $x \in E(\lambda)$, $y \in E(\mu)$, $x, y \neq 0$, we have

$$(\lambda - \mu) \langle x, y \rangle = \lambda \langle x, y \rangle - \mu \langle x, y \rangle = \langle Tx, y \rangle - \langle x, Ty \rangle = 0 .$$

This proves $\langle x, y \rangle = 0$. □

Lemma 11.5. *Let $T \in L(\mathcal{H})$ and $\alpha > 0$ with*

$$|\langle Tx, x \rangle| \leq \alpha \|x\|^2 \quad \textit{for all } x \in \mathcal{H} .$$

Then the following holds:

(i) $|\langle Tx, y\rangle + \langle Ty, x\rangle| \le 2\alpha\|x\|\|y\|$ *for all* $x, y \in \mathcal{H}$.

(ii) If further $\mathbb{K} = \mathbb{C}$, *then*

$$|\langle Tx, y\rangle| + |\langle Ty, x\rangle| \le 2\alpha\|x\|\|y\|$$

for all $x, y \in \mathcal{H}$.

Proof. (i). First of all, we observe that with

$$\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle = 2(\langle Tx, y\rangle + \langle Ty, x\rangle)$$

we obtain by the parallelogram identity

$$2|\langle Tx, y\rangle + \langle Ty, x\rangle| \le \alpha(\|x+y\|^2 + \|x-y\|^2) = 2\alpha(\|x\|^2 + \|y\|^2)$$

for all $x, y \in \mathcal{H}$. By substituting x by $c^{-1}x$ and y by cy, we have

$$|\langle Tx, y\rangle + \langle Ty, x\rangle| \le \alpha(c^{-2}\|x\|^2 + c^2\|y\|^2)\,.$$

Now, choose c by

$$c := \left(\tfrac{\|x\|}{\|y\|}\right)^{1/2}$$

for $y \neq 0$. This yields

$$|\langle Tx, y\rangle + \langle Ty, x\rangle| \le 2\alpha\|x\|\|y\|\,.$$

(ii). Substituting x by $\mathrm{e}^{\mathrm{i}t}x$, $t \in \mathbb{R}$, and multiplying (i) with $1 = |e^{\mathrm{i}s}|$, $s \in \mathbb{R}$, gives

$$|\mathrm{e}^{\mathrm{i}s}\langle T(\mathrm{e}^{\mathrm{i}t}x), y\rangle + \mathrm{e}^{\mathrm{i}s}\langle Ty, \mathrm{e}^{\mathrm{i}t}x\rangle| \le 2\alpha\|x\|\|y\|\,.$$

Thus

$$|\mathrm{e}^{\mathrm{i}(s+t)}\langle Tx, y\rangle + \mathrm{e}^{\mathrm{i}(s-t)}\langle Ty, x\rangle| \le 2\alpha\|x\|\|y\|\,.$$

For suitable $u, v \in \mathbb{R}$ such that

$$\mathrm{e}^{\mathrm{i}u}\langle Tx, y\rangle = |\langle Tx, y\rangle| \quad \text{and} \quad \mathrm{e}^{\mathrm{i}v}\langle Ty, x\rangle = |\langle Ty, x\rangle|$$

choose $s = \frac{1}{2}(u+v)$ and $t = \frac{1}{2}(u-v)$ such that $u = s+t$ and $v = s-t$. Then

$$|\langle Tx, y\rangle| + |\langle Ty, x\rangle| \le 2\alpha\|x\|\|y\|$$

for all $x, y \in \mathcal{H}$. □

Corollary 11.6. *Let* $\mathcal{H}$ *be a Hilbert space and* $T \in L(\mathcal{H})$. *Then*

$$\|T\| = \sup\{|\langle Tx, y\rangle| : \|x\| = \|y\| = 1\} = \inf\{c > 0 : |\langle Tx, y\rangle| \le c\|x\|\|y\|\,\forall x, y \in \mathcal{H}\}\,.$$

If T *is self-adjoint, then*

$$\|T\| = \sup\{|\langle Tx, x\rangle| : \|x\| = 1\} = \inf\{\alpha > 0 : |\langle Tx, x\rangle| \le \alpha\|x\|^2\}\,.$$

Proof. First, we have that

$$\sup_{\|x\|=1} \sup_{\|y\|=1} |\langle Tx, y\rangle| = \sup_{\|x\|=1} \|f_{Tx}\| = \sup_{\|x\|=1} \|Tx\| = \|T\|\,,$$

since the Riesz map is an isometry. With $|\langle Tx, y\rangle| \le c\|x\|\|y\|$ and thus

$$\left|\left\langle T\tfrac{x}{\|x\|}, \tfrac{y}{\|y\|}\right\rangle\right| \le c\,,$$

we also obtain that

$$\sup\{|\langle Tx, y\rangle| : \|x\| = \|y\| = 1\} = \inf\{c > 0 : |\langle Tx, y\rangle| \le c\|x\|\|y\|\}\,.$$

Now, let T be self-adjoint. We only need to prove $\|T\| = \alpha := \sup\{|\langle Tx, x\rangle| : \|x\| = 1\}$. Then the rest follows similarly as above. Obviously,

$$\alpha \le \sup\{|\langle Tx, y\rangle| : \|x\| = \|y\| = 1\} = \|T\|\,.$$

Moreover, since $|\langle Tx, x\rangle| \le \alpha\|x\|^2$ for all $x \in \mathcal{H}$, by Lemma 11.5 we have for $x, y \in \mathcal{H}$ with $\|x\| = \|y\| = 1$:

$$|\langle Tx, y\rangle| = \tfrac{1}{2}|\langle Tx, y\rangle + \langle x, Ty\rangle| \le \tfrac{1}{2}\begin{cases} |\langle Tx, y\rangle + \langle Ty, x\rangle| & \text{if } \mathbb{K} = \mathbb{R} \\ |\langle Tx, y\rangle| + |\langle Ty, x\rangle| & \text{if } \mathbb{K} = \mathbb{C} \end{cases} \le \alpha\,.$$

This proves $\|T\| = \sup\{|\langle Tx, y\rangle| : \|x\| = \|y\| = 1\} \le \alpha$. □

Lemma 11.7. *Let $\mathcal{H}$ be a Hilbert space and $K\colon \mathcal{H} \to \mathcal{H}$ be a compact and self-adjoint operator. Then $\|K\|$ or $-\|K\|$ is an eigenvalue of K.*

Proof. First, observe that

$$\langle Kx, x\rangle = \langle x, Kx\rangle = \overline{\langle Kx, x\rangle}$$

for all $x \in \mathcal{H}$, hence $\langle Kx, x\rangle \in \mathbb{R}$. By Corollary 11.6, there exists a sequence $(x_n)_n \subseteq \mathcal{H}$ such that

$$\|x_n\| = 1 \quad \text{and} \quad \lim_{n\to\infty} |\langle Kx_n, x_n\rangle| = \|K\|\,.$$

It is no restriction to assume $K \neq 0$ and that the real sequence $(\langle Kx_n, x_n\rangle)_n$ is convergent. Set

$$c := \lim_{n\to\infty} \langle Kx_n, x_n\rangle\,.$$

Then $|c| = \|K\|$. Since K is compact and $(x_n)_n$ is bounded, we can furthermore assume that $(Kx_n)_n$ is convergent. Next,

$$\begin{aligned} 0 &\le \|Kx_n - cx_n\|^2 = \|Kx_n\|^2 + \|cx_n\|^2 - 2c\langle Kx_n, x_n\rangle \le \|K\|^2 + c^2 - 2c\langle Kx_n, x_n\rangle \\ &= 2\|K\|^2 - 2c\underbrace{\langle Kx_n, x_n\rangle}_{\to c} \to 2\|K\|^2 - 2c^2 = 0\,. \end{aligned}$$

Hence, $\|Kx_n - cx_n\| \to 0$, $n \to \infty$. Since $(Kx_n)_n$ is convergent and $c \neq 0$, also $(x_n)_n$ is convergent. Set

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

We have $\|x\| = 1$ and

$$Kx - cx = \lim_{n\to\infty} (Kx_n - cx_n) = 0\,.$$

Thus, $c \in \{-\|K\|, \|K\|\}$ is an eigenvalue of K. □

Theorem 11.8 (Spectral Theorem for compact self-adjoint operators)**.** *Let $\mathcal{H}$ be a Hilbert space, and let $0 \neq K\colon \mathcal{H} \to \mathcal{H}$ be a compact self-adjoint operator. Then there exist sequences $(\lambda_n)_n \subseteq \mathbb{R}$ and $(x_n)_n \subseteq \mathcal{H}$ (either finite or infinite sequences) such that*

(i) *the numbers λ_n are ordered by $|\lambda_1| \geq |\lambda_2| \geq \cdots$, $\lambda_n \neq 0$, and $\lim_{n\to\infty} \lambda_n = 0$ (if $(\lambda_n)_n$ is infinite),*

(ii) *the sequence $(x_n)_n$ forms an orthonormal system in $\mathcal{H}$ and $Kx_n = \lambda_n x_n$,*

(iii) *if $\lambda \neq 0$ is an eigenvalue of K, then λ appears in $(\lambda_n)_n$ exactly $\dim E(\lambda)$ times and*

(iv) *for each $x \in \mathcal{H}$,*
$$Kx = \sum_n \lambda_n \langle x, x_n\rangle x_n\,.$$

Proof. By Lemma 11.7, K has an eigenvalue λ_1 with $|\lambda_1| = \|K\|$ and $\lambda_1 \in \mathbb{R}$. Let x_1 be an eigenvector associated to λ_1 with $\|x_1\| = 1$. Now set
$$\mathcal{H}_1 := \operatorname{span}\{x_1\}\,.$$
Since $x \in \mathcal{H}_1^\perp$ implies that
$$\langle Kx, x_1\rangle = \langle x, Kx_1\rangle = \langle x, \lambda_1 x_1\rangle = 0\,,$$
we have
$$K(\mathcal{H}_1^\perp) \subseteq \mathcal{H}_1^\perp\,.$$
The restriction $K|_{\mathcal{H}_1^\perp} : \mathcal{H}_1^\perp \to \mathcal{H}_1^\perp$ is still a compact self-adjoint operator.

<u>Case $K|_{\mathcal{H}_1^\perp} = 0$.</u> Let $x = y + z$, $y \in \mathcal{H}_1$, $z \in \mathcal{H}_1^\perp$. Then
$$Kx = Ky + Kz = K(\langle y, x_1\rangle x_1) = \langle y, x_1\rangle Kx_1 = \lambda_1\langle y, x_1\rangle x_1 = \lambda_1 \langle x, x_1\rangle x_1\,.$$

<u>Case $K|_{\mathcal{H}_1^\perp} \neq 0$.</u> By applying Lemma 11.7 to $K|_{\mathcal{H}_1^\perp}$, there exists some $\lambda_2 \in \mathbb{R}$ with
$$|\lambda_2| = \|K|_{\mathcal{H}_1^\perp}\| \leq \|K\| = |\lambda_1|\,.$$
Let $x_2 \in \mathcal{H}_1^\perp$ be an eigenvector to λ_2 with $\|x\|_2 = 1$, i.e. $Kx_2 = \lambda_2 x_2$, and set
$$\mathcal{H}_2 = \operatorname{span}\{x_1, x_2\}\,.$$
$\mathcal{H}_2$ is a two-dimensional, closed subspace of $\mathcal{H}$. As before we obtain that $K|_{\mathcal{H}_2^\perp} : \mathcal{H}_2^\perp \to \mathcal{H}_2^\perp$ is compact and self-adjoint.

<u>Case $K|_{\mathcal{H}_2^\perp} = 0$.</u> Let $x = y + z$, $y \in \mathcal{H}_2$, $z \in \mathcal{H}_2^\perp$. Then
$$\begin{aligned} Kx = Ky + Kz &= K(\langle y, x_1\rangle x_1 + \langle y, x_2\rangle x_2) \\ &= \lambda_1\langle y, x_1\rangle x_1 + \lambda_2\langle y, x_2\rangle x_2 \\ &= \lambda_1\langle x, x_1\rangle x_1 + \lambda_2\langle x, x_2\rangle x_2\,. \end{aligned}$$

<u>Case $K|_{\mathcal{H}_2^\perp} \neq 0$.</u> Continue this process.

<u>Case 1</u> If at some point, $K|_{\mathcal{H}_n^\perp} = 0$, then
$$Kx = \sum_{i=1}^n \lambda_i \langle x, x_i\rangle x_i \quad \forall x \in \mathcal{H}\,. \tag{11.1}$$

In this case, (i), (ii) and (iv) are satisfied. It remains to show (iii). For this, let $\lambda \neq 0$ be an eigenvalue of K and $x \in E(\lambda)$. Then

$$x = \tfrac{1}{\lambda} K x = \sum_{i=1}^{n} \tfrac{\lambda_i}{\lambda} \langle x, x_i \rangle x_i \,.$$

Moreover, equation (11.1) shows that $K(\mathcal{H}) = \operatorname{span}\{x_1, \ldots, x_n\}$. Hence, by (ii) and since $x = \lambda^{-1} K x \in K(\mathcal{H})$, we also have

$$x = \sum_{i=1}^{n} \langle x, x_i \rangle x_i \,.$$

This implies that $\langle x, x_i \rangle = 0$ for each i with $\lambda \neq \lambda_i$. Thus

$$E(\lambda) = \operatorname{span}\{x_i : \lambda_i = \lambda\} \,.$$

This shows that λ appears $\dim E(\lambda) = n_\lambda$ times in $(\lambda_i)_{i=1}^n$.
<u>Case 2</u> Now assume that the process does not stop. This yields sequences (λ_n) and (x_n) with $\|x_n\| = 1$, $K x_n = \lambda_n x_n$ and $x_n \perp x_m$ for $n \neq m$. Now, (ii) is satisfied. For (i) assume that

$$|\lambda_n| \geq \delta > 0$$

for infinitely many $n \in \mathbb{N}$. This implies

$$\|K x_n - K x_m\|^2 = \|K x_n\|^2 + \|K x_m\|^2 = \lambda_n^2 + \lambda_m^2 \geq 2\delta^2 \,.$$

But this contradicts the compactness of K (no convergent subsequence).
For (iv) define $L = \overline{\operatorname{span}\{x_n : n \in \mathbb{N}\}}$. Then, for $x \in L^\perp$, we have

$$\langle K x, x_n \rangle = \langle x, K x_n \rangle = \lambda_n \langle x, x_n \rangle = 0 \,,$$

thus $K(L^\perp) \subseteq L^\perp$. Now let $x \in L^\perp$. Then, as $x \in \mathcal{H}_n^\perp$ for each n,

$$|\langle K x, x \rangle| \leq \left\| K|_{\mathcal{H}_n^\perp} \right\| \|x\|^2 = |\lambda_{n+1}| \|x\|^2 \to 0, \;\; n \to \infty \,.$$

By Corollary 11.6, $\langle K x, x \rangle = 0$ for all $x \in L^\perp$ implies

$$K|_{L^\perp} = 0.$$

For each $x \in \mathcal{H}$ we write $x = y + z$, $y \in L$, $z \in L^\perp$. Then

$$K x = K y + K z = K y = K \left(\sum_{n \in \mathbb{N}} \langle y, x_i \rangle x_i \right) = \sum_{n \in \mathbb{N}} \lambda_i \langle y, x_i \rangle x_i = \sum_{n \in \mathbb{N}} \lambda_i \langle x, x_i \rangle x_i \,.$$

(iii) can be shown exactly as in the first case. □

Remark 11.9. (1) The orthonormal system $(x_n)_n$ is an orthonormal basis of $(\ker K)^\perp = \overline{K(\mathcal{H})}$, and if P_0 is the orthogonal projection of $\mathcal{H}$ onto $\ker K$, then

$$x = P_0 x + \sum_n \langle x, x_n \rangle x_n, \quad x \in \mathcal{H} \,.$$

(2) Extending $(x_n)_n$ by an orthonormal basis of $\ker K$ yields an orthonormal basis of $\mathcal{H}$, which consists of eigenvectors of K. Hence $\mathcal{H}$ is the direct sum of eigenspaces.

(3) For each eigenvalue $\lambda \neq 0$, let P_λ be the orthogonal projection onto $E(\lambda)$. Then by Theorem 11.8,

$$K = \sum_{\lambda \in \sigma(K)} \lambda P_\lambda \,.$$

This is the *spectral decomposition* of K.

(4) Kx, $x \in \mathcal{H}$, is completely determined by the eigenvalues and eigenvectors of K.

Index

References

[1] H. W. Alt, *Lineare Funktionalanalysis. Eine anwendungsorientierte Einführung*, 6. Aufl., Springer, 2012.

[2] J. B. Conway, *A course in functional analysis*, 2. Aufl., Springer, 1990.

[3] H. Heuser, *Funktionalanalysis. Theorie und Anwendung*, 4. Aufl., Teubner, 2006

[4] W. Rudin, *Functional Analysis*, McGraw-Hill, 1991.

[5] D. Werner, *Funktionalanalysis*, 7. Aufl., Springer, 2011.

[6] O. Christensen, *Frames and bases: an introduction course*, Birkhäuser, 2008