Functional Analysis I

Lecture Notes

Martin Genzel

Sommersemester 2019

TECHNISCHE UNIVERSITÄT BERLIN
Institut für Mathematik

Acknowledgments

Many thanks go to Martin Schäfer for very carefully proofreading these lecture notes.

Contents

1 Metric Spaces

1.1 Basic Definitions

Metric spaces are arbitrary sets with the only structure being a distance measure. Such a distance measure should certainly be zero only if the two points are equal, it should be symmetric, and satisfy a triangle inequality. And this is exactly how it is defined, as we will see now.

Throughout this lecture, notice that $\mathbb{K}$ 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 $x, y, z \in X$, the following conditions are satisfied:

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

(2) $d(x, y) = d(y, x)$.

(3) $d(x, z) \leq d(x, y) + d(y, z)$. (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 $Y \subseteq X$, then $d|_{Y\times Y}$ is the *induced metric on* Y.

We next define what we mean by open and closed sets. These notions are based on the concept of open balls. An open ball is exactly what one would expect, since it consists of those points whose distance from a center point is less than a particular value called the radius of the ball. The definition of openness then also gives rise to notions such as neighborhood, etc.

Definition 1.2 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. $U \subseteq X$ is called *open*, if for each $x \in U$ there exists some $\varepsilon > 0$ such that $U_\varepsilon(x) \subset U$.

(2) A set $A \subseteq X$ is *closed*, if $X \backslash A$ 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 $E \subseteq X$, then $x \in E$ is an *interior point of E*, if there exists some open set $U \subseteq E$ with $x \in U$. Then U is 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 $U \cap E \neq \varnothing$ 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 $\overline{E} = X$.

Next we introduce the notion of convergence, which is also a natural extension from $\mathbb{K}^n$ (with the Euclidean metric) to metric spaces. Notice that basically the Euclidean metric is replaced by a general metric.

Definition 1.3 Let (X,d) be a metric space.

(1) A sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ *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 $x = \lim\limits_{n\to\infty} x_n$. x is called the *limit* of $(x_n)_{n\in\mathbb{N}}$.

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

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

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

1.2 Theorem of Baire

Our first main theorem is the Theorem of Baire, which shows that quite surprisingly an arbitrary countable intersection of open and dense sets is again dense. Note however that this theorem only holds in complete metric spaces, showing also that completeness is a quite strong assumption.

This theorem – better to say a corollary from it – will be one key ingredient for some fundamental theorems in functional analysis as we will later see.

Theorem 1.4 (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 $\bigcap\limits_{n\in\mathbb{N}} D_n$ is dense in X.*

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

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

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

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

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 $x_1, \ldots, x_n, r_1, \ldots, r_n$ are already chosen ($n \geq 1$). Since D_n is open and dense, also $D_n \cap U_{r_n}(x_n) \neq \varnothing$ is open. Hence there exists $x_{n+1} \in X$ and $r_{n+1} > 0$ with

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

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

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

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. ■

As already indicated, we next show that indeed the Theorem of Baire is in general false if the metric space is not complete. That certainly does not mean that there cannot be any artificial not complete metric spaces in which the theorem is true for a specific sequence of open and dense subsets. But in general, the statement of the theorem is false once the completeness assumption is deleted.

Remark 1.5 (1) Theorem 1.4 is in general false if X is not complete. As an example choose $X = \mathbb{Q} = \{q_1, q_2, \ldots\}$ and $D_n = X \backslash \{q_n\}$, $n \in \mathbb{N}$, which are open and dense. We immediately see that however

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

(2) 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).

$\Diamond$

We now draw one corollary from the Theorem of Baire, which will be used extensively in the sequel.

Corollary 1.6 *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*

$$\mathring{A}_n \neq \varnothing.$$

Proof. Towards a contradiction, assume that

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

Then $X \backslash A_n$ are open and dense for all $n \in \mathbb{N}$. By Baire's Theorem 1.4, $\bigcap_{n=1}^{\infty} (X \backslash A_n)$ is dense in X. But $\bigcap_{n=1}^{\infty} (X \backslash A_n) = X \backslash \bigcup_{n=1}^{\infty} A_n = \varnothing$, a contradiction. ∎

1.3 Compactness and Continuity

Compactness turned out to be a quite useful property in Analysis I, and the generalization to arbitrary metric spaces – once the notion of openness is defined – is not too surprising. The notion of totally boundedness which we will also define is though not necessary in the situation of $\mathbb{K}^n$. The reason for this is that totally boundedness is precisely the missing concept to derive an equivalence between compactness and completeness (see the next theorem). Now, since $\mathbb{K}^n$ is complete, this notion equals compactness. But in a general metric space, totally boundedness turns out to be a highly useful property.

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

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

(2) 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. $(A, d|_{A \times A})$ 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, which means that a subset $A \subseteq (X, d)$ of some metric space is compact or totally bounded, if the metric space $(A, d|_{A \times A})$ is compact or totally bounded, respectively.

It is quite straightforward to prove that every compact metric space is totally bounded. As already indicated before, we will now show that the notions of compactness and total boundedness coincide for complete metric spaces; in fact, compactness is equivalent to completeness and total boundedness. Notice however that this does not imply that these two properties coincide for all subsets of a complete metric space (see Corollary 1.10).

Theorem 1.8 *Let (X, d) be a metric space. Then the following conditions are equivalent.*

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

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

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

Proof. (i) $\Rightarrow$ (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 $(x_n)_{n \in \mathbb{N}} \subset X$ satisfying

(a) $U_{2^{-n}}(x_n)$ is not covered by finitely many $U \in \mathfrak{U}$.

(b) $U_{2^{-n}}(x_n) \cap U_{2^{-(n+1)}}(x_{n+1}) \neq \varnothing$.

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 $U \in \mathfrak{U}$. 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 $U_{2^{-(n+1)}}(y) \cap U_{2^{-n}}(x_n) \neq \varnothing$, the set $U_{2^{-(n+1)}}(y)$ is covered by finitely many $U \in \mathfrak{U}$. Then this is also true for $U_{2^{-n}}(x_n)$, a contradiction. Hence there exists $x_{n+1} \in X$ such that $U_{2^{-(n+1)}}(x_{n+1})$ is not covered by finitely many $U \in \mathfrak{U}$ and $U_{2^{-n}}(x_n) \cap U_{2^{-(n+1)}}(x_{n+1}) \neq \varnothing$.

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 $x = \lim_{n \to \infty} x_n$ with the property $d(x, x_n) \leq \frac{1}{2^{n-2}}$.

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 $n \in \mathbb{N}$ if $2^{-n} < \frac{\varepsilon}{8}$, and for these $n \in \mathbb{N}$ also $U_{2^{-n}}(x_n) \subset U$ which is a contradiction to (a) (the 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

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

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

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

Thus $A_N = \varnothing$, a contradiction. This proves that $\bigcap_{n\in\mathbb{N}} A_n \neq \varnothing$. 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 $d(x_{n_k}, x) \le \frac{1}{k}$ [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 subsequence, 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

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

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

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

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

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

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 there is $k_0 \in \mathbb{N}$ such that $d(x_{n_k}, x) < \frac{\varepsilon}{2}$ for all $k > k_0$, hence $d(x_{n_k}, x_{n_l}) < \varepsilon$ for all $k, l > k_0$, a contradiction. ∎

Certainly, when introducing a new property for metric spaces, we have to check when and how subspaces inherit this property. In this sense, (i) is not too surprising. But (ii) is quite neat showing that this property is even inherited by a larger space, namely the closure.

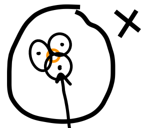
Week 2
Dien
16.4

Lemma 1.9 *Let (X, d) be a metric space, and let $A \subset X$, ~~$A \neq \varnothing$.~~*

(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 $A \cap U_{\frac{\varepsilon}{2}}(x_j) \neq \varnothing$ 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 $\{y_1, \ldots, y_n\}$ is an $\varepsilon-$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

Randpunkt
(limit point)

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

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

The next result seems to be a bit unexciting. However, this observation will help us in the proof of the Theorem of Arzela-Ascoli. Its proof is a beautiful combination of all previously derived results for totally boundedness.

Corollary 1.10 *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)$\Rightarrow$(ii). Since $\overline{A}$ is compact, by Theorem 1.8, $\overline{A}$ is totally bounded. By Lemma 1.9, A is totally bounded.

(ii)$\Rightarrow$(i). Since A is totally bounded, by Lemma 1.9, $\overline{A}$ is totally bounded. Since X is complete, $\overline{A}$ is also complete. Hence Theorem 1.8 implies that $\overline{A}$ is compact. ∎

The next notions are very well known from Analysis 1, and their extension to the metric space setting is – as for the notion of convergence – quite straightforward and natural.

Notice the combination of algebraic properties such as bijectivity and analytic properties such as continuity in the notion of a homeomorphism. The interplay between these two "worlds", namely algebra and analysis, will often be at the heart of upcoming results.

Definition 1.11 Let (X, d) and (X', d') be metric spaces, and let $f: X \to X'$.

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

isometry + bijectivity

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

ein delta für alle X

We next aim to relate relative compactness, i.e., the compactness of the closure, of a set of continuous real functions to pointwise relative compactness of these functions. Recall that the relatively compact sets in $\mathbb{R}$ are the bounded sets by the Theorem of Heine-Borel. If a set of continuous real functions is relatively compact, it is not too difficult to 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 direction, we require a new property of a family of functions, namely equicontinuity.

compact + closed = bounded + closed

Equicontinuity can be thought of as continuity in a point, which is uniform with respect to the family of functions.

continuous

Definition 1.12 A family of functions $F \subset C(X)$ is *equicontinuous in* $x \in X$, if, for each $\varepsilon > 0$, there exists a neighbourhood U of x with

$$|f(x) - f(y)| < \varepsilon \quad \text{for all } y \in U \text{ and } f \in F.$$

F is called *equicontinuous*, if it is equicontinuous in each $x \in X$.

Abschluss ist kompakt = relativ kompakt

$f_1, f_2, f_3, \dots$

U

$f_n(y) - f_n(x) < \varepsilon$

Armed with this notion, we can now state and prove the already announced Theorem of Arzela-Ascoli.

Theorem 1.13 (Arzelà-Ascoli) *Let X be a compact metric space and $F \subset C(X)$. Then the following are equivalent.*

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

How do compact sets look like?

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

(ii) F is equicontinuous and pointwise bounded.

↳ $\sup\{|f(x)| : f \in F\} < \infty \quad \forall x \in X$

Proof. (i) $\Rightarrow$ (ii). This is an exercise.

(ii) $\Rightarrow$ (i). For $x \in X$ we write $F(x) := \{f(x) : f \in F\}$. Let F be equicontinuous and $F(x) \subset \mathbb{K}$ bounded for all $x \in X$. Since $C(X)$ is complete, by Corollary 1.10, it remains

Key idea: discretize

compact means you can discretize

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

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

(equicontinuous)

Let now $x_1, \ldots, x_n \in X$ be chosen such that $\underbrace{X = \bigcup_{i=1}^{n} U_{x_i}}_{\text{compact}}$ 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 has finite elements

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

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

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

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 $f \in F_\varphi$.

to show

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. $\square$

bounded $\Leftarrow$ F bounded

$\Downarrow$ complete

compact

Pick $h_\varphi \in F_\varphi \ \forall \varphi \in \Phi$

$\Rightarrow \{h_\varphi : \varphi \in \Phi\}$ are an ε-net for F.

2 Normed Spaces

2.1 Key Definition

vector space = linear space

In metric spaces we can quantify the distance between two elements. If we now turn to vector spaces (also referred to as linear spaces), we would in addition like to quantify the distance of an element (a vector) to the zero vector. This is precisely the interpretation of the norm of a vector.

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 on* E, and $(E, \|\cdot\|)$ a *normed space*, if, for all $x, y \in E$, $\lambda \in \mathbb{K}$, the following conditions are satisfied:

   a) $\|x\| = 0 \Leftrightarrow x = 0$,

   b) $\|\lambda x\| = |\lambda| \cdot \|x\|$, and

   c) $\|x + y\| \leq \|x\| + \|y\|$.

   $d_{\|\cdot\|} = \|x-y\|$

   E is called a *Banach space*, if $(E, d_{\|\cdot\|})$ 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

$$\alpha \|x\|_1 \leq \|x\|_2 \leq \beta \|x\|_1 \text{ for all } x \in E.$$

Since the set of metric spaces should include the set of normed spaces, we need to check whether each normed space also possesses a metric. Moreover, we are interested in analytic properties of the vector space operations $+ \colon E \times E \mapsto E, (x, y) \mapsto x + y$ and $\cdot \colon \mathbb{K} \times E \mapsto E, (\lambda, y) \mapsto \lambda \cdot y$. Using the metric given by the norm, we can speak of continuity properties, and in fact show that those maps are continuous.

Week 2 Wed. 17.4

This and other easy-to-see results are contained in the following remark.

Remark 2.2 Let E be a linear space.

(1) If $\|\cdot\|$ is a norm on E, then $d_{\|\cdot\|}(x, y) = \|x - y\|$ defines a metric on E. gezeigt im Tut

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

   im metrischen Raum gilt das nicht

(3) $\big| \|x\| - \|y\| \big| \leq \|x - y\|$. In particular $\|\cdot\| \colon E \to \mathbb{R}$ is Lipschitz-continuous.

(4) The algebraic operations

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

are continuous, with $E \times E$ and $\mathbb{K} \times E$ metrized by the respecive norms

$$\|(x, y)\| = \max\{\|x\|, \|y\|\} \quad \text{and} \quad \|(\lambda, x)\| = \max\{|\lambda|, \|x\|\}.$$

This is due to the estimates

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

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

2.2 Quotient Spaces as Normed Spaces

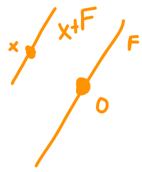
don't require any norm

Quotient spaces are interesting in their own right, since one might want to factor out the kernel of a map to derive an injective map. They also – in a certain sense – "bridge the gap" between a subspace and the full space. We will later use this property quite often by showing that a space has a certain property if and only if a subspace as well as the associated quotient space exhibit this property.

Let's start with the definition.

Definition 2.3 (Quotient space) Let E be a linear space, $F \subset E$ a subspace. Then

$$x \sim y :\Leftrightarrow x - y \in F \qquad (x, y \in E)$$



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

$$E/F := \{\underbrace{x + F}_{\text{affine subspace}} : x \in E\}.$$

Via

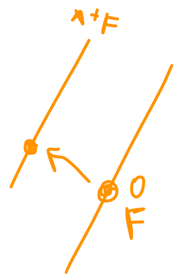
$$[x]_\sim + [y]_\sim := [x + y]_\sim \qquad (x + F) + (y + F) := ((x + y) + F)$$

and

$$\lambda[x]_\sim := [\lambda x]_\sim \qquad \lambda(x + F) := ((\lambda x) + F)$$

the space E/F becomes a linear space.

First, we need to equip a quotient space with a norm. For this, we require a *closed* subspace. Check for yourself what happens, if the subspace is not closed. Why is the defined map $\|\cdot\|: E/F \to [0, \infty)$ in this case not a norm?



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

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

x is fixed

defines a norm on E/F. Moreover, if E is a Banach space, so is E/F. ← F is closed (due to norm)

Proof. We start by proving the the norm properties.

because of inf

(i) Let $\|x + F\| = 0$, this implies that there exists $(y_n)_{n \in \mathbb{N}} \subset F$, such that

$$\|x - y_n\| \to 0, \; n \to \infty.$$

Since F is closed, we have $x \in F$ and thus

$$x + F = F = [0]_\sim.$$

(ii) We first observe that

$$\|\lambda(x + F)\| = \|(\lambda x) + F\| = \inf\{\|\lambda x + y\| : y \in F\}.$$

lecturer gave the main idea

We now need to distinguish two cases. For $\lambda = 0$, we have

$$\|\lambda(x + F)\| = \inf\{\|y\| : y \in F\} = 0 = |\lambda| \|x + F\|.$$

And for $\lambda \neq 0$, we obtain

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

where in the second step $\lambda y \in F$ was used.

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

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

This yields

$$\begin{aligned} \|(x + F) + (y + F)\| &= \|(x + y) + F\| \\ &\leq \|x + z_1 + y + z_2\| \\ &\leq \|x + F\| + \|y + F\| + \varepsilon. \end{aligned}$$

Since this estimate is valid for all $\varepsilon > 0$ the triangle inequality follows.

to show: E/F is Bsp

Next, we show that E being a Banach space implies that E/F is Banach. For this, let E be complete and let $(x_n + F)_{n\in\mathbb{N}}$ be a Cauchy-sequence in E/F, i.e., for all $\varepsilon > 0$ there exists some $N \in \mathbb{N}$ such that

this should converge

$$\|(x_n - x_m) + F\| \leq \varepsilon \quad \text{for all } n, m \geq N.$$

Hence, there exists a subsequence $(n_i)_{i\in\mathbb{N}} \subseteq \mathbb{N}$ with $n_i < n_{i+1}$ for all $i \in \mathbb{N}$, such that

increasing

$$\|x_{n_{i+1}} - x_{n_i} + F\| < 2^{-i}.$$

F is closed!

As a consequence, for every $i \in \mathbb{N}$, there exists some $y_i \in F$ such that

def. of norm

$$\|x_{n_{i+1}} - x_{n_i} + y_i\| < 2^{-i}.$$

E/F is BSp?

Let us define

$$z_1 := 0$$
$$z_{i+1} := \underbrace{y_i + z_i}_{\in F} \qquad i \geq 1.$$

Idee: mit Elementen aus E arbeiten

Then we have

$y_i = z_{i+1} - z_i$

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

Now, we define $\overset{\in E}{\eta_i} := x_{n_i} + z_i$, which yields $\|\eta_{i+1} - \eta_i\| < 2^{-i}$. This in turn implies

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

showing that $(\eta_n)_{n\in\mathbb{N}}$ is a Cauchy-sequence in E. Since E is by assumption a Banach space, the sequence $(\eta_n)_{n\in\mathbb{N}}$ converges.

We now set $\lim_{n\to\infty} \eta_n =: x$ to obtain

Norm von E/F Norm von F

$$\|(x_{n_i} + F) - (x + F)\| = \|(x_{n_i} - x) + F\| \overset{z_i \in F}{\leq} \|x_{n_i} + z_i - x\| = \|\eta_i - x\| \to 0.$$

Thus we obtain a convergent subsequence. But this implies that the Cauchy-sequence is convergent itself. □

Now we are ready to derive our first result showing a relation between properties of the full space E and those of a closed subspace $F \subset E$ and its associated quotient space E/F. One advantage of such a result is the fact that it is often easier to prove properties – such as being Banach in this case – for the "smaller" spaces.

Theorem 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.*

F and E/F should be complete

Highlight

$E = F \oplus G \iff E = F + G,\ F \cap G = \{0\}$

Proof. Let $(x_n)_{n\in\mathbb{N}} \subset E$ be a Cauchy-sequence in E. Since, for any $n, m \in \mathbb{N}$,

To show: E is Banach

$E \subset F$

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

the sequence $(x_n + F)_n \subset E/F$ is a Cauchy-sequence in E/F. Using the hypothesis that E/F is a Banach space, we can set $x + F := \lim_{n\to\infty} x_n + F$. We then obtain

$$\inf\{\|x_n - x + y\| : y \in F\} = \|(x_n - x) + F\| \to 0 \quad \text{as } n \to \infty.$$

Thus there exists $(y_n)_{n\in\mathbb{N}} \subset F$ with $\|x_n - x + y_n\| \to 0$.

Bsp

Since

show y_n is a CSq

$$\begin{aligned}\|y_n - y_m\| &= \|y_n + x_n - x - x_n + x_m - y_m - x_m + x\| \\ &\leq \|y_n + x_n - x\| + \|x_n - x_m\| + \|y_m + x_m - x\| \to 0 \quad \text{as } m, n \to \infty,\end{aligned}$$

we can conclude that the sequence $(y_n)_{n\in\mathbb{N}}$ is in fact a Cauchy sequence. Now we use that also the space F is assumed to be a Banach space, which implies that $(y_n)_{n\in\mathbb{N}}$ converges. Setting $y := \lim_{n\to\infty} y_n \in F$, we derive

$$\|x_n - x + y\| \leq \|x_n + y_n - x\| + \|y - y_n\| \to 0 \quad \text{as } n \to \infty,$$

Thus $(x_n)_{n\in\mathbb{N}}$ converges to $x - y$. □

2.3 Life is easier in Finite-Dimensional Spaces <3

Indeed the situation is much easier – but also less exciting – in finite-dimensional spaces than in infinite-dimensional ones. One instance of this is the following theorem. Another one is the fact that in finite-dimensional spaces all norms are equivalent (see Theorem 2.9).

Theorem 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. This implies that $E \cong \mathbb{R}$ and E is a Banach space.

$d(x,y)=d($[illegible]$)$,

Next assume that the claim holds for n and let E be such that $\dim E = n + 1$. Then we choose $x \in E \setminus \{0\}$ and set $F := \operatorname{span}\{x\}$. Since $\dim F = 1$, we know that F is complete, hence closed. We can compute the dimension of the quotient space E/F to be

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

By the assumption in the induction step, the space E/F is complete.

Finally, application of Theorem 2.5 implies that E is a Banach space. ■

The next two lemmata are a bit technical, but prepare the stage for Theorem 2.9.

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*

$$|\lambda| \leq M\|\lambda x + y\| \quad \textit{and} \quad \|y\| \leq M'\|\lambda x + y\|.$$

Proof. Since $x \notin F$, it follows that $\|x + F\| \neq 0$. This allows us to set

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

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

$$|\lambda| = M|\lambda|\|x + F\| = M\|\lambda x + F\| \leq M\|\lambda x + y\|$$

as well as

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

This proves the lemma. ∎

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

$$\|Tx\|_X \leq c\|x\|_E.$$

Proof. We prove the statement by induction over $n = \dim E$. For this, let first $n = 1$ and choose $x_0 \in E$ with $\|x_0\| = 1$. Then

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

The case $n = 1$ is proven by the observation that, for $x = \lambda x_0 \in E$, we have

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

Next assume that the statement holds for $\dim E = n$. Let E be a normed space with $\dim E = n + 1$, and choose an n-dimensional subspace $F \subset E$. Moreover, let $x_0 \in E \setminus F$. Then, for each $x \in E$, there exist $\lambda \in \mathbb{K}$ and $y \in F$ such that

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

i.e., $E = F \dot{+} \operatorname{span}\{x_0\}$. Since $\dim F = n$, by assumption, we can choose $c' > 0$ with

$$\|Ty\|_X \leq c'\|y\|_E \quad \text{for all } y \in F.$$

By Theorem 2.6, the subspace F is closed. Application of Lemma 2.7 now yields

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

The lemma is proven. ∎

The next theorem states several results which show that many results known from Analysis 1 also hold in the situation of a finite-dimensional normed space. This makes life easy in finite-dimensional spaces. For instance, in arbitrary normed spaces, it is always crucial to know which norm the space is equipped with. But in a finite-dimensional space – as part (i) shows – all norms are equivalent. Hence with respect to properties such as convergence, it is of no importance which norm is used.

Theorem 2.9 *(i) Each two norms on a finite-dimensional space are equivalent.*

(ii) 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

$$\|Tx\| \leq c\|x\| \quad \textit{and} \quad \|T^{-1}y\| \leq d\|y\|.$$

(iii) Let E be a normed space with $\dim E < \infty$ and $A \subset E$ a subset. Then the following conditions are equivalent:

(a) A is compact.

(b) A is bounded and closed.

Proof. (i). Consider the linear map

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

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

$$\|T^{-1}x\|_1 \leq c\|x\|_2 \quad \text{and} \quad \|Tx\|_2 \leq d\|x\|_1$$

for some $c, d > 0$. Since $Tx = x$, this proves the claim.

(ii). Since T^{-1} is linear, Lemma 2.8 implies the claim.

(iii). It is well known that the statement holds for $\mathbb{K}^n$ with the Euclidean norm. Letting $\dim E = n$, there exists a bijective linear map $T\colon \mathbb{K}^n \to E$. By (ii), there exist $c, d > 0$ with $\|Tx\| \leq c\|x\|$ and $\|T^{-1}y\| \leq d\|y\|$. This implies that A is bounded (closed, compact) if and only if $T^{-1}A$ is bounded (closed, compact) in $\mathbb{K}^n$, which finishes the proof by using Heine-Borel. ∎

As the final result of this chapter, we derive several equivalence conditions for a normed space to be finite-dimensional. Of particular importance – and also quite intriguing, since it might be a bit unintuitive – is part (iii), which will also be used in the sequel.

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

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

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

(iii) *$K_1(0)$ is compact.*

Proof. (i)⇒(ii). This follows from Theorem 2.9(iii).

(ii)⇒(iii). Let K be a compact neighborhood of 0. Then there exists $\delta > 0$ with $K_\delta(0) \subset K$. $K_\delta(0)$ 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 $y_1, \ldots, y_n$ of $K_1(0)$. Let now $F := \operatorname{span}(y_1, \ldots, y_n)$. Notice that it suffices to show that $E = F$.

Towards a contradiction, assume that $F \neq E$. Then there exists $x_0 \in E \backslash F$. Since F is finite-dimensional, hence closed, we have

$$\operatorname{dist}(x_0, F) = \inf_{y \in F} \|x_0 - y\| > 0.$$

Next, let $y_0 \in F$ be such that

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

Defining

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

for all $y \in F$, we obtain

$$\|x - y\| = \|x_0 - y_0\|^{-1} \|x_0 - (y_0 + \|x_0 - y_0\| y)\|. \tag{2.1}$$

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

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

By (2.1),

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

Thus

$$x \notin \bigcup_{i=1}^{n} U_{\frac{1}{2}}(y_i),$$

a contradiction to $x \in K_1(0)$. ■

3 Linear Operators and Dual Spaces

3.1 The Space of Linear and Bounded Operators

Whereas in the last chapter we focussed more on the normed spaces themselves, we turn now to the study of linear operators on such spaces. The first quite surprising result about such operators is the following lemma. It states that considering linear operators, continuity is equivalent to boundedness. Also, continuity needs to be checked only in one point. Recalling results on continuous functions from Analysis 1, this is certainly quite unexpected. But it shows how strong the requirement to the operator being linear is.

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

(i) T is continuous on E.

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

(iii) 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 some $\delta > 0$ such that, for all $x \in E$ with $\|x - x_0\| \leq \delta$, we have

$$\|Tx - Tx_0\| \leq 1.$$

Now let $y \in E$, $y \neq 0$. Further, set

$$x := x_0 + \frac{\delta y}{\|y\|}.$$

Then $\|x - x_0\| \leq \delta$, and hence

$$\frac{\delta}{\|y\|}\|Ty\| = \left\|T\left(\frac{\delta y}{\|y\|}\right)\right\| = \|Tx_0 - Tx\| \leq 1,$$

which implies $\|Ty\| \leq \frac{1}{\delta}\|y\|$.

(iii)$\Rightarrow$(i). Let $\varepsilon > 0$ and assume that $\|T(x - x_0)\| \leq c\|x - x_0\|$. Then, if $\|x - x_0\| \leq \frac{\varepsilon}{c}$ we have $\|T(x - x_0)\| \leq c\|x - x_0\| < \varepsilon$. This in turn implies

$$T(K_{\frac{\varepsilon}{c}}(x_0)) \subseteq K_\varepsilon(Tx_0),$$

showing that T is continous. ∎

We next introduce the set of linear and bounded operators from E to F with E and F being normed spaces. The next lemma will show, that this object is even a linear space. Certainly, one then aims to endow this space with a norm to turn it into a normed space. This is achieved by the so-called operator norm, which we also introduce in the next definition.

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

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

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

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

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

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

The following lemma then shows that we indeed obtain a normed linear space. The other properties such as the numerous ways to write the norm will turn out to be highly useful later on.

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

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

$$\|T\| = \sup_{\|x\|=1} \|Tx\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \inf\{c : \|Tx\| \leq c\|x\| \quad \forall x \in E\}.$$

(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). For $T_1, T_2, T \in L(E, F)$, $\lambda \in \mathbb{K}$ and arbitrary $x \in E$, we have

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

and

$$\|(\lambda T)x\| = |\lambda| \, \|Tx\| \leq |\lambda| \, \|T\|\|x\|.$$

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.

This computation also provides us some insight into the norm property. Since it is clear that $\|T\| = 0$ implies $T = 0$, we can already conclude that $\|\cdot\|$ is a norm on $L(E, F)$.

We next check the equivalent forms to write the norm. For this, observe that

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

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

$$\|Ty\| \leq c \quad \text{for all } \|y\| \leq 1 \quad \text{implies} \quad \|T\| \leq c.$$

This shows that

$$\inf\{c : \|Tx\| \leq c\|x\| \quad \forall x \in E\} \geq \|T\|. \tag{3.1}$$

On the other hand, for each $x \neq 0$, we have

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

Thus $\|T\| \in \{c : \|Tx\| \leq c\|x\| \quad \forall x \in E\}$, which implies

$$\|T\| \geq \inf\{c : \|Tx\| \leq c\|x\| \quad \forall x \in E\},$$

and with (3.1), we obtain

$$\inf\{c : \|Tx\| \leq c\|x\| \quad \forall x \in E\} = \|T\|.$$

(ii). This is an exercise. ■

We will next introduce a very special space among the spaces $L(E, F)$, namely when choosing $F = \mathbb{K}$.

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^*.

Since $L(E, F)$ is a normed space, so is E^*. But we certainly would also like to know whether E^* forms a Banach space. Interestingly, we can show an even stronger result for the more general spaces $L(E, F)$.

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 $(T_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $L(E,F)$. Then, since

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

also $(T_n x)_{n\in\mathbb{N}}$ is a Cauchy sequence in F for each $x \in E$. We now use the completeness of F to conclude that, for every $x \in E$, there exists some $Tx := \lim_{n\to\infty} T_n x$. This then defines a mapping $T\colon E \to F$. We however do not know any functional analytic properties of this map. Thus we next need to prove that $T \in L(E,F)$.

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

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

Next we prove that T is bounded. For this, let $\varepsilon > 0$. Then there exists $N_\varepsilon \in \mathbb{N}$ such that $\|T_n - T_m\| < \varepsilon$ for all $m, n \geq N_\varepsilon$. This implies

$$\|T_n x - T_m x\| \leq \|T_n - T_m\| \|x\| < \varepsilon \|x\| \quad \text{for all } m, n \geq N_\varepsilon, x \in E.$$

Letting $n \to \infty$ we obtain

$$\|Tx - T_m x\| \leq \varepsilon \|x\| \quad \text{for all } m \geq N_\varepsilon. \tag{3.2}$$

Hence, in particular,

$$\|Tx - T_{N_\varepsilon x}\| \leq \varepsilon \|x\|.$$

This yields

$$\|Tx\| \leq \varepsilon\|x\| + \|T_{N_\varepsilon} x\| \leq (\varepsilon + \|T_{N_\varepsilon}\|) \cdot \|x\| \quad \text{for all } x \in E,$$

which proves $T \in L(E,F)$.

Finally, it remains to show that $T_n \to T$ in the operator norm. By using (3.2), we obtain $\|T - T_m\| \leq \varepsilon$, which implies $\lim_{m\to\infty} T_m = T$. ∎

3.2 Some Useful Insights into Linear Operators

We start with a result, which shows uniqueness of extension of a linear bounded operator from a subspace to its closure. This is a key property, which will later on be shown in more generality, in particular, in the Hahn-Banach theorem.

Lemma 3.6 *Let E be a normed space over $\mathbb{K}$, let $L \subset E$ be a subspace, let F be a Banach space, and let $T \in L(L,F)$. Then there exists a unique $S \in L(\overline{L}, F)$ with $S|_L = T$. Moreover,* *it satisfies* $\|S\| = \|T\|$.

Proof. Let $x \in \bar{L}$, and let $(x_n)_{n\in\mathbb{N}} \subset L$ be a sequence which converges to x. We first observe that

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

This implies that $(Tx_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in F. Since F is assumed to be complete, the sequence $(Tx_n)_{n\in\mathbb{N}}$ converges. At this point, we are not allowed to define the operator S by $Sx := \lim_{n\to\infty} Tx_n$, since we do not know whether this is well-defined. To check this, let $(y_n)_{n\in\mathbb{N}}$ be any other sequence, which also converges to x. Then

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

implying that $\lim_{n\to\infty} Tx_n = \lim_{n\to\infty} Ty_n$. Thus, now we are allowed to define

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

To check that $S|_L = T$, notice that for $x \in L$ we can just choose $x_n := x$ for all $n \in \mathbb{N}$ as "defining sequence". This immediately implies $S|_L = T$.

The linearity of S is also easily proven by

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

and analogously $S(\lambda x) = \lambda Sx$.

Next, we prove $\|S\| = \|T\|$. We first notice that $\|S\| \geq \|S|_L\| = \|T\|$. To prove the reverse inequality, let $0 \neq x \in \bar{L}$ 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, $\|S\| = \|T\|$. We also proved along the way that S is bounded. Thus $S \in L(\bar{L}, F)$.

It only remains to prove the uniqueness of S. Consider another continuous linear operator R with $R|_L = T$, and $x \in \bar{L}$. For any sequence converging to x, using the continuity of S, we then have

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

The proof is finished. ∎

Having two operators, one is certainly interested to have weak and easy-to-check conditions for them to coincide. The next result, which is a direct consequence from the previous lemma – in particular the "uniqueness-part" –, provides such a condition.

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

Next, we study the existence of a linear inverse operator, which is in addition also continuous. We note that the equivalent condition for its existence looks somehow like a "reversed" version of the boundedness condition for a linear operator. Recall that every linear operator is continuous if and only if it is bounded.

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

(i) There exists a linear, continuous, inverse operator

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

(ii) There exists $c > 0$ such that $c\|x\| \leq \|Tx\|$ for all $x \in E$.

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

$$\|T^{-1}y\| \leq \gamma\|y\| \quad \text{for all } y \in T(E).$$

For an arbitrary $x \in E$, we then set $y := Tx$ to obtain

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

Finally, with $c := \frac{1}{\gamma}$, we have shown (ii).

(ii)⇒(i). We observe that (ii) implies the injectivity of T, since $x \in \ker(T)$ implies $\|x\| = 0$. Thus $T^{-1}\colon T(E) \to E$ exists. Now letting $y = Tx$ in (ii) ensures the existence of some $c > 0$ with

$$c\|T^{-1}y\| \leq \|y\| \quad \text{for all } y \in T(E).$$

Thus the inverse operator is indeed continuous. ∎

The graph of a function is a well-known object. Its generalization to operators is in fact not a surprise. Related to this is the notion of a closed operator. Both notions will come into play in the closed graph theorem, which we will state and prove later in this lecture.

Definition 3.9 (Graph of T) Let E and F be normed spaces, let $L \subseteq E$ be a subspace, and let $T\colon L \to F$ be a linear operator.

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

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

(ii) T is called *closed*, if G_T is closed.

Studying the notion of closedness of an operator, the next lemma provides us with a useful equivalent condition, which is easier to check.

Lemma 3.10 *Let E and F be normed spaces, let $L \subseteq E$ be a subspace, and let $T\colon L \to F$ be a linear operator. Then the following conditions are equivalent.*

(i) T is closed.

(ii) If $(x_n)_{n\in\mathbb{N}} \subseteq L$ converges to $x \in E$ and $(Tx_n)_{n\in\mathbb{N}}$ to $y \in F$, then $x \in L$ and $y = Tx$.

Proof. In general, notice that since

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

it follows that $x_n \to x$ and $Tx_n \to y$ implies

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

(i) $\Rightarrow$ (ii). Because of G_T being closed, $(x, y) \in G_T$. Thus $x \in L$ and $y = Tx$.

(ii) $\Rightarrow$ (i). Now consider a sequence $((x_n, y_n))_n$ in G_T which converges to some $(x, y) \in E \times F$ in. Because of the convergence of $y_n = Tx_n$ and (ii), it follows that $x \in L$ and $y = Tx$. This shows that $(x, y) \in G_T$, hence closedness of G_T. ∎

We finish this part with a short remark about sufficient conditions for closedness of an operator.

Remark 3.11 If L is closed and T is continuous, then T is closed. In particular, each $T \in L(E, F)$ is closed.
This can be proven as follows: If $(x_n, Tx_n) \to (x, y)$, then $x \in L$, since L closed. Continuity of T now implies $Tx_n \to Tx$. This shows $(x, y) \in G_T$, hence closedness of G_T and thus of T. ◊

3.3 Back to Dual Spaces

We will finish this chapter with some more results on dual spaces. We start with an interesting example of a normed space, for which we are able to compute its dual space explicitly up to an isometric isomorphism. Even more intriguing is the fact that the dual of this normed space, which is ℓ_p for $1 \leq p < \infty$, is again of the type of a sequence space. Moreover, the dual for $p = 2$ is the normed space ℓ_2 itself making this a self-dual space.

Theorem 3.12 *Let $1 \le p < \infty$, and define q so that $\frac{1}{p} + \frac{1}{q} = 1$, i.e.*

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

For each $y = (y_n) \in \ell_q$, 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, we have $\ell_q \cong \ell_p^*$.

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

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

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

$$\|f_y(x)\| \le \|x\|_p \|y\|_q.$$

Thus, f_y is bounded. We conclude that $f_y \in \ell_p^*$.

We now claim that $\|f_y\| = \|y\|_q$. We already showed that $\|f_y\| \le \|y\|_q$. To prove the inverse inequality, we consider the cases $p = 1$, $p > 1$ separately.

$\underline{p = 1}$: In this case $q = \infty$. Let $\varepsilon > 0$. Then there exists an $n \in \mathbb{N}$ such that

$$|y_n| \ge \|y\|_\infty - \varepsilon.$$

Next set $x = e_n \in \ell_1$. We then obtain $f_y(x) = y_n$. Since $\|x\| = 1$, it follows that

$$\|f_y\| \ge \|y\|_\infty - \varepsilon.$$

Recalling that ε is arbitrary, we conclude $\|f_y\| \ge \|y\|_\infty$.

$\underline{p > 1}$: Define $x = (x_n)_{n \in \mathbb{N}}$ 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 $x \in \ell_p$ with $\|x\|_p = \|y\|_q^{q/p}$. We now compute $f_y(x)$ by

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

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

Summarizing, we finished our proof that $y \mapsto f_y$ is isometric.

Next, we will show surjectivity of $y \mapsto f_y$. For this, let $f \in \ell_p^*$ and set

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

To prove $y = (y_n)_{n\in\mathbb{N}} \in \ell_q$, we again treat the two cases $p = 1$, $p > 1$ separately.

$\underline{p = 1}$: For all $n \in \mathbb{N}$,

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

This implies $\|y\|_\infty \leq \|f\|$ for any $y \in \ell_\infty$.

$\underline{p > 1}$: For all $m \in \mathbb{N}$, we have

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

We have further

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

and thus

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

yielding

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

Letting $m \to \infty$, we obtain $\|y\|_q \leq \|f\|$ for any $y \in \ell_q$.

To finish the surjectivity part, we are finally required to show that $f = f_y$. For this, let $x = \sum_{n=1}^{m} x_n e_n$. Then, by definition of f_y,

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

Thus f coincides with f_n on the dense linear subspace $\operatorname{span}(e_n)_{n\in\mathbb{N}}$. Hence, by Lemma 3.6, noting that $\mathbb{K}$ is a Banach space, $f = f_y$. The proof is finished. ■

Studying operators between normed spaces, one might wonder whether there is a natural transition to their respective duals. Indeed, one can define such a "dual operator", which will be our task for the last part of the chapter.

Lemma 3.13 *Let E and F be normed spaces, and let $T \in L(E,F)$. Then the operator $T^*\colon F^* \to E^*$ defined by*

$$T^*\varphi = \varphi \circ T \quad \textit{for all } \varphi \in F^*,$$

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

Proof. The operator T^* is obviously linear. Moreover, for any $x \in E$, we have

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

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

The operator T^*, which we already studied, has to still be named.

Definition 3.14 Let E and F be normed spaces, and let $T \in L(E,F)$. Then $T^*\colon F^* \to E^*$, $T^*\varphi = \varphi \circ T$, is called the *dual operator* of T.

4 The Hahn-Banach Theorem

4.1 Statement, Proof, and some first Corollaries

Our first goal is to state and prove the Hahn-Banach Theorem (Theorem 4.6), which is a statement about extensions of elements from a dual space of a subspace to the dual space of the large space itself. Interestingly, this can be done so that the operator norm remains the same.

However, as typically in life, there is no free lunch, and we have to work before to prepare the proof of this result. This will be accomplished in stages. More precisely, each preparatory result which bring us one step closer to the general result.

The first result requires the definition of an algebraic dual space and sublinear functionals, which is a weaker version of dual space and functionals.

Definition 4.1

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

(2) Let E be an $\mathbb{R}$-vector space. Then $\varrho\colon E \to \mathbb{R}$ is a *sublinear functional* on E, if, for all $x, y \in E$ and $\lambda > 0$,

$$\varrho(x+y) \leq \varrho(x) + \varrho(y) \quad \text{and} \quad \varrho(\lambda x) = \lambda \varrho(x).$$

We now show that in the setting of algebraic dual spaces and sublinear functionals, an extension from a subspace $F \subset E$ to a subspace $L = F + \mathbb{R}x_0$ (not yet to the large space E) is possible. Notice that we also for now just consider $\mathbb{R}$-vector spaces.

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

$$f(x) \leq \varrho(x) \quad \textit{for all } x \in F.$$

Then there exists some $\ell \in L'$ such that $\ell|_F = f$ and $\ell(x) \leq \varrho(x)$ for all $x \in L$.

Proof. Since x_0 is not an element of F, each element $y \in L$ has a unique representation $y = x + \lambda x_0$, for some $x \in F$ and $\lambda \in \mathbb{R}$. Aiming to define $\ell \in L'$, we first observe that

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

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

$$f(x) + \lambda\gamma \le \varrho(x + \lambda x_0) \quad \text{for all } x \in F,\ \lambda \in \mathbb{R}. \tag{4.1}$$

For this, for $x, y \in F$, we first consider

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

This implies

$$f(y) - \varrho(y - x_0) \le \varrho(x + x_0) - f(x) \quad \text{for all } x, y \in F. \tag{4.2}$$

Defining

$$A := \{f(x) - \varrho(x - x_0) : x \in F\} \quad \text{and} \quad B := \{\varrho(x + x_0) - f(x) : x \in F\},$$

by (4.2), we obtain

$$\sup A \le \inf B.$$

Now we choose some $\gamma \in [\sup A, \inf B]$. By this choice, we can conclude that

$$f(x) - \gamma \le \varrho(x - x_0) \quad \text{and} \quad f(x) + \gamma \le \varrho(x + x_0) \quad \text{for all } x \in F.$$

Letting $\lambda > 0$ be arbitrarily chosen, we then obtain

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

and

$$f(x) + \lambda\gamma = \lambda\left(f(\lambda^{-1}x) + \gamma\right) \le \lambda\varrho(\lambda^{-1}x + x_0) = \varrho(x + \lambda x_0).$$

This proves (4.1). ∎

The next step deals with the extension to the whole space E, but again still in the setting of algebraic dual spaces and sublinear functionals. Its proof relies on Zorn's Lemma.

Lemma 4.3 *Let E be an $\mathbb{R}$-vector space, let $F \subseteq E$ be a linear subspace, and let ϱ be a sublinear functional on E. Further let $f \in F'$ with*

$$f(x) \le \varrho(x) \quad \textit{for all } x \in F.$$

Then there exists some $\ell \in E'$ with $\ell|_F = f$ and $\ell(x) \leq \varrho(x)$ for all $x \in E$.

Proof. First, set

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

Our goal is now to prove the existence of a pair of the form $(E,\ell) \in \mathcal{L}$, which then completes the proof.

To prepare the application of Zorn's Lemma, we first have to show that $\mathcal{L}$ is not empty. But this can be easily settled, since $(F,f) \in \mathcal{L}$. Next, we define a partial ordering on $\mathcal{L}$ by

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

and state the following

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

Proof of Claim. We first set

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

Observe that $\widetilde{L}$ is a linear subspace. To show this, we can argue as follows: For any $x,y \in \widetilde{L}$, there exist L_x, L_y with $x \in L_x$, $y \in L_y$ and $(L_x,\ell_x),(L_y,\ell_y) \in \mathscr{K}$ for some ℓ_x, ℓ_y, since $\mathscr{K}$ is linearly ordered. Since $\mathscr{K}$ is a chain, without loss of generality we can assume that $L_x \subset L_y$. But this already implies that $x+y \in L_y$, hence $x+y \in \widetilde{L}$. It remains to show that $\lambda x \in L$ for any $\lambda \in \mathbb{R}$, but this is obvious.

We next define $\tilde{\ell}\colon \widetilde{L} \to \mathbb{R}$ by

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

Once we have shown that $\tilde{\ell}$ is well-defined and an element of $\widetilde{L}'$, the claim is proven.

We start by showing that $\tilde{\ell}$ is well-defined. For this, let $x \in \widetilde{L}$. By the definition of $\widetilde{L}$, there exists some $(L_1,\ell_1) \in \mathscr{K}$ such that $x \in L_1$ and $\ell_1 \in L_1'$. Let now $(L_2,l_2) \in \mathscr{K}$ be another pair with $x \in L_2$. Since $\mathscr{K}$ is a chain, one of these two pairs has to be larger with respect to the chosen ordering on $\mathcal{L}$. Without loss of generality, assume that $(L_1,\ell_1) \leq (L_2,\ell_2)$. Noting that $x \in L_1 \subset L_2$, we can conclude that $\ell_2(x) = \ell_1(x)$.

Finally, we aim to prove that $\tilde{\ell} \in \widetilde{L}'$. Let now $x_i \in L_i$, $\alpha_i \in \mathbb{R}$, $(L_i,\ell_i) \in \mathscr{K}$, for $i = 1,2$. Again using the fact that $\mathscr{K}$ is a chain, without loss of generality, we assume that $(L_1,\ell_1) \leq (L_2,\ell_2)$. Then $x_1 \in L_2$, which implies $\alpha_1 x_1 + \alpha_2 x_2 \in L_2$. Hence, we obtain

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

This finishes the proof of the claim. $\diamond$

By Zorn's Lemma, there exists a maximal element $(L, \ell) \in \mathcal{L}$ of $\mathcal{L}$. It remains to show that $L = E$. Towards a contradiction, assume that there exists some $x_0 \in E \backslash L$. By Lemma 4.2, there exists $g \in (L + \mathbb{R}x_0)'$ with $g|_L = \ell$ and $g(x) \leq \varrho(x)$ for all $x \in (L + \mathbb{R}x_0)$. But this implies $(L, \ell) < (L + \mathbb{R}x_0, g)$, which is a contraction to the fact that (L, ℓ) is a maximal element. Thus $L = E$, which proves the lemma. ∎

We next extend the result to an arbitrary vector space, i.e., from $\mathbb{R}$ to $\mathbb{K}$. For this, we also need to adapt the notion of a sublinear functional, leading to the definition of a seminorm.

Definition 4.4 Let E be an vector space over $\mathbb{K}$. Let $\varrho\colon E \to \mathbb{R}$ be a *seminorm* on E, if, for all $x, y \in E$ and $\lambda \in \mathbb{K}$,

$$\varrho(x+y) \leq \varrho(x) + \varrho(y) \quad \text{and} \quad \varrho(\lambda x) = |\lambda| \varrho(x).$$

As already announced, the next result will be an extension to arbitrary vector spaces.

Theorem 4.5 *Let E be a vector space over $\mathbb{K}$, let F be a linear subspace, and let $\varrho\colon E \to \mathbb{R}$ be a seminorm on E. Further, let $f \in F'$ be such that*

$$|f(x)| \leq \varrho(x) \quad \textit{for all } x \in F.$$

Then there exists some $\ell \in E'$ with $\ell|_F = f$ and $|\ell(x)| \leq \varrho(x)$ for all $x \in E$.

Proof. First, again consider $\mathbb{K} = \mathbb{R}$. In this case,

$$f(x) \leq \varrho(x) \quad \text{for all } x \in F \quad \text{and} \quad \varrho(\alpha x) = \alpha \varrho(x) \quad \text{for all } x \in E, \alpha \geq 0.$$

By Lemma 4.3, there exists some $\ell \in E'$ with

$$\ell|_F = f \text{ and } \ell(x) \leq \varrho(x) \quad \text{for all } x \in E.$$

Since also

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

we can conclude that $|\ell(x)| \leq \varrho(x)$.

The case $\mathbb{K} = \mathbb{C}$ will be discussed in the exercises. ∎

Finally, it is now time for the Hahn-Banach Theorem, for which we are now properly prepared.

Theorem 4.6 (Hahn-Banach Theorem) *Let E be a normed space, and let F be a linear* *subspace of E. Then, for each $f \in F^*$, there exists some $\ell \in E^*$ with*

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

Proof. Let ϱ be defined by

$$\varrho(x) := \|f\|\|x\|.$$

Then ϱ is a seminorm, since the properties are inherited from the norm properties of $\|\cdot\|$. By definition of ϱ, we have

$$|f(x)| \leq \varrho(x) \quad \text{for all } x \in F.$$

Theorem 4.5 then shows the existence of some $\ell \in E'$ with

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

This proves, in particular, that $\ell \in E^*$ and $\|\ell\| \leq \|f\|$.

It remains to prove that $\|\ell\| \geq \|f\|$. But this follows immediately from $\ell|_F = f$. Hence $\|\ell\| = \|f\|$, and the theorem is proved. ∎

We will now draw two direct corollaries from the Hahn-Banach Theorem. The first one shows, in particular, that, for each $x \in E$, $x \neq 0$, there exists some $f \in E^*$ with $\|f\| = 1$ and $f(x) = \|x\|$.

Corollary 4.7 *Let E be a normed space. Then, for each $x \in E$,*

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

Moreover, this supremum is attained.

Proof. Let $x \in E$, and set

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

On the one-dimensional subspace $F := \operatorname{span}\{x\}$, we define a functional f by

$$f(\lambda x) := \lambda\|x\| \quad \text{for all } \lambda \in \mathbb{K}.$$

It follows immediately that $f \in F^*$ with $\|f\| = 1$. Hence, by the Hahn-Banach Theorem, there exists some $\ell \in E^*$ with $\ell|_F = f$ and $\|\ell\| = 1$. In particular, this implies

$$|\ell(x)| = |f(x)| = \|x\|.$$

This proves $\|x\| = S$ as well as that the supremum is attained. ∎

The second corollary shows that a normed functional can be chosen in an even more sophisticated way, separating a subspace from a single point.

Corollary 4.8 *Let E be a normed space, let F be a linear subspace of E, and let $x \in E$ be such that*

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

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

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

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

Proof. We will now explicitly construct this functional. For this, let $G := F + \mathbb{K}x$, and let $g \colon G \to \mathbb{K}$ be defined by

$$g(y + \lambda x) = \lambda\delta \quad \text{for all } y \in F, \lambda \in \mathbb{K}.$$

Since $x \notin F$ implies $G = F \oplus \mathbb{K}x$, the functional g is well defined. Moreover, g is linear, $g|_F = 0$, and $g(x) = \delta$.

We now claim $\|g\| = 1$. Since, for every $y + \lambda x \in G$,

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

we conclude that $\|g\| \leq 1$. To prove the reverse inequality, observe that, for every $\varepsilon > 0$, there exists some $z_\varepsilon \in F$ with $\delta \leq \|z_\varepsilon + x\| \leq \delta + \varepsilon$. This implies

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

hence

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

This proves $\|g\| = 1$.

By Theorem 4.6, there exists a unique extension of g to E^*, finishing the first part of the corollary.

For the 'in particular part', choose $F = \{0\}$. Then the claim follows immediately. ∎

4.2 Applications to Dual Spaces

This section will present two further applications of the Hahn-Banach Theorem. First, we will discuss two cases in which we can explicitly compute the dual space (up to an isometric isomorphism), see Theorem 4.12. And, second, we will show that also the converse of Theorem 3.5 is true.

For the first result, we require two novel concepts for subspaces which we will next introduce.

Definition 4.9 Let E be a normed space, let $M \subset E$ be an arbitrary subset of E, and let $L \subset E^*$. Then the *annihilator of M in E^** is defined by

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

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

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

Annihilators were just defined as certain subsets of a normed space or its dual. But in fact, they are even closed subspaces, hence have more algebraic structure.

Remark 4.10 The annihilators are closed linear subspaces of E^* and E, respectively. This follows from the continuity of $\ell \mapsto \ell(x)$, $x \mapsto \ell(x)$. $\Diamond$

Computing the annihilator of an annihilator, we even derive a subspace of a particularly nice structure, as shown in the following result.

Lemma 4.11 *Let E be a normed space, and let $\varnothing \neq M \subseteq E$. Then $(M^\perp)_\perp$ is the* closed linear hull *of M, i.e., the smallest closed linear subspace of E which contains M.*

Proof. We first show that M is indeed contained in $(M^\perp)_\perp$. For this, let $x \in M$. This implies $\ell(x) = 0$ for all $\ell \in M^\perp$, hence $x \in (M^\perp)_\perp$. Thus, we already proved $M \subseteq (M^\perp)_\perp$.

Now let F be the closed linear hull of M. By Remark 4.10, we have $F \subseteq (M^\perp)_\perp$. Now towards a contradiction, assume that there exists $x \in (M^\perp)_\perp \setminus F$. By Corollary 4.8, there exists some $\ell \in (M^\perp)_\perp^*$ with $\ell|_F = 0$ and $\ell(x) \neq 0$. Theorem 4.6 then implies the existence of some $f \in E^*$ with $f|_{(M^\perp)_\perp} = \ell$. Since $f|_F = \ell|_F = 0$ and $M \subseteq F$, we can conclude that $f \in M^\perp$. But $f(x) \neq 0$, which is a contradiction to $x \in (M^\perp)_\perp \setminus F$. ■

The next result shows that the dual space of a subspace $F \subseteq E$ can be determined to be the quotient space $E^*/F^\perp$, and the dual space of a quotient space $(E/F)^*$ can be identified with the annihilator $F^\perp$ of F. This is actually a quite deep result, since it gives us precise information about the previously seemingly quite inaccessible dual space for large classes of spaces.

Theorem 4.12 *Let E be a normed space over $\mathbb{K}$, and let $F \subset E$ be a linear subspace.*

(i) The linear operator

$$\Phi\colon E^*/F^\perp \to F^*, \quad \Phi(f + F^\perp) = f|_F, \text{ where } f \in E^*,$$

is an isometric isomorphism.

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

$$\Phi\colon (E/F)^* \to F^\perp, \quad (\Phi f)(x) = f(x+F), \text{ where } x \in E,\ f \in (E/F)^*,$$

is an isometric isomorphism.

Proof. (i). We need to check all properties necessary for Φ to form an isometric isomorphism. For this, first consider the map $T\colon E^* \to F^*$ defined by $Tf := f|_F$ for $f \in E^*$. This map satisfies $\ker T = F^\perp$, implying that Φ is well-defined, linear, and injective.

Further, by Theorem 4.6, for each $\ell \in F^*$, there exists some $f \in E^*$ with $f|_F = \ell$. This shows that Φ is surjective.

Finally, to show isometry, 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

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

To prove the reverse inequality notice that, for all $g \in F^\perp$,

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

This implies $\|\Phi(f + F^\perp)\| \le \|f + F^\perp\|$, thereby finishing the proof of part (i).

(ii). We first have to show that indeed Φ maps $(E/F)^*$ to $F^\perp$. For this, first we observe that $\Phi f\colon E \to \mathbb{K}$, $x \mapsto f(x+F)$, is linear. Since

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

it follows that $\Phi f \in E^*$. Now let $x \in F$. Then $(\Phi f)(x) = f(F) = 0$, and hence $\Phi f \in F^\perp$. Therefore, summarizing, indeed we have $\Phi\colon (E/F)^* \to F^\perp$.

We now have to check the properties of a map being an isometric isomorphism. Obviously, 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 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 $f \in (E/F)^*$. Noting that $\Phi f = g$ implies surjectivity.

It remains to show that Φ is also isometric. For this, notice that $|(\Phi f)(x)| \le \|f\| \cdot \|x\|$ for all $x \in E$, implies $\|\Phi f\| \le \|f\|$ for all $f \in (E/F)^*$. To prove the reverse inequality notice that, for each $\varepsilon > 0$, there exists some $x \in E$ 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 some $y \in F$ with $\|x + y\| \leq 1 + \varepsilon$. This implies $\left\| \frac{x+y}{1+\varepsilon} \right\| \leq 1$, and hence

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

Thus $\|\Phi f\| \geq \frac{\|f\| - \varepsilon}{1+\varepsilon}$, which yields $\|\Phi f\| \geq \|f\|$. This proves that Φ is isometric, and also part (ii) is proved. ∎

As announced, the second application of the Hahn-Banach Theorem, or, more precisely, Corollary 4.7, shows the converse direction of Theorem 3.5.

Theorem 4.13 *Let E and F be normed spaces with $E \neq \{0\}$. If $L(E, F)$ is complete, then so is F.*

Proof. First, choose $x_0 \in E$ with $\|x_0\| = 1$. Then, by Corollary 4.7, 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

$$T_n x := f(x) y_n, \quad x \in E.$$

Since

$$\|T_n x\| \leq \|f\| \|x\| \|y_n\| = \|y_n\| \|x\| \quad \text{for all } x \in E,$$

it follows that $T_n \in L(E, F)$. Moreover,

$$\|T_n x - T_m x\| = |f(x)| \|y_n - y_m\| \leq \|y_n - y_m\| \|x\|$$

implies that

$$\|T_n - T_m\| \leq \|y_n - y_m\| \quad \text{for all } m, n \in \mathbb{Z}.$$

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. ∎

4.3 Reflexive Spaces

Starting from a normed space, we introduced its dual space. It seems natural to ask what happens if this procedure is iterated, i.e., which structure the dual space of the dual space – later called bi-dual (space) – exhibits. Those normed spaces for which their bi-dual has a special structural relation with the original space are called reflexive, and will be the focus of this section.

We begin by analyzing a certain map from a normed space to its bi-dual.

Lemma 4.14 *Let E be a normed space over $\mathbb{K}$. For each $x \in E$, let $\hat{x}\colon E^* \to \mathbb{K}$ be defined by $\hat{x}(\ell) = \ell(x)$, $\ell \in E^*$. Then the map*

$$\Lambda_E\colon E \to (E^*)^*, \quad \Lambda_E x = \hat{x},$$

is an isometric linear operator.

Proof. We first notice that Λ_E is linear. Moreover, for all $x \in E$, 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 $x \in E$. ∎

Let's now fix some notation first...

Definition 4.15 Let E be a normed space over $\mathbb{K}$. The map Λ_E defined in Lemma 4.14 is called *canonical map* or *canonical embedding* of E in $E^{**} := (E^*)^*$. E is *reflexive*, if Λ_E is surjective. Furthermore, E^{**} is the *bi-dual* of E, and $\overline{\Lambda_E(E)}$ is the *completion* of E.

A necessary condition for a space to be reflexive is completeness as the next remark states.

Remark 4.16 Only Banach spaces can be reflexive, since the bi-dual is complete (Theorem 3.5), and hence by the definition of reflexivity so is the original space.

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. Moreover, notice that due to $\dim E^{**} = \dim E$ finite-dimensional spaces are always reflexive. ♢

Introducing a new property, the first question has to be whether a subspace inherits this property. This question will be answered by the following result, which also provides an equivalent condition for reflexivity (in terms of the dual space).

Theorem 4.17 *Let E be a normed space over $\mathbb{K}$.*

(i) If E is reflexive and $F \subset E$ a closed 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

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

it follows that $\psi \in E^{**}$. E being reflexive then implies that there exists some $y \in E$ – which will turn out to be our seeked point in F – 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 each $f \in F^*$, there exists some $\ell \in E^*$ with $\ell|_F = f$. Thus

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

which shows that F is reflexive.

(ii). Let now E be reflexive. We have to prove 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 $f(x) := u(\hat{x})$, $x \in E$. Obviously, $f \in E^*$. Next, let $\varphi \in E^{**}$, and let $x \in E$ be chosen such that $\hat{x} = \varphi$. This yields

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

hence E^* is reflexive.

For the converse direction, let E^* be reflexive. The proven direction shows that then also 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). ∎

We already had several results showing a relation between properties of the full space E and those of a closed subspace $F \subset E$ and its associated quotient space E/F. The next result studies the property of being reflexive.

Theorem 4.18 *Let E be a Banach space and $F \subset E$ a closed linear subspace. Then the following conditions are equivalent.*

(i) E is reflexive.

(ii) F and E/F are reflexive.

Proof. (i)⇒(ii). By (i) and Theorem 4.17(i), also F is reflexive. Then Theorem 4.17(ii) implies that E^* is reflexive, hence $F^\perp$ is reflexive. By Theorem 4.12(ii), $F^\perp$ is isometrically isomorphic to $(E/F)^*$ which is therefore also reflexive. Finally, applying Theorem 4.17(ii), we can conclude that E/F is reflexive.

(ii)⇒(i). For this, let $\varphi \in E^{**}$. We now aim to construct some $z \in E$ with $\hat{z} = \varphi$. It will turn out that $z = x + y$, where x arises from reflexivity of E/F and y from reflexivity of F.

We start by recalling the isometric isomorphism

$$\Phi\colon (E/F)^* \to F^\perp \subset E^*, \quad (\Phi u)(x) = u(x+F), \text{ where } u \in (E/F)^*, x \in E,$$

from Theorem 4.12. Then we can define $\psi \in (E/F)^{**}$ by

$$\psi(u) := \varphi(\Phi u), \quad u \in (E/F)^*.$$

Since E/F is reflexive, there exists some $x \in E$ with $\widehat{x+F} = \psi$, i.e.,

$$\varphi(\Phi u) = \psi(u) = (\widehat{x+F})(u) = u(x+F) = (\Phi u)(x) = \hat{x}(\Phi u), \quad u \in (E/F)^*.$$

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

To utilize the reflexivity of F, we next define a suitable $\varrho \in F^{**}$. For each $f \in F^*$, choose some $g \in E^*$ with $g|_F = f$ and $\|g\| = \|f\|$. Then define

$$\varrho(f) := (\varphi - \hat{x})(g).$$

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,

$$|\varrho(f)| \leq \|\varphi - \hat{x}\| \cdot \|g\| = \|\varphi - \hat{x}\| \cdot \|f\|.$$

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

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

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. ∎

5 Fundamental Theorems of Functional Analysis

We are now ready to prove several of the most fundamental theorems in functional analysis.

5.1 Open Mapping and Closed Graph Theorem

Recall that an operator is continuous, if pre-images of open sets of open. Certainly, also the property that open sets are mapped to open sets is important. Hence we will later coin such a map an *open* map. The first main theorem we are heading towards – the so-called open mapping theorem (Theorem 5.4) – provides a surprisingly weak sufficient condition for an operator to be open.
For its proof, we require some prerequisites, which are the following two lemmata. Both show that under certain conditions on the spaces and the linear operator specific sub-properties of openness are satisfied.

Lemma 5.1 *Let E be a normed space, let F be a Banach space, and let $T \in L(E, F)$ be surjective. Then*

$$K_1(0_F) \subset \overline{T(K_r(0_E))} \qquad \text{for some } r > 0.$$

Proof. Since T is surjective, we have

$$F = \bigcup_{n=1}^{\infty} T(K_n(0_E)).$$

According to Corollary 1.6, which was a corollary from the Theorem of Baire, the fact that F is complete implies the existence of some $m \in \mathbb{N}$ with

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

Therefore, there exist $y_0 \in F$ and $s > 0$ with

$$K_s(y_0) \subset \overline{T(K_m(0_E))} =: A. \tag{5.1}$$

Now let $y \in K_s(0_F)$. Then $\|(y+y_0)-y_0\| \le s$ and $\|(y_0-y)-y_0\| \le s$, and hence

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

Since $z_1, z_2, z \in A$ implies $-z, \frac{1}{2}(z_1+z_2) \in A$, we can conclude that

$$y = \tfrac{1}{2}((y+y_0)-(y_0-y)) \in A.$$

Thus, (5.1) leads to

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

Hence, we conclude

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

Choosing $r := \frac{m}{s}$ proves the claim. ■

It is interesting to notice that the previous lemma required F to be a Banach space, whereas the following lemma requires E to be Banach. In the end, for the open mapping theorem, we need both conditions – and the lemmata show nicely, why this is the case.

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

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

Then we have

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

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

$$\left\| y - \sum_{k=1}^{n} \frac{y_k}{2^{k-1}} \right\| \le \frac{1}{2^n} \quad \text{for all } n \in \mathbb{N}.$$

By hypothesis, there exists some $y_1 \in T(K_r(0_E))$ with $\|y-y_1\| \le \frac{1}{2}$, which gives the induction start. Now assume that $y_1, \ldots, y_n$ are already constructed appropriately, i.e., that

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

Since this vector is in the closure of $T(K_r(0_E))$, there exists some $y_{n+1} \in T(K_r(0_E))$ with

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

This shows that such a sequence $(y_n)_{n\in\mathbb{N}} \subset T(K_r(0_E))$ exists.

Next, 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. Again using the completeness of E, we can define

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

This vector satisfies

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

Thus $x \in K_{2r}(0_E)$. Since T is continuous, we can conclude for the image Tx that

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

and therefore $y \in T(K_{2r}(0_E))$. This shows $K_1(0_F) \subset T(K_{2r}(0_E))$.

It remains to prove 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 some $x \in E$ with $Tx = \frac{y}{\|y\|}$, which in turn implies $T(\|y\|x) = y$, thereby finishing the proof. ∎

Now it is time to formally state the notion of openness of an operator.

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

Finally, we a ready for the open mapping theorem. Provided that both spaces E and F are Banach spaces – recall the comment between the lemmata – a linear, bounded operator $T : E \to F$ is open, if it has the additional property of being surjective. After our hard work, its proof is then quite easy and short.

Theorem 5.4 (Open Mapping Theorem) *Let E and F be Banach spaces, and let $T \in L(E, F)$ be surjective. Then T is open.*

Proof. Let $U \subset E$ be open and $x \in U$. The openness implies that there exists some $t > 0$ with $K_t(x) \subset U$. Then

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

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

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

Application of Lemma 5.2 yields

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

Thus

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

and hence

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

The theorem is proved. ∎

Let us next take a look at some corollaries to get a feeling how the open mapping theorem can be applied.

Corollary 5.5

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

(ii) If $\|\cdot\|_1$ and $\|\cdot\|_2$ are Banach space norms on E and if $\|x\|_1 \leq c\|x\|_2$ 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, $\left(T^{-1}\right)^{-1}(U)$ is open for each open $U \subset E$. But this is already the continuity of T^{-1}.

(ii). We consider $T\colon (E, \|\cdot\|_1) \to (E, \|\cdot\|_2)$, $Tx = x$. Application of part (i) shows the claim. ∎

Finally, we can ask whether the completeness condition is really necessary. The next example shows that it indeed is.

Example 5.6 In general, the hypothesis of completeness cannot be removed. For this, consider $E = C[0, 1]$ with $\|\cdot\|_\infty$-norm, and $F = \{g \in C^1[0, 1] : g(0) = 0\}$ with $\|\cdot\|_\infty$-norm. Then define $T : E \to F$ by

$$(Tf)(t) := \int_0^t f(s)\mathrm{d}s.$$

Let us quickly check the necessary properties: T is obviously linear. T is bounded, since $\|Tf\|_\infty \leq \|f\|_\infty$, and it is injective, since $Tf = 0$ implies $f = 0$. The operator T is also surjective, since for $g \in F$ we have $T(g') = g$.

But T^{-1} is not continuous, i.e., Corollary 5.5 does not hold. For this, choose $f_n(t) = t^n$. Then $(Tf_n)(t) = \frac{1}{n+1}t^{n+1}$, hence

$$\|Tf_n\|_\infty = \tfrac{1}{n+1}, \quad \text{but } \|T^{-1}(Tf_n)\|_\infty = \|f_n\|_\infty = 1.$$

Thus, we can conclude that in case of E being a Banach space and F being only a normed space, bijectivity of $T \in L(E,F)$ does *not* imply that T^{-1} is continuous. $\diamondsuit$

The closed graph theorem, our next fundamental result, aims to derive sufficient properties for a linear operator $T\colon E \to F$ to be bounded, hence for $T \in L(E,F)$. Again, as in the open mapping theorem, the space E as well as F are both assumed to be Banach spaces. Intriguingly, the sufficient property is for T to be closed (recall Definition 3.9).

As we will see in the proof, it can be regarded as a corollary from the open mapping theorem, since the proof relies on Corollary 5.5.

Theorem 5.7 (Closed Graph Theorem) *Let E and F be Banach spaces, and let $T\colon E \to F$ be a closed linear operator. Then T is bounded.*

Proof. Since E and F are Banach spaces, also $E \times F$ is a Banach space. The operator T is assumed to be closed, hence the graph G_T is closed. This implies that G_T is a Banach space.

Now define $S\colon G_T \to E$ by $S(x, Tx) = x$. This operator S is linear, bijective and continuous, since

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

By Corollary 5.5, we can conclude that also $S^{-1}\colon E \to G_T$ is continuous. Thus

$$\|Tx\| \leq \|(x,Tx)\| = \|S^{-1}(x)\| \leq \|S^{-1}\|\,\|x\| \quad \text{for all } x \in E.$$

The theorem is proved. ■

5.2 Uniform Boundedness Principle and the Banach-Steinhaus Theorem

The next fundamental result will be the uniform boundedness principle. It does what it says: It shows under certain conditions that pointwise boundedness implies uniform boundedness for a family of linear, bounded operators $T : E \to F$. The sufficient condition for this to work is for E to be a Banach space. We will see how this condition will be explored in the proof.

Theorem 5.8 (Uniform Boundedness Principle) *Let E be a Banach space, let F be a normed space, and let $\mathcal{T} \subset L(E, F)$. If $\mathcal{T}$ is 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 $\|T\| < M$ for all $T \in \mathcal{T}$.*

Proof. We start by defining the following sets:

$$E_n := \{x \in E : \|Tx\| \leq n \quad \forall T \in \mathcal{T}\}, \quad n \in \mathbb{N}.$$

Since $\mathcal{T}$ is pointwise bounded, it follows that

$$E = \bigcup_{n=1}^{\infty} E_n.$$

Let us now show that each E_n is closed. For this, fix some n and let $x = \lim\limits_{j\to\infty} x_j$ with $x_j \in E_n$. 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.

We now employ Corollary 1.6 – the corollary from Baire's Theorem – to show that there exists some $n_0 \in \mathbb{N}$ with

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

This implies $K_r(x) \subset E_{n_0}$ for some $x \in E_{n_0}$, $r > 0$. Next let $y \in E$ with $\|y\| \leq r$. Then, we have

$$y + x \in K_r(x) = x + K_r(0) \subset E_{n_0},$$

which in turn yields

$$\|Ty\| = \|T(y+x) - Tx\| \leq \|T(y+x)\| + \|Tx\| \leq 2n_0 \quad \text{for all } T \in \mathcal{T}. \tag{5.2}$$

Finally, let $y \in E$, $y \neq 0$ be arbitrary. Then we obtain

$$\frac{r}{\|y\|}\|Ty\| = \left\| T\left(\frac{ry}{\|y\|}\right)\right\| \overset{(5.2)}{\leq} 2n_0.$$

This can be reformulated as $\|Ty\| \leq \frac{2n_0}{r}\|y\|$, which implies

$$\|T\| \leq \frac{2n_0}{r}.$$

The theorem is proved. ∎

The following corollary is a first example of the Uniform Boundedness Principle in action.

Corollary 5.9 *Let E be a Banach space, let F be a normed space, and let $T_n \in L(E, F)$. Further, assume that $(T_n x)_{n\in\mathbb{N}}$ is convergent in F for every $x \in E$, and define $T\colon E \to F$ by*

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

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

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

Proof. By definition, T obviously is linear. The convergence hypothesis implies that $(\|T_n x\|)_{n\in\mathbb{N}}$ is bounded. To show that also T is bounded, notice that, by Theorem 5.8, it follows that $\|T_n\| \leq M$ for all $n \in \mathbb{N}$. Hence, for all $x \in E$,

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

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 we can conclude that

$$\|Tx\| = \lim_{k\to\infty} \|T_{n_k} x\| \leq \|x\| \lim_{k\to\infty} \|T_{n_k}\|,$$

which implies

$$\|T\| \leq \lim_{k\to\infty} \|T_{n_k}\|.$$

But this yields $\|T\| \leq \liminf_{n\to\infty} \|T_n\|$, thereby finishing the proof. ∎

As in Example 5.6, we can ask whether the Banach property is really necessary in the Uniform Boundedness Principle. To answer this question, we now construct a counterexample.

Example 5.10 Consider the space (notice that this is not a Banach space)

$$E := \{x = (x_n)_{n\in\mathbb{N}} \in \ell_1 : x_n = 0 \text{ for almost all } n \in \mathbb{N}\},$$

$F := \mathbb{K}$, and the family of operators

$$\mathcal{T} := \{f_n : n \in \mathbb{N}\}, \quad \text{where } f_n(x) = n x_n \quad \text{for all } x \in E.$$

Since $x_n = 0$ from some $n \geq N$ on, we conclude that $\mathcal{T}$ is pointwise bounded. But we have

$$\|f_n\| = n, \quad n \in \mathbb{N}.$$

Thus, in general, Theorem 5.8 does not hold, if E is *not* a Banach space ◊

We are now heading towards our last fundamental result of this chapter, which is the Banach-Steinhaus Theorem. For this, we require a slight adaption of Lemma 3.6, which we state now first.

Lemma 5.11 *Let E be a normed space, and let F be a Banach space. Further, let E_0 be a dense linear subspace of E, and let $T_0 \in L(E_0, F)$. Then there exists a unique $T \in L(E, F)$ with*

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

The Banach-Steinhaus Theorem studies – roughly speaking – the relation between pointwise convergence of a sequence of operators and global properties of this sequence. Notice that in the two statements the role of the normed space and the Banach space are interchanged.

Theorem 5.12 (Banach-Steinhaus Theorem)

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

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

(ii) Let E be a normed space, and let F be 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

$$Tx = \lim_{n\to\infty} T_n x \quad \text{for all } x \in E.$$

Proof. (i). This is Corollary 5.9.

(ii). Let $y \in E_0$. Then, by condition (b), we can set

$$T_0 y := \lim_{n\to\infty} T_n y.$$

Obviously, T_0 is linear and we have

$$\|T_0 y\| = \lim_{n\to\infty} \|T_n y\| \leq \underbrace{\sup_{n\in\mathbb{N}} \|T_n\|}_{<\infty \text{ by (a)}} \|y\|.$$

This shows that $T_0 \in L(E_0, F)$.

By Lemma 5.11, there then exists some $T \in L(E, F)$ with $T|_{E_0} = T_0$. It remains to show that $T_n x \to Tx$ as $n \to \infty$. For this, let $x \in E$, let $\varepsilon > 0$, and let $y \in E_0$ and $N \in \mathbb{N}$ with

$$\|x - y\| \le \varepsilon \quad \text{and} \quad \|T_n y - T_0 y\| \le \varepsilon \quad \forall n \ge N. \quad \text{(by (b))}$$

Then, for all $n \ge N$,

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

This indeed implies $T_n x \to Tx$ as $n \to \infty$. The proof is finished. ∎

5.3 Taking a Functional Viewpoint

At the end of this chapter, we will prove two results which show that properties of a subset of a normed space can be investigated by applying functionals from the dual space. The great advantage is that the functionals bring the subsets to $\mathbb{K}$, which is certainly much easier to deal with. For the proofs, we require the Uniform Boundedness Principle and the Banach-Steinhaus Theorem.

Let's start with the first result, which deals with bounded subsets.

Theorem 5.13 *Let E be a normed space, and let $M \subset E$. Then the following conditions are equivalent.*

(i) M is bounded.

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

Proof. (i)⇒(ii). This follows from the fact that, for all $x \in M$,

$$\|x\| < c \quad \text{implies} \quad \|f(x)\| \le c\|f\|.$$

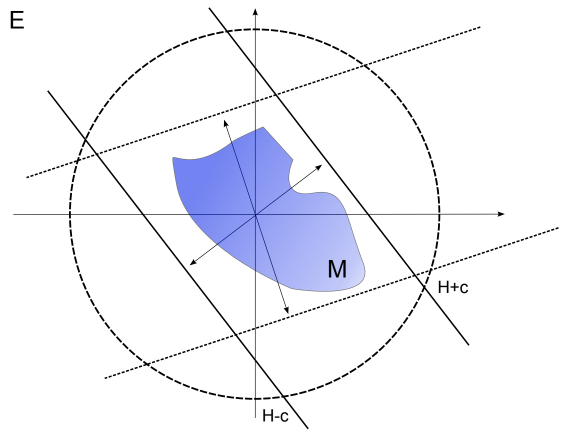
(ii)⇒(i). Consider the set

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

Since $\hat{M}(f) = f(M)$, by (ii), the family $\hat{M} \subset E^{**}$ is pointwise bounded. By Theorem 5.8, there then exists a uniform bound for $\hat{M}$. Since $\hat{}\, : E \to E^{**}$ is an isometry, we can conclude that M is bounded. ∎

There exists a nice geometric interpretation of Theorem 5.13, which we will now discuss.

Remark 5.14 Suppose that for every closed hyperplane H in E, i.e., the kernel of some $f \in E^*$, $f \neq 0$, there exists some c with M between $H + c$ and $H - c$. Then M is already contained in a ball. $\Diamond$



The next result takes a similar viewpoint as Theorem 5.13 for the study of linear, bounded operators.

Theorem 5.15 *Let E and F be normed spaces, and let $T\colon E \to F$ be linear. Then the following conditions are equivalent.*

(i) $T \in L(E, F)$.

(ii) For all $f \in F^*$*, we have* $f \circ T \in E^*$.

Proof. We immediately show the equivalence of (i) and (ii). For this, first note that boundedness of T is equivalent to boundedness of $T(K_1(0))$ in F. By Theorem 5.13, this can be equivalently reformulated as saying that $f(T(K_1(0)))$ is bounded for all $f \in F^*$. Finally, note that this is equivalent to $f \circ T \in E^*$ for all $f \in F^*$. $\blacksquare$

6 Weak Convergence and Weak Topology

6.1 Weak and Weak*-Convergence

At the end of the last chapter we already saw the benefit of analyzing properties by applying all functionals from the dual space. This viewpoint enables us to derive a weaker form of convergence, which is the so-called weak convergence. We will later see that extensions of this concept allow us to generalize well-known results from the finite-dimensional setting to the infinite-dimensional setting.

But let's start slowly with first stating the definition of weak convergence.

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

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

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

As always, we first check some basic properties of this new notion.

Remark 6.2 (a) The observation

$$|f(x_n) - f(x)| = |f(x_n - x)| \leq \|f\|\|x_n - x\| \to 0 \quad \text{as } n \to \infty,$$

shows that 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. Thus convergence implies weak convergence.

(b) The weak limit of a weakly 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 $x = y$. This can be shown as follows:

Since $(x_n)_{n\in\mathbb{N}}$ converges weakly to x and y, we obtain $f(x) = f(y)$. Thus $f(x-y) = 0$ for all $f \in E^*$, and, by Corollary 4.7, we have $x = y$.

(c) Notice that

$$f(x_n + \lambda y_n) = f(x_n) + \lambda f(y_n) \to f(x) + \lambda f(y) = f(x + \lambda y).$$

This shows that if $(x_n)_{n\in\mathbb{N}} \subset E$ and $(y_n)_{n\in\mathbb{N}} \subset E$ are weakly convergent sequences with $x_n \xrightarrow{w} x$ and $y_n \xrightarrow{w} y$, 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.

(d) The converse of (a) is in general not true. A counterexample can be given by choosing $E = C[0,1]$, endowed with the L_2-norm, and $f_n(t) = \sin(nt)$, $n \in \mathbb{N}$. Then

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

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

$\diamondsuit$

We are also interested in sequences in the dual space of a normed space. We could certainly just use the concept of weak convergence by applying elements from the bidual. Although $\hat{x}(f) = f(x)$, notice that the definition we choose below is slightly different, since we use only functionals of the type $\hat{x}$. This leads to the notion of weak*-convergence.

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

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

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

Some of the properties studied in Remark 6.2 are also true in this setting.

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

We are now interested to derive characterizing conditions for a sequence to be weak convergent or weak* convergent, which are hopefully easier to check. Indeed the next result shows that under certain circumstances it suffices to check those properties on a dense subspace. This certainly calls for the Banach-Steinhaus Theorem to be used in the proof.

Theorem 6.5 *(i) Let E be a normed space, let $(x_n)_{n\in\mathbb{N}} \subset E$, and let $x \in E$. Then the following conditions are equivalent.*

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

(b) We have

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

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

$$f(x_n) \to f(x) \quad \textit{for all } f \in D.$$

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

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

(b) *We have*

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

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

$$f_n(x) \to f(x) \quad \textit{for all } x \in D.$$

Proof. (i).(a)⇒(b). Since (x_n) converges weakly to x, it follows that

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

This implies (II). Moreover, $\{f(x_n) : n \in \mathbb{N}\} \subset \mathbb{K}$ is bounded for each $f \in E^*$. Hence, by Theorem 5.13, also $\{x_n : n \in \mathbb{N}\}$ is bounded, yielding (I).

(b)⇒(a). For this, consider the canonical embedding of E in E^{**}. Then, by (II),

$$\hat{x}_n(f) \to \hat{x}(f) \quad \text{for all } f \in D \quad \text{and} \quad \|\hat{x}_n\| = \|x_n\|.$$

By (I) and by the Banach-Steinhaus Theorem (Theorem 5.12) applied to $(\hat{x}_n)_{n\in\mathbb{N}}$ and $\hat{x} \in L(E^*, \mathbb{K})$,

$$\hat{x}_n(f) \to \hat{x}(f), \text{ as } n \to \infty \quad \text{for all } f \in E^*.$$

This implies

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

(ii). This is the Banach-Steinhaus Theorem (Theorem 5.12). ∎

We finish this part with an interesting statement, which shows that bounded sequences in a dual space do not necessarily contain a convergent subsequence, but at least a weak*-convergent subsequence. Notice that this result generalizes the result in the finite-dimensional setting that bounded sequences contain a convergent subsequence. This is a first instance, which shows us first, that direct generalizations from the finite-dimensional setting to the infinite-dimensional realm are often not possible, and second, that the just introduced concepts of weak and weak*-convergence provide a means to nevertheless enable such generalizations.

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 $\{x_1, x_2, x_3, \ldots\}$ be a dense countable subset of E, and let $(f_n)_{n\in\mathbb{N}} \subset E^*$ be a bounded sequence in E^*, i.e., there exists some $C > 0$ with $\|f_n\| \le C$ for all $n \in \mathbb{N}$. This implies that the sequence

$$(f_n(x_1))_{n\in\mathbb{N}} \subset \mathbb{K}$$

is bounded. Hence there exists a convergent subsequence

$$(f_{n,1}(x_1))_{n\in\mathbb{N}} \subset \mathbb{K}.$$

Now we apply this sequence in E^* to the second element of E, and observe that

$$(f_{n,1}(x_2))_{n\in\mathbb{N}} \subset \mathbb{K}$$

is bounded. Hence this sequence has again a convergent subsequence

$$(f_{n,2}(x_2))_{n\in\mathbb{N}} \subset \mathbb{K}.$$

We continue this way and select the diagonal sequence

$$g_n := f_{n,n}, \quad n \in \mathbb{N},$$

which is a subsequence of $(f_n)_{n\in\mathbb{N}}$. By construction, $(g_n(x_m))_{n\in\mathbb{N}}$ is convergent for any $m \in \mathbb{N}$. The Banach-Steinhaus Theorem (Theorem 5.12) now implies that $(g_n)_{n\in\mathbb{N}}$ is pointwise convergent to some $g \in E^*$. This finally shows $g_n \xrightarrow{w^*} g$, and the theorem is proved. ■

Separability was not required before, and it is natural to ask whether it is indeed necessary. The next example provides a counterexample, showing that in general the separability condition of E can not be omitted.

Example 6.7 Consider the normed $E = \ell_\infty$, which is not separable. For $x = (x_n)_{n\in\mathbb{N}} \in E$ define the maps

$$f_n(x) := x_n,\ n \in \mathbb{N}.$$

Then $\|f_n\| = 1$ for all $n \in \mathbb{N}$, which implies boundedness of the sequence $(f_n)_{n\in\mathbb{N}} \subset E^*$.

Let now $(f_{n_k})_{k\in\mathbb{N}}$ be any subsequence of $(f_n)_{n\in\mathbb{N}}$. Then, for $x = (x_n)_n \in E$ defined by

$$x_n = \begin{cases} 1, & n = n_k \text{ and } k \text{ is even,} \\ 0, & \text{otherwise,} \end{cases}$$

it follows 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 thus (f_n) does not possess a weak*-convergent subsequence.

This shows that, in general, Theorem 6.6 does not hold without the condition of separability. ◊

6.2 Weak and Weak* Topology

To derive results such as in Section 6.3 – again generalizing well-known results from the finite-dimensional setting –, we now need to delve a bit deeper into weak and weak*-convergence. For this, we need to first consider so-called topological spaces, which are basically sets equipped with even less structure than metric spaces.

If given a set equipped with a metric, the most basic objects one can imagine are open subsets. But we can also – more generally – just call some subsets 'open'. Certainly, this selected set of open sets needs to satisfy certain properties such as that arbitrary unions of open sets shall again be open. If those properties, which we will now specify in detail, are fulfilled, this set of open sets is called 'topology'.

Although a set with a topology is in a sense minimally structured, we still can define concepts such as continuity, convergence, etc. This will lead us later back to weak and weak*-convergence, which will then be based on certain topologies. But having the topology at hand, we can also define concepts such as compactness adapted to it. And it is this key idea, which will allow us to derive results such as the Theorem of Alaoglu (Theorem 6.31).

After this long introduction, let's now start with the definition of a topology and a topological space. We will also define a basis for a topology, which in a sense generates it.

Definition 6.8 (1) A *topology* $\mathcal{T}$ on a set $X \neq \varnothing$ is a family of subsets of X with the following properties:

(T_1) $\varnothing, X \in \mathcal{T}$.

(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 $\bigcap_{i=1}^{r} S_i \in \mathcal{T}$.

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*.

(2) 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 $\mathcal{B}$.

A topological space can be interpreted as being a set, where each point is equipped with a certain neighborhood structure. Formally, the (open) neighborhoods of a point are those elements of the topology that contain that point, i.e., the open sets localized at that point. Similarly, a neighborhood basis is a localized version of the concept of a basis.

Definition 6.9 Let $x \in X$ be a point in a topological space $(X, \mathcal{T})$.

(1) An *(open) neighborhood* of x is an element $U \in \mathcal{T}$ containing x, i.e., satisfying $x \in U$. The collection of all (open) neighborhoods of x is denoted by $\mathcal{U}_x$.

(2) A collection $\mathcal{B}_x \subset \mathcal{U}_x$ is called a *neighborhood basis* for x, if for each $U \in \mathcal{U}_x$ there

exists $V \in \mathcal{B}_x$ with $V \subset U$.

A characteristic property of a topology is the degree of separation of points it provides. A topological space, where every two distinct points possess disjoint open neighborhoods, is called a Hausdorff (topological) space.

Definition 6.10 A topological space $(X, \mathcal{T})$ is called a *Hausdorff (topological) space* if for each $x, y \in X$, $x \neq y$, there exist $U, V \in \mathcal{T}$ such that $x \in U$, $y \in V$, and $U \cap V = \varnothing$.

Since metric spaces were so far our most general spaces, we are certainly eager to know whether the concept of topological spaces is even more general. And in fact, each metric space is a Hausdorff space, in a sense made precise by the following remark.

Remark 6.11 If (X, d) is a metric space, then the set of open subsets (w.r.t. d) is a topology on X, called the *topology induced* by d. The set of open balls $U_\varepsilon(x)$ is a basis for $\mathcal{T}$. Further, it is easy to verify that the induced topology is a Hausdorff topology. $\Diamond$

As mentioned before, a topology allows us to lift notions, already introduced in the context of metric spaces, to a more abstract level, such as for example convergence, continuity, and compactness.

Definition 6.12 (1) Let $(X, \mathcal{T})$ be a topological space. A sequence $(x_n)_{n\in\mathbb{N}}$ in X is said to *converge* to $x \in X$ if for every $U \in \mathcal{T}$ with $x \in U$ there is $N \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq N$. In this case, x is called a *limit* of $(x_n)_{n\in\mathbb{N}}$.

(2) Let X and Y be topological spaces. Then $f\colon X \to Y$ is *continuous*, if

$$f^{-1}(U) = \{x \in X : f(x) \in U\}$$

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*.

(3) $\mathfrak{U} \subset \mathcal{T}$ is an *open covering* of X, if $X \subset \bigcup_{U\in\mathfrak{U}}$. X is called *compact*, if every open covering contains a finite subcover.

Compared to metric spaces, general topological spaces may possess various pathological properties. For example, the limit of a convergent sequence in a topological space need not be unique. Also, the relation between compactness and closedness of sets is more involved, since in general topologies the compactness of a subset need not imply its closedness.

Restricting to Hausdorff spaces, however, topological spaces behave in many ways similar to metric spaces.

Remark 6.13 In a Hausdorff space, every convergent sequence has a unique limit. $\Diamond$

Moreover, we have the following result.

Lemma 6.14 *Compact subsets of Hausdorff spaces are closed.*

Proof. This will be proved in the tutorials. ■

Let us next take a short look at the properties of continuous functions between topological spaces. We first observe that the convergence of sequences is preserved under continuous mappings.

Remark 6.15 Let $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ be topological spaces, and assume that the function $f\colon X \to Y$ is continuous. Then, if $(x_n)_{n\in\mathbb{N}} \subset X$ is a convergent sequence (w.r.t. $\mathcal{T}$) with limit $x \in X$, the sequence $(f(x_n))_{n\in\mathbb{N}} \subset Y$ also converges (w.r.t. $\mathcal{T}'$), with limit $f(x) \in Y$. However, in general topological spaces, even if they are Hausdorff, the opposite implication is not always fulfilled: In contrast to metric spaces, a function which preserves the convergence of sequences is not necessarily continuous. ◊

The following relation between continuous functions and compact sets is well-known for $\mathbb{K}^n$. It is though quite fascinating that it is still true in this very general setting

Remark 6.16 A continuous function $f\colon X \to Y$ maps compact sets to compact sets.

This can be shown as follows: 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 of subsets $f^{-1}(U_1), \dots, f^{-1}(U_n)$. Thus, $U_1, \dots, U_n$ is a finite subcover of $\mathfrak{U}$. ◊

Another later quite useful result related to compact sets is a seemingly weaker version, which though turns out to imply compactness. In this sense the next lemma shows that it is sufficient to consider only special open coverings when checking compactness.

Lemma 6.17 *Let $(X, \mathcal{T})$ be a topological space, let $\gamma \subset \mathcal{T}$, and let $\mathcal{B}$ be the set of all finite intersections of sets in γ. Assume further 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. Towards a contradiction, assume that there exists an open covering $\mathfrak{U}_0 \subset \mathcal{T}$ of X without a finite subcover. Then we define

$$\Omega := \{\mathfrak{U} \subset \mathcal{T} : \mathfrak{U} \supset \mathfrak{U}_0 \text{ and } \mathfrak{U} \text{ does not contain a finite subcover}\}.$$

We observe that $\Omega \neq \varnothing$, since $\mathfrak{U}_0 \in \Omega$. Also, Ω is partially ordered by inclusion "$\subset$". We now first show that Ω satisfies the hypothesis of Zorn's Lemma.

For this, 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 $\mathfrak{U}_i \in K$, $U_i \in \mathfrak{U}_i$, $i = 1, \ldots, r$, with

$$X = \bigcup_{i=1}^{r} U_i.$$

Since K is a chain, there exists some $\mathfrak{U}_{i_0} \in K$ with $\mathfrak{U}_i \subset \mathfrak{U}_{i_0}$, $i = 1, \ldots, r$. Thus $\mathfrak{U}_{i_0}$ contains the finite subcover $\{U_i\}_{i=1}^{r}$, a contradiction. Therefore, $\tilde{\mathfrak{U}}$ does not contain a finite subcover, which implies $\tilde{\mathfrak{U}} \in \Omega$. This shows that Ω satisfies the hypothesis of Zorn's Lemma.

By Zorn's Lemma, there exists a maximal element $\mathcal{M} \in \Omega$, i.e., no element in Ω is larger than $\mathcal{M}$ in the sense that $\mathcal{M}$ is a proper subset. We next claim that

$$A, B \in \mathcal{T},\ A \notin \mathcal{M},\ B \notin \mathcal{M} \quad \text{implies} \quad A \cap B \notin \mathcal{M}. \tag{6.1}$$

Since $A \notin \mathcal{M}$ and $\mathcal{M}$ is maximal, it follows that $\mathcal{M} \cup \{A\}$ contains a finite subcover $\{A, M_1, \ldots, M_n\}$. Similarly, $\mathcal{M} \cup \{B\}$ contains a finite subcover $\{B, M_1', \ldots, M_m'\}$. This implies $\{A \cap B, M_1, \ldots, M_n, M_1', \ldots, M_m'\}$ is a finite subcover of $\mathcal{M} \cup \{A \cap B\}$ of X, and therefore $A \cap B \notin \mathcal{M}$, showing (6.1).

By induction, we can now conclude that $A_1, \ldots, A_n \notin \mathcal{M}$ implies

$$\bigcap_{i=1}^{n} A_i \notin \mathcal{M}.$$

Notice that once we proved

$$\bigcup_{M \in \mathcal{M}} M = \bigcup_{M \in \gamma \cap \mathcal{M}} M, \tag{6.2}$$

we are done. The reason for this is that since $\mathcal{M}$ covers X, (6.2) 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 $\mathcal{M} \in \Omega$. Thus our initial assumption that there exists an open covering $\mathfrak{U}_0 \subset \mathcal{T}$ of X without a finite subcover is false, and hence X is compact.

It remains to prove (6.2). For this, let $x \in M \in \mathcal{M}$. Then there exist $S_1, \ldots, S_r \in \gamma$, 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 $\mathcal{M} \cup \{M\} = \mathcal{M}$, it follows that $\mathcal{M} \cup \{S_1 \cap \ldots \cap S_r\} \in \Omega$. Since $\mathcal{M}$ is the maximal element of Ω and (6.1), 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 shows (6.2), which finishes the proof. ■

Our goal is now, given a collection of subsets, to introduce a topology which contains those sets. This will allow us later to ensure continuity of certain operators. The following result describes the construction of such a topology.

Lemma 6.18 *Let $X \neq \varnothing$, 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. Further, 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. This will be proven in the tutorials. ∎

We next fix some terminology and notations related to the construction described in Lemma 6.18.

Definition 6.19 Let $X \neq \varnothing$ and $\gamma \subset \mathcal{P}(X)$.

(1) Then the topology defined in Lemma 6.18 is denoted by $\mathcal{T}(\gamma)$.

(2) Let $(X, \mathcal{T})$ be a topological space. Then a set $\gamma \subset \mathcal{P}(X)$ is called a *subbasis* for $\mathcal{T}$ if $\mathcal{T} = \mathcal{T}(\gamma)$.

Intuitively, it seems clear that the topology $\mathcal{T}(\gamma)$ we just constructed from γ is the smallest topology containing γ. The next lemma states this formally.

Lemma 6.20 *(i) The intersection of arbitrarily many topologies on a set X is again a topology on X.*

(ii) For a set $\gamma \subset \mathcal{P}(X)$, we have

$$\mathcal{T}(\gamma) = \bigcap \{\mathcal{T} : \mathcal{T} \text{ topology on } X \text{ with } \gamma \subset \mathcal{T}\}.$$

That is, $\mathcal{T}(\gamma)$ is the smallest topology containing γ.

Proof. This is an exercise. ∎

Finally, we arrive at the point where we can introduce the topology, which will eventually lead to weak convergence as defined in the beginning of this chapter. It will be defined depending on certain mappings, as to ensure their continuity.

Definition 6.21 Let X be a set, let I be an index set, let $(X_i, \mathcal{T}_i)$ be topological spaces, and let $f_i \colon X \to X_i$, $i \in I$. The *weak topology with respect to the mappings f_i* is defined as $\mathcal{T}(\gamma)$, where

$$\gamma := \left\{ f_i^{-1}(V) : V \in \mathcal{T}_i, i \in I \right\}.$$

Indeed the mappings are continuous with respect to this topology as the next result shows. Furthermore, the weak topology can also be defined by just choosing subbases.

Lemma 6.22 *Let X be a set, let I be an index set, let $(X_i, \mathcal{T}_i)$ be topological spaces, and let $f_i\colon X \to X_i$, $i \in I$. For each $i \in I$, let γ_i be a subbasis of $\mathcal{T}_i$. Then the weak topology $\mathcal{T}$ with respect to the mappings 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 first claim is an exercise. The second claim is an almost immediate consequence of Lemma 6.20. ■

By choosing the mappings f_i in a smart way, we will now define the weak topology on E and the weak* topology on E^*, leading to weak and weak*-convergence.

Definition 6.23 Let E be a normed space over $\mathbb{K}$.

(1) 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 $\sigma(E, E^*)$.

(2) 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 $\sigma(E^*, E)$.

Notice that the weak* topology is not defined by using all functionals $f\colon E^* \to \mathbb{K}$, $f \in E^{**}$, but only those of type $\hat{x}$. Hence, we have the following relation.

Remark 6.24 We have

$$\sigma(E^*, E) \subset \sigma(E^*, E^{**}),$$

which follows from the fact that each $\hat{x}$ is an element of E^{**}, and hence $\sigma(E^*, E^{**})$-continuous. ◊

For the proofs of the next section, we require a subbasis for the weak and weak*-topology, related to Lemma 6.22.

Lemma 6.25 *Let E be a normed space over $\mathbb{K}$.*

(i) A subbasis for the weak topology $\sigma(E, E^*)$ on E is given by the sets

$$\{x \in E : |f(x) - a| < \varepsilon\}, \quad f \in E^*, a \in \mathbb{K}, \varepsilon > 0.$$

(ii) For the weak*-topology $\sigma(E^*, E)$ on E^*, a subbasis is given by

$$\{f \in E^* : |f(x) - a| < \varepsilon\}, \quad x \in E, a \in \mathbb{K}, \varepsilon > 0.$$

Proof. (i). This is a consequence of Lemma 6.22. According to this result, a subbasis is given by

$$f^{-1}(U_\varepsilon(a)), \quad f \in E^*, a \in \mathbb{K}, \varepsilon > 0,$$

since $U_\varepsilon(a), a \in \mathbb{K}, \varepsilon > 0$ is a basis of the standard topology on $\mathbb{K}$. Now

$$f^{-1}(U_\varepsilon(a)) = \{x \in E : |f(x) - a| < \varepsilon\}.$$

(ii). This is proven by a similar argument. ∎

Corollary 6.26 *(1) Let $x \in E$ be a point of a normed space E. The collection of sets*

$$\mathcal{U}(x, f_1, \ldots, f_n, \varepsilon) := \{y \in E : |f_i(y) - f_i(x)| < \varepsilon \quad \text{for all } i \in \{1, \ldots, n\}\},$$

with $n \in \mathbb{N}$, $\varepsilon > 0$, and $f_1, \ldots, f_n \in E^*$, constitutes a neighborhood basis for x with respect to $\sigma(E, E^*)$.

(2) Analogously, for a functional $f \in E^*$ in the dual E^* of a normed space E, a neighborhood basis for f with respect to $\sigma(E^*, E)$ is provided by the sets

$$\mathcal{U}(f, x_1, \ldots, x_n, \varepsilon) := \{g \in E^* : |f(x_i) - g(x_i)| < \varepsilon \quad \text{for all } i \in \{1, \ldots, n\}\},$$

where $n \in \mathbb{N}$, $\varepsilon > 0$, and $x_1, \ldots, x_n \in E$.

As an immediate consequence of Corollary 6.26 and the Hahn-Banach Theorem 4.6, we can deduce the following important fact.

Lemma 6.27 *Let E be a normed space and E^* its dual. The weak topologies $\sigma(E, E^*)$ and $\sigma(E^*, E)$ are Hausdorff topologies.*

In particular, the limits of convergent sequences in $(E, \sigma(E, E^*))$ and $(E^*, \sigma(E^*, E))$ are unique. The following lemma finally draws a connection between the notions of weak and weak* convergence introduced in Section 6.1 and the weak and weak* topologies investigated in this section.

Lemma 6.28 *Let E be a normed space, E^* its dual.*

(1) A sequence $(x_n)_{n\in\mathbb{N}}$ in E is weakly convergent to $x \in E$ in the sense of Definition 6.1 if and only if $(x_n)_{n\in\mathbb{N}}$ is convergent to x with respect to $\sigma(E, E^*)$.

(2) A sequence $(f_n)_{n\in\mathbb{N}}$ in E^* is weak-convergent to $f \in E^*$ in the sense of Definition 6.3 if and only if $(f_n)_{n\in\mathbb{N}}$ is convergent to f with respect to $\sigma(E^*, E)$.*

We finish this section by introducing yet another topology, again to ensure that certain maps are continuous. This time we focus on a product space and aim to enforce continuity of the associated projections.

Definition 6.29 Let I be an index set, and let $(X_i, \mathcal{T}_i), i \in I$, be topological spaces. Further, denote by $X = \prod_{i\in I} X_i$ the cartesian product of the X_i, and let p_i be the canonical projection from X onto X_i, i.e., $p_j((x_i)_{i\in I}) = x_j$ for each j. Then the weak topology with respect to the p_i is called the *product topology* on X.

Equipped with this topology, the famous Theorem of Tychonoff states that the product space is compact if and only if all of the single spaces are.

Theorem 6.30 (Tychonoff) *Let I be an index set, and let $(X_i, \mathcal{T}_i)$ be topological spaces. Further, let $\mathcal{T}$ denote the product topology on $X = \prod_{i\in I} X_i$. Then the following conditions are equivalent.*

(i) $(X, \mathcal{T})$ is compact.

(ii) $(X_i, \mathcal{T}_i)$ is compact for all $i \in I$.

Proof. (i) $\Rightarrow$ (ii). Let $(X, \mathcal{T})$ be compact. By definition of the product topology, each p_i is continuous. Thus $X_i = p_i(X)$ is compact for all $i \in I$.

(ii) $\Rightarrow$ (i). Now let each $(X_i, \mathcal{T}_i)$ be compact. By definition, $\gamma = \{p_i^{-1}(V) : V \in \mathcal{T}_i, i \in I\}$ is a subbasis for $\mathcal{T}$. Towards a contradiction, assume that $(X, \mathcal{T})$ is not compact. Then, by Lemma 6.17, there exists $W \subset \gamma$, which is an open covering of X, but does not contain a finite subcover. Next, for each $i \in I$, set

$$W_i := \{V \in \mathcal{T}_i : p_i^{-1}(V) \in W\}.$$

We first observe that W_i is not an open covering of X_i, since otherwise, by (ii), there would exist $V_{i_1}, \ldots, V_{i_n} \in W_i$ with $X_i = \bigcup_{k=1}^n V_{i_k}$, which in turn would imply

$$X = p_i^{-1}(X_i) = \bigcup_{k=1}^n \underbrace{p_i^{-1}(V_{i_k})}_{\in W},$$

a contradiction.

Now, for each $i \in I$, choose some $x_i \in X_i \setminus \bigcup_{V\in W_i} V$, and set $x := (x_i)_{i\in I} \in X$. Since W is an open covering of X, there exists $V \in W$ such that $x \in V$. Also $W \subset \gamma$ implies that there exists $i \in I$ such that $V = p_i^{-1}(\tilde{V})$ for some $\tilde{V} \in W_i$. But $p_i(x) = x_i \notin \tilde{V}$ implies $x \notin V$, a contradiction. ∎

6.3 Two Fundamental Results using Weak and Weak* Topology

The well-known Theorem of Heine-Borel states that in $\mathbb{K}^n$ a set is compact if and only if it is bounded and closed. In Theorem 2.9(iii), we showed that this is true for arbitrary finite-dimensional normed spaces. However, the generalization to the infinite-dimensional setting turns out to be extremely delicate. And – as we just saw – requires extensive preparations, since defining compactness, boundedness, and closedness with respect to the norm does not work at all. It is therefore even more fascinating, that considering subsets of a dual space and equipping it with the weak* topology, we can indeed derive a result in the spirit of the Theorem of Heine-Borel.

This theorem, which is due to Leonidas Alaoglu, also provides a generalization of Theorem 2.10, which showed that compactness of the unit ball is equivalent to the normed space being finite-dimensional. He could show that $E_1^* := K_1(0_{E^*})$ is not compact with respect to the norm, but at least $\sigma(E^*, E)$-compact.

Theorem 6.31 (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, $E_1^* := K_1(0_{E^*})$ is $\sigma(E^*, E)$-compact.

Proof. We start by setting $c := \sup\{\|f\| : f \in M\}$, and, for each $x \in E$, define

$$A_x := \{z \in \mathbb{K} : |z| \leq c\|x\|\}.$$

By Theorem 6.30 and using the fact that $A_x \subset \mathbb{K}$, $x \in E$, is bounded and closed, the product $A = \prod_{x \in E} A_x$, equipped with the product topology $\mathcal{T}_A$, is compact.

Now we define the mapping

$$\varphi\colon M \to A, \quad \varphi(f) = (f(x))_{x \in E}.$$

To check its properties, notice first that φ is well-defined, since

$$|f(x)| \leq \|f\|\|x\| \leq c\|x\| \quad \text{implies} \quad f(x) \in A_x.$$

Evidently, f is also injective, since $\varphi(f) = \varphi(g)$ implies $f(x) = g(x)$ for all $x \in E$, thus $f = g$. To prove that φ is continuous, we first conclude from Lemmata 6.22 and 6.25 that the collection of sets of the form

$$V(a, x, \varepsilon) = \{g \in E^* : \ |g(x) - a| < \varepsilon\}, \quad a \in \mathbb{K}, x \in E, \varepsilon > 0,$$

is a subbasis for $\sigma(E^*, E)$ and the sets

$$W(a, x, \varepsilon) = \{(z_y)_{y \in E} : \ |z_x - a| < \varepsilon\}, \quad a \in \mathbb{K}, x \in E, \varepsilon > 0,$$

form a subbasis for $\mathcal{T}_A$. Moreover,

$$\varphi(V(a, x, \varepsilon) \cap M) = W(a, x, \varepsilon) \cap \varphi(M). \tag{6.3}$$

Together with the injectivity of φ, this implies that φ is a homeomorphism between M and $\varphi(M)$, i.e., both $\varphi : M \to \varphi(M)$ and $\varphi^{-1} : \varphi(M) \to M$ are continuous[1].

[1] One easily proves that $f\colon X \to Y$ is continuous if for a subbasis γ, $f^{-1}(V)$ is open in X for all $V \in \gamma$.

We will next show that $\varphi(M)$ is closed in A. For this, let $a = (a_x)_{x \in E} \in \overline{\varphi(M)}$, and define the functional

$$f \colon E \to \mathbb{K}, \quad x \mapsto a_x.$$

To prove 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, since a is contained in all the sets by definition and finite intersections of open sets are open. We will use the fact shown in the tutorials that

$$x \in \overline{A} \iff U \cap A \neq \varnothing \quad \text{for all } U \text{ open with } x \in U.$$

This implies there exists $g \in M$ such that $\varphi(g) \in W$. Hence, we obtain

$$\begin{aligned} \left| a_{\lambda x + \mu y} - (\lambda a_x + \mu a_y) \right| &\leq \left| a_{\lambda x + \mu y} - g(\lambda x + \mu y) \right| + |\lambda g(x) - \lambda a_x| + \left| \mu g(y) - \mu a_y \right| \\ &\leq \varepsilon + |\lambda| \, \varepsilon + |\mu| \, \varepsilon. \end{aligned}$$

Since ε was chosen arbitrarily, we conclude that $a_{\lambda x + \mu y} = \lambda a_x + \mu a_y$, which implies linearity of f. 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 $a_1, \ldots, a_n \in \mathbb{K}$, $x_1, \ldots x_n \in E$, and $\varepsilon > 0$ such that

$$f \in \bigcap_{k=1}^{n} V(a_k, x_k, \varepsilon) \subset U,$$

since the sets of this type form a basis of $\sigma(E^*, E)$. This implies

$$|a_{x_k} - a_k| = |f(x_k) - a_k| < \varepsilon \quad \text{for all } k = 1, \ldots, n,$$

and thus

$$a \in \bigcap_{k=1}^{n} W(a_k, x_k, \varepsilon).$$

Since $a \in \overline{\varphi(M)}$, it follows that

$$\varphi(M) \cap \bigcap_{k=1}^{n} W(a_k, x_k, \varepsilon) \neq \varnothing.$$

By (6.3), it follows that

$$M \cap \bigcap_{k=1}^{n} V(a_k, x_k, \varepsilon) \neq \varnothing.$$

Since M is $\sigma(E^*, E)$-closed, we conclude that $f \in M$, arguing that any neighborhood U of f has non-empty intersection with M, thus $f \in \overline{M} = M$. Thus $a = \varphi(f) \in \varphi(M)$, and hence $\varphi(M)$ is closed.

Finally, since A is compact, also $\varphi(M)$ as a closed subset is compact. Since φ is a homeomorphism from M to $\varphi(M)$, we can conclude that M is compact in $\sigma(E^*, E)$.

For the 'in particular'-part, notice that it is sufficient to prove that E_1^* is weak*-closed, or equivalently, that
$$\{f \in E^* : \|f\| > 1\}$$
is open. For this, let $f \in E^*$ have norm larger than 1. Then there exists $x \in E$, $\|x\| = 1$, satisfying $|f(x)| > 1$. Since
$$V(f(x), x, |f(x)| - 1) \cap E_1^* = \varnothing,$$
the claim is proven. ∎

We next aim to derive a corollary from this result showing that $\hat{E}_1$ is $\sigma(E^{**}, E^*)$-dense in E_1^{**}, which is quite interesting by itself, but will also help us in the proof of our final theorem. For this corollary however, we need the following somehow technical preparatory lemma.

Lemma 6.32 *Let E be a normed space, let $\varphi \in E_1^{**}$, let $f_1, \ldots, f_n \in E^*$, and define $h\colon E \to \mathbb{R}$ by*
$$h(x) = \sum_{i=1}^{n} |\varphi(f_i) - f_i(x)|^2.$$
Then we have
$$\inf_{x \in E_1} h(x) = 0.$$

Proof. Let $\tau = \inf_{x \in E_1} h(x)$, and let $(x_k)_k \subset E_1$ with $\tau = \lim_{k\to\infty} h(x_k)$. Notice that the sequence $(f_1(x_k))_k$ is bounded, and hence contains a convergent subsequence $(f_1(x_{k_j}))_j$. Continuing this process, we can assume without loss of generality that $(f_i(x_k))_k$ converges for all $i = 1, \ldots, n$. Hence we can set
$$\alpha_i := \lim_{k\to\infty} f_i(x_k) \quad \text{and} \quad \delta_i := \varphi(f_i) - \alpha_i.$$
Then we have
$$\tau = \lim_{k\to\infty} h(x_k) = \lim_{k\to\infty} \sum_{i=1}^{n} |\varphi(f_i) - f_i(x_k)|^2 = \sum_{i=1}^{n} |\delta_i|^2.$$

Next, let
$$f := \sum_{i=1}^{n} \overline{\delta_i} f_i \in E^*.$$
We will first show that $\|f\| \leq \operatorname{Re} \sum_{i=1}^{n} \overline{\delta_i} \alpha_i$. For this, let $x \in E_1$ be arbitrary. Then, for $t \in [0, 1]$, we have
$$\tau \leq h((1-t)x_k + tx) = \sum_{i=1}^{n} \left|\varphi(f_i) - (1-t) f_i(x_k) - t f_i(x)\right|^2$$

$$= \sum_{i=1}^{n} |\varphi(f_i) - f_i(x_k)|^2 - 2t \operatorname{Re} \sum_{i=1}^{n} (f_i(x_k) - f_i(x)) \overline{(\varphi(f_i) - f_i(x_k))}$$
$$+ t^2 \sum_{i=1}^{n} |f_i(x_k) - f_i(x)|^2,$$

and hence as $k \to \infty$

$$-2t \operatorname{Re} \sum_{i=1}^{n} (f_i(x) - \alpha_i)\overline{\delta}_i + t^2 \sum_{i=1}^{n} |f_i(x) - \alpha_i|^2 \geq 0.$$

Division by t for $t > 0$ and $t \to 0$ yields $-2 \operatorname{Re} \sum_{i=1}^{n} (f_i(x) - \alpha_i)\overline{\delta}_i \geq 0$, that is

$$\operatorname{Re} f(x) \leq \operatorname{Re} \sum_{i=1}^{n} \overline{\delta}_i \alpha_i.$$

Since $|f(x)| = \lambda f(x) = f(\lambda x)$ with $|\lambda| = 1$, this implies $|f(x)| = \operatorname{Re} f(\lambda x) \leq \operatorname{Re} \sum_{i=1}^{n} \overline{\delta}_i \alpha_i$, and thus $\|f\| \leq \operatorname{Re} \sum_{i=1}^{n} \overline{\delta}_i \alpha_i$.

Finally, since $f(x_k) = \sum_{i=1}^{n} \overline{\delta}_i f_i(x_k) \to \sum_{i=1}^{n} \overline{\delta}_i \alpha_i$ and $x_k \in E_1$, we even have

$$\left| \sum_{i=1}^{n} \overline{\delta}_i \alpha_i \right| \leq \|f\| \leq \operatorname{Re} \sum_{i=1}^{n} \overline{\delta}_i \alpha_i.$$

Therefore, we can conclude

$$\|f\| = \sum_{i=1}^{n} \overline{\delta}_i \alpha_i.$$

With $\|\varphi\| \leq 1$, we finally obtain that

$$\tau = \sum_{i=1}^{n} |\delta_i|^2 = \sum_{i=1}^{n} \overline{\delta}_i (\varphi(f_i) - \alpha_i) = \varphi(f) - \|f\| \leq 0,$$

which shows $\inf_{x \in E_1} h(x) = 0$. ∎

This allows us to prove the announced corollary quite easily.

Corollary 6.33 *Let E be a normed space. Then $\hat{E}_1$ is $\sigma(E^{**}, E^*)$-dense in E_1^{**}.*

Proof. For $\varphi \in E_1^{**}$, $\varepsilon > 0$, and $f_1, \dots, f_n \in E^*$, define

$$\mathcal{U}(\varphi, f_1, \dots, f_n, \varepsilon) := \{\psi \in E^{**} : |\psi(f_i) - \varphi(f_i)| < \varepsilon \; \forall i = 1, \dots, n\}. \tag{6.4}$$

Similar to the sets from Corollary 6.26, the sets $\mathcal{U}(\varphi, f_1, \dots, f_n, \varepsilon)$ are a basis for the neighborhoods of φ in $\sigma(E^{**}, E^*)$ in the sense of Definition 6.9. By Lemma 6.32, there exists some $x \in E_1$ such that

$$|\varphi(f_i) - \hat{x}(f_i)| \leq \sqrt{h(x)} < \varepsilon \quad \text{for all } i = 1, \ldots, n.$$

Hence $\hat{x} \in \mathcal{U}(\varphi, f_1, \ldots, f_n, \varepsilon) \cap \hat{E}_1$. ∎

The last theorem considers the weak-compactness of the unit ball in a normed space, and shows that this is actually equivalent to the normed space being reflexive. This can be interpreted as yet another generalization of Theorem 2.10,

Theorem 6.34 *Let E be a normed space. Then the following conditions are equivalent.*

(i) E is reflexive.

(ii) $E_1 := K_1(0_E)$ is $\sigma(E, E^*)$-compact.

Proof. (i)⇒(ii). Recall the sets $\mathcal{U}(x, f_1, \ldots, f_n, \varepsilon)$ introduced in Corollary 6.26, as well as the sets defined in (6.4). Since

$$\Lambda_E(\mathcal{U}(x, f_1, \ldots, f_n, \varepsilon)) = \hat{E} \cap \mathcal{U}(\hat{x}, f_1, \ldots, f_n, \varepsilon),$$

we can conclude that

$$\Lambda_E \colon (E, \sigma(E, E^*)) \to (E^{**}, \sigma(E^{**}, E^*))$$

is continuous and a homeomorphism. By Theorem 6.31, E_1^{**} is $\sigma(E^{**}, E^*)$-compact. Hence, by (i), the set $E_1 = \Lambda_E^{-1}(E_1^{**})$ is $\sigma(E, E^*)$-compact.

(ii)⇒(i). By (ii), the set $\hat{E}_1$ is $\sigma(E^{**}, E^*)$-compact. This implies that $\hat{E}_1$ is closed. By Corollary 6.33, $\hat{E}_1$ is also $\sigma(E^{**}, E^*)$-dense in E_1^{**}. This implies that

$$\hat{E}_1 = E_1^{**},$$

and thus $\hat{E} = E^{**}$, which is (i). ∎

7 Hilbert Spaces and the Riesz Representation Theorem

In the sequence

$$\text{Topological Spaces} \supseteq \text{Metric Spaces} \supseteq \text{Normed Spaces}$$

each time structural properties are added. However, we are still not in the situation of being able to employ the property of orthogonality such as in the case $\mathbb{R}^n$, preventing us from, for instance, using the concept of orthonormal bases. Therefore, we now specialize again by requiring additional structural properties leading to so-called Hilbert spaces, which are very convenient spaces to work in as we will see.

7.1 The Structure of a Hilbert Space

The starting point of the definition of a Hilbert space is the concept of a scalar product, which will later on allow us to talk about orthogonality.

Definition 7.1 (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) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$,

(b) $\langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle$, and

(c) $\langle x, y \rangle = \overline{\langle y, x \rangle}$,

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*.

(2) 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).

The Cauchy-Schwarz Inequality is a tremendously useful formula, which we will exploit extensively. It shows how the absolute value of the scalar product of two vectors can be upper bounded by two scalar products of each vector with itself. It is interesting to notice, that we can even prove an equivalent condition for equality to hold. This will be particularly beneficial, since often the Cauchy-Schwarz Inequality provides a sufficient, but quite weak upper bound.

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*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle \langle y, y\rangle \quad \text{for all } x, y \in E.$$

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 sesquilinearity and hermiticity of $\langle \cdot , \cdot \rangle$, for all $x, y \in E$, we have

$$\begin{aligned} 0 \leq \left\| \|y\|^2 x - \langle x, y\rangle y \right\|^2 &= \|y\|^4 \|x\|^2 - 2 \operatorname{Re} \left\langle \|y\|^2 x, \langle x, y\rangle y \right\rangle + |\langle x, y\rangle|^2 \|y\|^2 \\ &= \|y\|^4 \|x\|^2 - \|y\|^2 |\langle x, y\rangle|^2 \\ &= \|y\|^2 \left(\|x\|^2 \|y\|^2 - |\langle x, y\rangle|^2 \right). \end{aligned}$$

This proves the Cauchy-Schwarz inequality. This computation also shows that, in the positive definite case, equality implies that x and y are linearly dependent. The converse can also be easily seen, since if x and y are linearly dependent, it follows that $|\langle x, y\rangle|^2 = \|x\|^2 \|y\|^2$. ∎

We claimed at the beginning of the chapter that Hilbert spaces are a special case of normed spaces. For this, to have a chance to hold true, we require the definition of a norm from the building blocks of a Hilbert space, i.e., from a scalar product. And indeed, we can easily define a norm as the next lemma shows. Note though that this is certainly not the only norm possible.

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. We first prove the norm properties. For this, we first observe that, by positive definiteness,

$$\|x\| = 0 \quad \Longleftrightarrow \quad x = 0.$$

Next, for each $x \in E$ and $\lambda \in \mathbb{K}$, we have

$$\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{|\lambda|^2 \langle x, x \rangle} = |\lambda| \|x\|.$$

Third, by Lemma 7.2,

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

Thus, $\|\cdot\|$ is indeed a norm.

To show continuity, we observe that

$$\begin{aligned} |\langle x, y \rangle - \langle x_0, y_0 \rangle| &\leq |\langle x - x_0, y \rangle| + |\langle x_0, y - y_0 \rangle| \\ &\leq \|x - x_0\| \|y\| + \|x_0\| \|y - y_0\| \\ &\leq \|x - x_0\| \|y - y_0\| + \|x - x_0\| \|y_0\| + \|x_0\| \|y - y_0\|. \end{aligned}$$

The last term is small, if both $\|x - x_0\|$ and $\|y - y_0\|$ are small, thereby implying continuity. ■

Now it is time to finally provide the definition of a Hilbert space.

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 a *Hilbert space*.

Notice that this immediately shows that a Hilbert space, equipped with the associated norm, is a Banach space.

We already mentioned that a Hilbert space can certainly be equipped with many norms. However, there is an easy way to check whether a norm arises from a scalar product as in Lemma 7.3, namely by verifying the parallelogram identity. This identity is also a useful tool, if our norm is known to be constructed from a scalar product.

Lemma 7.5 *Let $(E, \|\cdot\|)$ be a normed space. Then the following conditions are equivalent.*

(i) There exists a scalar product $\langle \cdot , \cdot \rangle$ on E such that

$$\|x\| = \sqrt{\langle x, x \rangle} \quad \textit{for all } x \in E.$$

(ii) For all $x, y \in E$, the parallelogram identity

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

holds.

Proof. (i) $\Rightarrow$ (ii). This is an easy calculation.

(ii) $\Rightarrow$ (i). First, we consider $\mathbb{K} = \mathbb{C}$, and claim that

$$\langle x, y\rangle := \tfrac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right), \quad x, y \in E,$$

defines an inner product. For this, $\langle x, y\rangle = \overline{\langle y, x\rangle}$ is immediate.

Second, to prove that $\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$, we have to show

$$\begin{aligned}\|x+x'+y\|^2 &- \|x+x'-y\|^2 + i\|x+x'+iy\|^2 - i\|x+x'-iy\|^2 \\ &= \|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2 \\ &\quad + \|x'+y\|^2 - \|x'-y\|^2 + i\|x'+iy\|^2 - i\|x'-iy\|^2.\end{aligned}$$

By (ii), for any $z \in E$,

$$\begin{aligned}\|x+x'+z\|^2 &= \tfrac{1}{2}\|x+x'+z\|^2 + \tfrac{1}{2}\|x+x'+z\|^2 \\ &= \|x+z\|^2 + \|x'\|^2 - \tfrac{1}{2}\|x+z-x'\|^2 + \|x'+z\|^2 + \|x\|^2 - \tfrac{1}{2}\|x'+z-x\|^2.\end{aligned}$$

Now we choose $z = y$ and $z = -y$ and substract, which yields

$$\|x+x'+y\|^2 - \|x+x'-y\|^2 = \|x+y\|^2 + \|x'+y\|^2 - (\|x-y\|^2 + \|x'-y\|^2).$$

In a similar way, we also choose $z = iy$ and $z = -iy$ and substract. This then proves

$$\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle.$$

Third, we will show that

$$\langle \lambda x, y\rangle = \lambda\langle x, y\rangle \quad \text{for all } \lambda \in \mathbb{C}, x, y \in E. \tag{7.1}$$

We already proved that

$$\langle mx, y\rangle = m\langle x, y\rangle \quad \text{for all } m \in \mathbb{N}, x, y \in E.$$

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

$$\left\langle \frac{m}{n}x, y\right\rangle = \frac{m}{n}\langle x, y\rangle.$$

Recall that

$$\lambda \mapsto \langle \lambda x, y \rangle - \lambda \langle x, y \rangle$$

is continuous (since $\|\cdot\|$ is continuous). Since we just showed that this function is zero on $\mathbb{Q}$, it must be zero on all of $\mathbb{R}$. Finally, $\langle ix, y \rangle = i\langle x, y \rangle$ follows by definition of $\langle \cdot\,,\cdot \rangle$, which proves (7.1).

For $\mathbb{K} = \mathbb{R}$ we define

$$\langle x, y \rangle := \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right).$$

Then, similar arguments as above show that $\langle \cdot\,,\cdot \rangle$ is an inner product inducing the norm $\| \cdot \|$. ■

Let us finish this part with an interesting construction. Given an inner product space, it shows how this can be slightly enlarged to yield a Hilbert space.

Remark 7.6 Let $(E, \langle \cdot\,,\cdot \rangle)$ be an inner product space, and let $\Lambda_E \colon E \to E^{**}$ be the canonical embedding of E into E^{**}. We now equip E with $\|\cdot\| = \sqrt{\langle \cdot\,,\cdot \rangle}$. The completion of $(E, \|\cdot\|)$ is then given by

$$\mathcal{H} := \overline{\Lambda_E(E)} \subset E^{**}.$$

As Λ_E is isometric, the parallelogram identity also holds for $\Lambda_E(E)$. And since the norm on E^{**} is continuous, this identity holds for $\mathcal{H}$ as well. On $\Lambda_E(E)$, we then set

$$\langle \hat{x}, \hat{y} \rangle := \langle x, y \rangle, \quad x, y \in E.$$

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*. ◊

7.2 The Concept of Orthogonality

As anticipated, we now have the concept of orthogonality available. This allows us to define notions such as orthogonal complement even for any inner product space.

Definition 7.7 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 a set $M \subset \mathcal{H}$, we define by

$$M^{\perp} := \{y \in \mathcal{H} : \langle y, x \rangle = 0 \quad \text{for all } x \in M\}$$

the *orthogonal complement of M*.

The first lemma shows that the orthogonal complement is in fact a linear space and that it is automatically closed – even if the original set does not possess any of those properties. Moreover, not too surprisingly this time, the intersection of a set and its orthogonal complement is trivial.

Lemma 7.8 *Let $\mathcal{H}$ be an inner product space, and let $M \subset \mathcal{H}$. Then $M^\perp$ is a closed linear subspace of $\mathcal{H}$, and*

$$M \cap M^\perp = \{0\} \text{ if } 0 \in M, \text{ otherwise } M \cap M^\perp = \varnothing.$$

Proof. The first part follows directly from Definition 7.7. The second part follows from the properties of the inner product $\langle \cdot , \cdot \rangle$. ■

We are next aiming for a result which shows that each Hilbert space – notice that now we require the completeness property – can be decomposed in a closed subspace and its orthogonal complement. For the proof of this result, we need a certain uniqueness property of a vector with minimal norm, which we will discuss first.

Lemma 7.9 *Let $\mathcal{H}$ be a Hilbert space, and let $K \subset \mathcal{H}$, $K \neq \varnothing$, be 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 $(x_n)_{n \in \mathbb{N}} \subset K$ with

$$\|x_n\| \xrightarrow{n \to \infty} d.$$

We now first show that $(x_n)_{n \in \mathbb{N}}$ forms a Cauchy-sequence, so that we can exploit completeness of a Hilbert space. For this, notice that since K is convex, we have

$$\tfrac{1}{2}(x_n + x_m) \in K \quad \text{for all } m, n \in \mathbb{N}.$$

Thus,

$$\|x_n + x_m\| \geq 2d.$$

By the parallelogram identity, it follows that

$$0 \leq \|x_n - x_m\|^2 = 2(\|x_n\|^2 + \|x_m\|^2) - \|x_n + x_m\|^2 \leq 2(\|x_n\|^2 + \|x_m\|^2) - 4d^2. \tag{7.2}$$

By definition of $(x_n)_{n \in \mathbb{N}} \subset K$, for each $\varepsilon > 0$, there exists some N_ε with

$$\|x_n\|^2 \leq d^2 + \tfrac{\varepsilon}{4} \quad \text{for all } n \geq N_\varepsilon.$$

Using (7.2), this implies

$$0 \leq \|x_n - x_m\|^2 \leq 2(2d^2 + \tfrac{\varepsilon}{2}) - 4d^2 = \varepsilon \quad \text{for all } n, m \geq N_\varepsilon,$$

showing $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Since $\mathcal{H}$ is a Hilbert space, we can set

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

Notice that $x \in K$, since K is closed. Moreover,

$$\|x\| = \lim_{n\to\infty} \|x_n\| = d.$$

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

$$d^2 \leq \left\|\frac{1}{2}(x+y)\right\|^2 = 2\left(\left\|\frac{x}{2}\right\|^2 + \left\|\frac{y}{2}\right\|^2\right) - \left\|\frac{1}{2}(x-y)\right\|^2 = \frac{d^2}{2} + \frac{d^2}{2} - \left\|\frac{x-y}{2}\right\|^2.$$

This implies $x = y$. ■

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*, as is custom. By Lemma 7.8, $M + M^\perp = M \oplus M^\perp$.

The following result now shows that each Hilbert space can be decomposed as $M \oplus M^\perp$. But notice that although $M^\perp$ is always a closed subspace, those properties have to be satisfied for M as well for this decomposition to hold true.

Proposition 7.10 *Let M be a closed subspace of the Hilbert space $\mathcal{H}$. Then*

$$\mathcal{H} = M \oplus M^\perp.$$

Proof. Let $x \in \mathcal{H}$. We aim to decompose x as $x = x_1 + x_2$ with x_1 being in M and $x_2 \in M^\perp$. First, we set

$$K := \{x - y : y \in M\} = x - M.$$

M is closed, hence so is K. Since, for all $y_1, y_2 \in M$, we have

$$\lambda(x - y_1) + (1-\lambda)(x - y_2) = x - (\lambda y_1 + (1-\lambda) y_2) \in K,$$

it follows that K is convex. By Lemma 7.9, there exists a unique $x_2 \in K$ such that

$$\|x_2\| = \inf_{y\in K} \|y\|.$$

By definition of K, $x_2 = x - x_1$ for some $x_1 \in M$. Hence, $x = x_1 + x_2$ with $x_1 \in M$.

It now remains to show that $x_2 \in M^\perp$. For this, let $y \in M \setminus \{0\}$ be arbitrary. Then

$$\|x_2\|^2 \leq \|\underbrace{x_2 - \lambda y}_{\in K - M = K}\|^2 \quad \text{for all } \lambda \in \mathbb{K}. \tag{7.3}$$

Now we choose $\lambda := \frac{\langle x_2, y\rangle}{\langle y, y\rangle}$. Inserting this in (7.3) yields

$$0 \le \left\| x_2 - \frac{\langle x_2, y\rangle}{\langle y, y\rangle} y \right\|^2 - \|x_2\|^2 = -\frac{|\langle x_2, y\rangle|^2}{\langle y, y\rangle} \le 0.$$

This implies $\langle x_2, y\rangle = 0$, hence $x_2 \in M^\perp$. ∎

7.3 Dual Spaces of Hilbert Spaces

We now turn to study dual spaces of Hilbert spaces. Remember that we already showed the relation $\ell_q \cong \ell_p^*$ for $1 \le p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Setting $p = 2$, we observe that in this situation $\ell_2 \cong \ell_2^*$. And in fact, ℓ_2 is a Hilbert space as we will see a bit later.

It is quite fascinating that this self-duality holds for any Hilbert space. Moreover, this relation is in addition given by a conveniently simple map $\mathcal{H} \to \mathcal{H}^*$, the so-called Riesz map. The famous Riesz Representation Theorem provides us now with the precise statement.

Theorem 7.11 (Riesz Representation Theorem) *Let $\mathcal{H}$ be a Hilbert space. For $y \in \mathcal{H}$, we define*

$$f_y \colon \mathcal{H} \to \mathbb{K}, \quad f_y(x) := \langle x, y\rangle.$$

Then $f_y \in \mathcal{H}^*$ and $\|f_y\| = \|y\|$ for all $y \in \mathcal{H}$.

Conversely, for each $f \in \mathcal{H}^*$, there exists a unique $y \in \mathcal{H}$ such that

$$f = f_y.$$

Finally, the Riesz map

$$y \mapsto f_y, \quad \mathcal{H} \to \mathcal{H}^*,$$

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 $\|f_y\| = \|y\|$. Moreover, we also proved $f_y \in \mathcal{H}^*$.

Now, let $f \in \mathcal{H}^*$, $f \neq 0$, be arbitrary. We aim to construct some $y \in \mathcal{H}$ such that $f(x) = \langle x, y\rangle$ for all $x \in \mathcal{H}$. For this, denote by N the kernel $\ker(f)$ of f, and let $z' \in \mathcal{H}$

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 $z \in N^{\perp}$ (Proposition 7.10). Then

$$1 = f(z') = f(x) + f(z) = f(z).$$

For arbitrary $x \in \mathcal{H}$, we have $x - f(x)z \in \ker(f)$, which implies

$$f(x) = \langle f(x)z, \|z\|^{-2}z\rangle = \langle (x - f(x)z) + f(x)z, \|z\|^{-2}z\rangle = \langle x, \|z\|^{-2}z\rangle = \langle x, y\rangle = f_y(x),$$

where $y := \|z\|^{-2}z$.

The claimed properties of the Riesz map follow from straightforward computations. ■

We saw before that some statements are also true, if we merely have an inner product space available. To some extent, this is also the case for the Riesz Representation Theorem.

Remark 7.12 Note that the first claim in Theorem 7.11 also holds, if $\mathcal{H}$ is only 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. ■

Ideally, we would like the dual of a Hilbert space to form a Hilbert space itself, such as the dual of a normed space is also a normed space. The next theorem ensures this structural property. Notice that the definition of the inner product is seemingly reversed, which is though easily explained by the conjugate linearity of the Riesz map, as we will see in the proof.

Theorem 7.13 *Let $\mathcal{H}$ be a Hilbert space. Moreover, for each $f \in \mathcal{H}^*$, 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. We need to prove inner product properties of $\langle \cdot, \cdot\rangle_{\mathcal{H}^*}$. For this, notice that, for each $f \neq 0$, we have $y_f \neq 0$, and thus

$$\langle f, f\rangle = \langle y_f, y_f\rangle > 0.$$

Further, for all $f, g \in \mathcal{H}$ and $\lambda \in \mathbb{K}$,

$$\langle f, g\rangle = \langle y_g, y_f\rangle = \overline{\langle y_f, y_g\rangle} = \overline{\langle g, f\rangle},$$

$$\langle f_1 + f_2, g\rangle = \langle y_g, y_{f_1+f_2}\rangle = \langle y_g, y_{f_1} + y_{f_2}\rangle = \langle y_g, y_{f_1}\rangle + \langle y_g, y_{f_2}\rangle = \langle f_1, g\rangle + \langle f_2, g\rangle,$$

and

$$\langle \lambda f, g\rangle = \langle y_g, y_{\lambda f}\rangle = \langle y_g, \overline{\lambda} y_f\rangle = \lambda\langle y_g, y_f\rangle = \lambda\langle f, g\rangle.$$

For $f \in \mathcal{H}^*$, we also 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, by definition of this inner product, the completeness of $\mathcal{H}^*$ follows from the completeness of $\mathcal{H}$. ∎

Discussing the structure of dual spaces naturally leads to the question of the structure of bi-duals in the sense of reflexivity. Applying the Riesz Representation Theorem twice shows that Hilbert spaces are indeed reflexive.

Corollary 7.14 *Hilbert spaces are reflexive.*

Proof. Let $\varphi \in \mathcal{H}^{**}$ be arbitrary. By Theorems 7.11 and 7.13, 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 and again exploiting both theorems,

$$\hat{y}_{f_\varphi}(f) = f(y_{f_\varphi}) = \langle y_{f_\varphi}, y_f\rangle = \langle f, f_\varphi\rangle = \varphi(f).$$

Thus $\hat{y}_{f_\varphi} = \varphi$. ∎

7.4 Applications of the Riesz Representation Theorem

We will now draw two quite different corollaries from the Riesz Representation Theorem. We are already familiar with the first, which is – except for the uniqueness – Theorem 4.6. However, interestingly, we can now provide a significantly shorter and maybe more insightful proof for the situation of Hilbert spaces.

Corollary 7.15 *Let $\mathcal{H}$ be a Hilbert space, let $L \subset \mathcal{H}$ be a linear subspace, and let $g \in L^*$. Then there exists a* unique *$f \in \mathcal{H}^*$ with*

$$f|_L = g \text{ and } \|f\| = \|g\|.$$

Proof. Apart from uniqueness this follows from the Theorem of Hahn-Banach, Theorem 4.6. We will now show a direct proof for Hilbert spaces.

First, define $\overline{g}\colon \overline{L} \to \mathbb{K}$ by continuously extending g, which implies $\|\overline{g}\| = \|g\|$. Thus, without loss of generality, L is closed, hence a Hilbert space.

By the Riesz Representation Theorem, there exists some $y \in L$ with

$$g(x) = \langle x, y \rangle \quad \text{for all } x \in L,$$

and $\|g\| = \|y\|$. Defining $f\colon \mathcal{H} \to \mathbb{K}, x \mapsto \langle x, y \rangle$, we obtain $f|_L = g$ and $\|f\| = \|y\| = \|g\|$.

To prove uniqueness, let $f' \in \mathcal{H}^*$ with $f'|_L = g$ and $\|f'\| = \|g\|$. By the Riesz Representation Theorem,

$$f'(x) = \langle x, y' \rangle \quad \text{for some } y' \in \mathcal{H} \text{ with } \|y'\| = \|f'\|.$$

Since $f'|_L = g$, it follows that $0 = \langle x, y' - y \rangle$ for all $x \in L$. This implies $y' - y \in L^\perp$ and thus $y' = y + z$ with $z \in L^\perp$. Thus, we obtain

$$\|f'\|^2 = \|y'\|^2 = \|y + z\|^2 = \|y\|^2 + \|z\|^2 = \|g\|^2 + \|z\|^2,$$

and hence $z = 0$. This implies $y = y'$ and in turn $f = f'$. ∎

The second corollary from the Riesz Representation Theorem is of tremendous importance, for instance, for the definition of so-called adjoint operators, which will follow later. Also sesquilinear forms of the type $\varphi(x, y) := \langle Ax, y \rangle$ appear in the discretization of partial differential equations.

Notice the almost equivalence of both parts of the statement.

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., it is linear in x and conjugate linear in y. Further,*

$$\sup_{x,y \neq 0} \frac{|\varphi(x, y)|}{\|x\| \|y\|} = \|A\|. \tag{7.4}$$

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 (7.4).

Proof. Given some operator $A \in L(\mathcal{H})$, the map φ defined by $\varphi(x, y) := \langle Ax, y \rangle$ is easily shown to be sesquilinear. Next, we have

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{x \neq 0} \frac{\|f_{Ax}\|}{\|x\|} = \sup_{x \neq 0} \|x\|^{-1} \sup_{y \neq 0} \frac{|\langle Ax, y \rangle|}{\|y\|} = \sup_{x,y \neq 0} \frac{|\langle Ax, y \rangle|}{\|x\| \|y\|} = \sup_{x,y \neq 0} \frac{|\varphi(x, y)|}{\|x\| \|y\|},$$

proving the first part of the claim.

Conversely, let φ be a sesquilinear form with $M := \sup_{x,y\neq 0} \frac{|\varphi(x,y)|}{\|x\|\|y\|} < \infty$. Then define

$$f_x(y) := \overline{\varphi(x,y)}.$$

Since f_x is linear and

$$|f_x(y)| = |\varphi(x,y)| \leq M\|x\|\|y\| \quad \text{for all } y \in \mathcal{H},$$

it follows that $f_x \in \mathcal{H}^*$. By the Riesz Representation Theorem, there exists a unique $z_x \in \mathcal{H}$ with $f_x(y) = \overline{\varphi(x,y)} = \langle y, z_x\rangle$ and $\|f_x\| = \|z_x\|$. Now define the operator

$$A\colon \mathcal{H} \to \mathcal{H}, \quad x \mapsto z_x.$$

We next show that A indeed satisfies the anticipated properties. For this, the map $x \mapsto f_x$ is conjugate linear as well as $f_x \mapsto z_x$, wherefore their composition $x \mapsto z_x$ is linear. Also,

$$\varphi(x,y) = \overline{f_x(y)} = \overline{\langle y, z_x\rangle} = \langle z_x, y\rangle = \langle Ax, y\rangle \quad \text{for all } x, y \in \mathcal{H},$$

which, in particular, implies

$$\|Ax\|^2 = \varphi(x, Ax) \leq M\|x\|\|Ax\| \quad \text{for all } x \in \mathcal{H}.$$

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

$$\langle (A-B)x, y\rangle = 0 \quad \text{for all } x, y \in \mathcal{H}.$$

In particular, with $y := (A-B)x$ we conclude that $(A-B)x = 0$ for all $x \in \mathcal{H}$, and therefore $A = B$. ■

We finish this chapter with a discussion of ℓ_2 as already announced. But at the same time, we will extend such a sequence space to arbitrary index sets, even uncountable ones.

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 common space ℓ_2 as defined in Chapter 1. In this case, the Riesz Representation Theorem coincides with Theorem 3.12 (stating that $\ell_p \cong \ell_q^*$) for $p = q = 2$. ◊

8 Orthogonality and Bases

In finite-dimensional spaces, orthonormal bases are a tremendously useful tool for decomposing or expanding vectors in a controlled manner, also leading to a notion of dimension. Since in Hilbert spaces we have an inner product at hand, we are now certainly aiming for introducing such concepts in this more general setting as well. As already observed before, such generalizations require particular attention.

8.1 Summability

A first obstacle we face is the problem that sets of vectors might have uncountable index sets. We still would like and also have to regard them as candidates for orthonormal systems or bases. Since decompositions or expansions will require summation over such sets, we have to first introduce a general notion of summability for an arbitrary set $\{x_i : i \in I\}$. Since this notion does not rely on an inner product structure, we will state its definition in the general setting of normed spaces.

Definition 8.1 Let E be a normed space, let $\{x_i : i \in I\} \subset E$, and let $x \in E$. 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

$$\left\| \sum_{i \in J} x_i - x \right\| \leq \varepsilon.$$

Obviously, $\{x_i : i \in I\}$ can only be summable to at most one $x \in E$, and we write

$$x = \sum_{i \in I} x_i.$$

Furthermore, we call $\{x_i : i \in I\}$ *summable*, if it is summable to some $x \in E$.

First, we need to check how the summability behaves with respect to the vector space operations. The following remark shows that the interaction is as good as one can hope for.

Remark 8.2 Let $\lambda \in \mathbb{K}$. If $\{x_i : i \in I\}$ and $\{y_i : i \in I\}$ are summable, so are $\{\lambda x_i : i \in I\}$ and $\{x_i + y_i : i \in I\}$. Moreover, we have

$$\sum_{i \in I} \lambda x_i = \lambda \sum_{i \in I} x_i$$

as well as

$$\sum_{i \in I} (x_i + y_i) = \sum_{i \in I} x_i + \sum_{i \in I} y_i.$$

$\Diamond$

We next provide two equivalent conditions for summability in Banach spaces, which are often more convenient to check. In particular, in (i) one should carefully compare the equivalent condition to the definition of summability, since the difference is marginal, but crucial.

Lemma 8.3 *Let E be a Banach space and $\{x_i : i \in I\} \subset E$.*

(i) $\{x_i : i \in I\}$ 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 $I_\varepsilon \cap J = \varnothing$, we have

$$\left\| \sum_{i \in J} x_i \right\| < \varepsilon.$$

(ii) $\{x_i : i \in I\}$ is summable to some $x \in E$ if and only if there exists a countable set $J \subset I$ with $x_i = 0$ for all $i \in I \setminus J$ 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 $x \in E$. Then, for every $\varepsilon > 0$, there exists some finite $I_\varepsilon \subset I$ satisfying

$$\left\| \sum_{i \in L} x_i - x \right\| < \frac{\varepsilon}{2} \quad \text{for all finite } L \supset I_\varepsilon.$$

Let now $J \subset I$ be finite with $J \cap I_\varepsilon = \varnothing$. Then we obtain

$$\left\| \sum_{i \in J} x_i \right\| \underset{J \cap I_\varepsilon = \varnothing}{=} \left\| \sum_{i \in J \cup I_\varepsilon} x_i - \sum_{i \in I_\varepsilon} x_i \right\| \le \left\| \sum_{i \in J \cup I_\varepsilon} x_i - x \right\| + \left\| \sum_{i \in I_\varepsilon} x_i - x \right\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

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, for each $n \in \mathbb{N}$, let $J_n \subset I$ be a finite set such that, for each finite $J \subset I$ with $J \cap J_n = \varnothing$, we have

$$\left\| \sum_{i \in J} x_i \right\| < \frac{1}{n}.$$

Next consider the sequence

$$(y_n)_{n \in \mathbb{N}} = \left(\sum_{i \in J_1 \cup \cdots \cup J_n} x_i \right)_{n \in \mathbb{N}}.$$

We observe that , for all $m \geq n$, we have

$$\|y_m - y_n\| = \left\| \sum_{i \in \bigcup_{k=1}^m J_k \setminus \bigcup_{k=1}^n J_k} x_i \right\| \leq \frac{1}{n},$$

since $\bigcup_{k=1}^m J_k \setminus \bigcup_{k=1}^n J_k$ is finite and disjoint from J_n. This implies that $(y_n)_{n \in \mathbb{N}}$ constitutes a Cauchy sequence. Let now $y = \lim_{n \to \infty} y_n$. Further, for each $\varepsilon > 0$, let $n \in \mathbb{N}$ be such that

$$\|y - y_n\| < \frac{\varepsilon}{2} \quad \text{and} \quad \left\| \sum_{i \in L} x_i \right\| < \frac{\varepsilon}{2}$$

for all finite sets $L \subset I$ with $L \cap J_n = \varnothing$. Note, that the choice $n > \frac{2}{\varepsilon}$ ensures the second condition. Setting $I_\varepsilon := \bigcup_{k=1}^n J_k$, we then obtain, for each finite $L \supset I_\varepsilon$, that

$$\left\| y - \sum_{i \in L} x_i \right\| \leq \left\| y - \sum_{i \in I_\varepsilon} x_i \right\| + \left\| \sum_{i \in L \setminus I_\varepsilon} x_i \right\| < \|y - y_n\| + \frac{\varepsilon}{2} < \varepsilon,$$

since $(L \setminus I_\varepsilon) \cap J_n = \varnothing$. This yields $y = \sum_{i \in I} x_i$.

(ii). Let $\{x_i : i \in I\}$ be summable to some $x \in E$. By (i), for each $n \in \mathbb{N}$, there exists a finite set $J_n \subset I$ with $\left\| \sum_{i \in J} x_i \right\| \leq \frac{1}{n}$ for all finite $J \subset I$ with $J \cap J_n = \varnothing$.
We next claim that the set $\{x_i : x_i \neq 0\}$ is contained in $\bigcup_{n \in \mathbb{N}} J_n$. To prove this, let $x_i \notin \bigcup_{n \in \mathbb{N}} J_n$ and take $J = \{i\}$. Then $J \cap J_n = \varnothing$ 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, $x_i = 0$, which implies that $\{x_i : x_i \neq 0\}$ is contained in a countable set, and hence is countable itself.
Let now $\mathbb{N} \to \bigcup_{n \in \mathbb{N}} J_n$, $k \mapsto i_k$ be a bijection (if $\bigcup_{n \in \mathbb{N}} J_n$ is infinite). Then we claim that

$$x = \lim_{n \to \infty} \sum_{k=1}^n x_{i_k}.$$

For this, for each $\varepsilon > 0$, 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 let $N \in \mathbb{N}$ be chosen such that

$$I_\varepsilon \cap \{i \in I : x_i \neq 0\} \subset \{i_k : k = 1, \dots, N\},$$

which is possible due to the bijectivity of $k \mapsto i_k$. Together with the injectivity, we can conclude that, for every $m \geq N$,

$$\left\| x - \sum_{k=1}^{m} x_{i_k} \right\| < \varepsilon.$$

The converse direction of statement (ii) is obvious if $J = \{i \in I : x_i \neq 0\}$ is finite. If J is countable, yet infinite, let us assume that the series $\sum_{j \in J} x_j$ converges to some $x \in E$, independent of the order of summation. Yet, towards a contradiction, we also assume that $\{x_i : i \in I\}$ is not summable to $x \in E$. Then there exists some $\varepsilon > 0$ such that, for every finite $I' \subset I$, there exists a finite set $I'' \supseteq I'$ with

$$\left\| x - \sum_{i \in I''} x_i \right\| \geq \varepsilon.$$

Let us fix a bijection $\mathbb{N} \to J$, $l \to j_l$. Then set $I_1' := \{j_1\}$ and let $I_1'' \subset J$ be some superset of I_1', say $I_1'' = \{i_1, \dots i_{n_1}\}$ with pairwise unequal indices $i_1, \dots, i_{n_1}$ contained in J, satisfying

$$\left\| x - \sum_{k=1}^{n_1} x_{i_k} \right\| \geq \varepsilon.$$

Then set $I_2' := I_1'' \cup \{j_r\}$, where $r \in \mathbb{N}$ is minimally chosen with $j_r \notin I_1''$. Furthermore, let $I_2'' := I_1'' \cup \{i_{n_1+1}, \dots, i_{n_2}\}$ be a superset of I_2', again contained in J, with

$$\left\| x - \sum_{k=1}^{n_2} x_{i_k} \right\| \geq \varepsilon.$$

Continuing this process, thereby never choosing some $i \in J$ twice, we obtain by $k \mapsto i_k$ another bijection $\mathbb{N} \to J$. Notice that $k \mapsto i_k$ is surjective due to $j_k \in I_k''$. Finally, for all $m \in \mathbb{N}$,

$$\left\| x - \sum_{k=1}^{n_m} x_{i_k} \right\| \geq \varepsilon,$$

thus $\sum_{k=1}^{n} x_{i_k} \nrightarrow x$, a contradiction. ∎

This lemma allows to derive a very convenient condition for the summability of nonnegative numbers, in terms of boundedness of one set.

Corollary 8.4 *Let $a_i \geq 0$ for $i \in I$. 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 the implication $S < \infty$ 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 $J \supset I_\varepsilon$, yielding that $\{a_i : i \in I\}$ is summable to S. ■

Our focus before was on Hilbert spaces, hence we now need to study how summability behaves with respect to an inner product. The next remark shows that again the interaction is as expected most convenient.

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 $y \in \mathcal{H}$,

$$\left\langle \sum_{i \in I} x_i, y \right\rangle = \sum_{i \in I} \langle x_i, y \rangle .$$

This can be shown as follows: Let $J \subset I$ be finite and $\|x - \sum_{j \in J} x_j\| < \varepsilon$. Then, for every $y \in \mathcal{H}$, we have

$$\left| \langle x, y \rangle - \sum_{j \in J} \langle x_j, y \rangle \right| \leq \|y\| \left\| x - \sum_{j \in J} x_j \right\| \leq \|y\| \varepsilon.$$

This proves the claim. ◊

The last result on summability studies a family of pairwise orthogonal vectors. For them, maybe not completely unexpected, summability in the Hilbert space is equivalent to summability of their norm squares in $\mathbb{R}$. The second condition is certainly quite convenient to check, but notice that this requires pairwise orthogonality.

Lemma 8.6 *Let $\mathcal{H}$ be a Hilbert space, and let $\{x_i : i \in I\}$ be a family of pairwise orthogonal elements in $\mathcal{H}$. Then the following conditions are equivalent.*

(i) $\{x_i : i \in I\}$ *is summable in $\mathcal{H}$.*

(ii) $\{\|x_i\|^2 : i \in I\}$ *is summable in $\mathbb{R}$.*

Moreover, we have $\|\sum_{i\in I} x_i\|^2 = \sum_{i\in I}\|x_i\|^2$. In particular, summability is not equivalent to absolute convergence in infinite dimensions.

Proof. (i)⇒(ii). For an arbitrary $\varepsilon > 0$, let $I_\varepsilon \subseteq I$ be a finite subset, which satisfies

$$\left\|\sum_{j\in J} x_j\right\| < \varepsilon \quad \text{for all finite } J \subset I \text{ with } J \cap I_\varepsilon = \varnothing.$$

Since by "Pythagoras", for $x, y \in \mathcal{H}$ with $\langle x, y\rangle = 0$, we have

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

it follows that

$$\sum_{j\in J}\|x_j\|^2 = \left\|\sum_{j\in J} x_j\right\|^2 < \varepsilon^2$$

for all finite $J \subset I$ with $J \cap I_\varepsilon = \varnothing$. This implies (ii).

(ii)⇒(i). If $\left\{\|x_i\|^2 : i \in I\right\}$ is summable in $\mathbb{R}$, then, for all $\varepsilon > 0$, there exists some finite subset $I_\varepsilon \subset I$, such that, for every finite $J \subset I$ with $J \cap I_\varepsilon = \varnothing$, we have

$$\left\|\sum_{j\in J} x_j\right\|^2 = \sum_{j\in J}\|x_j\|^2 < \varepsilon.$$

Thus, $\{x_i : i \in I\}$ is summable in $\mathcal{H}$. ∎

8.2 Orthonormal Basis and Dimension

Finally, we arrive at the definition of an orthonormal basis for an inner product space; notice that this definition certainly does not require any completeness properties. The definition of an orthonormal system in infinite dimensions is still quite straightforward, whereas the definition of a basis might be worth thinking about.

Definition 8.7 Let $\mathcal{H}$ be an inner product space, and let $M \subset \mathcal{H}$. Then the set M is called an *orthonormal system* (ONS), if

$$\langle x, y\rangle = \delta_{x,y} \quad \text{for all } x, y \in M.$$

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 $M = N$.

Although the definition for a basis might take some time getting used to, the well-known equivalent condition in finite dimensions of the complement being $\{0\}$ still holds.

Lemma 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 orthonormal system with $M \subsetneq N$. The completeness now yields $y = 0$.

Conversely, assume that $M^{\perp} = \{0\}$, and let $N \supset M$ be an orthonormal system. Further, suppose that there exists some $y \in N \setminus M$. Then $\langle y, x\rangle = 0$ for all $x \in M$, which implies $y = 0$. But $\|y\| = 1$, since N forms an orthonormal system. Thus $M = N$. ■

The next theorem states two famous and highly useful (in)equalities for orthonormal systems $(x_i)_{i\in I}$ in terms of the ℓ_2 norm of the sequence of inner products $(\langle x, x_i\rangle)$ for some $x \in \mathcal{H}$. We remark that those inner products will later (Theorem 8.13), in the situation of an orthonormal basis, be shown to be the expansion coefficients of said x. In this sense, part (ii) can be interpreted as showing that the Hilbert space norm of a vector equals the ℓ_2 norm of its expansion coefficients.

Theorem 8.9 *Let $\mathcal{H}$ be an inner product space, and let $(x_i)_{i\in I}$ be an orthonormal system.*

(i) Bessel inequality. *For all $x \in \mathcal{H}$,*

$$\sum_{i\in I} |\langle x, x_i\rangle|^2 \leq \|x\|^2.$$

(ii) Parseval identity. *For all $x \in \mathcal{H}$,*

$$\sum_{i\in I} |\langle x, x_i\rangle|^2 = \|x\|^2 \quad \Longleftrightarrow \quad x = \sum_{i\in I} \langle x, x_i\rangle\, x_i.$$

Proof. (i). For all finite $J \subset I$, we have

$$\begin{aligned}
0 &\leq \left\| x - \sum_{i\in J} \langle x, x_i\rangle x_i \right\|^2 \\
&= \|x\|^2 - \sum_{i\in J} \overline{\langle x, x_i\rangle}\langle x, x_i\rangle - \sum_{i\in J}\langle x, x_i\rangle\overline{\langle x, x_i\rangle} + \sum_{i,j\in J} \langle x, x_i\rangle\overline{\langle x, x_j\rangle}\langle x_i, x_j\rangle \\
&= \|x\|^2 - \sum_{i\in J} |\langle x, x_i\rangle|^2.
\end{aligned}$$

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). Let $x \in \mathcal{H}$. First observe that for finite $J \subset I$ it holds

$$\left\| x - \sum_{i\in J} \langle x, x_i\rangle x_i \right\|^2 = \|x\|^2 - \sum_{i\in J} |\langle x, x_i\rangle|^2.$$

The relation

$$\|x\|^2 = \sum_{i\in I} |\langle x, x_i\rangle|^2$$

is equivalent to the fact that, for all $\varepsilon > 0$ there exists a finite $I_\varepsilon \subset I$ such that for all finite $J \supset I_\varepsilon$ we have

$$\|x\|^2 - \sum_{i\in J} |\langle x, x_i\rangle|^2 \leq \varepsilon.$$

This in turn is equivalent to the statement that, for all $\varepsilon > 0$ there is a finite $I_\varepsilon \subset I$ such that for all finite $J \supset I_\varepsilon$,

$$\left\| x - \sum_{i\in J} \langle x, x_i\rangle x_i \right\|^2 \leq \varepsilon.$$

Finally, this is by definition equivalent to the summability to x of $\{\langle x, x_i\rangle x_i : i \in I\}$. ∎

Since orthonormal bases can be regarded as a means to decompose a vector in a Hilbert space, one might also want to consider a general decomposition of Hilbert spaces. A natural approach to this problem is the concept of an orthogonal sum.

Definition 8.10 The *orthogonal sum* $\mathcal{H}$ of Hilbert spaces $\mathcal{H}_i$, $i \in I$, denoted by

$$\mathcal{H} = \bigoplus_{i\in I} \mathcal{H}_i,$$

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).

The newly defined object of an orthogonal sum of Hilbert spaces is in fact again an inner product space as the following remark shows.

Remark 8.11 $\mathcal{H} = \oplus_{i\in I} H_i$ is a linear space, since, for each $(x_i)_{i\in I}, (y_i)_{i\in I} \in \mathcal{H}$,

$$\begin{aligned}
\sum_{i\in I} \|x_i + y_i\|^2 &\leq \sum_{i\in I} \|x_i\|^2 + 2\sum_{i\in I} |\langle x_i, y_i\rangle| + \sum_{i\in I} \|y_i\|^2 \\
&\leq \sum_{i\in I} \|x_i\|^2 + 2\sum_{i\in I} \|x_i\|\|y_i\| + \sum_{i\in I} \|y_i\|^2 \\
&\leq \sum_{i\in I} \|x_i\|^2 + 2\left(\sum_{i\in I} \|x_i\|^2\right)^{1/2} \left(\sum_{i\in I} \|y_i\|^2\right)^{1/2} + \sum_{i\in I} \|y_i\|^2.
\end{aligned}$$

Moreover, we can define an inner product $\langle \cdot, \cdot \rangle$ on $\mathcal{H}$ by

$$\langle (x_i)_i, (y_i)_i \rangle = \sum_{i\in I} \langle x_i, y_i\rangle.$$

It is straightforward to check that this is indeed an inner product. ◊

After these preparations, it doesn't come as a surprise that an orthogonal sum of Hilbert spaces is again a Hilbert space.

Lemma 8.12 *Let $\mathcal{H}_i$, $i \in I$, be Hilbert spaces.*

(i) $\mathcal{H} = \oplus_{i\in I} \mathcal{H}_i$ is a Hilbert space.

(ii) We can interpret each $\mathcal{H}_i$ as a subspace of $\mathcal{H}$.

Proof. (i). By Remark 8.11, it remains to show that $\mathcal{H} = \oplus_{i\in I} \mathcal{H}_i$ is complete. For this, 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 $N_\varepsilon \in \mathbb{N}$ such that

$$\|x_n - x_m\|^2 = \|(x_{n,i})_i - (x_{m,i})_i\|^2 \leq \varepsilon \quad \text{for all } n, m \geq N_\varepsilon.$$

This implies that, for all $j \in I$, we have

$$\|x_{n,j} - x_{m,j}\|^2 \leq \sum_{i\in I} \|x_{n,i} - x_{m,i}\|^2 = \|(x_{n,i})_i - (x_{m,i})_i\|^2 \leq \varepsilon. \tag{8.1}$$

Hence $(x_{n,i})_n$ is a Cauchy-sequence in $\mathcal{H}_i$, which is complete. Now, let

$$x_i := \lim_{n\to\infty} x_{n,i}.$$

By (8.1), for each finite $J \subset I$ and $n \geq N_\varepsilon$, we obtain

$$\sum_{i\in J} \|x_{n,i} - x_i\|^2 = \lim_{m\to\infty} \sum_{i\in J} \|x_{n,i} - x_{m,i}\|^2 \leq \varepsilon,$$

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

$$\sum_{i\in I} \|x_{n,i} - x_i\|^2 \leq \varepsilon \quad \text{for all } n \geq N_\varepsilon.$$

This finally implies that $x := (x_i)_i \in \mathcal{H}$ 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, yielding the claim. ∎

We finally arrive at the main theorem of this section, which provides us with several equivalent conditions for an orthonormal system to be complete. It might be interesting to compare part (iii) with the previous result and part (iv) with Theorem 8.9(ii).

Theorem 8.13 *For an orthonormal system $\{x_i : i \in I\}$ in a Hilbert space $\mathcal{H}$, the following conditions are equivalent.*

(i) $\{x_i : i \in I\}$ *is complete.*

(ii) $\operatorname{span}\{x_i : i \in I\}$ *is dense in* $\mathcal{H}$.

(iii) If $\mathcal{H}_i := \operatorname{span} x_i$, $i \in I$, *then* $\oplus_{i\in I} \mathcal{H}_i$ *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 Proposition 7.10,

$$\mathcal{H} = \overline{L} \oplus \overline{L}^{\perp}.$$

But $\overline{L}^{\perp} \subset \{x_i : i \in I\}^{\perp} = \{0\}$ by Lemma 8.8. This then implies $\overline{L} = \mathcal{H}$.

(ii)⇒(iii). Let F denote the linear subspace of all $(\lambda_i x_i)_{i\in I} \in \oplus_{i\in I} \mathcal{H}_i$ with $\lambda_i \neq 0$ for an at most finite number of indices $i \in I$. Since

$$\left\| \sum_{i\in I} \lambda_i x_i \right\|^2 = \sum_{i,j\in I} \lambda_i \overline{\lambda}_j \langle x_i, x_j\rangle = \sum_{i\in I} |\lambda_i|^2 = \|(\lambda_i x_i)_{i\in I}\|^2,$$

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. By definition of F, it is dense in $\oplus_{i\in I} \mathcal{H}_i$. Hence 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 as well. By (ii), also L is dense in $\mathcal{H}$ and $\varphi(\oplus_{i\in I} \mathcal{H}_i) \subset \mathcal{H}$ is complete, hence closed. This yields $\varphi(\oplus_{i\in I} \mathcal{H}_i) = \mathcal{H}$.
It remains to prove that, for all $x = (\lambda_i x_i)_i \in \oplus_{i\in I} \mathcal{H}_i$, 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 an arbitrary finite subset $J \supset I_\varepsilon$, $J \subset I$, define

$$y_i = \begin{cases} \lambda_i x_i & : i \in J, \\ 0 & : \text{otherwise.} \end{cases}$$

With setting $y = (y_i)_{i \in I}$, we obtain

$$\|x - y\|^2 = \sum_{i \notin J} |\lambda_i|^2 < \varepsilon,$$

which implies that

$$\varepsilon > \|\varphi(x) - \varphi(y)\|^2 = \left\| \varphi(x) - \sum_{i \in J} \lambda_i x_i \right\|^2.$$

(iii)$\Rightarrow$(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)$\Rightarrow$(v). By (iv),

$$\langle x, y \rangle = \left\langle \sum_{i \in I} \langle x, x_i \rangle x_i, y \right\rangle = \sum_{i \in I} \langle x, x_i \rangle \langle x_i, y \rangle.$$

(v)$\Rightarrow$(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 Lemma 8.8. ■

For finite-dimensional spaces, it is easy to show that an orthonormal basis always exist, and that each orthonormal basis for the same space has the same cardinality. The same is true for arbitrary – in particular, infinite-dimensional – Hilbert spaces. However, as we will see, this result again requires the usage of Zorn's lemma.
Also, notice that the theorem is stated for orthogonal bases. But by normalization, the same is true for orthonormal bases.

Theorem 8.14 *Every Hilbert space $\mathcal{H} \neq \{0\}$ possesses an orthogonal basis, and all orthogonal bases have the same cardinality.*

Proof. We prove the statement for orthonormal bases. Let γ be the family of all orthonormal systems in $\mathcal{H}$. Observe that $\gamma \neq \varnothing$, since there exists some $x \in \mathcal{H}$, $x \neq 0$, with $\|x\| = 1$. Next we order γ by inclusion. To be able to apply Zorn's lemma, we need to show that each chain possesses a maximal element. For this, 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 we can assume that $S_1 \subset S_2$, hence $x_1, x_2 \in S_2$, yielding

$$\langle x_1, x_2 \rangle = 0.$$

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. This finishes the first claim.

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 some $x_y \in B$ with

$$\langle x_y, y \rangle \neq 0.$$

By definition we then have $y \in C_{x_y}$. 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

$$|C| \leq |M| \leq |B| \cdot \aleph_0 = |B|.[1]$$

By symmetry, the cardinalities coincide. ∎

After having established that the cardinality of all orthonormal bases for a Hilbert space coincide, we can define this cardinality as its dimension.

[1] $\aleph_0$ is the cardinality of $\mathbb{N}$.

Definition 8.15 Let $\mathcal{H}$ be a Hilbert space over $\mathbb{K}$. Then the *(Hilbert) dimension* of $\mathcal{H}$ ($\dim \mathcal{H}$) is the cardinality of some (and hence of each) orthonormal basis of $\mathcal{H}$.

We already discussed before the importance of two normed spaces being isometrically isomorphic. Now in the situation of Hilbert spaces, we are able to show that this is in fact equivalent to the dimensions being equal.

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) $\dim \mathcal{H}_1 = \dim \mathcal{H}_2$.

(ii) $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *are isometrically isomorphic.*

Proof. (i)⇒(ii). Let $\{x_i \colon i \in I\}$ and $\{y_i \colon i \in I\}$ be orthonormal bases of $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. Then the map

$$\varphi\colon \mathcal{H}_1 \to \mathcal{H}_2, \quad \sum_{i\in I} \lambda_i x_i \mapsto \sum_{i \in I} \lambda_i y_i, \ (\lambda_i)_i \in \ell_2(I),$$

satisfies 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 orthonormal system 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. ■

This immediately implies that intriguingly each Hilbert space is isometrically isomorphic to $\ell_2(I)$ with I being the index set of one (and hence all) of its orthonormal bases. It is worthwhile to consider this in the finite-dimensional situation.

Corollary 8.17 *Let $\mathcal{H}$ be a Hilbert space, and let I be an index set of an orthonormal basis of $\mathcal{H}$. Then $\mathcal{H}$ is isometrically isomorphic to $\ell^2(I)$.*

Proof. Notice that the unit vectors $e_j = (\delta_{ij})_{i \in I}$ form an orthonormal basis for $\ell^2(I)$. The claim now follows by applying Theorem 8.16. ■

We discussed extensively orthogonal systems, namely which properties they possess and when they constitute a basis. It is quite natural to ask how to actually construct orthogonal systems. The Schmidt Orthogonalization Method is a particularly beautiful way to derive an orthogonal system from a linearly independent set.

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*

$$\begin{aligned} y_1 &:= \|x_1\|^{-1} x_1, \\ y_{n+1} &:= \left\| x_{n+1} - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle y_i \right\|^{-1} \cdot \left(x_{n+1} - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle y_i \right), \; n \geq 1. \end{aligned}$$

Then $\{y_1, y_2, \dots\}$ is an orthonormal system 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 by induction that $\{y_1, \dots, y_n\}$ is an orthonormal system and

$$\operatorname{span}\{y_1, \dots, y_n\} = \operatorname{span}\{x_1, \dots, x_n\}.$$

For $n = 1$, there is nothing to prove.
Now assume that the claim holds for n. 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, we have

$$x_{n+1} - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle y_i \neq 0.$$

By construction, $\|y_{n+1}\| = 1$. Moreover, for all $k = 1, \dots, n$, we observe that

$$\begin{aligned} \langle y_{n+1}, y_k \rangle &= \left\| x_{n+1} - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle y_i \right\|^{-1} \cdot \left(\langle x_{n+1} - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle y_i, y_k \rangle \right) \\ &= \left\| x_{n+1} - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle y_i \right\|^{-1} \cdot \left(\langle x_{n+1}, y_k \rangle - \sum_{i=1}^{n} \langle x_{n+1}, y_i \rangle \underbrace{\langle y_i, y_k \rangle}_{=\delta_{ik}} \right) \\ &= 0. \end{aligned}$$

By induction hypothesis, $\operatorname{span}\{x_1, \dots, x_n\} = \operatorname{span}\{y_1, \dots, y_n\}$. Using the definition of y_{n+1}, it then follows that also $\operatorname{span}\{x_1, \dots, x_{n+1}\} = \operatorname{span}\{y_1, \dots, y_{n+1}\}$. The theorem is proved. ∎

Countability is among the infinite cardinalities certainly the most desirable. Hence one might ask when an infinite-dimensional Hilbert space possesses a countable orthonormal basis. Our next result provides a more or less conveniently checkable condition for this situation.

Theorem 8.19 *Let $\mathcal{H}$ be a Hilbert space. Then the following conditions are equivalent.*

(i) $\mathcal{H}$ possesses an at most countable orthonormal basis.

(ii) $\mathcal{H}$ 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

$$\operatorname{span}\{x_1, x_2, \dots\} = \operatorname{span}\{x_{n_1}, x_{n_2}, \dots\} =: L.$$

Application of the Schmidt Orthogonalization Method (Theorem 8.18) generates an orthonormal system $\{y_1, y_2 \dots\}$ with

$$\operatorname{span}\{y_1, y_2, \dots\} = L.$$

Since $D \subset L$, also L is dense in $\mathcal{H}$. Finally, by Theorem 8.13, the system $\{y_1, y_2, \dots\}$ is an orthonormal basis of $\mathcal{H}$.

(i)⇒(ii). Let $\{y_1, y_2, \dots\}$ be a countable orthonormal basis of $\mathcal{H}$. Now we set $M = \mathbb{Q}$, if $\mathbb{K} = \mathbb{R}$, and $M = \mathbb{Q} + i \cdot \mathbb{Q}$, if $\mathbb{K} = \mathbb{C}$. Then we define

$$D := \{\sum_{i=1}^{n} \lambda_i y_i \colon \lambda_i \in M, n \in \mathbb{N}\}$$

as a candidate for a countable dense subset of $\mathcal{H}$. Since M is countable, also D is countable. Next, let $x \in \mathcal{H}$ and $\varepsilon > 0$. By Theorem 8.13, there exist $n \in \mathbb{N}$ and $\mu_i \in \mathbb{K}$ satisfying that

$$\left\| x - \sum_{i=1}^{n} \mu_i y_i \right\| \le \tfrac{\varepsilon}{2}.$$

By definition of M, for each μ_i, we can find some $\lambda_i \in M$ with

$$|\lambda_i - \mu_i| \le \tfrac{\varepsilon}{2n}.$$

This implies

$$\left\| x - \sum_{i=1}^{n} \lambda_i y_i \right\| \le \left\| x - \sum_{i=1}^{n} \mu_i y_i \right\| + \left\| \sum_{i=1}^{n} (\mu_i - \lambda_i) y_i \right\| \le \varepsilon,$$

showing that D is dense in $\mathcal{H}$. ■

9 Compact Operators

This chapter is devoted to a particularly nice class of operators, so-called compact operators. In a certain sense, which we will make more precise later on, those operators are close to the finite-dimensional situation. Roughly speaking, they can be approximated by finite-dimensional operators and also inherit useful properties of linear operators in finite-dimensional Hilbert spaces. The famous and very deep spectral theorem, which provides a meaningful decomposition of operators, will also be proven at the end of this course first in the situation of compact operators.

9.1 Basic Properties

We start with the definition of compact operators, which also shows why they are called 'compact'.

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 $\mathcal{K}(E, E)$.

Certainly, one is interested in equivalent conditions for compactness of operators, which do not require to check compactness of the closure of the image under the operator for *all* bounded sets. And indeed – among other statements – the following result shows that it is sufficient to just consider the unit closed ball. It should however be mentioned that, despite the elegance of condition (b), it will be condition (c), which we will primarily exploit in the sequel.

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 $T \in \mathcal{K}(E,F)$, which implies that $\overline{T(K_1(0))}$ is compact. Thus, $T(K_1(0)) = \{Tx : \|x\| \leq 1\}$ is bounded. This in turn implies that $\sup\{\|Tx\| : \|x\| \leq 1\} < \infty$, 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 some $r > 0$ such that $B \subset rK_1(0)$. Since

$$\overline{T(B)} \subset \overline{T(rK_1(0))} = r\overline{T(K_1(0))},$$

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 1.8. This implies (a)⇔(c). ∎

The first part of the next result shows that $\mathcal{K}(E,F)$ is not only contained in $L(E,F)$, but is in fact a linear subspace. The second part shows that it is also closed under composition of operators.

Lemma 9.3 *Let E, F, and G be normed spaces.*

(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 $S,T \in \mathcal{K}(E,F)$ and $\alpha, \beta \in \mathbb{K}$, and let $(x_n)_n \subset E$ be bounded. Then $(Sx_n)_n$ contains a convergent subsequence $(Sx_{n_k})_k$. Since $(x_{n_k})_k$ is bounded, also $(Tx_{n_k})_k$ contains a convergent subsequence $(Tx_{n_{k_j}})_j$, implying that $(\alpha Sx_{n_{k_j}} + \beta Tx_{n_{k_j}})_j$ converges. By Lemma 9.2(ii), $\alpha S + \beta T$ is compact.

(ii). Let $(x_n)_n \in E$ be bounded. If T is compact, then $(Tx_n)_n$ contains a convergent subsequence $(Tx_{n_k})_k$. Thus $(STx_{n_k})_k$ converges. If S is compact, then $(STx_n)_n$ contains a convergent subsequence, since $(Tx_n)_n$ is bounded. Again applying Lemma 9.2(ii) proves the claim. ∎

Under certain conditions – and we will see how those are used in the proof – the space of compact operators is even closed in $L(E,F)$.

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 $T \in \overline{\mathcal{K}(E,F)}$, and let $\varepsilon > 0$. Then there exists some $S \in \mathcal{K}(E,F)$ with

$$\|S - T\| < \tfrac{\varepsilon}{3}.$$

Since S is compact, also $\overline{S(K_1(0))}$ is compact, and hence $\overline{S(K_1(0))}$ is totally bounded. By Lemma 1.9, also $S(K_1(0))$ is totally bounded. By definition of totally boundedness, there exist $x_1, \ldots, x_r \in K_1(0)$ such that

$$S(K_1(0)) \subset \bigcup_{j=1}^{r} K_{\varepsilon/3}(Sx_j).$$

Thus if $x \in K_1(0)$, then $Sx \in K_{\varepsilon/3}(Sx_j)$ for some j. This implies that

$$\|Tx - Tx_j\| \leq \|T - S\|\|x\| + \|Sx - Sx_j\| + \|S - T\|\|x_j\| < \varepsilon.$$

This in turn implies that $Tx \in K_\varepsilon(Tx_j)$ and

$$T(K_1(0)) \subset \bigcup_{j=1}^{r} K_\varepsilon(Tx_j).$$

Hence $T(K_1(0))$ is totally bounded, and by Corollary 1.10 its closure is compact. ∎

Of particular interest later on will be the relation between compactness of an operator and its dual operator. We obtain almost equivalence as the next result shows.

Theorem 9.5 *Let E, F be normed spaces, and let $T \in L(E, F)$. Also let $T^*\colon F^* \to E^*$, $(T^*f)(x) = f(Tx)$ be the dual operator.*

(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 $(f_n)_n \subset F^*$ be bounded. By Lemma 9.2, it is sufficient to prove that $(T^*f_n)_n$ contains a convergent subsequence. Set $Y := \overline{T(K_1(0))}$, which is compact in F, and consider

$$\mathcal{F} := \{f_n|_Y : n \in \mathbb{N}\} \subseteq C(Y).$$

With $C := \sup_n \|f_n\|$, we first observe that

$$|f_n(y)| \leq C\|y\|,$$

hence $\mathcal{F}$ is pointwise bounded, and that

$$|f_n(y_1) - f_n(y_2)| \leq C\|y_1 - y_2\|,$$

i.e., $\mathcal{F}$ is equicontinous. Then Arzelà-Ascoli implies that $\overline{\mathcal{F}} \subseteq (C(Y), \|\cdot\|_\infty)$ is compact. Thus there exists a convergent subsequence $(f_{n_k}|_Y)_k$, implying that, for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ with

$$\|f_{n_k}|_Y - f_{n_l}|_Y\|_\infty < \varepsilon \quad \text{for all } k, l \geq N.$$

This yields

$$\|T^*f_{n_k} - T^*f_{n_l}\| = \sup_{\|x\|=1} |f_{n_k}(Tx) - f_{n_l}(Tx)| < \varepsilon \quad \text{for all } k, l \geq N,$$

and thus $(T^*f_{n_k})_k$ is a Cauchy-sequence in E^*, hence converges.

(ii). Since T^* is compact, by (i), also T^{**} is compact. It is straightforward to prove 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, it follows that T is compact. ∎

We finish this first part with several examples of compact operators. It is recommended to pay particular attention to (3) concerning finite-dimensional operators.

Example 9.6 (1) The identity operator $\mathrm{I}\colon E \to E$, $x \mapsto x$ is compact if and only if E is finite-dimensional (by Theorem 2.10).

(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.
To show this, one can argue as follows: Let $(x_n)_n \subset E$ 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). In Theorem 9.9, we will however 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.
For the proof, we will utilize Arzelà-Ascoli, which implies that it is sufficient to prove that for each bounded $B \subset E$, $K(B)$ 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 each $\varepsilon > 0$ there exists some $\delta > 0$ such that

$$t \in [a, b], |s_1 - s_2| < \delta \quad \Longrightarrow \quad |k(s_1, t) - k(s_2, t)| < \varepsilon.$$

Hence, if $|s_1 - s_2| < \delta$, we obtain

$$|(Kf)(s_1) - (Kf)(s_2)| \leq \int_a^b |k(s_1, t) - k(s_2, t)||f(t)|\mathrm{d}t \leq (b-a)\varepsilon \|f\|_\infty.$$

This implies that $K(B)$ is equicontinuous, due to the boundedness of B.

$\Diamond$

9.2 Closeness to Finite-Dimensional Operators

As argued before and made already a bit more precise in Example 9.6(3), compact operators are quite close to finite-dimensional operators. To delve deeper into this topic, we first require the method of orthogonal projections, which we now have at hand in Hilbert spaces.

Definition 9.7 Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{L} \subset \mathcal{H}$ be a closed subspace. Since $\mathcal{H} = \mathcal{L} \oplus \mathcal{L}^\perp$, for each $x = u + v$ with $u \in \mathcal{L}$, $v \in \mathcal{L}^\perp$, we can define

$$P_\mathcal{L} x := u.$$

The operator $P_\mathcal{L} \colon \mathcal{H} \to \mathcal{H}$ is then called the *orthogonal projection onto* $\mathcal{L}$. More generally, a linear operator $P \colon \mathcal{H} \to \mathcal{H}$ with $P^2 = P$ is called a *projection*.

Before stating our main theorem of this section, we need some more properties of orthogonal projections.

Lemma 9.8 *Let $\mathcal{H}$ be a Hilbert space, let $\mathcal{L} \subset \mathcal{H}$ be a closed subspace, and let $P_\mathcal{L} \colon \mathcal{H} \to \mathcal{H}$ be the orthogonal projection onto $\mathcal{L}$. Then $P_\mathcal{L}$ is well-defined, linear and bounded. Moreover, we have* $\ker P_\mathcal{L} = \mathcal{L}^\perp$, $P_\mathcal{L}^2 = P_\mathcal{L}$ *and* $\|P_\mathcal{L}\| = 1$ *if* $\mathcal{L} \neq \{0\}$.

Proof. The claims follow directly from the definition of an orthogonal projection. ∎

Finally, we are in the position to formulate and prove the result about closeness of compact operators to finite-dimensional ones. Notice that we indeed require a Hilbert space, since the proof relies on orthogonal projections.

Theorem 9.9 *Let E be a normed space, let $\mathcal{H}$ be a Hilbert space, and let $T \in \mathcal{K}(E, \mathcal{H})$. Then there exist finite-dimensional operators $T_n \in L(E, \mathcal{H})$ with $\|T - T_n\| \to 0$.*

Proof. Since T is compact, the set $\overline{T(K_1(0))}$ is compact, and hence $T(K_1(0))$ is totally bounded. Thus, for each $n \in \mathbb{N}$ there exist $x_{n,1}, \ldots, x_{n,r_n}$ in $K_1(0)$ such that

$$T(K_1(0)) \subset \bigcap_{i=1}^{r_n} K_{1/n}(Tx_{n,i}).$$

Now set

$$\mathcal{L}_n := \operatorname{span}\{Tx_{n,i} : i = 1, \ldots, r_n\} \quad \text{and} \quad P_n := P_{\mathcal{L}_n}.$$

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 $x \in K_1(0)$, we have $Tx \in K_{1/n}(Tx_{n,i})$ for some i. Thus, for $x \in K_1(0)$, it follows that

$$\begin{aligned} \|Tx - P_n Tx\| &\leq \|Tx - Tx_{n,i}\| + \|Tx_{n,i} - P_n Tx_{n,i}\| + \|P_n Tx_{n,i} - P_n Tx\| \\ &\leq \|Tx - Tx_{n,i}\| + \|Tx_{n,i} - Tx\| \leq \tfrac{2}{n}. \end{aligned}$$

This implies that $\|T - P_n T\| \to 0$ as $n \to \infty$, i.e., the finite-dimensional operators $P_n T$ converge to T. ∎

9.3 Fredholm Operators

Fredholm operators are a particular type of operators which satisfy certain finite-dimensionality conditions on their kernel and range. These operators will turn out to be very useful, for instance, for analyzing compact operators.

To prepare Theorem 9.13, which is one of the two key results in this part, we start by introducing the notion of a complementary subspace.

Definition 9.10 Let E be a linear space, and let F, G be linear subspaces such that $E = F \dot{+} G$ (direct sum). Then G is called a *complementary subspace* to F in E.

Their relation to projections is discussed in the following remark.

Remark 9.11 Let E be a Banach space, and let $F, G \subset E$ be closed linear subspaces such that $E = F \dot{+} G$. Then the mapping $P : E \to F$, $x + y \mapsto x$, where $x \in F$, $y \in G$, is a continuous projection.
This can be argued as follows: It is clear that P is a projection, i.e., that $P^2 = P$. To show continuity, by the Closed Graph Theorem, it suffices to prove that P is closed. For this, let $(x_n)_n \subset F$ and $(y_n)_n \subset G$ 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 $u = x + y \in F \dot{+} G$. Therefore, $Pu = x$, which we were required to prove. ◊

Imposing finite dimensionality in a particular way ensures not only the existence of a complementary subspace, but even a closed one.

Lemma 9.12 *Let E be a normed space, and let F be 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 first that $\dim E/F < \infty$, and let $x_1, \ldots, x_n \in E$ be such that $\{x_1 + F, \ldots, x_n + F\}$ is a basis of E/F. Then, define

$$G := \operatorname{span}\{x_1, \ldots, x_n\} \subseteq E$$

as our complementary subspace. Since $\dim G < \infty$, we can conclude that G is closed. We now show that $E = F + G$ and $F \cap G = \{0\}$. For this, for each $x \in E$, we have that

$$x + F = \sum_{i=1}^{n} \lambda_i (x_i + F) \quad \text{with some } \lambda_1, \ldots, \lambda_n \in \mathbb{K}.$$

Set $g := \sum_{i=1}^{n} \lambda_i x_i \in G$ 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, since $x \in G$, we have

$$x = \sum_{i=1}^{n} \lambda_i x_i,$$

and since $x \in F$, we have

$$0 = x + F = \sum_{i=1}^{n} \lambda_i (x_i + F).$$

Since $(x_i + F)_i$ are linearly independent, we conclude that $\lambda_1 = \ldots = \lambda_n = 0$, and thus $x = 0$.

Assume now that $\dim F < \infty$, and let $\{x_1, \ldots, x_n\}$ be a basis of F. Then we choose $f_1, \ldots, f_n \in F^*$ such that

$$f_i(x_j) = \delta_{ij} \quad \forall i, j \in \{1, \ldots, n\},$$

which defines a basis of F^* (the so-called dual basis). By the Hahn-Banach Theorem, there then exist $\ell_1, \ldots, \ell_n \in E^*$ such that $\ell_i|_F = f_i$. Now we define $P \colon E \to E$ by

$$Px := \sum_{j=1}^{n} \ell_j(x) x_j.$$

It is obvious that P is linear, continuous, and satisfies $P|_F = \mathrm{I}|_F$ (since $P(x_k) = x_k$). Moreover,

$$P^2 x = \sum_{j=1}^{n} \ell_j(x) P x_j = \sum_{j=1}^{n} \ell_j(x) x_j = Px,$$

showing that $P^2 = P$. Define $G := \ker P$ as our complementary subspace. It then follows immediately that G is closed, $G \cap F = \{0\}$, and $E = F + G$, since $x = Px + (x - Px)$. ∎

We now arrive at our first main theorem. The conditions (i)–(iii) will in fact be the defining conditions of a Fredholm operator. Thus, this theorem shows that the operator $\mathrm{I}-K$ with K being a compact operator is a Fredholm operator. This operator reminds us slightly of the theory of eigenvalues. It will turn out that indeed this is precisely how we will utilize it – to be honest, also for the theory of so-called spectral values which is an extension of the concept of eigenvalues appropriate for the infinite-dimensional situation.

Theorem 9.13 *Let E be a Banach space, let $K\colon E \to E$ be a compact operator, and let $T := \mathrm{I}-K \in L(E)$. Then the following conditions hold.*

(i) $\dim(\ker T) < \infty$,

(ii) *$T(E)$ is closed in E, and*

(iii) $\dim(E/T(E)) < \infty$.

Proof. (i). Since

$$\mathrm{I}\,|_{\ker T} = K|_{\ker T},$$

the operator $\mathrm{I}\,|_{\ker T}$ 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

$$S\colon G \to T(E), \quad Sx := Tx,\ x \in G.$$

Obviously, S is continuous and bijective; in particular, we have $S(G) = T(E)$. Now let $(y_n)_n \subset T(E)$ with $y_n \to y$ as $n \to \infty$. Then

$$y_n = Sx_n = x_n - Kx_n, \quad \text{where } x_n = S^{-1}y_n \in G,\ n \in \mathbb{N}.$$

It remains to prove that $y \in T(E)$.

For this, suppose that $(x_n)_n$ possesses no bounded subsequence, which implies that $\|x_n\| \to \infty$ as $n \to \infty$. Now define

$$u_n := \tfrac{x_n}{\|x_n\|} \in G, \quad n \in \mathbb{N}.$$

Then, since $(y_n)_n$ is bounded,

$$Su_n = u_n - Ku_n = \tfrac{y_n}{\|x_n\|} \to 0 \quad \text{as } n \to \infty.$$

Since $(u_n)_n$ 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 that $v \in G$ due to the closedness of G. Further by continuity 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, $(x_n)_n$ possesses 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

$$x_{n_j} = Sx_{n_j} + Kx_{n_j} \to y + v \quad \text{as } j \to \infty$$

and thus $y = \lim_j Tx_{n_j} = T(\lim_j x_{n_j}) = T(y+v) \in T(E) = S(G)$.

(iii). Since $T(E)$ is closed, by Theorem 4.12, it follows that $(E/T(E))^*$ is isometrically isomorphic to $T(E)^\perp$. Further, we have

$$\begin{aligned} T(E)^\perp &= \{f \in E^* : f(Tx) = 0 \, \forall x \in E\} \\ &= \{f \in E^* : (T^* f)x = 0 \, \forall x \in E\} \\ &= \{f \in E^* : T^* f = 0\}. \end{aligned}$$

This proves $T(E)^\perp = \ker T^*$.

Since K^* is compact by Theorem 9.5, and $T^* = \mathrm{I} - K^*$, using (i) we obtain

$$\dim(\ker T^*) < \infty.$$

By our previous considerations, $(E/T(E))^*$ is isometrically isomorphic to $\ker T^*$, hence we can conclude that

$$\dim(E/T(E))^* < \infty,$$

which implies (iii). ∎

Finally, we are prepared for the definition of Fredholm operators. The definition of the index of a Fredholm operator might seem strange at first sight. However, we will later (Theorem 9.17) show that the map of some Fredholm operator to its index is even always a continuous map, having certain implications due to the fact that the index is always integer.

Definition 9.14 An operator $T \in L(E)$ which satisfies (i) - (iii) in Theorem 9.13 is called *Fredholm operator*. The integer

$$\mathrm{ind}(T) := \dim(\ker T) - \dim(E/T(E))$$

is called the *index* of the Fredholm operator T. Moreover, $\mathcal{F}(E)$ shall denote the set of all Fredholm operators on E.

Notice that Theorem 9.13 immediately implies that $\mathrm{I} - \mathcal{K}(E) \subseteq \mathcal{F}(E)$.

We are now on our way to our second main result on Fredholm operators, namely the aforementioned Theorem 9.17. Again we require some preparations. The next lemma provides us with conditions for invertibility of a linear and bounded operator, which will also be useful in various other situations. Notice that the sufficient condition is a closeness condition of the operator under consideration to an operator of which we know invertibility.

Lemma 9.15 *Let E and F be Banach spaces, and let $T \in L(E,F)$ be bijective. Further, let $T^{-1} \in L(F,E)$ be the inverse of T, and let $S \in L(E,F)$ be such that*

$$\|S - T\| < \|T^{-1}\|^{-1}.$$

Then S is also invertible.

Proof. Notice that we can expand S as

$$S = T(\mathrm{I} - T^{-1}(T-S)).$$

Based on this, we define

$$Q := T^{-1}(T-S).$$

For $c := \|T-S\|\|T^{-1}\| < 1$, we then obtain

$$\left\| \sum_{k=m+1}^{n} Q^k \right\| = \left\| \sum_{k=m+1}^{n} (T^{-1}(T-S))^k \right\| \leq \sum_{k=m+1}^{n} c^k.$$

Thus the geometric series $\sum_{k=0}^{\infty} Q^k$ is convergent in $L(E,F)$. Finally,

$$\sum_{k=0}^{\infty} Q^k T^{-1} S = \lim_{n\to\infty} \sum_{k=0}^{n} Q^k (\mathrm{I} - Q) = \lim_{n\to\infty} (\mathrm{I} - Q^{n+1}) = \mathrm{I},$$

and also

$$S \sum_{k=0}^{\infty} Q^k T^{-1} = \sum_{k=0}^{\infty} T(\mathrm{I}-Q)Q^k T^{-1} = T\left(\sum_{k=0}^{\infty} (\mathrm{I}-Q)Q^k\right) T^{-1} = \mathrm{I}.$$

This shows that S is invertible with $S^{-1} = \sum_{k=0}^{\infty} Q^k T^{-1}$. ∎

Before continuing, we quickly draw the following remark from the proof.

Remark 9.16 If E is a Banach space and $S \in L(E)$ with $\|S\| < 1$, then Lemma 9.15 implies that $\mathrm{I} - S$ is invertible. It also follows from the proof of this lemma that the inverse of $\mathrm{I} - S$ is given by

$$(\mathrm{I} - S)^{-1} = \sum_{k=0}^{\infty} S^k.$$

This series is called the *Neumann series*. ◊

We arrive at our second main result, which analyzes the set of Fredholm operators and the map to their indices.

Theorem 9.17 *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. We start by proving that $\mathcal{F}(E)$ is open in $L(E)$. For this, let $T \in \mathcal{F}(E)$ be fixed. By Lemma 9.12, there exist closed complementary subspaces L and M to $\ker T$ and $T(E)$, respectively. By definition of a Fredholm operator, we have $\dim M < \infty$, and the isomorphisms $\ker T \times L \to E$ and $T(E) \times M \to E$ (with $(x, y) \mapsto x + y$) are continuous. To each $S \in L(E)$, we next associate the operator

$$\tilde{S} : L \times M \to E, \quad \tilde{S}(x, y) := Sx + y,\ x \in L,\ y \in M.$$

Then, since

$$\|\tilde{S}(x,y)\| = \|Sx+y\| \leq \|S\|\|x\| + \|y\| \leq \max\{1, \|S\|\}(\|x\| + \|y\|) = \max\{1, \|S\|\}\|(x,y)\|,$$

we have $\tilde{S} \in L(L \times M, E)$. Moreover, for $S_1,\ S_2 \in L(E)$,

$$\|\tilde{S}_1(x,y) - \tilde{S}_2(x,y)\| = \|S_1x - S_2x\| \leq \|S_1 - S_2\|\|x\| \leq \|S_1 - S_2\|\|(x,y)\|.$$

This implies that the (nonlinear) mapping $L(E) \to L(L \times M, E)$, $S \mapsto \tilde{S}$, is continuous.

Let now $\tilde{T}$ denote the associated operator to our fixed $T \in \mathcal{F}(E)$. We notice that $\tilde{T}(x,y) = Tx + y = 0$ implies $Tx = 0$ and $y = 0$. Further, since $x \in L$, we have $x \in \ker T \cap L = \{0\}$, showing that $\tilde{T}$ is injective. To also prove surjectivity, let $u \in E$. Then $u = z + y$ with $z \in T(E)$ and $y \in M$. Now let $x \in L$ be such that $Tx = z$, which yields

$$\tilde{T}(x,y) = Tx + y = z + y = u.$$

Thus $\tilde{T}$ is bijective.

We now show that $S \in \mathcal{F}(E)$ if $S \in L(E)$ and if $\tilde{S}$ is a bijection. Indeed, for those operators S the following conditions hold:

(i) $S(L) = \tilde{S}(L \times \{0\})$ is closed, since L is closed and $\widetilde{S}^{-1}$ is continuous,

(ii) $\dim(E/S(L)) = \dim M < \infty$, in view of $S(L) = \tilde{S}(L \times \{0\})$,

(iii) $L \cap \ker S = \{0\}$, since $\tilde{S}$ is injective.

Statement (iii) implies that $\dim(\ker S) \leq \dim(\ker T) < \infty$. From (ii) we conclude that $\dim(E/S(E)) < \infty$. Moreover, from the closedness of $S(L)$, the property $\dim(E/S(L)) < \infty$, and $S(E) \supset S(L)$, it follows that $S(E)$ is closed. Hence, $S \in \mathcal{F}(E)$.

After this preparation, we can finally prove that $\mathcal{F}(E)$ is open in $L(E)$. By Lemma 9.15, the set of bijective operators in $L(L \times M, E)$ is open. Hence, since $\tilde{T}$ is bijective and $S \mapsto \tilde{S}$ is continuous, there exists some open neighborhood V of T in $L(E)$ such that $\tilde{S}$ is bijective for each $S \in V$. According to the above, this implies $S \in \mathcal{F}(E)$ and hence $\mathcal{F}(E)$ is open.

We next turn to showing continuity of the map $T \mapsto \operatorname{ind} T$, $\mathcal{F}(E) \to \mathbb{Z}$. Again, let $T \in \mathcal{F}(E)$ be fixed. We already showed that there exists a neighborhood $V \subset \mathcal{F}(E)$ of T such that

$$\dim(\ker S) \leq \dim(\ker T) \quad \text{for all } S \in V,$$

as well as

$$\dim(E/S(E)) \leq \dim(E/S(L)) = \dim M = \dim(E/T(E)) \quad \text{for all } S \in V.$$

We will now show that $\operatorname{ind}(S) = \operatorname{ind}(T)$ for all $S \in V$. Since $L \cap \ker S = \{0\}$ and $\dim(E/L) < \infty$, it follows that there exists a finite-dimensional subspace F of E such that E is isomorphic to $\ker S \times F \times L$. We deduce

$$\dim(\ker T) = \dim(E/L) = \dim(\ker S) + \dim F \quad \text{and} \quad \dim(S(E)/S(L)) = \dim F.$$

Now, (ii) yields $\dim(E/S(L)) = \dim M = \dim(E/T(E))$, and hence we have

$$\dim(E/S(E)) = \dim(E/S(L)) - \dim(S(E)/S(L)) = \dim(E/T(E)) - \dim F.$$

Finally, we obtain

$$\operatorname{ind}(S) = \dim(\ker S) - \dim(E/S(E)) = \dim(\ker T) - \dim(E/T(E)) = \operatorname{ind}(T),$$

and the theorem is proved. ∎

One might wonder about properties of the set of Fredholm operators with the same index. The next remark provides some insight into this question, by using the previous theorem.

Remark 9.18 Notice that Theorem 9.17 implies that for each $k \in \mathbb{Z}$ the 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 $\mathcal{F}(E)$. Furthermore, by Lemma 9.15, the set of invertible operators is a (proper) subset of the open set $\mathcal{F}_0(E)$ and is itself open. Finally, we have $\mathcal{F}(E) = \mathcal{F}_0(E)$, if E is finite-dimensional. ◊

We already discussed before that $\mathrm{I} - \mathcal{K}(E) \subseteq \mathcal{F}(E)$. The index of this set of Fredholm operators can now be easily computed.

Corollary 9.19 *Let E be a Banach space and $K \in \mathcal{K}(E)$. Then* $\operatorname{ind}(\mathrm{I} - K) = 0$.

Proof. By Theorem 9.17, the map

$$\mathbb{R} \to \mathcal{F}(E) \to \mathbb{Z}, \quad t \mapsto \mathrm{I} - tK \mapsto \operatorname{ind}(\mathrm{I} - tK)$$

is continuous, which implies

$$\operatorname{ind}(\mathrm{I}-K) = \operatorname{ind}(\mathrm{I}) = 0.$$

■

Again considering Fredholm operators of the form $T := \mathrm{I}-K$ with K being a compact operator, the final corollary provides a useful relation between the dimension of the kernel of this operator and its dual.

Corollary 9.20 *Let E be a Banach space, let $K \in \mathcal{K}(E)$, and define $T := \mathrm{I}-K$. Then*

$$\dim(\ker T) = \dim(\ker T^*).$$

Proof. By Theorem 9.5, the operator K^* is compact, and by Corollary 9.19, we have $\operatorname{ind}(T) = 0$. Since

$$\ker T^* = T(E)^{\perp} \quad \text{and} \quad T(E)^{\perp} \cong (E/T(E))^*,$$

we can conclude that

$$\dim(\ker(T^*)) = \dim(T(E)^{\perp}) = \dim((E/T(E))^*) = \dim(E/T(E)) = \dim(\ker T).$$

■

10 The Spectrum of an Operator

The concept of eigenvalues and eigenvectors is of fundamental importance for linear algebra. At this point of the course we already suspect that an extension of those notions to the infinite-dimensional situation is again not straightforward, but requires a more sophisticated treatment. This is the topic of this chapter.

10.1 Eigenvalues and Spectral Values

Certainly, the notion of eigenvalues and eigenvectors can also be formulated in the infinite-dimensional setting. However, it is too restrictive for our purposes. Analyzing the definition of an eigenvalue, we are led to a very natural generalization, which will be a spectral value. It is worth noticing that this definition also requires a notion of continuity, since it is required that for a certain operator no inverse exist in the space of linear and continuous operators.

Definition 10.1 Let E be a normed space over $\mathbb{K}$ and $T \in L(E)$.

(1) A scalar $\lambda \in \mathbb{K}$ is an *eigenvalue* of T, if $\ker(\lambda \mathrm{I} - T) \neq \{0\}$. The elements of $\ker(\lambda \mathrm{I} - T) \setminus \{0\}$ are called *eigenvectors* of T associated with λ, and

$$E(\lambda) := \{x \in E : Tx = \lambda x\}$$

is called the *eigenspace* of T with respect to λ.

(2) A scalar $\lambda \in \mathbb{K}$ is called a *spectral value* of T, if $\lambda \mathrm{I} - T$ 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 $\varrho(T) := \mathbb{K} \setminus \sigma(T)$ is called the *resolvent set* of T, and the elements of $\varrho(T)$ are called *regular values* of T.

The first question we need to ask is the relation of eigenvalues and spectral values. This and some other useful observations are summarized in the following remark.

Remark 10.2 (1) If λ is an eigenvalue of T, then it is also a spectral value of T. In the situation $\dim E < \infty$ also the converse is true. In fact, if $\lambda \in \sigma(T)$, then $\lambda \mathrm{I} - T$ is not bijective, hence not injective, and thus λ is an eigenvalue. This shows that in the finite-dimensional situation, not unexpectedly, also each spectral value is an eigenvalue.

(2) Let E be a Banach space. Then, as a direct consequence of the open mapping theorem, the inverse $(\lambda\,\mathrm{I} - T)^{-1}$ exists if and only if $\lambda\,\mathrm{I} - T$ is bijective.

(3) Let $\dim E = \infty$ and $T \in \mathcal{K}(E)$. We have $0 \in \sigma(T)$, since $0 \in \varrho(T)$ would imply the invertibility of T. It would follow that T is open, and thus that $\overline{T(K_1(0))}$ is a compact neighborhood of 0 in E. However, there are no compact neighborhoods of 0 if $\dim E = \infty$.

◊

We first need to check some basic properties of the spectrum such as whether it is bounded. The following lemma provides us with some insight, with its proof mainly relying on the technique of Neumann series.

Lemma 10.3 *Let E be a normed space over $\mathbb{K}$, and let $T \in L(E)$. Then $\sigma(T)$ is closed. Also, for every $\lambda \in \sigma(T)$, we have*

$$|\lambda| \leq \|T\|,$$

i.e., $\sigma(T)$ is bounded.

Proof. We first show that $\sigma(T)$ is closed. For this, let $\lambda_0 \in \varrho(T)$, and let $\lambda \in \mathbb{K}$ be such that

$$\|(\lambda\,\mathrm{I} - T) - (\lambda_0\,\mathrm{I} - T)\| = |\lambda - \lambda_0| < \|(\lambda_0\,\mathrm{I} - T)^{-1}\|^{-1}.$$

Lemma 9.15 then implies that $\lambda\,\mathrm{I} - T$ is (boundedly) invertible, and thus $\lambda \in \varrho(T)$. This argument shows that $\varrho(T)$ is open, implying that $\sigma(T)$ is closed.

To prove boundedness – more precisely, a bound of $\|T\|$ –, let $\lambda \in \mathbb{K}$ be such that $|\lambda| > \|T\|$. Then

$$\|(\lambda\,\mathrm{I} - T) - \lambda\,\mathrm{I}\| = \|T\| < |\lambda| = \|(\lambda\,\mathrm{I})^{-1})\|^{-1}.$$

Again, Lemma 9.15 implies that $\lambda\,\mathrm{I} - T$ is invertible, hence $\lambda \in \varrho(T)$. ∎

10.2 Two Famous Results on the Spectrum of an Operator on a Banach Space

In the situation of E being a Banach space, stronger results on the spectrum of a linear operator are possible. The next lemma, which can be regarded as a strengthening of Lemma 9.15 adapted to operators of the form $\lambda\,\mathrm{I} - T$, will be required in their proofs.

Lemma 10.4 *Let E be a Banach space, and let $T \in L(E)$.*

(i) If $|\lambda| > \|T\|$, then

$$(\lambda\,\mathrm{I} - T)^{-1} = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}} \quad \textit{and} \quad \|(\lambda\,\mathrm{I} - T)^{-1}\| \leq (|\lambda| - \|T\|)^{-1}.$$

(ii) If T is invertible and $S \in L(E)$ with

$$\|T - S\| \leq \varepsilon \|T^{-1}\|^{-1}, \quad \varepsilon < 1,$$

then S is invertible and $\|T^{-1} - S^{-1}\| \leq \frac{1}{1-\varepsilon}\|T^{-1}\|^2\|S - T\|$.

Proof. (i). It suffices to prove that, for $A = \frac{1}{\lambda}T$, we have

$$(\mathrm{I} - A)^{-1} = \sum_{n=0}^{\infty} A^n.$$

Observe first that $\|A\| < 1$, which implies

$$\sum_{n=0}^{\infty} \|A\|^n < \infty.$$

Thus $B_m := \sum_{n=0}^{m} A^n$ converges to some $B \in L(E)$ as $m \to \infty$. By definition of B_m, the following equation holds:

$$(\mathrm{I} - A)B_m = \mathrm{I} - A^{m+1} = B_m(\mathrm{I} - A).$$

Since $\|A^{m+1}\| \to 0$ as $m \to \infty$, we can let m tend to infinity and obtain

$$(\mathrm{I} - A)B = \mathrm{I} = B(\mathrm{I} - A).$$

This proves the first part of the claim.
For the second part, just consider

$$\|(\mathrm{I} - A)^{-1}\| \leq \sum_{n=0}^{\infty} \|A\|^n = \frac{1}{1 - \|A\|},$$

and multiply with $|\lambda|$.

(ii). This is proved similarly as Lemma 9.15. ∎

We are now prepared for the first of our anticipated two famous results, this one being due to Gelfand and Mazur. Its proof requires results and techniques from complex analysis, which we need to rely on.

Theorem 10.5 (Gelfand-Mazur, 1941) *Let E be a Banach space over $\mathbb{C}$ and $T \in L(E)$. Then $\sigma(T) \neq \varnothing$.*

Proof. Let $f \in L(E)^*$ and define

$$\varphi : \varrho(T) \to \mathbb{C}, \quad \varphi(\lambda) := f\left((\lambda\,\mathrm{I} - T)^{-1}\right).$$

Recall that $\varrho(T)$ is open. For $\lambda, \lambda_0 \in \varrho(T)$, we have

$$\begin{aligned}\varphi(\lambda) - \varphi(\lambda_0) &= f\left((\lambda\,\mathrm{I} - T)^{-1} - (\lambda_0\,\mathrm{I} - T)^{-1}\right)\\ &= (\lambda_0 - \lambda) f\left((\lambda\,\mathrm{I} - T)^{-1}(\lambda_0\,\mathrm{I} - T)^{-1}\right).\end{aligned}$$

By Lemma 10.4, the inversion of operators is continuous. This fact allows us to conclude that

$$\lim_{\lambda \to \lambda_0} \frac{\varphi(\lambda) - \varphi(\lambda_0)}{\lambda - \lambda_0} = -f\left((\lambda_0\,\mathrm{I} - T)^{-2}\right).$$

Hence φ is holomorphic on $\varrho(T)$.

Towards a contradiction, assume now that $\sigma(T) = \varnothing$. 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). Application of Liouville's Theorem (from complex analysis) yields that φ is constant, and hence zero. Since this is true for any $f \in L(E)^*$, Corollary 4.7 implies that $(\lambda\,\mathrm{I} - T)^{-1} = 0$ for each $\lambda \in \varrho(T)$, which is a contradiction. ∎

The next result, also due to Gelfand, is even more interesting, since it provides a very insightful formula for the smallest bound for the spectrum. In fact, the left hand side of the equation is called the *spectral radius* of T. As before, the proof requires methods from complex analysis.

Theorem 10.6 (Formula for the spectral radius, Gelfand 1941) *Let E be a Banach space over $\mathbb{C}$, and let $T \in L(E)$. Then*

$$\sup_{\lambda \in \sigma(T)} |\lambda| = \lim_{n \to \infty} \|T^n\|^{\frac{1}{n}}.$$

Proof. For all $\lambda \in \sigma(T)$,

$$T^n - \lambda^n\,\mathrm{I} = (T - \lambda\,\mathrm{I}) \left(\sum_{k=1}^{n} \lambda^{k-1} T^{n-k}\right) = \left(\sum_{k=1}^{n} \lambda^{k-1} T^{n-k}\right)(T - \lambda\,\mathrm{I}),$$

which implies $\lambda^n \in \sigma(T^n)$. Now Lemma 10.3 yields $|\lambda|^n \leq \|T^n\|$. Thus, if $r(T)$ denotes the spectral radius, then

$$r(T) \leq \liminf_{n \to \infty} \|T^n\|^{\frac{1}{n}}.$$

Next we consider

$$\varphi(\lambda) := f((\lambda\,\mathrm{I} - T)^{-1}), \quad f \in E^*, \lambda \in \varrho(T).$$

We know by the proof of Theorem 10.5 that φ is holomorphic on $\{\lambda \in \mathbb{C} : |\lambda| > r(T)\}$. Further $\{\lambda \in \mathbb{C} : |\lambda| > \|T\|\} \subset \{\lambda \in \mathbb{C} : |\lambda| > r(T)\}$. As seen from Lemma 10.4, utilizing methods of complex analysis, φ is therefore 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, we have

$$\sup_n \left| \tfrac{f(T^n)}{\lambda^n} \right| < \infty \quad \text{for all } |\lambda| > r(T), f \in L(E)^*.$$

By Theorem 5.13, for each $\lambda \in \mathbb{C}$ with $|\lambda| > r(T)$, there exists some $M_\lambda > 0$ such that

$$\left\| \tfrac{T^n}{\lambda^n} \right\| \leq M_\lambda \quad \text{for all } n \in \mathbb{N}.$$

Hence $\|T^n\|^{\frac{1}{n}} \leq (M_\lambda)^{\frac{1}{n}} |\lambda|$, which implies $\limsup_{n\to\infty} \|T^n\|^{\frac{1}{n}} \leq |\lambda|$ for all $\lambda \in \mathbb{C}$ with $|\lambda| > r(T)$, and we can conclude that

$$\limsup_{n\to\infty} \|T^n\|^{\frac{1}{n}} \leq r(T).$$

Thus, finally, we arrive at the claimed $r(T) = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}$. ∎

10.3 The Spectrum of Compact Operators

We are now interested in analyzing the spectrum of compact operators. As mentioned already several times, those are in a certain sense close to finite-dimensional operators or operators in finite-dimensional spaces. Thus, we expect them to also inherit some of their properties concerning eigenvalues and spectral values. Let us remind ourselves that by Remark 10.2 in the finite-dimensional situation eigenvalues and spectral values coincide.

We have to start with two technical lemmata.

Lemma 10.7 *Let E be a normed space, and let $T \in L(E)$. Further, let F and G be closed subspaces of E with $F \subset G$, $F \neq G$, and*

$$(\mathrm{I} - T)G \subset F.$$

Then there exists some $a \in G$ with $\|a\| = 1$ and

$$\|Ta - Tx\| \geq \tfrac{1}{2} \quad \text{for all } x \in F.$$

Proof. Let $b \in G \backslash F$, and consider

$$\alpha := \operatorname{dist}(b, F) = \inf_{x\in F} \|x - b\|.$$

Since F is closed, it follows that $\alpha > 0$. This implies that there exists some $y \in F$ such that

$$\|b - y\| < 2\alpha.$$

Now define

$$a := \frac{b-y}{\|b-y\|} \in G.$$

Then $\|a\| = 1$ and, for arbitrary $z \in F$, we have

$$\|z - a\| = \tfrac{1}{\|b-y\|} \| \underbrace{z\|b-y\| + y}_{\in F} - b\| \geq \frac{\alpha}{2\alpha} = \tfrac{1}{2}.$$

Thus, finally, for each $x \in F$, we obtain

$$\|Tx - Ta\| = \|\underbrace{x - (\mathrm{I} - T)x + (\mathrm{I} - T)a}_{\in F} - a\| \geq \tfrac{1}{2}.$$

∎

The second technical lemma is a result about linear independence of certain sets of eigenvectors, which should be well-known in the situation of matrices.

Lemma 10.8 *Let E be a linear space, let $T : E \to E$ be a linear operator, and let 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 will now prove the claim via induction over n.
Clearly, each set in $\mathcal{M}_1$ is linearly independent. Assume now that the claim is proven for $n-1$. Further, let $x_1, \ldots x_n \in M$, $Tx_i = \lambda_i x_i$, be such that

$$\sum_{i=1}^{n} \alpha_i x_i = 0. \tag{10.1}$$

This implies

$$\sum_{i=1}^{n} \alpha_i \lambda_i x_i = \sum_{i=1}^{n} \alpha_i T x_i = 0.$$

By (10.1), it follows that

$$0 = \alpha_1(\lambda_1 - \lambda_n)x_1 + \cdots + \alpha_{n-1}(\lambda_{n-1} - \lambda_n)x_{n-1}.$$

Since by induction hypothesis the vectors $x_1, \ldots, x_{n-1}$ are linearly independent, this implies

$$\alpha_i(\lambda_i - \lambda_n) = 0 \quad \forall i = 1, \ldots, n-1.$$

Using the fact that the λ_i are mutually distinct, this yields $\alpha_i = 0$ for $i = 1, \ldots, n-1$. Thus, finally, again by (10.1), also $\alpha_n = 0$. ∎

The next result provides us already with quite detailed information about the structure of the set of eigenvalues of a compact operator. As can be seen, indeed the set of eigenvalues behaves close to a set of eigenvalues in the finite-dimensional situation, since it is at most countable.

Theorem 10.9 *Let E be a normed space, and let $K\colon E \to E$ be 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 $(x_n)_n \subset E$, $x_n \neq 0$ for all n, with

$$Kx_n = \mu_n x_n,$$

and define

$$F_n := \operatorname{span}\{x_1, \ldots, x_n\}.$$

By Lemma 10.8, the set $\{x_1, \ldots, x_n\}$ is linearly independent, which implies that

$$F_n \subsetneq F_{n+1}.$$

Let now $y_n := \sum_{i=1}^n \alpha_i x_i \in F_n$. Then we obtain

$$\begin{aligned}
(K - \mu_n \mathrm{I})y_n &= \sum_{i=1}^n \alpha_i K x_i - \mu_n \sum_{i=1}^n \alpha_i x_i \\
&= \sum_{i=1}^n \alpha_i \mu_i x_i - \mu_n \sum_{i=1}^n \alpha_i x_i \\
&= \sum_{i=1}^n \alpha_i (\mu_i - \mu_n) x_i \in F_{n-1},
\end{aligned}$$

which yields

$$(\mathrm{I} - \tfrac{1}{\mu_n} K) F_n \subset F_{n-1}.$$

For each $n \in \mathbb{N}$, F_n is closed. Hence, by Lemma 10.7, there exists a sequence $(y_n)_{n\in\mathbb{N}} \subset E$ with

$$y_n \in F_n, \quad \|y_n\| = 1, \quad \text{and} \quad \|Ky_n - Ky_m\| \geq \tfrac{1}{2}|\mu_n| \quad \text{for all } m > n,\ n \in \mathbb{N}.$$

This implies

$$\|Ky_n - Ky_m\| \geq \tfrac{1}{2}\delta \quad \text{for all } m > n,\ n \in \mathbb{N}.$$

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. ∎

We are now able to derive a complete analysis of the spectrum of a compact operator on a Banach space. Again notice that indeed the properties are quite similar as in the finite-dimensional situation.

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 9.19, we have

$$\operatorname{ind}(\mathrm{I} - \tfrac{1}{\lambda}K) = 0.$$

This implies that $\mathrm{I} - \frac{1}{\lambda}K$ (and thus also $\lambda\,\mathrm{I} - K$) is injective if and only if it is surjective.

(ii). By definition $E(\lambda) = \ker(\mathrm{I} - \frac{1}{\lambda}K)$. Since $\mathrm{I} - \frac{1}{\lambda}K$ is a Fredholm operator, we obtain

$$\dim(E(\lambda)) = \dim(\ker(\mathrm{I} - \tfrac{1}{\lambda}K)) < \infty.$$

(iii). This follows from Theorem 10.9 and (i). ∎

11 Spectral Theory for Compact Operators

This chapter is devoted to one of the maybe most important theorems of functional analysis, the so-called spectral theorem. It provides us with a decomposition of a compact, self-adjoint operator in terms of its spectral values. In fact, the theorem we will state here (Theorem 11.9), is a special case of a much more general theorem for unbounded operators, which will be part of Functional Analysis II. The situation of compact operators we consider is again very close to the finite-dimensional situation, and we recall in Remark 11.4 the corresponding result.

11.1 Self-Adjoint Operators

Since we will focus on so-called self-adjoint operators, we first need to introduce those and analyze their spectrum. For this, we start with some considerations related to Theorem 7.16.

Remark 11.1 Let $\mathcal{H}$ be a Hilbert space, let $T \in L(\mathcal{H})$, and let $\varphi\colon \mathcal{H} \times \mathcal{H} \to \mathbb{K}$ be 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, there exists a unique operator $T^* \in L(\mathcal{H})$ with

$$\langle T^* x, y \rangle = \varphi(x,y).$$

This implies

$$\langle x, Ty \rangle = \langle T^* x, y \rangle \quad \text{and} \quad \langle Tx, y \rangle = \langle x, T^* y \rangle \quad \text{for all } x, y \in \mathcal{H}. \tag{11.1}$$

$\Diamond$

The specified operator T^* with the quite useful properties given by (11.1) will become the adjoint operator of T.

Definition 11.2 Let E be a Hilbert space, and let $T \in L(\mathcal{H})$. Then T^* from Remark 11.1 is called the *adjoint operator* to T. If $T = T^*$, then T is called *self-adjoint*.

There might be some confusion with the dual operator, since we also used the notation T^* for it. Intriguingly, there exists a close connection which justifies such a similar notation.

Remark 11.3 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

$$T^* = j^{-1}\tilde{T}^* j.$$

$\Diamond$

We next recall the corresponding result from the finite-dimensional situation to be able to compare it with the Spectral Theorem (Theorem 11.9).

Remark 11.4 Let $\mathcal{H}$ be a finite-dimensional inner product space, and let $T\colon \mathcal{H} \to \mathcal{H}$ be self-adjoint. By results from Linear Algebra, there exists 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 then have

$$Tx = \sum_{i=1}^{n} \lambda_i \langle x, u_i \rangle u_i.$$

$\Diamond$

A self-adjoint operator possesses only real eigenvalues. This result is not unexpected, since the same holds for self-adjoint matrices.

Lemma 11.5 *If $T \in L(\mathcal{H})$ is self-adjoint, then the eigenvalues of T are real. Moreover, the eigenspaces corresponding to two different eigenvalues are orthogonal.*

Proof. Let λ be an eigenvalue of T, and let x be an eigenvector to λ. Then we have

$$\lambda \|x\|^2 = \lambda \langle x, x \rangle = \langle Tx, x \rangle = \langle x, Tx \rangle = \overline{\lambda} \langle x, x \rangle = \overline{\lambda} \|x\|^2,$$

and hence $\lambda = \overline{\lambda}$.

Next, let $\lambda \neq \mu$ be two eigenvalues of T. For $x \in E(\lambda)$, $y \in E(\mu)$, $x, y \neq 0$, we obtain

$$(\lambda - \mu)\langle x, y \rangle = \lambda \langle x, y \rangle - \mu \langle x, y \rangle = \langle Tx, y \rangle - \langle x, Ty \rangle = 0,$$

which proves $\langle x, y \rangle = 0$. ∎

11.2 Preparatory Lemmata

The Spectral Theorem requires a bit of additional preparatory work, more precisely, three results. The second and third result are quite nice and interesting by themselves. However, the first result which we now state and prove is just technical and unintuitive.

Lemma 11.6 *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 conditions holds.

(i) $|\langle Tx, y\rangle + \langle Ty, x\rangle| \leq 2\alpha\|x\|\|y\|$ *for all* $x, y \in \mathcal{H}$.

(ii) If $\mathbb{K} = \mathbb{C}$, *then*

$$|\langle Tx, y\rangle| + |\langle Ty, x\rangle| \leq 2\alpha\|x\|\|y\|$$

for all $x, y \in \mathcal{H}$.

Proof. (i). We first observe that, for all $x, y \in \mathcal{H}$,

$$\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle = 2(\langle Tx, y\rangle + \langle Ty, x\rangle).$$

By the parallelogram identity, we can then conclude that

$$2|\langle Tx, y\rangle + \langle Ty, x\rangle| \leq \alpha(\|x+y\|^2 + \|x-y\|^2) = 2\alpha(\|x\|^2 + \|y\|^2)$$

for all $x, y \in \mathcal{H}$. By substituting x by $c^{-1}x$ and y by cy, where $c > 0$, we obtain

$$|\langle Tx, y\rangle + \langle Ty, x\rangle| \leq \alpha(c^{-2}\|x\|^2 + c^2\|y\|^2).$$

In case $y \neq 0$ – notice that we do not need to consider the case $y = 0$ – we now choose c by

$$c := \left(\tfrac{\|x\|}{\|y\|}\right)^{1/2},$$

which yields

$$|\langle Tx, y\rangle + \langle Ty, x\rangle| \leq 2\alpha\|x\|\|y\|.$$

(ii). We start with this last inequality, substitute x by $\mathrm{e}^{\mathrm{i}\,t}x$, $t \in \mathbb{R}$, and multiply (i) with $1 = |e^{\mathrm{i}\,s}|$, $s \in \mathbb{R}$, to obtain

$$|\mathrm{e}^{\mathrm{i}\,s}\langle T(\mathrm{e}^{\mathrm{i}\,t}x), y\rangle + \mathrm{e}^{\mathrm{i}\,s}\langle Ty, \mathrm{e}^{\mathrm{i}\,t}x\rangle| \leq 2\alpha\|x\|\|y\|.$$

This implies

$$|\mathrm{e}^{\mathrm{i}\,(s+t)}\langle Tx, y\rangle + \mathrm{e}^{\mathrm{i}\,(s-t)}\langle Ty, x\rangle| \leq 2\alpha\|x\|\|y\|.$$

For suitable $u, v \in \mathbb{R}$ satisfying

$$\mathrm{e}^{\mathrm{i}\,u}\langle Tx, y\rangle = |\langle Tx, y\rangle| \quad \text{and} \quad \mathrm{e}^{\mathrm{i}\,v}\langle Ty, x\rangle = |\langle Ty, x\rangle|,$$

we set $s := \frac{1}{2}(u+v)$ and $t := \frac{1}{2}(u-v)$, i.e., $u = s+t$ and $v = s-t$. Thus, finally,

$$|\langle Tx, y\rangle| + |\langle Ty, x\rangle| \le 2\alpha\|x\|\|y\|$$

for all $x, y \in \mathcal{H}$. ■

The next result is much more interesting, since it provides yet another equivalent form for the operator norm. Of particular importance is the fact that we can express it in terms of $|\langle Tx, y\rangle|$.

Corollary 11.7 *Let $\mathcal{H}$ be a Hilbert space, and let $T \in L(\mathcal{H})$. 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

$$\|T\| = \sup\{|\langle Tx, x\rangle| : \|x\| = 1\} = \inf\{\alpha > 0 : |\langle Tx, x\rangle| \le \alpha\|x\|^2 \; \forall x \in \mathcal{H}\}.$$

Proof. We start with the first claim. Since the Riesz map is an isometry, we obtain

$$\sup_{\|x\|=1}\sup_{\|y\|=1}|\langle Tx, y\rangle| = \sup_{\|x\|=1}\|f_{Tx}\| = \sup_{\|x\|=1}\|Tx\| = \|T\|.$$

Since

$$|\langle Tx, y\rangle| \le c\|x\|\|y\| \quad \iff \quad \left|\left\langle T\tfrac{x}{\|x\|}, \tfrac{y}{\|y\|}\right\rangle\right| \le c,$$

we also obtain that

$$\sup\{|\langle Tx, y\rangle| : \|x\| = \|y\| = 1\} = \inf\{c > 0 : |\langle Tx, y\rangle| \le c\|x\|\|y\|\}.$$

Turning to the second claim, let now T be self-adjoint. We remark that we only need to prove

$$\|T\| = \alpha := \sup\{|\langle Tx, x\rangle| : \|x\| = 1\},$$

since from this result the remaining claim follows similarly as above. First note that, obviously,

$$\alpha \le \sup\{|\langle Tx, y\rangle| : \|x\| = \|y\| = 1\} = \|T\|.$$

For the other direction, we first use the self-adjointness of T to estimate for $x, y \in \mathcal{H}$

$$|\langle Tx, y\rangle| = \tfrac{1}{2}|\langle Tx, y\rangle + \langle x, Ty\rangle| \le \tfrac{1}{2}\begin{cases} |\langle Tx, y\rangle + \langle Ty, x\rangle| & : \text{ if } \mathbb{K} = \mathbb{R}, \\ |\langle Tx, y\rangle| + |\langle Ty, x\rangle| & : \text{ if } \mathbb{K} = \mathbb{C}. \end{cases}$$

Since $|\langle Tx, x\rangle| \leq \alpha\|x\|^2$ for all $x \in \mathcal{H}$, Lemma 11.6 now implies that $\|T\| = \sup\{|\langle Tx, y\rangle| : \|x\| = \|y\| = 1\} \leq \alpha$. ∎

The third and last of the preparatory lemmata is also the most interesting one. Recall Lemma 10.3, which states that the absolute values of spectral values are bounded by $\|T\|$. The next result shows that in case of a compact and self-adjoint operator, this bound is even attained.

Lemma 11.8 *Let $\mathcal{H}$ be a Hilbert space, and let $K \colon \mathcal{H} \to \mathcal{H}$ be a compact and self-adjoint operator. Then $\|K\|$ or $-\|K\|$ is an eigenvalue of K.*

Proof. First, observe that

$$\langle Kx, x\rangle = \langle x, Kx\rangle = \overline{\langle Kx, x\rangle} \quad \text{for all } x \in \mathcal{H},$$

hence $\langle Kx, x\rangle \in \mathbb{R}$. By Corollary 11.7, there exists a sequence $(x_n)_n \subset \mathcal{H}$ such that

$$\|x_n\| = 1 \quad \text{and} \quad \lim_{n\to\infty} |\langle Kx_n, x_n\rangle| = \|K\|.$$

Notice that without loss of generality, we can assume that $K \neq 0$ and the real sequence $(\langle Kx_n, x_n\rangle)_n$ is convergent.

Next, we define

$$c := \lim_{n\to\infty} \langle Kx_n, x_n\rangle,$$

hence $|c| = \|K\|$. Since K is compact and $(x_n)_n$ is bounded, we can furthermore assume that $(Kx_n)_n$ is convergent. The computation

$$\begin{aligned} 0 &\leq \|Kx_n - cx_n\|^2 \\ &= \|Kx_n\|^2 + \|cx_n\|^2 - 2c\langle Kx_n, x_n\rangle \\ &\leq \|K\|^2 + c^2 - 2c\langle Kx_n, x_n\rangle \\ &= 2\|K\|^2 - 2c\,\underbrace{\langle Kx_n, x_n\rangle}_{\to c} \to 2\|K\|^2 - 2c^2 = 0. \end{aligned}$$

implies $\|Kx_n - cx_n\| \to 0$, $n \to \infty$. But since $(Kx_n)_n$ is convergent and $c \neq 0$, also $(x_n)_n$ is convergent. Finally, setting

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

we can conclude that $\|x\| = 1$ and

$$Kx - cx = \lim_{n\to\infty} (Kx_n - cx_n) = 0.$$

Thus, $c \in \{-\|K\|, \|K\|\}$ is an eigenvalue of K. ∎

11.3 The Spectral Theorem for Compact Self-Adjoint Operators

We are now ready for the previously already many times announced Spectral Theorem in the situation of compact, self-adjoint operators, which provides us with a decomposition of such operators in terms of their spectral values (for a finite-dimensional analog see Remark 11.4). The proof is a beauty by itself, and uses various previous results in an intruigingly constructive manner.

Notice that self-adjointness requires us to work in a Hilbert space. Moreover, compactness implies that spectral values and eigenvalues coincide except for zero, and self-adjointness ensures that those values are all in $\mathbb{R}$.

Theorem 11.9 (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 $(\lambda_n)_n \subset \mathbb{R}$ and $(x_n)_n \subset \mathcal{H}$ (either finite or infinite sequences) such that the following conditions hold.*

(i) $|\lambda_1| \geq |\lambda_2| \geq \cdots$, $\lambda_n \neq 0$ *for all* $n \in \mathbb{N}$, *and* $\lim_{n\to\infty} \lambda_n = 0$ *(if $(\lambda_n)_n$ is infinite).*

(ii) $(x_n)_n$ *forms an orthonormal system in $\mathcal{H}$ and* $Kx_n = \lambda_n x_n$.

(iii) If $\lambda \neq 0$ is an eigenvalue of K and $n_\lambda := \dim E(\lambda)$, then λ appears in $(\lambda_n)_n$ exactly n_λ times.

(iv) For each $x \in \mathcal{H}$,

$$Kx = \sum_n \lambda_n \langle x, x_n \rangle x_n.$$

Proof. We start by applying Lemma 11.8, which shows that 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 $\|x_1\| = 1$, and set

$$\mathcal{H}_1 := \operatorname{span}\{x_1\}.$$

Since $x \in \mathcal{H}_1^\perp$ implies

$$\langle Kx, x_1 \rangle = \langle x, Kx_1 \rangle = \langle x, \lambda_1 x_1 \rangle = 0,$$

we can conclude that

$$K(\mathcal{H}_1^\perp) \subset \mathcal{H}_1^\perp.$$

Observe that the restriction $K|_{\mathcal{H}_1^\perp} : \mathcal{H}_1^\perp \to \mathcal{H}_1^\perp$ is still a compact, self-adjoint operator. We now have to deal with two cases:

Case $K|_{\mathcal{H}_1^\perp} = 0$. Then, for $x = y + z$ with $y \in \mathcal{H}_1$ and $z \in \mathcal{H}_1^\perp$, we have

$$Kx = Ky + Kz = K(\langle y, x_1 \rangle x_1) = \langle y, x_1 \rangle Kx_1 = \lambda_1 \langle y, x_1 \rangle x_1 = \lambda_1 \langle x, x_1 \rangle x_1.$$

This finishes the proof.

Case $K|_{\mathcal{H}_1^\perp} \neq 0$. By applying Lemma 11.8 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 now $x_2 \in \mathcal{H}_1^\perp$ 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 $\mathcal{H}$, and as before we obtain that $K|_{\mathcal{H}_2^\perp} : \mathcal{H}_2^\perp \to \mathcal{H}_2^\perp$ is compact and self-adjoint. Again we are facing two cases:

Case $K|_{\mathcal{H}_2^\perp} = 0$. For $x = y + z$ with $y \in \mathcal{H}_2$ and $z \in \mathcal{H}_2^\perp$, we have

$$\begin{aligned} Kx &= Ky + Kz \\ &= K(\langle y, x_1\rangle x_1 + \langle y, x_2\rangle x_2) \\ &= \lambda_1\langle y, x_1\rangle x_1 + \lambda_2\langle y, x_2\rangle x_2 \\ &= \lambda_1\langle x, x_1\rangle x_1 + \lambda_2\langle x, x_2\rangle x_2. \end{aligned}$$

This finishes the proof.

Case $K|_{\mathcal{H}_2^\perp} \neq 0$. We continue the process as before. This will eventually lead us to the following two general cases:

Case 1 If we have $K|_{\mathcal{H}_n^\perp} = 0$ at some point, then

$$Kx = \sum_{i=1}^{n} \lambda_i \langle x, x_i\rangle x_i \quad \text{for all } x \in \mathcal{H}. \tag{11.2}$$

In this case, (i), (ii), and (iv) are satisfied, and to remains to show (iii). For this, let $\lambda \neq 0$ be an eigenvalue of K and $x \in E(\lambda)$, which implies

$$x = \tfrac{1}{\lambda}Kx = \sum_{i=1}^{n} \tfrac{\lambda_i}{\lambda}\langle x, x_i\rangle x_i. \tag{11.3}$$

Moreover, (11.2) shows that $K(\mathcal{H}) = \operatorname{span}\{x_1, \ldots, x_n\}$. Hence, by (ii) and since $x = \lambda^{-1}Kx \in K(\mathcal{H})$, we also have

$$x = \sum_{i=1}^{n} \langle x, x_i\rangle x_i.$$

Comparing this equation to (11.3), we can conclude by uniqueness that $\langle x, x_i\rangle = 0$ for each i with $\lambda \neq \lambda_i$. Thus

$$E(\lambda) = \operatorname{span}\{x_i : \lambda_i = \lambda\},$$

which shows that λ appears $\dim E(\lambda) = n_\lambda$ times in $(\lambda_i)_{i=1}^n$, proving (iii).

Case 2 Now assume that the process does not stop. This will yield sequences $(\lambda_n)_n$ and $(x_n)_n$ with $\|x_n\| = 1$, $Kx_n = \lambda_n x_n$, and $x_n \perp x_m$ for $n \neq m$. Thus (ii) is immediately satisfied.

For (i), assume that

$$|\lambda_n| \geq \delta > 0$$

for infinitely many $n \in \mathbb{N}$. This implies

$$\|Kx_n - Kx_m\|^2 = \|Kx_n\|^2 + \|Kx_m\|^2 = \lambda_n^2 + \lambda_m^2 \geq 2\delta^2.$$

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

$$\langle Kx, x_n \rangle = \langle x, Kx_n \rangle = \lambda_n \langle x, x_n \rangle = 0,$$

thus $K(L^\perp) \subset L^\perp$. Now let $x \in L^\perp$. Then, since $x \in \mathcal{H}_n^\perp$ for each n, we obtain

$$|\langle Kx, x \rangle| \leq \left\| K|_{\mathcal{H}_n^\perp} \right\| \|x\|^2 = |\lambda_{n+1}| \|x\|^2 \to 0 \quad \text{as } n \to \infty.$$

By Corollary 11.7, $\langle Kx, x \rangle = 0$ for all $x \in L^\perp$ now implies

$$K|_{L^\perp} = 0.$$

Finally, for each $x = y + z \in \mathcal{H}$ with $y \in L$ and $z \in L^\perp$, we obtain

$$Kx = Ky + Kz = Ky = K\left(\sum_{n \in \mathbb{N}} \langle y, x_i \rangle x_i \right) = \sum_{n \in \mathbb{N}} \lambda_i \langle y, x_i \rangle x_i = \sum_{n \in \mathbb{N}} \lambda_i \langle x, x_i \rangle x_i.$$

(iii) can be shown exactly as in the first case. The theorem is proven. ∎

We finish this chapter with some additional comments on the Spectral Theorem and its extensions.

Remark 11.10 (1) The orthonormal system $(x_n)_n$ is an orthonormal basis of $(\ker K)^\perp = \overline{K(\mathcal{H})}$. If P_0 is the orthogonal projection of $\mathcal{H}$ onto $\ker K$, then

$$x = P_0 x + \sum_n \langle x, x_n \rangle x_n \quad \text{for all } x \in \mathcal{H}.$$

(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 $\mathcal{H}$ can be interpreted as 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 11.9,

$$K = \sum_{\lambda \in \sigma(K)} \lambda P_\lambda.$$

This is the *spectral decomposition* of K, which will be shown in more generality in Functional Analysis II.

(4) The image Kx of some $x \in \mathcal{H}$ under the operator K is completely determined by the eigenvalues and eigenvectors of K.

$\lozenge$

References

[1] Hans Wilhelm Alt, *Lineare Funktionalanalysis. Eine anwendungsorientierte Einführung*, Springer, 2011.

[2] John B. Conway, *A course in functional analysis*, Springer, 1990.

[3] Christopher Heil, *A basis theory primer*, Birkhäuser, 2011.

[4] Harro Heuser, *Funktionalanalysis. Theorie und Anwendung*, Teubner, 2006.

[5] Walter Rudin, *Functional analysis*, McGraw-Hill, 1991.

[6] Dirk Werner, *Funktionalanalysis*, Springer, 2011.

Index