Skript zur Vorlesung

# Funktionalanalysis

von Ralph Chill
Universität Ulm

Wintersemester 2004/05

**Vorwort:**

Ich habe dieses Skript zur Vorlesung Funktionalanalysis an der Universität Ulm nach meinem besten Wissen und Gewissen geschrieben. Mit Sicherheit schlichen sich jedoch Druckfehler oder gar mathematische Ungenauigkeiten ein, die man beim ersten Schreiben eines Skripts nicht vermeiden kann. Möge man mir diese Fehler verzeihen.

Obwohl die Vorlesung auf Deutsch gehalten wird, habe ich mich entschieden, dieses Skript auf Englisch zu verfassen. Auf diese Weise wird eine Brücke zwischen der Vorlesung und der (meist englischsprachigen) Literatur geschlagen. Mathematik sollte jedenfalls unabhängig von der Sprache sein in der sie präsentiert wird.

Für Verbesserungen bin ich sehr dankbar.

# Contents

CHAPTER 1

# Some problems related to functional analysis

**Ordinary differential equations**

Let $f : D \to \mathbb{R}^N$ be continuous on a domain $D \subset \mathbb{R}^{1+N}$, and let $(t_0, y_0) \in D$. A basic problem in the theory of ordinary differential equations is to find a solution $y$ of the initial value problem

$$y'(t) = f(t, y(t)), \quad y(t_0) = y_0.$$

What is meant by a *solution*? Does a solution exist? Is it unique? How to prove existence and uniqueness?

**Integral equations**

Now let $k : [0,1] \times [0,1] \to \mathbb{R}$ and $g : [0,1] \to \mathbb{R}$ be continuous functions. A problem is to find a solution $y$ of the *integral equation*

$$y(t) = g(t) + \int_0^1 k(t,s)y(s)\, ds, \quad t \in [0,1].$$

The questions asked for the ordinary differential equation above can also be asked for this integral equation.

**Optimisation**

Let $j : [0,1] \times \mathbb{R} \to \mathbb{R}$ be a continuous, convex (in the second variable) function (which is bounded below). Find a function $u \in C$, where $C$ is a convenient convex set of real valued functions on the interval $[0,1]$ (e.g. the set of all continuous functions $v : [0,1] \to \mathbb{R}$ such that $\int_0^1 v = 1$), which minimizes the functional $J$ defined by

$$J(v) := \int_0^1 j(x, v(x))\, dx.$$

Does such a minimizer $u$ exist?

**Partial differential equations**

Let $\Omega \subset \mathbb{R}^N$ be an open bounded set, and let $f : \Omega \to \mathbb{R}$ be continuous and bounded. Find a solution $u$ of the partial differential equation

$$u(x) - \Delta u(x) = f(x), \quad x \in \Omega,$$

where $u$ satisfies in addition the *boundary condition* that $u = 0$ on the boundary $\partial\Omega$. Here $\Delta$ stands for the *Laplace operator*:

$$\Delta u(x) := \sum_{i=1}^N \frac{\partial^2}{\partial x_i^2} u(x).$$

Again we ask what is meant by a *solution*, how one can prove existence and uniqueness of a solution?

CHAPTER 2

# Primary on topology

It is the purpose of this introductory chapter to recall some basic facts about metric spaces, sequences in metric spaces, compact metric spaces, and continuous functions between metric spaces. Most of the material should be known, and if it is not known in the context of metric spaces, it has certainly been introduced on $\mathbb{R}^N$. The generalization to metric spaces should be straightforward, but it is nevertheless worthwhile to spend some time on the examples.

We also introduce some further notions from topology which may be new; see e.g. the definitions of density or of completion of a metric space.

## 1. Metric spaces

Definition 1.1. Let $M$ be a set. We call a function $d : M \times M \to \mathbb{R}_+$ a *metric* or a *distance* on $M$ if for every $x$, $y$, $z \in M$

(i) $d(x, y) = 0$ if and only if $x = y$,
(ii) $d(x, y) = d(y, x)$ (symmetry), and
(iii) $d(x, y) \leq d(x, z) + d(z, y)$ (triangle inequality).

A pair $(M, d)$ of a set $M$ and a metric $d$ on $M$ is called a *metric space*.

It will be convenient to write only $M$ instead of $(M, d)$ if the metric $d$ on $M$ is known, and to speak of a metric space $M$.

Example 1.2. (1) Let $M \subset \mathbb{R}^N$ and

$$d(x, y) := \sum_{i=1}^{N} |x_i - y_i|$$

or

$$d(x, y) := \left( \sum_{i=1}^{N} |x_i - y_i|^2 \right)^{\frac{1}{2}} .$$

Then $(M, d)$ is a metric space. The second metric is called the *euclidean metric*. Often, if the metric on $\mathbb{R}^N$ is not explicitly given, we mean the euclidean metric.

(2) Let $M \subset C([0, 1])$, the space of all continuous functions on the interval $[0, 1]$ and

$$d(f, g) := \sup_{x \in [0,1]} |f(x) - g(x)|.$$

Then $(M, d)$ is a metric space.

(3) Let $M$ be any set and

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

Then $(M,d)$ is a metric space. The metric $d$ is called the *discrete metric*.

(4) Let $(M,d)$ be a metric space. Then

$$d_1(x,y) := \frac{d(x,y)}{1 + d(x,y)}$$

and

$$d_2(x,y) := \min\{d(x,y), 1\}$$

define also metrics on $M$.

(5) Let $M = C(\mathbb{R})$, the space of all continuous functions on $\mathbb{R}$, and let

$$d_n(f,g) := \sup_{x \in [-n,n]} |f(x) - g(x)| \quad (n \in \mathbb{N})$$

and

$$d(f,g) := \sum_{n \in \mathbb{N}} 2^{-n} \frac{d_n(f,g)}{1 + d_n(f,g)}.$$

Then $(M,d)$ is a metric space. Note that the functions $d_n$ are not metrics for any $n \in \mathbb{N}$!

(6) Let $(M,d)$ be a metric space. Then any subset $\tilde{M} \subset M$ is a metric space for the *induced metric*

$$\tilde{d}(x,y) = d(x,y), \quad x,\, y \in \tilde{M}.$$

We may sometimes say that $\tilde{M}$ is a *subspace* of $M$, i.e. a subset and a metric space, but certainly this is not to be understood in the linear sense of vector spaces ($M$ need not be a vector space).

(7) Let $(M_n, d_n)$ be metric spaces ($n \in \mathbb{N}$). Then the cartesian product $M := \bigotimes_{n \in \mathbb{N}} M_n$ is a metric space for the metric

$$d(x,y) := \sum_{n \in \mathbb{N}} 2^{-n} \min\{d_n(x_n, y_n), 1\}.$$

Clearly, in a similar way, every finite cartesian product of metric spaces is a metric space.

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

(a) For every $x \in M$ and every $r > 0$ we define the *open ball* $B(x,r) := \{y \in M : d(x,y) < r\}$ with center $x$ and radius $r$.
(b) A set $O \subset M$ is called *open* if for every $x \in O$ there exists some $r > 0$ such that $B(x,r) \subset O$.
(c) A set $A \subset M$ is called *closed* if its complement $A^c = M \setminus A$ is open.
(d) A set $U \subset M$ is called a neighbourhood of $x \in M$ if there exists $r > 0$ such that $B(x,r) \subset U$.

REMARK 1.4. (a) The notions *open*, *closed*, *neighbourhood* depend on the set $M$!! For example, $M$ is always closed and open in $M$. The set $\mathbb{Q}$ is not closed in $\mathbb{R}$ (for the euclidean metric), but it is closed in $\mathbb{Q}$ for the induced metric! Therefore, one should always say in which metric space some given set is open or closed.
(b) Clearly, a set $O \subset M$ is open (in $M$) if and only if it is a neighbourhood of every of its elements.

LEMMA 1.5. *Let $(M, d)$ be a metric space. The following are true:*

(i) *Arbitrary unions of open sets are open. I.e.: if $(O_i)_{i \in I}$ is an arbitrary family of open sets (no restrictions on the index set $I$), then $\bigcup_{i \in I} O_i$ is open.*
(ii) *Arbitrary intersections of closed sets are closed. I.e.: if $(A_i)_{i \in I}$ is an arbitrary family of closed sets, then $\bigcap_{i \in I} A_i$ is closed.*
(iii) *Finite intersections of open sets are open.*
(iv) *Finite unions of closed sets are closed.*

PROOF. (i) Let $(O_i)_{i \in I}$ be an arbitrary family of open sets and let $O := \bigcup_{i \in I} O_i$. If $x \in O$, then $x \in O_i$ for some $i \in I$, and since $O_i$ is open, $B(x, r) \subset O_i$ for some $r > 0$. This implies that $B(x, r) \subset O$, and therefore $O$ is open.

(iii) Next let $(O_i)_{i \in I}$ be a finite family of open sets and let $O := \bigcap_{i \in I} O_i$. If $x \in O$, then $x \in O_i$ for every $i \in I$. Since the $O_i$ are open, there exist $r_i$ such that $B(x, r_i) \subset O_i$. Let $r := \min_{i \in I} r_i$ which is positive since $I$ is finite. By construction, $B(x, r) \subset O_i$ for every $i \in I$, and therefore $B(x, r) \subset O$, i.e. $O$ is open.

The proofs for closed sets are similar or follow just from the definition of closed sets and the above two assertions. □

EXERCISE 1.6. Determine all the open sets (closed sets) of a metric space $(M, d)$, where $d$ is the discrete metric.

EXERCISE 1.7. Show that a ball $B(x, r)$ in a metric space $M$ is always open. Show also that

$$\bar{B}(x, r) := \{y \in M : d(x, y) \leq r\}$$

is always closed.

DEFINITION 1.8 (Closure, interior, boundary). Let $(M, d)$ be a metric space and let $S \subset M$ be a subset. Then the set $\bar{S} := \bigcap \{A : A \subset M \text{ is closed and } S \subset A\}$ is called the *closure* of $S$. The set $S^\circ := \bigcup \{O : O \subset M \text{ is open and } O \subset S\}$ is called the *interior* of $S$. Finally, we call $\partial S := \{x \in M : \forall \varepsilon > 0\, B(x, \varepsilon) \cap S \neq \emptyset \text{ and } B(x, \varepsilon) \cap S^c \neq \emptyset\}$ the *boundary* of $S$.

By Lemma 1.5, the closure of a set $S$ is always closed (arbitrary intersections of closed sets are closed). By definition, $\bar{S}$ is the smallest closed set which contains $S$. Similarly, the interior of a set $S$ is always open, and by definition it is the largest open set which is contained in $S$. Note that the interior might be empty.

Exercise 1.9. Give an example of a metric space $M$ and some $x \in M$ to show that $\bar{B}(x, r)$ need not coincide with the closure of $B(x, r)$.

Exercise 1.10. Let $(M, d)$ be a metric space and consider the metrics $d_1$ and $d_2$ from Example 1.2 (4). Show that the set of all open subsets, closed subsets or neighbourhoods of $M$ is the same for the three given metrics.

The set of all open subsets is also called the *topology* of $M$. The three metrics $d$, $d_1$ and $d_2$ thus induce the same topology. Sometimes it is good to know that one can pass from a given metric $d$ to a finite metric ($d_1$ and $d_2$ take only values between 0 and 1) without changing the topology.

## 2. Sequences, convergence

Throughout the following, sequences will be denoted by $(x_n)$. Only when it is necessary, we precise the index $n$; usually, $n \geq 0$ or $n \geq 1$, but sometimes we will also consider finite sequences or sequences indexed by $\mathbb{Z}$.

Definition 2.1. Let $(M, d)$ be a metric space.

(a) We call a sequence $(x_n) \subset M$ a *Cauchy sequence* if for every $\varepsilon > 0$ there exists $n_0$ such that for every $n$, $m \geq n_0$ one has $d(x_n, x_m) < \varepsilon$.

(b) We say that a sequence $(x_n) \subset M$ *converges* to some element $x \in M$ if for every $\varepsilon > 0$ there exists $n_0$ such that for every $n \geq n_0$ one has $d(x_n, x) < \varepsilon$. If $(x_n)$ converges to $x$, we also write $\lim_{n\to\infty} x_n = x$ or $x_n \to x$ as $n \to \infty$.

Exercise 2.2. Let $C([0, 1])$ be the metric space from Example 1.2 (2). Show that a sequence $(f_n) \subset C([0, 1])$ converges to some $f$ for the metric $d$ if and only if it converges uniformly. We say that the metric $d$ induces the topology of *uniform convergence*.

Show also that a sequence $(f_n) \subset C(\mathbb{R})$ (Example 1.2 (5)) converges to some $f$ for the metric $d$ if and only if it converges uniformly on compact subsets of $\mathbb{R}$. In this example, we say that the metric $d$ induces the topology of *local uniform convergence*.

Exercise 2.3. Determine all Cauchy sequences and all convergent sequences in a discrete metric space.

Lemma 2.4. *Let $M$ be a metric space and $(x_n) \subset M$ be a sequence. Then:*

(1) $\lim_{n\to\infty} x_n = x$ *for some element $x \in M$ if and only if for every neighbourhood $U$ of $x$ there exists $n_0$ such that for every $n \geq n_0$ one has $x_n \in U$.*

(2) (Uniqueness of the limit) *If* $\lim_{n\to\infty} x_n = x$ *and* $\lim_{n\to\infty} x_n = y$, *then* $x = y$.

Lemma 2.5. *A set $A \subset M$ is closed if and only if for every sequence $(x_n) \subset A$ which converges to some $x \in M$ one has $x \in A$.*

PROOF. Assume first that $A$ is closed and let $(x_n) \subset A$ be convergent to $x \in M$. If $x$ does not belong to $A$, then it belongs to $A^c$ which is open. By definition, there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subset A^c$. Given this $\varepsilon$, there exists $n_0$ such that $x_n \in B(x, \varepsilon)$ for every $n \geq n_0$, a contradiction to the assumption that $x_n \in A$. Hence, $x \in A$.

On the other hand, assume that $\lim_{n\to\infty} x_n = x \in A$ for every convergent $(x_n) \subset A$ and assume in addition that $A$ is not closed or, equivalently, that $A^c$ is not open. Then there exists $x \in A^c$ such that for every $n \in \mathbb{N}$ the set $B(x, \frac{1}{n}) \cap A$ is nonempty. From this one can construct a sequence $(x_n) \subset A$ which converges to $x$, which is a contradiction because $x \in A^c$. □

LEMMA 2.6. *Let $(M, d)$ be a metric space, and let $S \subset M$ be a subset. Then*

$$\begin{aligned} \bar{S} &= \{x \in M : \exists (x_n) \subset S \text{ s.t. } \lim_{n\to\infty} x_n = x\} \\ &= \{x \in M : d(x, S) := \inf_{y\in S} d(x, y) = 0\}. \end{aligned}$$

PROOF. Let

$$A := \{x \in M : \exists (x_n) \subset S \text{ s.t. } \lim_{n\to\infty} x_n = x\}$$

and

$$B := \{x \in M : d(x, S) := \inf_{y\in S} d(x, y) = 0\}.$$

These two sets are clearly equal by the definition of the inf and the definition of convergence. Moreover, the set $B$ is closed by the following argument. Assume that $(x_n) \subset B$ is convergent to $x \in M$. By definition of $B$, for every $n$ there exists $y \in S$ such that $d(x_n, y_n) \leq 1/n$. Hence,

$$\limsup_{n\to\infty} d(x, y_n) \leq \limsup_{n\to\infty} d(x, x_n) + \limsup_{n\to\infty} d(x_n, y_n) = 0,$$

so that $x \in B$.

Clearly, $B$ contains $S$, and since $B$ is closed, $B$ contains $\bar{S}$. It remains to show that $B \subset \bar{S}$. If this is not true, then there exists $x \in B \setminus \bar{S}$. Since the complement of $\bar{S}$ is open in $M$, there exists $r > 0$ such that $B(x, r) \cap \bar{S} = \emptyset$, a contradiction to the definition of $B$. □

DEFINITION 2.7. A metric space $(M, d)$ is called *complete* if every Cauchy sequence converges.

EXERCISE 2.8. Show that the spaces $\mathbb{R}^N$, $C([0, 1])$ and $C(\mathbb{R})$ are complete. Show also that any discrete metric space is complete.

LEMMA 2.9. *A subspace $N \subset M$ of a complete metric space is complete if and only if it is closed in $M$.*

PROOF. Assume that $N \subset M$ is closed, and let $(x_n)$ be a Cauchy sequence in $N$. By the assumption that $M$ is complete, $(x_n)$ is convergent to some element $x \in M$. Since $N$ is closed, $x \in N$.

Assume on the other hand that $N$ is complete, and let $(x_n) \subset N$ be convergent to some element $x \in M$. Clearly, every convergent sequence is also a Cauchy sequence, and since $N$ is complete, $(x_n)$ converges to some element $y \in N$. By uniqueness of the limit, $x = y \in N$. Hence, $N$ is closed. □

## 3. Compactness

DEFINITION 3.1. We say that a metric space $(M, d)$ is *compact* if for every open covering there exists a finite subcovering, i.e. whenever $(O_i)_{i \in I}$ is a family of open sets (no restrictions on the index set $I$) such that $M = \bigcup_{i \in I} O_i$, then there exists a *finite* subset $I_0 \subset I$ such that $M = \bigcup_{i \in I_0} O_i$.

LEMMA 3.2. *A metric space $(M, d)$ is compact if and only if it is* sequentially compact*, i.e. if and only if every sequence $(x_n) \subset M$ has a convergent subsequence.*

PROOF. Assume that $M$ is compact and let $(x_n) \subset M$. Assume that $(x_n)$ does not have a convergent subsequence. Then for every $x \in M$ there exists $\varepsilon_x > 0$ such that $B(x, \varepsilon_x)$ contains only finitely many elements of $\{x_n\}$. Note that $(B(x, \varepsilon_x))_{x \in M}$ is an open covering of $M$ so that by the compactness of $M$ there exists a finite subset $N \subset M$ such that $M = \bigcup_{x \in N} B(x, \varepsilon_x)$. But this means that $(x_n)$ takes only finitely many values, and hence there exists even a constant subsequence which is in particular also convergent; a contradiction to the assumption on $(x_n)$.

On the other hand, assume that $M$ is sequentially compact and let $(O_i)_{i \in I}$ be an open covering of $M$. We first show that there exists $\varepsilon > 0$ such that for every $x \in M$ there exists $i_x \in I$ with $B(x, \varepsilon) \subset O_{i_x}$. If this were not true, then for every $n \in \mathbb{N}$ there exists $x_n$ such that $B(x_n, \frac{1}{n}) \not\subset O_i$ for every $i \in I$. Passing to a subsequence, we may assume that $(x_n)$ is convergent to some $x \in M$. There exists some $i_0 \in I$ such that $x \in O_{i_0}$, and since $O_{i_0}$ is open, we find some $\varepsilon > 0$ such that $B(x, \varepsilon) \subset O_{i_0}$. Let $n_0$ be such that $\frac{1}{n_0} < \frac{\varepsilon}{2}$. By the triangle inequality, for every $n \geq n_0$ we have $B(x_n, \frac{1}{n}) \subset B(x, \varepsilon) \subset O_{i_0}$, a contradiction to the construction of the sequence $(x_n)$.

Next we show that $M = \bigcup_{j=1}^{n} B(x_j, \varepsilon)$ for a finite family of $x_j \in M$. Choose any $x_1 \in M$. If $B(x_1, \varepsilon) = M$, then we are already done. Otherwise we find $x_2 \in M \setminus B(x_1, \varepsilon)$. If $B(x_1, \varepsilon) \cup B(x_2, \varepsilon) \neq M$, then we even find $x_3 \in M$ which does not belong to $B(x_1, \varepsilon) \cup B(x_2, \varepsilon)$, and so on. If $\bigcup_{j=1}^{n} B(x_j, \varepsilon)$ is never all of $M$, then we find actually a sequence $(x_j)$ such that $d(x_j, x_k) \geq \varepsilon$ for all $j \neq k$. This sequence can not have a convergent subsequence, a contradiction to sequential compactness.

Since every of the $B(x_j, \varepsilon)$ is a subset of $O_{i_{x_j}}$ for some $i_{x_j} \in I$, we have proved that $M = \bigcup_{j=1}^{n} O_{i_{x_j}}$, i.e. the open covering $(O_i)$ admits a finite subcovering. The proof is complete. □

LEMMA 3.3. *Any compact metric space is complete.*

Proof. Let $(x_n)$ be a Cauchy sequence in $M$. By the preceeding lemma, there exists a subsequence which converges to some $x \in M$. If a subsequence of a Cauchy sequence converges, then the sequence itself converges, too. $\square$

## 4. Continuity

Definition 4.1. Let $(M_1, d_1)$, $(M_2, d_2)$ be two metric spaces, and let $f : M_1 \to M_2$ be a function.

(a) We say that $f$ is *continuous in some point* $x \in M_1$ if

$$\forall \varepsilon > 0 \, \exists \delta > 0 \, \forall y \in B(x, \delta) : d_2(f(x), f(y)) < \varepsilon.$$

(b) We say that $f$ is *continuous* if it is continuous in every point.

(c) We say that $f$ is *uniformly continuous* if

$$\forall \varepsilon > 0 \, \exists \delta > 0 \, \forall x, \, y \in M_1 : d_1(x, y) < \delta \Rightarrow d_2(f(x), f(y)) < \varepsilon.$$

(d) We say that $f$ is *Lipschitz continuous* if

$$\exists L \geq 0 \, \forall x, \, y \in M : \, d_2(f(x), f(y)) \leq L \, d_1(x, y).$$

Lemma 4.2. *A function $f : M_1 \to M_2$ between two metric spaces is continuous in some point $x \in M_1$ if and only if it is* sequentially continuous *in $x$, i.e. if and only if for every sequence $(x_n) \subset M_1$ which converges to $x$ one has $\lim_{n \to \infty} f(x_n) = f(x)$.*

Proof. Assume that $f$ is continuous in $x \in M_1$ and let $(x_n)$ be convergent to $x$. Let $\varepsilon > 0$. There exists $\delta > 0$ such that for every $y \in B(x, \delta)$ one has $f(y) \in B(f(x), \varepsilon)$. By definition of convergence, there exists $n_0$ such that for every $n \geq n_0$ one has $x_n \in B(x, \delta)$. For this $n_0$ and every $n \geq n_0$ one has $f(x_n) \in B(f(x), \varepsilon)$. Hence, $\lim_{n \to \infty} f(x_n) = f(x)$.

Assume on the other hand that $f$ is sequentially continuous in $x$. If $f$ was not continuous in $x$ then there exists $\varepsilon > 0$ such that for every $n \in \mathbb{N}$ there exists $x_n \in B(x, \frac{1}{n})$ with $f(x_n) \notin B(f(x), \varepsilon)$. By construction, $\lim_{n \to \infty} x_n = x$. Since $f$ is sequentially continuous, $\lim_{n \to \infty} f(x_n) = f(x)$. But this is a contradiction to $f(x_n) \notin B(f(x), \varepsilon)$, and therefore $f$ is continuous. $\square$

Lemma 4.3. *A function $f : M_1 \to M_2$ between two metric spaces is continuous if and only if preimages of open sets are open, i.e. if and only if for every open set $O \subset M_2$ the preimage $f^{-1}(O)$ is open in $M_1$.*

Proof. Let $f : M_1 \to M_2$ be continuous and let $O \subset M_2$ be open. Let $x \in f^{-1}(O)$. Since $O$ is open, there exists $\varepsilon > 0$ such that $B(f(x), \varepsilon) \subset O$. Since $f$ is continuous, there exists $\delta > 0$ such that for every $y \in B(x, \delta)$ one has $f(y) \in B(f(x), \varepsilon)$. Hence, $B(x, \delta) \subset f^{-1}(O)$ so that $f^{-1}(O)$ is open.

On the other hand, if the preimage of every open set is open, then for every $x \in M_1$ and every $\varepsilon > 0$ the preimage $f^{-1}(B(f(x), \varepsilon))$ is open. Clearly, $x$ belongs to this preimage, and therefore there exists $\delta > 0$ such that $B(x, \delta) \subset f^{-1}(B(f(x), \varepsilon))$. This proves continuity. $\square$

Lemma 4.4. *Let $f : K \to M$ be a continuous function from a compact metric space $K$ into a metric space $M$. Then*

1. *The image $f(K)$ is compact.*
2. *The function $f$ is uniformly continuous.*

Proof. (1) Let $(O_i)_{i\in I}$ be an open covering of $f(K)$. Since $f$ is continuous, $f^{-1}(O_i)$ is open in $K$. Moreover, $(f^{-1}(O_i))_{i\in I}$ is an open covering of $K$. Since $K$ is compact, there exists a finite subcovering: $K = \bigcup_{i\in I_0} f^{-1}(O_i)$ for some finite $I_0 \subset I$. Hence, $(O_i)_{i\in I_0}$ is a finite subcovering of $f(K)$.

(2) Let $\varepsilon > 0$. Since $f$ is continuous, for every $x \in K$ there exists $\delta_x > 0$ such that for all $y \in B(x, \delta_x)$ one has $f(y) \in B(f(x), \varepsilon)$. By compactness, there exists a finite family $(x_i)_{1\le i\le n} \subset K$ such that $K = \bigcup_{i=1}^n B(x_i, \delta_{x_i}/2)$. Let $\delta = \min\{\delta_{x_i}/2 : 1 \le i \le n\}$ and let $x$, $y \in K$ such that $d(x,y) < \delta$. Since $x \in B(x_i, \delta_{x_i}/2)$ for some $1 \le i \le n$, we find that $y \in B(x_i, \delta_{x_i})$. By construction, $f(x)$, $f(y) \in B(f(x_i), \varepsilon)$ so that the triangle inequality implies $d(f(x), f(y)) < 2\varepsilon$. □

Lemma 4.5. *Any Lipschitz continuous function $f : M_1 \to M_2$ between two metric spaces is uniformly continuous.*

Proof. Let $L > 0$ be a Lipschitz constant for $f$ and let $\varepsilon > 0$. Define $\delta := \varepsilon/L$. Then, for every $x$, $y \in M$ such that $d_1(x,y) \le \delta$ one has

$$d_2(f(x), f(y)) \le L d_1(x,y) \le \varepsilon,$$

and therefore $f$ is uniformly continuous. □

## 5. Completion of a metric space

Definition 5.1. A subset $D \subset M$ of a metric space $(M,d)$ is called *dense in $M$* if $\bar{D} = M$. Equivalently, $D$ is dense in $M$ if for every $x \in M$ there exists $(x_n) \subset D$ such that $\lim_{n\to\infty} x_n = x$.

Lemma 5.2 (Completion). *Let $(M,d)$ be a metric space. Then there exists a complete metric space $(\hat{M}, \hat{d})$ and a continuous, injective $j : M \to \hat{M}$ such that*

$$d(x,y) = \hat{d}(j(x), j(y)), \quad x,\, y \in M,$$

*and such that the image $j(M)$ is dense in $\hat{M}$.*

Definition 5.3. Let $(M,d)$ be a metric space. A complete metric space $(\hat{M}, \hat{d})$ fulfilling the properties from Lemma 5.2 is called a *completion* of $M$.

Proof of Lemma 5.2. Let

$$\bar{M} := \{(x_n) \subset M : (x_n) \text{ is a Cauchy sequence}\}.$$

We say that two Cauchy sequences $(x_n)$, $(y_n) \subset \bar{M}$ are equivalent (and we write $(x_n) \sim (y_n)$) if $\lim_{n\to\infty} d(x_n, y_n) = 0$. Clearly, $\sim$ is an equivalence relation on $\bar{M}$.

We denote by $[(x_n)]$ the equivalence class in $\bar{M}$ of a Cauchy sequence $(x_n)$, and we let

$$\hat{M} := \bar{M}/\sim = \{[(x_n)] : (x_n) \in \bar{M}\}$$

be the set of all equivalence classes. If we define

$$\hat{d}([(x_n)], [(y_n)]) := \lim_{n\to\infty} d(x_n, y_n),$$

then $\hat{d}$ is well defined (the definition is independent of the choice of representatives) and it is a metric on $\hat{M}$. The fact that $\hat{d}$ is a metric and also that $(\hat{M}, \hat{d})$ is a complete metric space are left as exercises.

One also easily verifies that $j : M \to \hat{M}$ defined by $j(x) = [(x)]$ (the equivalence class of the constant sequence $(x)$) is continuous, injective and in fact isometric, i.e.

$$d(x, y) = \hat{d}(j(x), j(y))$$

for every $x$, $y \in M$. The proof is here complete. □

Lemma 5.4. *Let $(\hat{M}_i, \hat{d}_i)$ ($i = 1, 2$) be two completions of a metric space $(M, d)$. Then there exists a bijection $b : \hat{M}_1 \to \hat{M}_2$ such that for every $x$, $y \in \hat{M}_1$*

$$\hat{d}_1(x, y) = \hat{d}_2(b(x), b(y)).$$

Lemma 5.4 shows that up to isometric bijections there exists only one completion of a given metric space and it allows us to speak of *the* completion of a metric space.

Lemma 5.5. *Let $f : M_1 \to M_2$ be a uniformly (!) continuous function between two metric spaces. Let $\hat{M}_1$ and $\hat{M}_2$ be the completions of $M_1$ and $M_2$, respectively. Then there exists a unique continuous extension $\hat{f} : \hat{M}_1 \to \hat{M}_2$ of $f$.*

Proof. Since $f$ is uniformly continuous, it maps equivalent Cauchy sequences into equivalent Cauchy sequences (equivalence of Cauchy sequences is defined as in the proof of Lemma 5.2). Hence, the function $\hat{f}([(x_n)]) := [(f(x_n))]$ is well defined. It is easy to check that $\hat{f}$ is an extension of $f$ and that $\hat{f}$ is continuous (even uniformly continuous). □

The assumption of uniform continuity in Lemma 5.5 is necessary in general. The functions $f(x) = \sin(1/x)$ and $f(x) = 1/x$ on the open interval $(0, 1)$ do not admit continuous extensions to the closed interval $[0, 1]$ (which is the completion of $(0, 1)$).

CHAPTER 3

# Banach spaces and bounded linear operators

Throughout, let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

## 1. Normed spaces

DEFINITION 1.1. Let $X$ be a vector space over $\mathbb{K}$. A function $\|\cdot\| : X \to \mathbb{R}_+$ is called a *norm* if for every $x, y \in X$ and every $\lambda \in \mathbb{K}$

(i) $\|x\| = 0$ if and only if $x = 0$,
(ii) $\|\lambda x\| = |\lambda| \, \|x\|$, and
(iii) $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality).

A pair $(X, \|\cdot\|)$ of a vector space $X$ and a norm $\|\cdot\|$ is called a *normed space*.

Often, we will speak of a normed space $X$ if it is clear which norm is given on $X$.

EXAMPLE 1.2. (1) (Finite dimensional spaces) Let $X = \mathbb{K}^N$. Then

$$\|x\|_p := \left( \sum_{i=1}^{N} |x_i|^p \right)^{1/p}, \quad 1 \leq p < \infty,$$

and

$$\|x\|_\infty := \sup_{1 \leq i \leq N} |x_i|$$

are norms on $X$.

(2) (Sequence spaces) Let $1 \leq p < \infty$, and let

$$l^p := \{(x_n) \subset \mathbb{K} : \sum_n |x_n|^p < \infty\}$$

with norm

$$\|x\|_p := \left( \sum_n |x_n|^p \right)^{1/p}.$$

Then $(l^p, \|\cdot\|_p)$ is a normed space.

(3) (Sequence spaces) Let $X$ be one of the spaces

$$\begin{aligned} l^\infty &:= \{(x_n) \subset \mathbb{K} : \sup_n |x_n| < \infty\}, \\ c &:= \{(x_n) \subset \mathbb{K} : \lim_{n\to\infty} x_n \text{ exists}\}, \text{ or} \\ c_0 &:= \{(x_n) \subset \mathbb{K} : \lim_{n\to\infty} x_n = 0\}, \text{ or} \\ c_{00} &:= \{(x_n) \subset \mathbb{K} : \text{the set } \{n : x_n \neq 0\} \text{ is finite}\}, \end{aligned}$$

and let

$$\|x\|_\infty := \sup_n |x_n|.$$

Then $(X, \|\cdot\|_\infty)$ is a normed space.

(4) (Function spaces: continuous functions) Let $C([a,b])$ be the space of all continuous, $\mathbb{K}$-valued functions on a compact interval $[a,b] \subset \mathbb{R}$. Then

$$\|f\|_p := \left(\int_a^b |f(x)|^p \, dx\right)^{1/p}, \quad 1 \leq p < \infty,$$

and

$$\|f\|_\infty := \sup_{x\in[a,b]} |f(x)|$$

are norms on $C([a,b])$.

(5) (Function spaces: continuous functions) Let $K$ be a compact metric space and let $C(K)$ be the space of all continuous, $\mathbb{K}$-valued functions on $K$. Then

$$\|f\|_\infty := \sup_{x\in K} |f(x)|$$

is a norm on $C(K)$.

(6) (Function spaces: integrable functions) Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and let $X_p = L^p(\Omega)$ $(1 \leq p \leq \infty)$. Let

$$\|f\|_p := \left(\int_\Omega |f|^p \, d\mu\right)^{1/p}, \quad 1 \leq p < \infty,$$

or

$$\|f\|_\infty := \text{ess sup}|f(x)| := \inf\{c \in \mathbb{R}_+ : \mu(\{|f| > c\}) = 0\}.$$

Then $(X_p, \|\cdot\|_p)$ is a normed space.

(7) (Function spaces: differentiable functions) Let

$$C^1([a,b]) := \{f \in C([a,b]) : f \text{ is continuously differentiable}\}.$$

Then $\|\cdot\|_\infty$ and

$$\|f\|_{C^1} := \|f\|_\infty + \|f'\|_\infty$$

are norms on $C^1([a,b])$.

We will see more examples in the sequel.

LEMMA 1.3. *Every normed space is a metric space for the metric*

$$d(x,y) := \|x-y\|, \quad x,\, y \in X.$$

By the above lemma, also every subset of a normed space becomes a metric space in a natural way. Moreover, it is natural to speak of closed or open subsets (or linear subspaces!) of normed spaces, or of closures and interiors of subsets.

EXERCISE 1.4. Show that in a normed space $X$, for every $x \in X$ and every $r > 0$ the closed ball $\bar{B}(x,r)$ coincides with closure $\overline{B(x,r)}$ of the open ball.

Also the notion of continuity of functions between normed spaces (or between a metric space and a normed space) makes sense. The following is a first example of a continuous function.

LEMMA 1.5. *Given a normed space, the norm is a continuous function.*

This lemma is a consequence of the following lemma.

LEMMA 1.6 (Triangle inequality below). *Let $X$ be a normed space. Then, for every $x,\, y \in X$,*

$$\|x-y\| \geq \big|\|x\| - \|y\|\big|.$$

PROOF. The triangle inequality implies

$$\begin{aligned} \|x\| &= \|x-y+y\| \\ &\leq \|x-y\| + \|y\|, \end{aligned}$$

so that

$$\|x\| - \|y\| \leq \|x-y\|.$$

Changing the role of $x$ and $y$ implies

$$\|y\| - \|x\| \leq \|y-x\| = \|x-y\|,$$

and the claim follows. □

A notion which can not really be defined in metric spaces but in normed spaces is the following.

DEFINITION 1.7. A subset $B$ of a normed space $X$ is called *bounded* if

$$\sup\{\|x\| : x \in B\} < \infty.$$

It is easy to check that if $X$ is a normed space, and $M$ is a metric space, then the set $C(M;X)$ of all continuous functions from $M$ into $X$ is a vector space for the obvious addition and scalar mutliplication. If $M$ is in addition compact, then $f(M) \subset X$ is also compact for every such function, and hence $f(M)$ is necessarily bounded (every compact subset of a normed space is bounded!). So we can define a new example of a normed space.

EXAMPLE 1.8. (8) (Function spaces: vector-valued continuous functions) Let $(X, \|\cdot\|)$ be a normed space and let $K$ be a compact metric space. Let $E = C(K; X)$ be the space of all $X$-valued continuous functions on $K$. Then

$$\|f\|_\infty := \sup_{x\in K} \|f(x)\|$$

is a norm on $C(K; X)$.

Also the notions of Cauchy sequences and convergent sequences make sense in normed spaces. In particular, one can speak of a complete normed space, i.e. a normed space in which every Cauchy sequence converges.

DEFINITION 1.9. A complete normed space is called a *Banach space*.

EXAMPLE 1.10. The finite dimensional spaces, the sequence spaces $l^p$ $(1 \le p \le \infty)$, $c$, and $c_0$, and the function spaces $(C([a, b]), \|\cdot\|_\infty)$, $(L^p(\Omega), \|\cdot\|_p)$ are Banach spaces.

The spaces $(c_{00}, \|\cdot\|_\infty)$, $(C([a, b]), \|\cdot\|_p)$ $(1 \le p < \infty)$ are no Banach spaces.

It is not too difficult to check that if $X$ is a Banach space, then also $(C(K; X), \|\cdot\|_\infty)$ is a Banach space.

DEFINITION 1.11. We say that two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a real or complex vector space $X$ are equivalent if there exist two constants $c$, $C > 0$ such that for every $x \in X$

$$c\,\|x\|_1 \le \|x\|_2 \le C\,\|x\|_1.$$

LEMMA 1.12. *Let $\|\cdot\|_1$, $\|\cdot\|_2$ be two norms on a vector space $X$ (over $\mathbb{K}$). The following are equivalent:*

(1) *The norms $\|\cdot\|_1$, $\|\cdot\|_2$ are equivalent.*
(2) *A set $O \subset X$ is open for the norm $\|\cdot\|_1$ if and only if it is open for the norm $\|\cdot\|_2$ (and similarly for closed sets).*
(3) *A sequence $(x_n) \subset X$ converges to $0$ for the norm $\|\cdot\|_1$ if and only if it converges to $0$ for the norm $\|\cdot\|_2$.*

In other words, if two norms $\|\cdot\|_1$, $\|\cdot\|_2$ on a vector space $X$ are equivalent, then the open sets, the closed sets and the null sequences are the same. We also say that the two norms define the same *topology*. In particular, if $X$ is a Banach space for one norm then it is also a Banach space for the other (equivalent) norm.

EXERCISE 1.13. The norms $\|\cdot\|_\infty$ and $\|\cdot\|_p$ are not equivalent on $C([0, 1])$.

THEOREM 1.14. *Any two norms on a finite dimensional real or complex vector space are equivalent.*

PROOF. We may without loss of generality consider $\mathbb{K}^N$. Let $\|\cdot\|$ be a norm on $\mathbb{K}^N$ and let $(e_i)_{1\le i\le N}$ be the canonical basis of $\mathbb{K}^N$. For every $x \in \mathbb{K}^N$

$$\begin{aligned} \|x\| &= \|\sum_{i=1}^N x_i e_i\| \\ &\le \sum_{i=1}^N |x_i|\,\|e_i\| \\ &\le C\,\|x\|_1, \end{aligned}$$

where $C := \sup_{1\le i\le N} \|e_i\|$ and $\|\cdot\|_1$ is the norm from Example (1). By the triangle inequality from below, for every $x, y \in \mathbb{K}^N$,

$$|\|x\| - \|y\|| \le \|x-y\| \le C\,\|x-y\|_1.$$

Hence, the norm $\|\cdot\| : (\mathbb{K}^N, \|\cdot\|_1) \to \mathbb{R}_+$ is continuous (on $\mathbb{K}^N$ equipped with the norm $\|\cdot\|_1$). If $S := \{x \in \mathbb{K}^N : \|x\|_1 = 1\}$ denotes the unit sphere for the norm $\|\cdot\|_1$, then $S$ is compact and

$$c := \inf\{\|x\| : x \in S\} > 0,$$

since the infimum is attained by the continuity of $\|\cdot\|$. This implies

$$c\,\|x\|_1 \le \|x\| \text{ for every } x \in \mathbb{K}^N.$$

We have proved that every norm on $\mathbb{K}^N$ is equivalent to the norm $\|\cdot\|_1$. Hence, any two norms on $\mathbb{K}^N$ are equivalent. □

COROLLARY 1.15. *Any finite dimensional normed space is complete. Any finite dimensional subspace of a normed space is closed.*

PROOF. The space $(\mathbb{K}^N, \|\cdot\|_1)$ is complete (exercise!). If $\|\cdot\|$ is a second norm on $\mathbb{K}^N$ and if $(x_n)$ is a Cauchy sequence for that norm, then it is also a Cauchy sequence in $(\mathbb{K}^N, \|\cdot\|_1)$ (use that the norms $\|\cdot\|_1$ and $\|\cdot\|$ are equivalent), and therefore convergent in $(\mathbb{K}^N, \|\cdot\|_1)$. By equivalence of norms again, the sequence $(x_n)$ is also convergent in $(\mathbb{K}^N, \|\cdot\|)$, and therefore $(\mathbb{K}^N, \|\cdot\|)$ is complete.

Let $Y$ be a finite dimensional subspace of a normed space $X$, and let $(x_n) \subset Y$ be a convergent sequence with $x = \lim_{n\to\infty} x_n \in X$. Since $(x_n)$ is also a Cauchy sequence, and since $Y$ is complete, we find (by uniqueness of the limit) that $x \in Y$, and therefore $Y$ is closed (Lemma 2.5). □

DEFINITION 1.16. Let $(x_n)$ be a sequence in a normed space $X$. We say that the series $\sum_n x_n$ is *convergent* if the sequence $(\sum_{j\le n} x_j)$ of partial sums is convergent. We say that the series $\sum_n x_n$ is *absolutely convergent* if $\sum_n \|x_n\| < \infty$.

LEMMA 1.17. *Let $(x_n)$ be a sequence in a normed space $X$. If the series $\sum_n x_n$ is convergent, then necessarily $\lim_{n\to\infty} x_n = 0$.*

Note that in a normed space not every absolutely convergent series is convergent. In fact, the following is true.

LEMMA 1.18. *A normed space $X$ is a Banach space if and only if every absolutely convergent series converges.*

PROOF. Assume that $X$ is a Banach space, and let $\sum_n x_n$ be absolutely convergent. It follows easily from the triangle inequality that the corresponding sequence of partial sums is a Cauchy sequence, and since $X$ is complete, the series $\sum_n x_n$ is convergent.

On the other hand, assume that every absolutely convergent series is convergent. Let $(x_n)_{n\geq 1} \subset X$ be a Cauchy sequence. From this Cauchy sequence, one can extract a subsequence $(x_{n_k})_{k\geq 1}$ such that $\|x_{n_{k+1}} - x_{n_k}\| \leq 2^{-k}$, $k \geq 1$. Let $y_0 = x_{n_1}$ and $y_k = x_{n_{k+1}} - x_{n_k}$, $k \geq 1$. Then the series $\sum_{k\geq 0} y_k$ is absolutely convergent. By assumption, it is also convergent. But by construction, $(\sum_{l=0}^{k} y_l) = (x_{n_k})$, so that $(x_{n_k})$ is convergent. Hence, we have extracted a subsequence of the Cauchy sequence $(x_n)$ which converges. As a consequence, $(x_n)$ is convergent, and since $(x_n)$ was an arbitrary Cauchy sequence, $X$ is complete. □

LEMMA 1.19 (Riesz). *Let $X$ be a normed space and let $Y \subset X$ be a closed linear subspace. If $Y \neq X$, then for every $\delta > 0$ there exists $x \in X \setminus Y$ such that $\|x\| = 1$ and*

$$\operatorname{dist}(x, Y) = \inf\{\|x - y\| : y \in Y\} \geq 1 - \delta.$$

PROOF. Let $z \in X \setminus Y$. Since $Y$ is closed,

$$d := \operatorname{dist}(z, Y) > 0.$$

Let $\delta > 0$. By definition of the infimum, there exists $y \in Y$ such that

$$\|z - y\| \leq \frac{d}{1 - \delta}.$$

Let $x := \frac{z-y}{\|z-y\|}$. Then $x \in X \setminus Y$, $\|x\| = 1$, and for every $u \in Y$

$$\begin{aligned} \|x - u\| &= \|z - y\|^{-1} \|z - (y + \|z - y\|u)\| \\ &\geq \|z - y\|^{-1} d \quad \geq \quad 1 - \delta, \end{aligned}$$

since $(y + \|z - y\|u) \in Y$. □

THEOREM 1.20. *A normed space is finite dimensional if and only if every closed bounded set is compact.*

PROOF. If the normed space is finite dimensional, then every closed bounded set is compact by the Theorem of Heine-Borel. Note that by Theorem 1.14 it is not important which norm on the finite dimensional space is considered. By Lemma 1.12, the closed and bounded sets do not change.

On the other hand, if the normed space is infinite dimensional, then, by the Lemma of Riesz, one can construct inductively a sequence $(x_n) \subset X$ such that

$\|x_n\| = 1$ and $\operatorname{dist}(x_{n+1}, X_n) \geq \frac{1}{2}$ for every $n \in \mathbb{N}$, where $X_n = \operatorname{span}\{x_i : 1 \leq i \leq n\}$ (note that $X_n$ is closed by Corollary 1.15). By construction, $(x_n)$ belongs to the closed unit ball, but it can not have a convergent subsequence (even not a Cauchy subsequence). Hence, the closed unit ball is not compact. □

LEMMA 1.21 (Completion of a normed space). *For every normed space $X$ there exists a Banach space $\hat{X}$ and a linear injective $j : X \to \hat{X}$ such that $\|j(x)\| = \|x\|$ ($x \in X$) and $j(X)$ is dense in $\hat{X}$.* Up to isometry, *the Banach space $\hat{X}$ is unique. It is called the* completion *of $X$.*

PROOF. It suffices to repeat the proof of Lemma 5.2 and to note that the completion $\hat{X}$ of $X$ (considered as a metric space) carries in a natural way a linear structure: addition of - equivalence classes of - Cauchy sequences is their componentwise addition, and also multiplication of - an equivalence class - of a Cauchy sequence and a scalar is done componentwise. Moreover, for every $[(x_n)]$, one defines the norm

$$\|[(x_n)]\| := \lim_{n\to\infty} \|x_n\|.$$

Uniqueness of $\hat{X}$ follows from Lemma 5.4. □

## 2. Product spaces and quotient spaces

LEMMA 2.1 (Product spaces). *Let $(X_i)_{i\in I}$ be a finite (!) family of normed spaces, and let $\mathcal{X} := \bigotimes_{i\in I} X_i$ be the cartesian product. Then*

$$\|x\|_p := \left(\sum_{i\in I} \|x_i\|_{X_i}^p\right)^{1/p} \qquad (1 \leq p < \infty),$$

*and*

$$\|x\|_\infty := \sup_{i\in I} \|x_i\|_{X_i}$$

*define equivalent norms on $\mathcal{X}$. In particular, the cartesian product is a normed space.*

PROOF. The easy proof is left to the reader. □

LEMMA 2.2. *Let $(X_i)_{i\in I}$ be a finite family of normed spaces, and let $\mathcal{X} := \bigotimes_{i\in I} X_i$ be the cartesian product equipped with one of the equivalent norms $\|\cdot\|_p$ from Lemma 2.1. Then a sequence $(x^n) = ((x_i^n)_i) \subset \mathcal{X}$ converges (is a Cauchy sequence) if and only if $(x_i^n) \subset X_i$ is convergent (is a Cauchy sequence) for every $i \in I$.*

*As a consequence, $\mathcal{X}$ is a Banach space if and only if all the $X_i$ are Banach spaces.*

PROPOSITION 2.3 (Quotient space). *Let $X$ be a vector space (!) over $\mathbb{K}$, and let $Y \subset X$ be a linear subspace. Define, for every $x \in X$, the affine subspace*

$$x + Y := \{x + y : y \in Y\},$$

*and define the* quotient space *or* factor space

$$X/Y := \{x + Y : x \in X\}.$$

*Then $X/Y$ is a vector space for the addition*

$$(x + Y) + (z + Y) := (x + z + Y),$$

*and the scalar multiplication*

$$\lambda(x + Y) := (\lambda x + Y).$$

*The neutral element is $Y$.*

For the definition of quotient spaces, it is not important that we consider real or complex vector spaces.

Examples of quotient spaces are already known. In fact, $L^p$ is such an example. Usually, one defines

$$\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$$

to be the space of *all* mesurable functions $f : \Omega \to \mathbb{K}$ such that $\int_\Omega |f|^p \, d\mu < \infty$. Moreover,

$$N := \{f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu) : \int_\Omega |f|^p = 0\}.$$

Note that $N$ is a linear subspace of $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$, and that $N$ is the space of all functions $f \in \mathcal{L}^p$ which vanish almost everywhere. Then

$$L^p(\Omega, \mathcal{A}, \mu) := \mathcal{L}^p(\Omega, \mathcal{A}, \mu)/N.$$

PROPOSITION 2.4. *Let $X$ be a normed space and let $Y \subset X$ be a linear subspace. Then*

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

*defines a norm on $X/Y$ if and only if $Y$ is closed in $X$. If $X$ is a Banach space and $Y \subset X$ closed, then $X/Y$ is also a Banach space.*

PROOF. We have to check that $\|\cdot\|$ satisfies all properties of a norm. Recall that $0_{X/Y} = Y$, and that for all $x \in X$

$$\begin{aligned} & \|x + Y\| = 0 \\ \Leftrightarrow \quad & \inf\{\|x - y\| : y \in Y\} = 0 \\ \Leftrightarrow \quad & \exists (y_n) \subset Y : \lim_{n\to\infty} y_n = x \\ \Leftrightarrow \quad & (\Rightarrow \text{ if } Y \text{ closed}) : x \in Y \\ \Leftrightarrow \quad & x + Y = Y. \end{aligned}$$

Second, for every $x \in X$ and every $\lambda \in \mathbb{K} \setminus \{0\}$,

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

Third, for every $x$, $z \in X$,

$$\begin{aligned} \|(x+Y)+(z+Y)\| &= \|(x+z)+Y\| \\ &= \inf\{\|x+z-y\| : y \in Y\} \\ &= \inf\{\|x+z-y_1-y_2\| : y_1, \, y_2 \in Y\} \\ &\leq \inf\{\|x-y_1\| + \|z-y_2\| : y_1, \, y_2 \in Y\} \\ &\leq \inf\{\|x-y\| : y \in Y\} + \inf\{\|z-y\| : y \in Y\} \\ &= \|x+Y\| + \|z+Y\|. \end{aligned}$$

Hence, $X/Y$ is a normed space if $Y$ is closed.

Assume next that $X$ is a Banach space. Let $(x_n) \subset X$ be such that the series $\sum_{n\geq 1} x_n + Y$ converges absolutely, i.e. $\sum_{n\geq 1} \|x_n + Y\| < \infty$. By definition of the norm in $X/Y$, we find $(y_n) \subset Y$ such that $\|x_n - y_n\| \leq \|x_n + Y\| + 2^{-n}$. Replacing $(x_n)$ by $(\hat{x}_n) = (x_n - y_n)$, we find that $x_n + Y = \hat{x}_n + Y$ and that the series $\sum_{n\geq 0} \hat{x}_n$ is absolutely convergent. Since $X$ is complete, by Lemma 1.18, the limit $\sum_{n\geq 1} \hat{x}_n = x \in X$ exists. As a consequence,

$$\begin{aligned} \|(x+Y) - \sum_{k=1}^n (\hat{x}_k + Y)\| &= \|(x - \sum_{k=1}^n \hat{x}_k) + Y\| \\ &\leq \|x - \sum_{k=1}^n \hat{x}_k\| \quad \to \quad 0, \end{aligned}$$

i.e. the series $\sum_{n\geq 1} x_n + Y$ converges. By Lemma 1.18, $X/Y$ is complete. □

## 3. Bounded linear operators

In the following a linear mapping between two normed spaces $X$ and $Y$ will also be called a *linear operator* or just *operator*. If $Y = \mathbb{K}$, then we call linear operators also *linear functionals*. If $T : X \to Y$ is a linear operator between two normed spaces, then we denote by

$$\operatorname{Ker} T := \{x \in X : Tx = 0\}$$

its *kernel* or *null space*, and by

$$\operatorname{Rg} T := \{Tx : x \in X\}$$

its *range* or *image*. Observe that we simply write $Tx$ instead of $T(x)$, meaning that $T$ is applied to $x \in X$. The identity $X \to X$, $x \mapsto x$ is denoted by $I$.

LEMMA 3.1. *Let $T : X \to Y$ be a linear operator between two normed spaces $X$ and $Y$. Then the following are equivalent*

(1) *$T$ is continuous.*
(2) *$T$ is continuous in $0$.*
(3) *$TB$ is bounded in $Y$, where $B = B(0,1)$ denotes the unit ball in $X$.*
(4) *There exists a constant $C \geq 0$ such that for every $x \in X$*

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

PROOF. The implication (1)⇒(2) is trivial.

(2)⇒(3). If $T$ is continuous in 0, then there exists some $\delta > 0$ such that for every $x \in B(0,\delta)$ one has $Tx \in B(0,1)$ (so the $\varepsilon$ from the $\varepsilon$-$\delta$ definition of continuity is chosen to be 1 here). By linearity, for every $x \in B = B(0,1)$

$$\|Tx\| = \frac{1}{\delta}\|T(\delta x)\| \leq \frac{1}{\delta},$$

and this means that $TB$ is bounded.

(3)⇒(4). The set $TB$ being bounded in $Y$ means that there exists some constant $C \geq 0$ such that for every $x \in B$ one has $\|Tx\| \leq C$. By linearity, for every $x \in X \setminus \{0\}$,

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

(4)⇒(1). Let $x \in X$, and assume that $\lim_{n\to\infty} x_n = x$. Then

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

so that $\lim_{n\to\infty} Tx_n = Tx$. □

DEFINITION 3.2. We call a continuous linear operator $T : X \to Y$ between two normed spaces $X$ and $Y$ also a *bounded operator* (since it maps the unit ball of $X$ to a bounded subset of $Y$). The set of all bounded linear operators is denoted by $\mathcal{L}(X,Y)$. Special cases: If $X = Y$, then we write $\mathcal{L}(X,X) =: \mathcal{L}(X)$. If $Y = \mathbb{K}$, then we write $\mathcal{L}(X,\mathbb{K}) =: X'$.

LEMMA 3.3. *The set $\mathcal{L}(X,Y)$ is a vector space and*

$$\begin{aligned} \|T\| &:= \inf\{C \geq 0 : \|Tx\| \leq C\,\|x\| \text{ for all } x \in X\} \\ &= \sup\{\|Tx\| : \|x\| \leq 1\} \\ &= \sup\{\|Tx\| : \|x\| = 1\} \end{aligned} \tag{3.1}$$

*is a norm on $\mathcal{L}(X,Y)$.*

PROOF. We first show that the three quantities on the right-hand side of (3.1) are equal. In fact, the equality

$$\sup\{\|Tx\| : \|x\| \leq 1\} = \sup\{\|Tx\| : \|x\| = 1\}$$

is easy to check so that it remains only to show that

$$A := \inf\{C \geq 0 : \|Tx\| \leq C\,\|x\| \text{ for all } x \in X\} = \sup\{\|Tx\| : \|x\| = 1\} =: B.$$

If $C > A$, then for every $x \in X \setminus \{0\}$, $\|Tx\| \leq C\,\|x\|$ or $\|T\frac{x}{\|x\|}\| \leq C$. Hence, $C \geq B$ which implies that $A \geq B$. If $C > B$, then for every $x \in X \setminus \{0\}$, $\|T\frac{x}{\|x\|}\| \leq C$, and therefore $\|Tx\| \leq C\,\|x\|$. Hence, $C \geq A$ which implies that $A \leq B$.

Now we check that $\|\cdot\|$ is a norm on $\mathcal{L}(X,Y)$. First, for every $T \in \mathcal{L}(X,Y)$,

$$\begin{aligned} \|T\| = 0 \quad &\Leftrightarrow \quad \sup\{\|Tx\| : \|x\| \leq 1\} = 0 \\ &\Leftrightarrow \quad \forall x \in X,\ \|x\| \leq 1 : \ \|Tx\| = 0 \\ &\Leftrightarrow \quad (\|\cdot\| \text{ is a norm on } Y)\, \forall x \in X,\ \|x\| \leq 1 : \ Tx = 0 \\ &\Leftrightarrow \quad (\Rightarrow \text{ linearity of } T)\, \forall x \in X : Tx = 0 \\ &\Leftrightarrow \quad T = 0. \end{aligned}$$

Second, for every $T \in \mathcal{L}(X,Y)$ and every $\lambda \in \mathbb{K}$

$$\begin{aligned} \|\lambda T\| &= \sup\{\|(\lambda T)x\| : \|x\| \leq 1\} \\ &= \sup\{|\lambda|\,\|Tx\| : \|x\| \leq 1\} \\ &= |\lambda|\,\|T\|. \end{aligned}$$

Finally, for every $T$, $S \in \mathcal{L}(X,Y)$,

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

The proof is complete. □

Remark 3.4. (a) Note that the infimum on the right-hand side of (3.1) in Lemma 3.3 is always attained. Thus, for every operator $T \in \mathcal{L}(X,Y)$ and every $x \in X$,

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

This inequality shall be frequently used in the sequel! Note that on the other hand the suprema on the right-hand side of (3.1) are not always attained. (b) From Lemma 3.3 we can learn how to show that some operator $T : X \to Y$ is bounded and how to calculate the norm $\|T\|$. Usually (in most cases), one should prove in the first step some inequality of the form

$$\|Tx\| \leq C\,\|x\|, \quad x \in X,$$

because this inequality shows on the one hand that $T$ *is* bounded, and on the other hand it shows the estimate $\|T\| \leq C$. In the second step one should prove that the estimate $C$ was optimal by finding some $x \in X$ of norm $\|x\| = 1$ such that $\|Tx\| = C$, or by finding some sequence $(x_n) \subset X$ of norms $\|x_n\| \leq 1$ such that $\lim_{n\to\infty} \|Tx_n\| = C$, because this shows that $\|T\| = C$. Of course, the second step only works if one has not lost anything in the estimate of the first step. There are in fact many examples of bounded operators for which it is difficult to estimate their norm.

EXAMPLE 3.5. (1) (Shift-operator). On $l^p(\mathbb{N})$ consider the *left-shift operator*

$$Lx = L(x_n) = (x_{n+1}).$$

Then

$$\|L(x_n)\|_p = \left(\sum_n |x_{n+1}|^p\right)^{1/p} \le \left(\sum_n |x_n|^p\right)^{1/p},$$

so that $L$ is bounded and $\|L\| \le 1$. On the other hand, for $x = (0,1,0,0,\dots)$ one computes that $\|x\|_p = 1$ and $\|Lx\|_p = \|(1,0,0,\dots)\|_p = 1$, and one concludes that $\|L\| = 1$.

(2) (Shift-operator). Similarly, one shows that the *right-shift operator* $R$ on $l^p(\mathbb{N})$ defined by

$$Rx = R(x_n) = (0, x_0, x_1, \dots)$$

is bounded and $\|R\| = 1$. Note that actually $\|Rx\|_p = \|x\|_p$ for every $x \in l^p$.

(3) (Multiplication operator). Let $m \in l^\infty$ and consider on $l^p$ the *multiplication operator*

$$Mx = M(x_n) = (m_n x_n).$$

(4) (Functionals on $C$). Consider the linear functional $\varphi : C([0,1]) \to \mathbb{K}$ defined by

$$\varphi(f) := \int_0^{\frac{1}{2}} f(x)\,dx.$$

Then

$$|\varphi(f)| \le \int_0^{\frac{1}{2}} |f(x)|\,dx \le \frac{1}{2}\|f\|_\infty,$$

so that $\varphi$ is bounded and $\|\varphi\| \le \frac{1}{2}$. On the other hand, for the constant function $f = 1$ one has $\|f\|_\infty = 1$ and $|\varphi(f)| = \frac{1}{2}$, so that $\|\varphi\| = \frac{1}{2}$.

LEMMA 3.6. *Let $X$, $Y$, $Z$ be three Banach spaces, and let $T \in \mathcal{L}(X,Y)$ and $S \in \mathcal{L}(Y,Z)$. Then $ST \in \mathcal{L}(X,Z)$ and*

$$\|ST\| \le \|S\|\,\|T\|.$$

PROOF. The boundedness of $ST$ is clear since compositions of continuous functions are again continuous. To obtain the bound on $ST$, we calculate

$$\begin{aligned} \|ST\| &= \sup_{\|x\|\le 1} \|STx\| \\ &\le \sup_{\|x\|\le 1} \|S\|\,\|Tx\| \\ &\le \|S\|\,\|T\|. \end{aligned}$$

□

LEMMA 3.7. *If $Y$ is a Banach space then $\mathcal{L}(X,Y)$ is a Banach space.*

Proof. Assume that $Y$ is a Banach space and let $(T_n)$ be a Cauchy sequence in $\mathcal{L}(X,Y)$. By the estimate

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

the sequence $(T_n x)$ is a Cauchy sequence in $Y$ for every $x \in X$. Since $Y$ is complete, the limit $\lim_{n\to\infty} T_n x$ exists for every $x \in X$. Define $Tx := \lim_{n\to\infty} T_n x$. Clearly, $T : X \to Y$ is linear. Moreover, since any Cauchy sequence is bounded, we find that

$$\|Tx\| \leq \sup_n \|T_n x\| \leq C \, \|x\|$$

for some constant $C \geq 0$, i.e. $T$ is bounded. Moreover, for every $n \in \mathbb{N}$ we have the estimate

$$\begin{aligned} \|T - T_n\| &= \sup_{\|x\|\leq 1} \|Tx - T_n x\| \\ &\leq \sup_{\|x\|\leq 1} \sup_{m\geq n} \|T_m x - T_n x\| \\ &\leq \sup_{m\geq n} \|T_m - T_n\|. \end{aligned}$$

Since that right-hand side of this inequality becomes arbitrarily small for large $n$, we see that $\lim_{n\to\infty} T_n = T$ exists, and so we have proved that $\mathcal{L}(X,Y)$ is a Banach space. □

Remark 3.8. The converse of the statement in Lemma 3.7 is also true, i.e if $\mathcal{L}(X,Y)$ is a Banach space then necessarily $Y$ is a Banach space. For the proof, however, one has to know that there are nontrivial operators in $\mathcal{L}(X,Y)$ as soon as $Y$ is nontrivial (i.e. $Y \neq \{0\}$). For this, we need the Theorem of Hahn-Banach and its consequences discussed in Chapter 5.

Corollary 3.9. *The space $X' = \mathcal{L}(X,\mathbb{K})$ of all bounded linear functionals on $X$ is always a Banach space.*

Definition 3.10. Let $X$, $Y$ be two normed spaces.

(1) We call $T \in \mathcal{L}(X,Y)$ an *isomorphism* if $T$ is bijective and $T^{-1} \in \mathcal{L}(Y,X)$.
(2) We call $T \in \mathcal{L}(X,Y)$ an *isometry* if $\|Tx\| = \|x\|$ for every $x \in X$.
(3) We say that $X$ and $Y$ are *isomorphic* (and we write $X \cong Y$) if there exists an isomorphism $T \in \mathcal{L}(X,Y)$.
(4) We say that $X$ and $Y$ are *isometrically isomorphic* if there exists an isometric isomorphism $T \in \mathcal{L}(X,Y)$.

Remark 3.11.

(1) Two norms $\|\cdot\|_1$, $\|\cdot\|_2$ on a $\mathbb{K}$ vector space $X$ are equivalent if and only if the identity operator $I : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ is an isomorphism.
(2) Saying that two *normed* spaces $X$ and $Y$ are isomorphic means that they are not only 'equal' as vector spaces (in the sense that we find a bijective linear operator) but also as normed spaces (i.e. the bijection is continuous as well as its inverse).

(3) If $T \in \mathcal{L}(X,Y)$ and $S \in \mathcal{L}(Y,Z)$ are isomorphisms, then $ST \in \mathcal{L}(X,Z)$ is an isomorphism and $(ST)^{-1} = T^{-1}S^{-1}$.
(4) Every isometry $T \in \mathcal{L}(X,Y)$ is clearly injective. If it is also surjective, then $T$ is an isometric isomorphism, i.e. the inverse $T^{-1}$ is also bounded (even isometric).
(5) Clearly, if $T \in \mathcal{L}(X,Y)$ is isometric, then it is an isometric isomorphism from $X$ onto $\operatorname{Rg} T$, and we may say that $X$ is *isometrically embedded* into $Y$ (via $T$).

EXAMPLE 3.12. The right-shift operator from Example 3.5 (2) is isometric, but not surjective. In particular, $l^p$ is isometrically isomorphic to a proper subspace of $l^p$.

EXERCISE 3.13. Show that the spaces $(c, \|\cdot\|_\infty)$ of all convergent sequences and $(c_0, \|\cdot\|_\infty)$ of all null sequences are isomorphic.

EXERCISE 3.14. Show that $(c_0, \|\cdot\|_\infty)$ is (isometrically) isomorphic to a linear subspace of $(C([0,1]), \|\cdot\|_\infty)$, i.e. find an isometry $T : c_0 \to C([0,1])$.

LEMMA 3.15 (Neumann series). *Let $X$ be a Banach space and let $T \in \mathcal{L}(X)$ be such that $\|T\| < 1$. Then $I - T$ is boundedly invertible, i.e. it is an isomorphism. Moreover, $(I-T)^{-1} = \sum_{n\geq 0} T^n$.*

PROOF. Since $X$ is a Banach space, $\mathcal{L}(X)$ is also a Banach space by Lemma 3.7. By assumption on $\|T\|$, the series $\sum_{n\geq 0} T^n$ is absolutely convergent, and hence, by Lemma 1.18, it is convergent to some element $S \in \mathcal{L}(X)$. Moreover,

$$(I-T)S = \lim_{n\to\infty}(I-T)\sum_{k=0}^{n} T^k = \lim_{n\to\infty}(I - T^{k+1}) = I,$$

and similarly, $S(I-T) = I$. □

COROLLARY 3.16. *Let $X$ and $Y$ be two Banach spaces. Then the set $\mathcal{I}(X,Y)$ of all isomorphisms in $\mathcal{L}(X,Y)$ is open, and the mapping $T \mapsto T^{-1}$ is continuous from $\mathcal{I}(X,Y)$ onto $\mathcal{I}(Y,X)$.*

PROOF. Let $\mathcal{I} \subset \mathcal{L}(X,Y)$ be the set of all isomorphisms, and assume that $\mathcal{I}$ is not empty (if it is empty, then it is also open). Let $T \in \mathcal{I}$. Then for every $S \in B(T, \frac{1}{\|T^{-1}\|})$ we have

$$S = T + S - T = T(I + T^{-1}(S-T)),$$

and since $\|T^{-1}(S-T)\| \leq \|T^{-1}\|\,\|S-T\| < 1$, the operator $I + T^{-1}(S-T) \in \mathcal{L}(X)$ is an isomorphism by Lemma 3.15. As a composition of two isomorphisms, $S \in \mathcal{I}$, and hence $\mathcal{I}$ is open. The continuity is also a direct consequence of the above representation of $S$ (and thus of its inverse), using the Neumann series. □

## 4. Calculus on Banach spaces

Definition 4.1. Let $X$, $Y$ be two Banach spaces, and let $U \subset X$ be open. A function $f : U \to Y$ is called

(a) *differentiable* in $x_0 \in U$ if there exists a bounded linear operator $T \in \mathcal{L}(X,Y)$ and some function $r : X \to Y$ such that

$$f(x) = f(x_0) + T(x - x_0) + r(x - x_0), \quad x \in U, \tag{4.1}$$

and

$$\lim_{\|h\| \to 0} \frac{\|r(h)\|}{\|h\|} = 0. \tag{4.2}$$

(b) *differentiable* if it is differentiable in every point $x_0 \in U$.

If $f$ is differentiable in a point $x_0 \in U$, then $T \in \mathcal{L}(X,Y)$ is uniquely determined. We write $Df(x_0) := f'(x_0) := T$ and call $Df(x_0) = f'(x_0)$ the *derivative* of $f$ in $x_0$.

Lemma 4.2. *If a function $f : U \to Y$ is differentiable in $x \in U$, then it is continuous in $x$. In particular, every differentiable function is continuous.*

Proof. Let $(x_n) \subset U$ be convergent to $x$. By definition (equation (4.1)),

$$\begin{aligned} \|f(x_n) - f(x)\| &= \|f'(x)(x - x_n) + r(x - x_n)\| \\ &\leq \|f'(x)\| \, \|x - x_n\| + \|r(x - x_n)\| \quad \to \quad 0, \end{aligned}$$

as $n \to \infty$ (use also (4.2)). □

Definition 4.3. Let $X$, $Y$ be two Banach spaces, and let $U \subset X$ be open. A function $f : U \to Y$ is called *continuously differentiable* if it is differentiable and if $f' : U \to \mathcal{L}(X,Y)$ is continuous. We denote by

$$C^1(U;Y) := \{f : U \to Y : f \text{ differentiable and } f' \in C(U;\mathcal{L}(X,Y))\}$$

the space of all continuously differentiable functions. Moreover, for $k \geq 2$, we denote by

$$C^k(U;Y) := \{f : U \to Y : f \text{ differentiable and } f' \in C^{k-1}(U;\mathcal{L}(X,Y))\}$$

the space of all $k$ times continuously differentiable functions.

Theorem 4.4 (Local inverse theorem). *Let $X$, $Y$ be two Banach spaces, $U \subset X$ open, $f \in C^1(U;Y)$, and $x \in U$. Suppose that $f'(x) \in \mathcal{L}(X,Y)$ is an isomorphism. Then there exists a neighbourhood $V \subset U$ of $x$ and a neighbourhood $W \subset Y$ of $f(x)$ such that the restriction $f : V \to W$ is* homeomorphic, *i.e. continuous, bijective and the inverse $f^{-1} : W \to V$ is also continuous.*

Definition 4.5. Let $X_i$ $(1 \leq i \leq n)$ and $Y$ be Banach spaces. Let $U \subset \bigotimes_{i=1}^n X_i$ be open. We say that a function $f : U \to Y$ is in $a = (a_i)_{1 \leq i \leq n} \in U$ *partially differentiably* with respect to the $i$-th coordinate if the function

$$f_i : U_i \subset X_i \to Y, \quad x_i \mapsto f(a_1, \dots, x_i, \dots, a_n)$$

is differentiable in $a_i$. We write $\frac{\partial f}{\partial x_i}(a) := f_i'(a_i) \in \mathcal{L}(X_i,Y)$.

Theorem 4.6 (Implicit function theorem). *Let $X_1$, $X_2$, $Y$ be three Banach spaces and let $U \subset X_1 \times X_2$ be open. Let $f \in C^1(U;Y)$, $a = (a_1, a_2) \in U$, and suppose that $f(a) = 0$ and that $\frac{\partial f}{\partial x_2}(a) \in \mathcal{L}(X_2, Y)$ is an isomorphism. Then there exist neighbourhoods $U_1 \subset X_1$ of $a_1$ and $U_2 \subset X_2$ of $a_2$ and a function $g \in C^1(U_1; X_2)$ such that $g(a_1) = a_2$ and*

$$\{x = (x_1, x_2) \in U_1 \times U_2 : f(x) = 0\} = \{(x_1, g(x_1)) : x_1 \in U_1\}.$$

## 5. * Newton's method

Theorem 5.1 (Newton's method). *Let $X$ and $Y$ be two Banach spaces, $U \subset X$ an open set. Let $f \in C^1(U;Y)$ and assume that there exists $\bar{x} \in U$ such that (i) $f(\bar{x}) = 0$ and (ii) $f'(\bar{x}) \in \mathcal{L}(X,Y)$ is an isomorphism. Then there exists a neighbourhood $V \subset U$ of $\bar{x}$ such that for every $x_0 \in V$ the operator $f'(x_0)$ is an isomorphism, the sequence $(x_n)$ defined iteratively by*

$$x_{n+1} = x_n - f'(x_n)^{-1} f(x_n), \quad n \geq 0, \tag{5.1}$$

*remains in $V$ and $\lim_{n\to\infty} x_n = \bar{x}$.*

Proof. By Corollary 3.16 and continuity, there exists a neighbourhood $\tilde{V} \subset U$ of $\bar{x}$ such that $f'(x)$ is isomorphic for all $x \in \tilde{V}$. Next, it will be useful to define the auxiliary function $\varphi : \tilde{V} \to X$ by

$$\varphi(x) := x - f'(x)^{-1} f(x), \quad x \in \tilde{V}.$$

Since $f(\bar{x}) = 0$, we find that for every $x \in \tilde{V}$

$$\begin{aligned} \varphi(x) - \varphi(\bar{x}) &= x - f'(x)^{-1}(f(x) - f(\bar{x})) - \bar{x} \\ &= x - \bar{x} - f'(x)^{-1}(f'(\bar{x})(x - \bar{x}) + r(x - \bar{x})), \end{aligned}$$

so that by the continuity of $f'(\cdot)^{-1}$

$$\lim_{x\to\bar{x}} \frac{\|\varphi(x) - \varphi(\bar{x})\|}{\|x - \bar{x}\|} = 0.$$

Hence, there exists $r > 0$ such that $V := B(\bar{x}, r) \subset \tilde{V} \subset U$ and such that for every $x \in V$

$$\|\varphi(x) - \bar{x}\| = \|\varphi(x) - \varphi(\bar{x})\| \leq \frac{1}{2} \|x - \bar{x}\|.$$

This implies that for every $x_0 \in V$ one has $\varphi(x_0) \in V$ and if we define iteratively $x_{n+1} = \varphi(x_n) = \varphi^{n+1}(x_0)$, then

$$\|x_n - \bar{x}\| \leq \left(\frac{1}{2}\right)^n \|x_0 - \bar{x}\| \to 0 \text{ as } n \to \infty.$$

□

CHAPTER 4

# Hilbert spaces

Let $H$ be a vector space over $\mathbb{K}$.

## 1. Inner product spaces

DEFINITION 1.1. A function $(\cdot,\cdot) : H \times H \to \mathbb{K}$ is called a *scalar product* or *inner product* if for every $x$, $y$, $z \in H$ and every $\lambda \in \mathbb{K}$

(i) $(x,x) \geq 0$ for every $x \in H$ and $(x,x) = 0$ if and only if $x = 0$,
(ii) $(x,y) = \overline{(y,x)}$,
(iii) $(\lambda x + y, z) = \lambda(x,z) + (y,z)$.

A pair $(H,(\cdot,\cdot))$ of a vector space over $\mathbb{K}$ and a scalar product is called an *inner product space*.

EXAMPLE 1.2. (1) On the space $H = \mathbb{K}^N$,

$$(x,y) := \sum_{i=1}^{N} x_i \bar{y}_i$$

defines a scalar product.

(2) On the space $H = l^2 := \{(x_n) \subset \mathbb{K} : \sum |x_n|^2 < \infty\}$,

$$(x,y) := \sum_n x_n \bar{y_n}$$

defines a scalar product.

(3) On the space $H = C([0,1])$, the Riemann integral

$$(f,g) := \int_0^1 f(x)\overline{g(x)}\, dx$$

defines a scalar product.

(4) On the space $H = L^2(\Omega)$, the integral

$$(f,g) := \int_\Omega f\bar{g}\, d\mu$$

defines a scalar product.

LEMMA 1.3. *Let $(\cdot,\cdot)$ be a scalar product on a vector space $H$. Then, for every $x$, $y$, $z \in H$ and $\lambda \in \mathbb{K}$*

(iv) $(x, \lambda y + z) = \bar{\lambda}(x,y) + (x,z)$.

Proof.

$$(x, \lambda y + z) = \overline{(\lambda y + z, x)} = \bar{\lambda}\,\overline{(y, x)} + \overline{(z, x)} = \bar{\lambda}(x, y) + (x, z).$$

□

In the following, if $H$ is an inner product space, then we put

$$\|x\| := \sqrt{(x, x)}, \quad x \in H.$$

Lemma 1.4 (Cauchy-Schwarz inequality). *Let $H$ be an inner product space. Then, for every $x$, $y \in H$,*

$$|(x, y)| \leq \|x\| \, \|y\|,$$

*and equality holds if and only if $x$ and $y$ are colinear.*

Proof. Let $\lambda \in \mathbb{K}$. Then

$$\begin{aligned} 0 &\leq (x + \lambda y, x + \lambda y) \\ &= (x, x) + (\lambda y, x) + (x, \lambda y) + |\lambda|^2 (y, y) \\ &= (x, x) + \lambda \overline{(x, y)} + \bar{\lambda}(x, y) + |\lambda|^2 (y, y), \end{aligned}$$

i.e.

$$0 \leq \|x + \lambda y\|^2 = \|x\|^2 + 2\mathrm{Re}\, \bar{\lambda}(x, y) + |\lambda|^2 \, \|y\|^2. \tag{1.1}$$

Assuming that $y \neq 0$ (for $y = 0$ the Cauchy-Schwarz inequality is trivial), we may put $\lambda := -(x, y)/\|y\|^2$. Then

$$\begin{aligned} 0 &\leq (x - \frac{(x, y)}{\|y\|^2} y, x - \frac{(x, y)}{\|y\|^2} y) \\ &= \|x\|^2 - \frac{|(x, y)|^2}{\|y\|^2}, \end{aligned}$$

which is the Cauchy-Schwarz inequality. The calculation also shows that equality holds if and only if $x = \lambda y$, i.e. if $x$ and $y$ are colinear. □

Lemma 1.5. *Every inner product space $H$ is a normed linear space for the norm*

$$\|x\| = \sqrt{(x, x)}, \quad x \in H.$$

Proof. Properties (i) and (ii) in the definition of a norm follow from the properties (i) and (iii) (together with Lemma 1.3) in the definition of a scalar product. The only difficulty is to show that $\|\cdot\|$ satisfies the triangle inequality. This, however, follows from putting $\lambda = 1$ in (1.1) and estimating with the Cauchy-Schwarz inequality:

$$\|x + y\|^2 \leq (\|x\| + \|y\|)^2.$$

□

Definition 1.6. A complete inner product space is called a *Hilbert space*.

EXAMPLE 1.7. The spaces $\mathbb{K}^N$ (with euclidean scalar product), $l^2$ and $L^2(\Omega)$ are Hilbert spaces. More examples are given by the Sobolev spaces defined below.

LEMMA 1.8 (Completion of an inner product space). *Let $H$ be an inner product space. Then there exists a Hilbert space $K$ and a bounded linear operator $j : H \to K$ such that for every $x$, $y \in H$*

$$(x, y)_H = (j(x), j(y))_K,$$

*and such that $j(H)$ is dense in $K$. The Hilbert space $K$ is unique* up to isometry. *It is called the* completion *of $H$.*

LEMMA 1.9 (Parallelogram identity). *Let $H$ be an inner product space. Then for every $x$, $y \in H$*

$$\|x + y\|^2 + \|x - y\|^2 = 2\,(\|x\|^2 + \|y\|^2).$$

PROOF. The parallelogram identity follows immediately from (1.1) by putting $\lambda = \pm 1$ and adding up. □

EXERCISE 1.10 (von Neumann). Show that a norm satisfying the parallelogram identity comes from a scalar product, i.e. the parallelogram identity characterises inner product spaces.

DEFINITION 1.11. A subset $K$ of a vector space $X$ (over $\mathbb{K}$) is called *convex* if for every $x$, $y \in K$ and every $t \in [0, 1]$ one has $tx + (1 - t)y \in K$.

THEOREM 1.12. *Given a nonempty closed, convex subset $K$ of a Hilbert space $H$, and given a point $x \in H$, there exists a unique $y \in K$ such that*

$$\|x - y\| = \inf\{\|x - z\| : z \in K\}.$$

PROOF. Let $d := \inf\{\|x - z\| : z \in K\}$, and choose $(y_n) \in K$ such that

$$\lim_{n\to\infty} \|x - y_n\| = d. \tag{1.2}$$

Applying the parallelogram identity to $(x - y_n)/2$ and $(x - y_m)/2$, we obtain

$$\|x - \frac{y_n + y_m}{2}\|^2 + \frac{1}{4}\|y_n - y_m\|^2 = \frac{1}{2}(\|x - y_n\|^2 + \|x - y_m\|^2).$$

Since $K$ is convex, $\frac{y_n+y_m}{2} \in K$ and hence $\|x - \frac{y_n+y_m}{2}\|^2 \geq d^2$. Using this and (1.2), the last identity implies that $(y_n)$ is a Cauchy sequence. Since $H$ is complete, $y := \lim_{n\to\infty} y_n$ exists. Since $K$ is closed, $y \in K$. Moreover, $\|x - y\| = \lim_{n\to\infty} \|x - y_n\| = d$, so that $y$ is a minimizer for the distance to $x$. To see that there is only on such minimizer, suppose that $y' \in K$ is a second one, and apply the parallelogram identity to $x - y$ and $x - y'$. □

DEFINITION 1.13. Let $H$ be an inner product space. We say that two vectors $x$, $y \in H$ are *orthogonal* (and we write $x \perp y$), if $(x, y) = 0$. Given a subset $S \subset H$, we define the *orthogonal space* $S^\perp := \{y \in H : x \perp y$ for all $x \in S\}$. If $S = K$ is a linear subspace of $H$, then we call $K^\perp$ also the *orthogonal complement* of $K$.

Theorem 1.14. *Let $H$ be a Hilbert space, $S \subset H$ be a subset and $K$ a closed linear subspace. Then*

(i) *$S^\perp$ is a closed linear subspace of $H$,*
(ii) *$K$ and $K^\perp$ are complementary subspaces, i.e. every $x \in H$ can be decomposed uniquely as a sum of an $x_0 \in K$ and an $x_1 \in K^\perp$,*
(iii) *$(K^\perp)^\perp = K$ and $(S^\perp)^\perp = \overline{\operatorname{span}} S$.*
(iv) *$\operatorname{span} S$ is dense in $H$ if and only if $S^\perp = \{0\}$.*

Proof. (i) It follows from the bilinearity of the scalar product that $S^\perp$ is a linear subspace of $H$. Let $(y_n) \in S^\perp$ be convergent to some $y \in H$. Then, for every $x \in S$, by the Cauchy-Schwarz inequality,

$$(x, y) = \lim_{n\to\infty} (x, y_n) = 0,$$

i.e. $y \in S^\perp$ and therefore $S^\perp$ is closed.

(ii) For every $x \in H$ we let $x_0 \in K$ be the unique element (Theorem 1.12) such that

$$\|x - x_0\| = \inf\{\|x - y\| : y \in K\}.$$

Put $x_1 = x - x_0$. For every $y \in K$ and every $\lambda \in \mathbb{K}$, by the minimum property of $x_0$,

$$\begin{aligned} \|x_1\|^2 &\leq \|x_1 - \lambda y\|^2 \\ &= \|x_1\|^2 - 2\mathrm{Re}\, \bar{\lambda}(x_1, y) + |\lambda|^2 \|y\|^2. \end{aligned}$$

This implies that $(x_1, y) = 0$, i.e. $x_1 \in K^\perp$. Every decomposition $x = x_0 + x_1$ with $x_0 \in K$ and $x_1 \in K^\perp$ is unique since $x \in K \cap K^\perp$ implies $(x, x) = 0$, i.e $x = 0$.

(iii) and (iv) follow immediately from (i) and (ii). □

Lemma 1.15 (Pythagoras). *Let $H$ be an inner product space. Whenever $x$, $y \in H$ are orthogonal, then*

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

Proof. The claim follows from (1.1) and putting $\lambda = 1$. □

Definition 1.16. Let $X$ be a normed space. We call an operator $P : X \to X$ a *projection* if $P^2 = P$.

Lemma 1.17. *Let $X$ be a normed space and let $P \in \mathcal{L}(X)$ be a bounded projection. Then the following are true:*

(1) *$Q = I - P$ is a projection.*
(2) *Either $P = 0$ or $\|P\| \geq 1$.*
(3) *The kernel $\ker P$ and the range $\operatorname{Rg} P$ are closed in $X$.*
(4) *Every $x \in X$ can be decomposed uniquely as a sum of an $x_0 \in \ker P$ and an $x_1 \in \operatorname{Rg} P$, and $X \cong \ker P \oplus \operatorname{Rg} P$.*

PROOF. (1) $Q^2 = (I - P)^2 = I - 2P + P^2 = I - P = Q$.

(2) follows from $\|P\| = \|P^2\| \leq \|P\|^2$.

(3) Since $\{0\}$ is closed in $X$ and since $P$ is continuous, $\ker P = P^{-1}(\{0\})$ is closed. Similarly, $\operatorname{Rg} P = \ker(I - P)$ is closed.

(4) For every $x \in X$ we can write $x = Px+(I-P)x = x_1+x_2$ with $x_1 \in \operatorname{Rg} P$ and $x_2 \in \ker P$. The decomposition is unique since if $x \in \ker P \cap \operatorname{Rg} P$, then $x = Px = 0$. This proves that the *vector* spaces $X$ and $\ker P \oplus \operatorname{Rg} P$ are isomorphic. That they are also isomorphic as normed spaces follows from the continuity of $P$. □

LEMMA 1.18. *Let $H$ be a Hilbert space and $K \subset H$ be a closed linear subspace. For every $x \in H$ we let $x_1 = Px$ be the unique element in $K$ which minimizes the distance to $x$ (Theorem 1.12). Then $P : H \to H$ is a bounded projection satisfying* $\operatorname{Rg} P = K$. *Moreover,* $\ker P = K^\perp$. *We call $P$ the* orthogonal projection *onto $K$.*

## 2. Orthogonal decomposition

DEFINITION 2.1. We call a metric space *separable* if there exists a countable dense subset.

EXAMPLE 2.2. The space $\mathbb{R}^N$ (or $\mathbb{C}^N$) is separable: one may take $\mathbb{Q}^N$ as an example of a dense countable subset. It is not too difficult to see that subsets of separable metric spaces are separable (note, however, that in general the dense subset has to be constructed carefully), and that finite products of separable metric spaces are separable.

LEMMA 2.3. *A normed space $X$ is separable if and only if there exists a sequence $(x_n) \subset X$ such that* $\operatorname{span}\{x_n : n \in \mathbb{N}\}$ *is dense in $X$ (such a sequence is in general called a* total *sequence).*

PROOF. If $X$ is separable, then there exists a sequence $(x_n) \subset X$ such that $\{x_n : n \in \mathbb{N}\}$ is dense. In particular, the larger set $\operatorname{span}\{x_n : n \in \mathbb{N}\}$ is dense.

If, one the other hand, there exists a total sequence $(x_n) \subset X$, , and if we put $D = \mathbb{Q}$ in the case $\mathbb{K} = \mathbb{R}$ and $D = \mathbb{Q} + i\mathbb{Q}$ in the case $\mathbb{K} = \mathbb{C}$, then the set

$$\{\sum_{i=1}^{m} \lambda_i x_{n_i} : m \in \mathbb{N},\, \lambda_i \in D,\, n_i \in \mathbb{N}\}$$

is dense in $X$ (in fact, the closure contains all finite linear combinations of the $x_n$, i.e. it contains $\operatorname{span}\{x_n : n \in \mathbb{N}\}$). It is an exercise to show that this set is countable. The claim follows. □

COROLLARY 2.4. *The space* $(C([0,1]), \|\cdot\|_\infty)$ *is separable.*

PROOF. By Weierstrass' theorem, the subspace of all polynomials is dense in $C([0,1])$ (every continuous function $f : [0,1] \to \mathbb{R}$ can be uniformly approximated by polynomials). The polynomials, however, are the linear span of the monomials $f_n(t) = t^n$. The claim follows from Lemma 2.3. □

Corollary 2.5. *The space $l^p$ is separable if $1 \le p < \infty$. The space $c_0$ is separable.*

Proof. Let $e_n = (\delta_{nk})_k \in l^p$ be the $n$-th unit vector in $l^p$ (here $\delta_{nk}$ denotes the Kronecker symbol: $\delta_{nk} = 1$ if $n = k$ and $\delta_{nk} = 0$ otherwise). Then $span\{e_n : n \in \mathbb{N}\} = c_{00}$ (the space of all finite sequences) is dense in $l^p$ if $1 \le p < \infty$. The claim for $l^p$ follows from Lemma 2.3. The argument for $c_0$ is similar. $\square$

Lemma 2.6. *The space $l^\infty$ is not separable.*

Proof. The set $\{0,1\}^{\mathbb{N}} \subset l^\infty$ of all sequences taking only values 0 or 1 is uncountable. Moreover, whenever $x, y \in \{0,1\}^{\mathbb{N}}$, $x \neq y$, then

$$\|x - y\|_\infty = 1.$$

Hence, the balls $B(x, \frac{1}{2})$ with centers $x \in \{0,1\}^{\mathbb{N}}$ and radius $\frac{1}{2}$ are mutually disjoint. If $l^\infty$ was separable, i.e. if there exists a dense countable set $D \subset l^\infty$, then in each $B(x, \frac{1}{2})$ there exists at least one element $y \in D$, a contradiction. $\square$

Definition 2.7. Let $H$ be an inner product space. A family $(e_l)_{l \in I} \subset H$ is called

(i) an *orthogonal system* if $(e_l, e_k) = 0$ whenever $l \neq k$,
(ii) an *orthonormal system* if it is an orthogonal system and $\|e_l\| = 1$ for every $l \in I$, and
(iii) an *orthonormal basis* if it is an orthonormal system and span$\{e_l : l \in I\}$ is dense in $H$.

Lemma 2.8 (Gram-Schmidt process). *Let $(x_n)$ be a sequence in an inner product space $H$. Then there exists an orthonormal system $(e_n)$ such that* $\text{span}\{x_n\} = \text{span}\{e_n\}$.

Proof. Passing to a subsequence, if necessary, we may assume that the $(x_n)$ are linearly independent.

Let $e_1 := x_1/\|x_1\|$. Then $e_1$ and $x_1$ span the same linear subspace. Next, assume that we have constructed an orthonormal system $(e_k)_{1 \le k \le n}$ such that

$$\text{span}\{x_k : 1 \le k \le n\} = \text{span}\{e_k : 1 \le k \le n\}.$$

Let $e'_{n+1} := x_{n+1} - \sum_{k=1}^n (x_{n+1}, e_k)e_k$. Since the $x_n$ are linearly independent, we find $e'_{n+1} \neq 0$. Let $e_{n+1} := e'_{n+1}/\|e'_{n+1}\|$. By construction, for every $1 \le k \le n$, $(e_{n+1}, e_k) = 0$, and

$$\text{span}\{x_k : 1 \le k \le n+1\} = \text{span}\{e_k : 1 \le k \le n+1\}.$$

Proceeding inductively, the claim follows. $\square$

Corollary 2.9. *Every separable inner product space admits an orthonormal basis.*

Example 2.10. Consider the inner product space $C([-1,1])$ equiped with the scalar product $(f,g) = \int_{-1}^1 f(t)\overline{g(t)}\, dt$ and resulting norm $\|\cdot\|_2$. Let

$f_n(t) := t^n$ $(n \geq 0)$, so that $\operatorname{span}\{f_n\}$ is the space of all polynomials on the interval $[-1,1]$. Applying the Gram-Schmidt process to the sequence $(f_n)$ yields a orthonormal sequence $(p_n)$ of polynomials. The $p_n$ are called *Legendre polynomials*.

Recall that the space of all polynomials is dense in $C([-1,1])$ by Weierstrass' theorem (even for the uniform norm; *a fortiori* also for the norm $\|\cdot\|_2$). Hence, the Legendre polynomials form an orthonormal basis in $C([-1,1])$.

LEMMA 2.11 (Bessel's inequality). *Let $H$ be an inner product space, $(e_n)_{n\in\mathbb{N}} \subset H$ an orthonormal system. Then, for every $x \in H$,*

$$\sum_{n\in\mathbb{N}} |(x,e_n)|^2 \leq \|x\|^2.$$

PROOF. Let $N \in \mathbb{N}$. Put $x_N = x - \sum_{n=1}^N (x,e_n)e_n$ so that $x_N \perp e_n$ for every $1 \leq n \leq N$. By Pythagoras (Lemma 1.15),

$$\begin{aligned} \|x\|^2 &= \|x_N\|^2 + \|\sum_{n=1}^N (x,e_n)e_n\|^2 \\ &= \|x_N\|^2 + \sum_{n=1}^N |(x,e_n)|^2 \\ &\geq \sum_{n=1}^N |(x,e_n)|^2. \end{aligned}$$

Since $N$ was arbitrary, the claim follows. □

LEMMA 2.12. *Let $H$ be a (separable) Hilbert space, $(e_n)_{n\in\mathbb{N}} \subset H$ an orthonormal system. Then:*

(i) *For every $x \in H$, the series $\sum_{n\in\mathbb{N}} (x,e_n)e_n$ converges.*
(ii) *$P : H \to H$, $x \mapsto \sum_{n\in\mathbb{N}} (x,e_n)e_n$ is the orthogonal projection onto $\overline{\operatorname{span}}\{e_n : n \in \mathbb{N}\}$.*

PROOF. (i) Let $x \in H$. Since $(e_n)$ is an orthonormal system, by Pythagoras (Lemma 1.15), for every $l > k \geq 1$,

$$\begin{aligned} \|\sum_{n=1}^l (x,e_n)e_n - \sum_{n=1}^k (x,e_n)e_n\|^2 &= \|\sum_{n=k+1}^l (x,e_n)e_n\|^2 \\ &= \sum_{n=k+1}^l |(x,e_n)|^2. \end{aligned}$$

Hence, by Bessel's inequality, the sequence $(\sum_{n=1}^l (x,e_n)e_n)$ of partial sums forms a Cauchy sequence. Since $H$ is complete, the series $\sum_{n\in\mathbb{N}} (x,e_n)e_n$ converges.

(ii) is an exercise. □

THEOREM 2.13. *Let $H$ be a (separable) Hilbert space, $(e_n)_{n\in\mathbb{N}}$ an orthonormal system. Then the following are equivalent:*

(i) *$(e_n)_{n\in\mathbb{N}}$ is an orthonormal basis.*
(ii) *If $x \perp e_n$ for every $n \in \mathbb{N}$, then $x = 0$.*
(iii) *$x = \sum_{n\in\mathbb{N}}(x, e_n)e_n$ for every $x \in H$.*
(iv) *$(x, y) = \sum_{n\in\mathbb{N}}(x, e_n)(e_n, y)$ for every $x$, $y \in H$.*
(v) *(Parseval's identity) For every $x \in H$,*

$$\|x\|^2 = \sum_{n\in\mathbb{N}} |(x, e_n)|^2.$$

PROOF. (i)⇒(ii) follows from Theorem 1.14.

(ii)⇒(iii) follows from Lemma 2.12 (i). In fact, let $x_0 = \sum_{n\in\mathbb{N}}(x, e_n)e_n$ (which exists by Lemma 2.12 (i)). Then $(x - x_0, e_n) = 0$ for every $n \in \mathbb{N}$, and by assumption (ii), this implies $x = x_0$.

(iii)⇒(iv) follows when multiplying $x$ scalarly with $y$, applying also the Cauchy-Schwarz inequality for the sequences $((x, e_l))$, $((e_l, y)) \in l^2$.

(iv)⇒(v) follows from putting $x = y$.

(v)⇒(i). Let $x \in \operatorname{span}\{e_n : n \in \mathbb{N}\}^\perp$. Then Parseval's identity implies $\|x\|^2 = 0$, i.e. $x = 0$. By Theorem 1.14, $\operatorname{span}\{e_n : n \in \mathbb{N}\}$ is dense in $H$, i.e. $(e_n)$ is an orthonormal basis. □

DEFINITION 2.14. A bounded linear operator $U \in \mathcal{L}(H, K)$ between two Hilbert spaces is called a *unitary operator* if it is invertible and for every $x$, $y \in H$,

$$(x, y)_H = (Ux, Uy)_K.$$

Two Hilbert spaces $H$ and $K$ are *unitarily equivalent* if there exists a unitary operator $U \in \mathcal{L}(H, K)$.

COROLLARY 2.15. *Every infinite dimensional separable Hilbert space $H$ is unitarily equivalent to $l^2$.*

PROOF. Choose an orthonormal basis $(e_n)_{n\in\mathbb{N}}$ of $H$ (which exists by Corollary 2.9), and define $U : H \to l^2$ by $U(x) = ((x, e_n))_{n\in\mathbb{N}}$. Then $(x, y)_H = (U(x), U(y))_{l^2}$ by Theorem 2.13; in particular, $U$ is bounded, isometric and injective. The fact that $U$ is surjective, i.e. that $\sum_n c_n e_n$ converges for every $c = (c_n) \in l^2$, follows as in the proof of Lemma 2.12 (i). □

Clearly, if a sequence $(e_n)$ in a Hilbert space $H$ is an orthonormal basis, then necessarily $H$ is separable by Lemma 2.3. Hence, the equivalent statements of Theorem 2.13 are only satisfied in separable Hilbert spaces. In most of the applications (if not all!), we will only deal with separable Hilbert spaces so that Theorem 2.13 is sufficient for our purposes.

However, what is true in general Hilbert spaces? The following sequence of results generalizes the preceeding results to arbitrary Hilbert spaces.

Definition 2.16. Let $X$ be a normed space, $(x_i)_{i\in I}$ be a family. We say that the series $\sum_{i\in I} x_i$ converges *unconditionally* if the set $I_0 := \{i \in I : x_i \neq 0\}$ is countable, and for every bijective $\varphi : \mathbb{N} \to I_0$ the series $\sum_{n=1}^{\infty} x_{\varphi(n)}$ converges.

Corollary 2.17 (Bessel's inequality, general case). *Let $H$ be an inner product space, $(e_l)_{l\in I} \subset H$ an orthonormal system. Then, for every $x \in H$, the set $\{l \in I : (x, e_l) \neq 0\}$ is countable and*

$$\sum_{l\in I} |(x, e_l)|^2 \leq \|x\|^2. \tag{2.1}$$

Proof. By Bessel's inequality, the sets $\{l \in I : |(x, e_l)| \geq 1/n\}$ must be finite for every $n \in \mathbb{N}$. The countability of $\{l \in I : (x, e_l) \neq 0\}$ follows. The inequality (2.1) is then a direct consequence of Bessel's inequality. □

Lemma 2.18. *Let $H$ be a Hilbert space, $(e_l)_{l\in I} \subset H$ an orthonormal system. Then:*

(i) *For every $x \in H$, the series $\sum_{l\in I}(x, e_l)e_l$ converges unconditionally.*
(ii) *$P : H \to H$, $x \mapsto \sum_{l\in I}(x, e_l)e_l$ is the orthogonal projection onto $\overline{\operatorname{span}}\{e_l : l \in I\}$.*

Corollary 2.19. *Every Hilbert space admits an orthonormal basis.*

Proof. If $H$ is separable, the claim follows directly from the Gram-Schmidt process and has already been stated in Corollary 2.9. In general, one may argue as follows:

The set of all orthonormal systems in $H$ forms a partially ordered set by inclusion. Given a totally ordered collection of orthonormal systems, the union of all vectors contained in all systems in this collection forms a supremum. By Zorn's lemma, there exists an orthonormal system $(e_l)_{l\in I}$ which is maximal. It follows from Bessel's inequality (2.1) that this system is actually an orthonormal basis. □

Theorem 2.13 remains true for arbitrary Hilbert spaces when replacing the countable orthonormal system $(e_n)_{n\in\mathbb{N}}$ by an arbitrary orthonormal system $(e_l)_{l\in I}$.

## 3. * Fourier series

In the following we will identify the space $L^1(0, 2\pi)$ with

$$L^1_{2\pi}(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{C} \text{ measurable, } 2\pi\text{-periodic} : \int_0^{2\pi} |f|\, d\lambda < \infty\}.$$

Similarly, we identify $L^2(0, 2\pi)$ with $L^2_{2\pi}(\mathbb{R})$, and we define

$$C_{2\pi}(\mathbb{R}) := \{f \in C(\mathbb{R}) : f \text{ is } 2\pi\text{-periodic}\}.$$

Definition 3.1. For every $f \in L^1(0,2\pi) = L^1_{2\pi}(\mathbb{R})$ and every $n \in \mathbb{Z}$ we call

$$\hat{f}(n) := \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}\,dt$$

the *n-th Fourier coefficient* of $f$. The sequence $\hat{f} = (\hat{f}(n))$ is called the *Fourier transform* of $f$. The formal series $\frac{1}{\sqrt{2\pi}}\sum_{n\in\mathbb{Z}} \hat{f}(n)e^{in\cdot}$ is called the *Fourier series* of $f$.

Lemma 3.2. *For every $f \in L^1(0,2\pi) = L^1_{2\pi}(\mathbb{R})$ we have $\hat{f} \in l^\infty(\mathbb{Z})$ and the* Fourier transform $\hat{}\,: L^1(0,2\pi) \to l^\infty$ *is a bounded, linear operator. More precisely,*

$$\|\hat{f}\|_\infty \leq \frac{1}{2\pi}\,\|f\|_1, \quad f \in L^1(0,2\pi).$$

Proof. For every $f \in L^1(0,2\pi)$ and every $n \in \mathbb{Z}$,

$$|\hat{f}(n)| = \frac{1}{2\pi}\Big|\int_0^{2\pi} f(t)e^{-int}\,dt\Big| \leq \frac{1}{2\pi}\int_0^{2\pi} |f(t)|dt.$$

This proves that $\hat{f} \in l^\infty$ and the required bound on $\|\hat{f}\|_\infty$. Linearity of $\hat{}$ is clear. □

Lemma 3.3 (Riemann-Lebesgue). *For every $f \in L^1(0,2\pi) = L^1_{2\pi}(\mathbb{R})$ we have $\hat{f} \in c_0(\mathbb{Z})$, i.e.*

$$\lim_{|n|\to\infty} |\hat{f}(n)| = 0.$$

Proof. Let $f \in L^1(0,2\pi) = L^1_{2\pi}(\mathbb{R})$ and $n \in \mathbb{Z}$, $n \neq 0$. Then

$$\begin{aligned}
\hat{f}(n) &= \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}\,dt \\
&= \frac{1}{4\pi}\int_0^{2\pi} f(t)e^{-int}(1 - e^{i\pi\frac{n}{n}})\,dt \\
&= \frac{1}{4\pi}\int_0^{2\pi} f(t)(e^{-int} - e^{-in(t-\frac{\pi}{n})})\,dt \\
&= \frac{1}{4\pi}\int_0^{2\pi} (f(t) - f(t+\frac{\pi}{n}))e^{-int}\,dt,
\end{aligned}$$

so that

$$|\hat{f}(n)| \leq \frac{1}{4\pi}\int_0^{2\pi} |f(t) - f(t+\frac{\pi}{n})|\,dt.$$

Hence, if $f = 1_O \in L^1(0,2\pi)$ for some open set $O \subset [0,2\pi]$, then $\hat{f} \in c_0(\mathbb{Z})$ by Lebesgue dominated convergence theorem. On the other hand, since span $\{1_O : O \subset [0,2\pi] \text{ open}\}$ is dense in $L^1(0,2\pi)$, since the Fourier transform is bounded with values in $l^\infty(\mathbb{Z})$ (Lemma 3.2), and since $c_0(\mathbb{Z})$ is a closed subspace of $l^\infty(\mathbb{Z})$, we find that $\hat{f} \in c_0(\mathbb{Z})$ for every $f \in L^1(0,2\pi)$. □

REMARK 3.4. At the end of the proof of the Lemma of Riemann-Lebesgue, we used the following general principle: if $T \in \mathcal{L}(X, Y)$ is a bounded linear operator between two normed linear spaces $X$, $Y$, and if $M \subset X$ is dense, then $\operatorname{Rg} T \subset \overline{T(M)}$. We used in addition that $c_0(\mathbb{Z})$ is closed in $l^\infty(\mathbb{Z})$.

THEOREM 3.5. *Let $f \in C_{2\pi}(\mathbb{R})$ be differentiable in some point $s \in \mathbb{R}$. Then*

$$f(s) = \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{ins}.$$

PROOF. Note that for $f_s(t) := f(s + t)$,

$$\hat{f}_s(n) = \frac{1}{2\pi}\int_0^{2\pi} f(s+t)e^{-int}\,dt = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-in(t-s)}\,dt = e^{ins}\hat{f}(n).$$

Hence, replacing $f$ by $f_s$, if necessary, we may without loss of generality assume that $s = 0$. Moreover, replacing $f$ by $f - f(0)$, if necessary, we may without loss of generality assume that $f(0) = 0$. We hence have to show that if $f$ is differentiable in 0 and if $f(0) = 0$, then $\sum_{n\in\mathbb{Z}} \hat{f}(n) = 0$.

Let $g(t) := \frac{f(t)}{1-e^{it}}$. Since $f$ is differentiable in 0, $f(0) = 0$, and since $f$ is $2\pi$-periodic, the function $g$ belongs to $C_{2\pi}(\mathbb{R})$. By the Lemma of Riemann-Lebesgue, $\hat{g} \in c_0(\mathbb{Z})$. Note that

$$\hat{f}(n) = \frac{1}{2\pi}\int_0^{2\pi} g(t)(1-e^{it})e^{-int}\,dt = \hat{g}(n) - \hat{g}(n-1).$$

Hence,

$$\begin{aligned}\sum_{k=-n}^{n} \hat{f}(k) &= \sum_{k=-n}^{n} \hat{g}(k) - \hat{g}(k-1) \\ &= \hat{g}(n) - \hat{g}(-n-1) \to 0 \qquad (n \to \infty).\end{aligned}$$

This is the claim. □

COROLLARY 3.6. *For every $f \in C^1_{2\pi}(\mathbb{R}) := C_{2\pi}(\mathbb{R}) \cap C^1(\mathbb{R})$ and every $t \in \mathbb{R}$*

$$f(t) = \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{int}.$$

REMARK 3.7. We will see that the convergence in the preceeding corollary is even uniform in $t \in \mathbb{R}$.

Throughout the following, we equip the space $L^2(0, 2\pi) = L^2_{2\pi}(\mathbb{R})$ with the scalar product given by

$$(f, g) := \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\,dt,$$

which differs from the usual scalar product by the factor $\frac{1}{2\pi}$.

LEMMA 3.8. *The space $C^1_{2\pi}(\mathbb{R})$ is dense in $L^2_{2\pi}(\mathbb{R})$.*

Proof. We first prove that $C([0,2\pi])$ is dense in $L^2(0,2\pi) = L^2_{2\pi}(\mathbb{R})$. For this, consider first a characteristic function $f = 1_{(a,b)} \in L^2(0,2\pi)$. Let $(g_n) \subset C([0,2\pi])$ be defined by

$$g_n(t) := \begin{cases} 1, & t \in [a,b], \\ 1 + n(t-a), & t \in [a-1/n, a), \\ 1 - n(t-b), & t \in (b, b+1/n], \\ 0, & \text{else}. \end{cases}$$

It is then easy to see that $\lim_{n\to\infty} \|f - g_n\|_{L^2} = 0$, so that $f = 1_{(a,b)} \in \overline{C([0,2\pi])}^{\|\cdot\|_{L^2}}$.

In the second step, consider a characteristic function $f = 1_A$ of an arbitrary Borel set $A \in \mathcal{B}([0,2\pi])$, and let $\varepsilon > 0$. By outer regularity of the Lebesgue measure, there exists an open set $O \supset A$ such that $\lambda(O \setminus A) < \varepsilon^2$. Recall that $O$ is the countable union of mutually disjoint intervals. Since $O$ has finite measure, there exist finitely many (mutually disjoint) intervals $(a_n, b_n) \subset O$ ($1 \leq n \leq N$) such that $\lambda(O \setminus \bigcup_{n=1}^N (a_n, b_n)) \leq \varepsilon^2$. By the preceeding step, for every $1 \leq n \leq N$ there exists $g_n \in C([0,2\pi])$ such that $\|1_{(a_n,b_n)} - g_n\|_2 \leq \frac{\varepsilon}{N}$. Let $g := \sum_{n=1}^N g_n \in C([0,2\pi])$. Then

$$\begin{aligned} \|f - g\|_2 &\leq \|1_A - 1_O\|_2 + \|1_O - 1_{\bigcup_{n=1}^N (a_n,b_n)}\|_2 + \|1_{\bigcup_{n=1}^N (a_n,b_n)} - g\|_2 \\ &\leq \varepsilon + \varepsilon + \|\sum_{n=1}^N (1_{(a_n,b_n)} - g_n)\|_2 \\ &\leq 3\varepsilon. \end{aligned}$$

This proves $1_A \in \overline{C([0,2\pi])}^{\|\cdot\|_{L^2}}$ for every Borel set $A \in \mathcal{B}([0,2\pi])$. Since $\overline{\operatorname{span}}\,\{1_A : A \in \mathcal{B}([0,2\pi])\} = L^2(0,2\pi)$, we find that $C([0,2\pi])$ is dense in $L^2(0,2\pi)$.

It remains to show that $C^1_{2\pi}(\mathbb{R})$ is dense in $C([0,2\pi])$ for the norm $\|\cdot\|_2$. So let $f \in C([0,2\pi])$ and let $\varepsilon > 0$. By Weierstrass' theorem, there exists a function $g_0 \in C^\infty([0,2\pi])$ (even a polynomial!) such that $\|f - g_0\|_\infty \leq \varepsilon$. Let $g_1 \in C^1([0,2\pi])$ be such that $g_1(2\pi) = g_1'(2\pi) = 0$, $g_1(0) = g_0(2\pi) - g_0(0)$ and $g_1'(0) = g_0'(2\pi) - g_0'(0)$ and $\|g_1\|_2 \leq \varepsilon$. Such a function $g_1$ exists: it suffices for example to consider functions for which the derivative is of the form

$$g_1'(t) = \begin{cases} g_0(2\pi) - g_0(0) + ct, & t \in [0, h_1], \\ g_0(2\pi) - g_0(0) + ch_1 + d(t - h_1), & t \in (h_1, h_2), \\ 0, & t \in [h_2, 2\pi], \end{cases}$$

with appropriate constants $0 \leq h_1 \leq h_2$ and $c$, $d \in \mathbb{C}$. Having chosen $g_1$, we let $g = g_0 + g_1$ and we calculate that

$$\|f - g\|_2 \leq \|f - g_0\|_2 + \|g_1\|_2 \leq 2\varepsilon.$$

Since $g$ extends to a function in $C^1_{2\pi}(\mathbb{R})$, we have thus proved that $C^1_{2\pi}(\mathbb{R})$ is dense in $L^2_{2\pi}(\mathbb{R})$. $\square$

REMARK 3.9. An adaptation of the above proof actually shows that for every $1 \le p < \infty$ and every compact interval $[a,b] \subset \mathbb{R}$, the space $C([a,b])$ is dense in $L^p(a,b)$. A further application of Weierstrass' theorem actually shows that the space of all polynomials is dense in $L^p(a,b)$. In particular, we may obtain the following result.

COROLLARY 3.10. *The space $L^p(a,b)$ is separable if $1 \le p < \infty$. The space $L^\infty(a,b)$ is* not *separable.*

COROLLARY 3.11. *Let $e_n(t) := e^{int}$, $n \in \mathbb{Z}$, $t \in \mathbb{R}$. Then $(e_n)_{n\in\mathbb{Z}}$ is an orthonormal basis in $L^2_{2\pi}(\mathbb{R})$.*

PROOF. The fact that $(e_n)_{n\in\mathbb{Z}}$ is an orthonormal system in $L^2_{2\pi}(\mathbb{R})$ is an easy calculation. We only have to prove that $\operatorname{span}\{e_n : n \in \mathbb{Z}\}$ is dense in $L^2_{2\pi}(\mathbb{R})$. Note that $\hat{f}(n) = (f, e_n)$ for every $f \in L^2_{2\pi}(\mathbb{R})$ and every $n \in \mathbb{Z}$. By Lemma 2.12, we know that for every $f \in L^2_{2\pi}(\mathbb{R})$

$$g := \sum_{n\in\mathbb{Z}} \hat{f}(n) e_n \text{ exists in } L^2_{2\pi}(\mathbb{R}).$$

In particular, a subsequence of $(\sum_{n=-k}^{k} \hat{f}(n)e_n)$ converges almost everywhere to $g$. But by Corollary 3.6 we know that $(\sum_{n=-k}^{k} \hat{f}(n)e_n)$ converges pointwise everywhere to $f$ if $f \in C^1_{2\pi}(\mathbb{R})$. As a consequence, for every $f \in C^1_{2\pi}(\mathbb{R})$,

$$\lim_{k\to\infty} \sum_{n=-k}^{k} \hat{f}(n) e_n = f \text{ in } L^2_{2\pi}(\mathbb{R}),$$

so that $\operatorname{span}\{e_n : n \in \mathbb{Z}\}$ is dense in $(C^1_{2\pi}(\mathbb{R}), \|\cdot\|_{L^2_{2\pi}})$. Since $C^1_{2\pi}(\mathbb{R})$ is dense in $L^2_{2\pi}(\mathbb{R})$ by Lemma 3.8, we find that $(e_n)_{n\in\mathbb{Z}}$ is an orthonormal basis in $L^2_{2\pi}(\mathbb{R})$. $\square$

THEOREM 3.12 (Plancherel). *For every $f \in L^2_{2\pi}(\mathbb{R})$ we have $\hat{f} \in l^2(\mathbb{Z})$ and the Fourier transform $\hat{}: L^2_{2\pi}(\mathbb{R}) \to l^2(\mathbb{Z})$ is an isometric isomorphism. Moreover, for every $f \in L^2_{2\pi}(\mathbb{R})$,*

$$\sum_{n\in\mathbb{Z}} \hat{f}(n) e_n = f \text{ in } L^2_{2\pi}(\mathbb{R}),$$

*i.e. the Fourier series of $f$ converges to $f$ in the $L^2$ sense.*

PROOF. By Corollary 3.11, the sequence $(e_n)_{n\in\mathbb{Z}}$ is an orthonormal basis in $L^2_{2\pi}(\mathbb{R})$. Moreover, recall that for every $f \in L^2_{2\pi}(\mathbb{R})$ and every $n \in \mathbb{Z}$, $\hat{f}(n) = (f, e_n)$. Hence, by Theorem 2.13, $\hat{f} \in l^2(\mathbb{Z})$, $f = \sum_{n\in\mathbb{Z}} \hat{f}(n)e_n$, and $\|f\|_{L^2_{2\pi}} = \|\hat{f}\|_{l^2}$ (the last property being Parseval's identity). $\square$

Corollary 3.13. *Let $f \in C_{2\pi}(\mathbb{R})$ be such that $\hat{f} \in l^1(\mathbb{Z})$. Then*

$$\sum_{n\in\mathbb{Z}} \hat{f}(n)e_n = f \text{ in } C_{2\pi}(\mathbb{R}),$$

*i.e. the Fourier series of $f$ converges uniformly to $f$.*

Proof. Note that for every $n \in \mathbb{Z}$, $\|e_n\|_\infty = 1$. The assumption $\hat{f} \in l^1(\mathbb{Z})$ therefore implies that the series $\sum_{n\in\mathbb{Z}} \hat{f}(n)e_n$ converges absolutely in $C_{2\pi}(\mathbb{R})$, i.e. for the uniform norm $\|\cdot\|_\infty$. Since $(C_{2\pi}(\mathbb{R}), \|\cdot\|_\infty)$ is complete, the series $\sum_{n\in\mathbb{Z}} \hat{f}(n)e_n$ converges uniformly to some element $g \in C_{2\pi}(\mathbb{R})$. By Plancherel, $g = f$. □

Remark 3.14. The assumption $\hat{f} \in l^1(\mathbb{Z})$ in Corollary 3.13 is essential. For general $f \in C_{2\pi}(\mathbb{R})$, the Fourier series $\sum_{n\in\mathbb{Z}} \hat{f}(n)e_n$ need not not converge uniformly. Questions regarding the convergence of Fourier series (which type of convergence? for which function?) can go deeply into the theory of harmonic analysis and answers are sometimes quite involved. The $L^2$ theory gives in this context satisfactory answers with relatively easy proofs (see Plancherel's theorem). For continuous functions we state the following result without giving a proof.

Theorem 3.15 (Féjer). *For every $f \in C_{2\pi}(\mathbb{R})$ one has*

$$\lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \sum_{n=-k}^{k} \hat{f}(n)e_n = f \text{ in } C_{2\pi}(\mathbb{R}),$$

*i.e. the Fourier series of $f$ converges in the* Césaro mean *uniformly to $f$.*

## 4. Linear functionals on Hilbert spaces

In this section, we start to discuss bounded functionals on Banach spaces and Hilbert spaces. The case of Hilbert spaces is considerably easy but it has far reaching consequences.

Theorem 4.1 (Riesz-Fréchet). *Let $H$ be a Hilbert space. Then for every bounded linear functional $\varphi \in H'$ there exists a unique $y \in H$ such that*

$$\varphi(x) = (x, y) \quad \forall x \in H.$$

Proof. *Uniqueness.* Let $y_1$, $y_2 \in H$ be two elements such that

$$\varphi(x) = (x, y_1) = (x, y_2) \quad \forall x \in H.$$

Then $(x, y_1 - y_2) = 0$ for every $x \in H$, in particular also for $x = y_1 - y_2$. This implies $\|y_1 - y_2\|^2 = 0$, i.e. $y_1 = y_2$.

*Existence.* We may assume that $\varphi \neq 0$ since the case $\varphi = 0$ is trivial. Let $\tilde{y} \in (\ker\varphi)^\perp \setminus \{0\}$. Since $H \neq \ker\varphi$ and since $\ker\varphi$ is closed, such a $\tilde{y}$ exists. Next, let

$$y := \overline{\varphi(\tilde{y})}/\|\tilde{y}\|^2\, \tilde{y}.$$

Note that $\varphi(y) = \|y\|^2 = (y, y)$. Recall that every $x \in H$ can be uniquely written as $x = x_0 + \lambda y$ with $x_0 \in \ker \varphi$ and $\lambda \in \mathbb{K}$ so that $\lambda y \in (\ker \varphi)^\perp$. Note that $(\ker \varphi)^\perp$ is one-dimensional. Hence, for every $x \in H$,

$$\begin{aligned} \varphi(x) &= \varphi(x_0 + \lambda y) \\ &= \varphi(x_0) + \lambda \varphi(y) \\ &= \lambda \varphi(y) \\ &= \lambda (y, y) \\ &= (\lambda y, y) \\ &= (x_0, y) + (\lambda y, y) \\ &= (x, y). \end{aligned}$$

The claim is proved. □

COROLLARY 4.2. *Let $J : H \to H'$ be the mapping which maps to every $y \in H$ the functional $Jy \in H'$ given by $Jy(x) = (x, y)$. Then $J$ is antilinear (linear in case $\mathbb{K} = \mathbb{R}$), isometric and bijective.*

PROOF. The fact that $J$ is isometric follows from the Cauchy-Schwarz inequality. Antilinearity (or linearity in case $\mathbb{K} = \mathbb{R}$) follows from the sesquilinearity (resp. bilinearity) of the scalar product on $H$. Since $J$ is isometric, it is injective. The surjectivity of $J$ follows from Theorem 4.1. □

REMARK 4.3. The theorem of Riesz-Fréchet allows us to identify any (real) Hilbert space $H$ with its dual space $H'$. Note, however, that there are situations in which one does not identify $H'$ with $H$. This is for example the case when $V$ is a second Hilbert space which embeds continuously and densely into $H$, i.e. for which there exists a bounded, injective $J : V \to H$ with dense range.

## 5. Sobolev spaces

**5.1. Preparation.** Let $\Omega \subset \mathbb{R}^N$ be an *open* set. For every continuous function $\varphi \in C(\Omega)$ we define the *support*

$$\operatorname{supp} \varphi := \overline{\{x \in \Omega : \varphi(x) \neq 0\}}.$$

The support is by definition always closed. Next we let

$$\mathcal{D}(\Omega) := C_c^\infty(\Omega) := \{\varphi \in C^\infty(\Omega) : \operatorname{supp} \varphi \subset \Omega \text{ is compact}\}$$

be the space of *test functions* on $\Omega$, and

$$L^1_{loc}(\Omega) := \{f : \Omega \to \mathbb{K} \text{ measurable} : \int_K |f| < \infty \, \forall K \subset \Omega \text{ compact}\}$$

the space of *locally integrable functions*.

For every $f \in L^1_{loc}(\mathbb{R}^N)$ and every $\varphi \in \mathcal{D}(\mathbb{R}^N)$ we define the *convolution* $f * \varphi$ by

$$\begin{aligned} f * \varphi(x) &:= \int_{\mathbb{R}^N} f(x-y)\varphi(y)\,dy \\ &= \int_{\mathbb{R}^N} f(y)\varphi(x-y)\,dy. \end{aligned}$$

LEMMA 5.1. *For every $f \in L^1_{loc}(\mathbb{R}^N)$ and every $\varphi \in \mathcal{D}(\mathbb{R}^N)$ one has $f * \varphi \in C^\infty(\mathbb{R}^N)$ and for every $1 \le i \le N$,*

$$\frac{\partial}{\partial x_i}(f * \varphi) = f * \frac{\partial \varphi}{\partial x_i}.$$

PROOF. Let $e_i \in \mathbb{R}^N$ be the $i$-th unit vector. Then

$$\lim_{h \to 0} \frac{1}{h}(\varphi(x + he_i) - \varphi(x)) = \frac{\partial \varphi}{\partial x_i}(x)$$

uniformly in $x \in \mathbb{R}^N$ (note that $\varphi$ has compact support). Hence, for every $x \in \mathbb{R}^N$

$$\begin{aligned} & \frac{1}{h}(f * \varphi(x + he_i) - f * \varphi(x)) \\ = & \frac{1}{h} \int_{\mathbb{R}^N} f(y)(\varphi(x + he_i - y) - \varphi(x-y))\,dy \\ \to & \int_{\mathbb{R}^N} f(y) \frac{\partial \varphi}{\partial x_i}(x-y)\,dy. \end{aligned}$$

□

The following theorem is proved in courses on measure theory. We omit the proof.

THEOREM 5.2 (Young's inequality). *Let $f \in L^p(\mathbb{R}^N)$ and $\varphi \in \mathcal{D}(\mathbb{R}^N)$. Then $f * \varphi \in L^p(\mathbb{R}^N)$ and*

$$\|f * \varphi\|_p \le \|f\|_p \, \|\varphi\|_1.$$

THEOREM 5.3. *For every $1 \le p < \infty$ and every open $\Omega \subset \mathbb{R}^N$ the space $\mathcal{D}(\Omega)$ is dense in $L^p(\Omega)$.*

PROOF. The technique of this proof (*regularization* and *truncation*) is important in the theory of partial differential equations, distributions and Sobolev spaces. The first step (regularization) is based on Lemma 5.1. The truncation step is in this case relatively easy.

*Regularization.* Let $\varphi \in \mathcal{D}(\mathbb{R}^N)$ be a positive function such that $\|\varphi\|_1 = \int_{\mathbb{R}^N} \varphi = 1$. One may take for example the function

$$\varphi(x) := \begin{cases} c\, e^{1/(1-|x|^2)} & \text{if } |x| < 1, \\ 0 & \text{otherwise,} \end{cases} \tag{5.1}$$

with an appropriate constant $c > 0$. Then let $\varphi_n(x) := n^N \varphi(nx)$, so that $\|\varphi_n\|_1 = \int_{\mathbb{R}^N} \varphi_n = 1$ for every $n \in \mathbb{N}$.

Let $f \in L^p(\mathbb{R}^N)$. By Lemma 5.1 and Young's inequality (Theorem 5.2), for every $n \in \mathbb{N}$, $f_n := f * \varphi_n \in C^\infty(\mathbb{R}^N) \cap L^p(\mathbb{R}^N)$ and $\|f_n\|_p \leq \|f\|_p$. Hence, for every $n \in \mathbb{N}$ the operator $T_n : L^p(\mathbb{R}^N) \to L^p(\mathbb{R}^N)$, $f \mapsto f * \varphi_n$ is linear and bounded and $\|T_n\| \leq 1$. Moreover, if $f = 1_I$ for some bounded interval $I = (a_1, b_1) \times \cdots \times (a_N, b_N) \subset \Omega$, then

$$\begin{aligned}
\|f_n - f\|_p^p &= \int_{\mathbb{R}^N} \left| \int_{\mathbb{R}^N} f(x-y)\varphi(ny) n^N \, dy - f(x) \right|^p dx \\
&= \int_{\mathbb{R}^N} \left| \int_{\mathbb{R}^N} (f(x - \frac{y}{n}) - f(x))\varphi(y) \, dy \right|^p dx \\
&\leq \int_{\mathbb{R}^N} \left( \int_{\mathbb{R}^N} |f(x - \frac{y}{n}) - f(x)| \varphi(y) \, dy \right)^p dx \to 0
\end{aligned}$$

as $n \to \infty$ by Lebesgue's dominated convergence theorem. In other words, $\lim_{n\to\infty} \|T_n f - f\|_p = 0$ for every $f = 1_I$ with $I$ as above. Since span $\{1_I : I \subset \mathbb{R}^N \text{ bounded interval}\}$ is dense in $L^p(\mathbb{R}^N)$, we find that $\lim_{n\to\infty} \|T_n f - f\|_p = 0$ for every $f$ from a dense subset $M$ of $L^p(\mathbb{R}^N)$. Since the $T_n$ are bounded, we conclude that $T_n f \to f$ in $L^p(\mathbb{R}^N)$ for every $f \in L^p(\mathbb{R}^N)$ (see the Lemma 5.4 below). This proves that $L^p \cap C^\infty(\mathbb{R}^N)$ is dense in $L^p(\mathbb{R}^N)$.

*Truncation.* Now we consider a general open set $\Omega \subset \mathbb{R}^N$ and prove the claim. Let $\varphi \in \mathcal{D}(\mathbb{R}^N)$ be a positive test function such that $\operatorname{supp} \varphi \subset \overline{B(0,1)}$ and $\int_{\mathbb{R}^N} \varphi = 1$ (one may take for example the function from (5.1)). Then let $\varphi_n(x) := n^N \varphi(nx)$.

For every $n \in \mathbb{N}$ we let

$$K_n := \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) \geq \frac{1}{n}\} \cap \overline{B(0,n)},$$

so that $K_n \subset \Omega$ is compact for every $n \in \mathbb{N}$.

Now let $f \in L^p(\Omega) \subset L^p(\mathbb{R}^N)$ and $\varepsilon > 0$. Let

$$f1_{K_n}(x) = \begin{cases} f(x) & \text{if } x \in K_n, \\ 0 & \text{if } x \in \Omega \setminus K_n. \end{cases}$$

By Lebesgue's dominated convergence theorem (since $\bigcup_n K_n = \Omega$),

$$\|f - f1_{K_n}\|_p^p = \int_\Omega |f|^p (1 - 1_{K_n})^p \to 0 \text{ as } n \to \infty.$$

In particular, there exists $n \in \mathbb{N}$ such that $\|f - f1_{K_n}\|_p \leq \varepsilon$.

For every $m \geq 4n$ we define $g_m := (f1_{K_n}) * \varphi_m \in L^p \cap C^\infty(\mathbb{R}^N)$; note that we here consider $L^p(\Omega)$ as a subspace of $L^p(\mathbb{R}^N)$ by extending functions in $L^p(\Omega)$ by 0 outside $\Omega$. However, since $g_m = 0$ outside $K_{2n}$, we find that actually $g_m \in \mathcal{D}(\Omega)$. By the first step (regularisation), there exists $m \geq 4n$ so large that $\|g_m - f 1_{K_n}\|_p \leq \varepsilon$. For such $m$ we have $\|f - g_m\|_p \leq 2\varepsilon$, and the claim is proved. □

In the proof of the preceeding lemma, we have used the following fundamental principle.

Lemma 5.4. *Let $X$ and $Y$ be two normed spaces, let $(T_n) \in \mathcal{L}(X,Y)$ be a bounded sequence of bounded operators. Assume that there exists a dense set $M \subset X$ such that $\lim_{n\to\infty} T_n x$ exists for every $x \in M$. Then $\lim_{n\to\infty} T_n x =: Tx$ exists for every $x \in X$ and $T \in \mathcal{L}(X,Y)$.*

Proof. Define $Tx := \lim_{n\to\infty} T_n x$ for every $x \in \operatorname{span} M$. Then

$$\|Tx\| = \lim_{n\to\infty} \|T_n x\| \leq \sup_n \|T_n\| \, \|x\|,$$

i.e. $T : \operatorname{span} M \to Y$ is a bounded linear operator. Since $M$ is dense in $X$, $T$ admits a unique bounded extension $T : X \to Y$.

Let $x \in X$ and $\varepsilon > 0$. Since $M$ is dense in $X$, there exists $y \in M$ such that $\|x - y\| \leq \varepsilon$. By assumption, there exists $n_0$ such that for every $n \geq n_0$ we have $\|T_n y - Ty\| \leq \varepsilon$. Hence, for every $n \geq n_0$,

$$\begin{aligned} \|T_n x - Tx\| &\leq \|T_n x - T_n y\| + \|T_n y - Ty\| + \|Ty - Tx\| \\ &\leq \sup_n \|T_n\| \, \|x - y\| + \varepsilon + \|T\| \, \|x - y\| \\ &\leq \varepsilon(\sup_n \|T_n\| + 1 + \|T\|), \end{aligned}$$

and therefore $\lim_{n\to\infty} T_n x = Tx$. □

Lemma 5.5. *Let $f \in L^1_{loc}(\Omega)$ be such that*

$$\int_\Omega f\varphi = 0 \quad \forall \varphi \in \mathcal{D}(\Omega).$$

*Then $f = 0$.*

Proof. We first assume that $f \in L^1(\Omega)$ is real and that $\Omega$ has finite measure. By Theorem 5.3, for every $\varepsilon > 0$ there exists $g \in \mathcal{D}(\Omega)$ such that $\|f - g\|_1 \leq \varepsilon$. By assumption, this implies

$$|\int_\Omega g\varphi| = |\int_\Omega (f - g)\varphi| \leq \varepsilon \|\varphi\|_\infty \quad \forall \varphi \in \mathcal{D}(\Omega).$$

Let $K_1 := \{x \in \Omega : g(x) \geq \varepsilon\}$ and $K_2 := \{x \in \Omega : g(x) \leq -\varepsilon\}$. Since $g$ is a test function, the sets $K_1$, $K_2$ are compact. Since they are disjoint and do not touch the boundary of $\Omega$,

$$\inf\{|x - y|, |x - z|, |y - z| : x \in K_1, \, y \in K_2, \, z \in \partial\Omega\} =: \delta > 0.$$

Let $K_i^\delta := \{x \in \Omega : \operatorname{dist}(x, K_i) \leq \delta/4\}$ $(i = 1, 2)$. Then $K_1^\delta$ and $K_2^\delta$ are two compact disjoint subsets of $\Omega$. Let

$$h(x) := \begin{cases} 1 & \text{if } x \in K_1^\delta, \\ -1 & \text{if } x \in K_2^\delta, \\ 0 & \text{else}, \end{cases}$$

choose a positive test function $\varphi \in \mathcal{D}(\mathbb{R}^N)$ such that $\int_{\mathbb{R}^N} \varphi = 1$ and $\operatorname{supp} \varphi \subset B(0, \delta/8)$, and let $\psi := h * \varphi$. Then $\psi \in \mathcal{D}(\Omega)$, $-1 \leq \psi \leq 1$, $\psi = 1$ on $K_1$ and $\psi = -1$ on $K_2$. Let $K := K_1 \cup K_2$. Then

$$\int_K |g| = \int_K g\psi \leq \varepsilon + \int_{\Omega \setminus K} |g\psi| \leq \varepsilon + \int_{\Omega \setminus K} |g|.$$

Hence,

$$\int_\Omega |g| = \int_K |g| + \int_{\Omega \setminus K} |g| \leq \varepsilon + 2 \int_{\Omega \setminus K} |g| \leq \varepsilon(1 + 2\,|\Omega|),$$

which implies

$$\int_\Omega |f| \leq \int_\Omega |f - g| + \int_\Omega |g| \leq 2\,\varepsilon(1 + |\Omega|).$$

Since $\varepsilon > 0$ was arbitrary, we find that $f = 0$.

The general case can be obtained from the particular case ($f \in L^1$ and $|\Omega| < \infty$) by considering first real and imaginary part of $f$ separately, and then by considering $f 1_B$ for all closed (compact) balls $B \subset \Omega$. □

**5.2. Sobolev spaces in one dimension.** Recall the fundamental rule of partial integration: if $f$, $g \in C^1([a, b])$ on some compact interval $[a, b]$, then

$$\int_a^b f g' = f(b)g(b) - f(a)g(a) - \int_a^b f' g.$$

In particular, for every $f \in C^1([a, b])$ and every $\varphi \in \mathcal{D}(a, b)$

$$\int_a^b f \varphi' = - \int_a^b f' \varphi, \tag{5.2}$$

since $\varphi(a) = \varphi(b) = 0$.

Definition 5.6 (Sobolev spaces). Let $-\infty \leq a < b \leq \infty$ and $1 \leq p \leq \infty$. We define

$$W^{1,p}(a, b) := \{u \in L^p(a, b) : \exists g \in L^p(a, b)\, \forall \varphi \in \mathcal{D}(a, b) : \int_a^b u\varphi' = - \int_a^b g\varphi\}.$$

The space $W^{1,p}(a, b)$ is called (first) *Sobolev space*. If $p = 2$, then we also write $H^1(a, b) := W^{1,2}(a, b)$.

By Lemma 5.5, the function $g \in L^p(a,b)$ is uniquely determined if it exists. In the following, we will write $u' := g$, in accordance with (5.2). We equip $W^{1,p}(a,b)$ with the norm

$$\|u\|_{W^{1,p}} := \|u\|_p + \|u'\|_p,$$

and if $p = 2$, then we define the scalar product

$$(u,v)_{H^1} := \int_a^b u\bar{v} + \int_a^b u'\bar{v}',$$

which actually yields the norm $\|u\|_{H^1} = (\|u\|_2^2 + \|u'\|_2^2)^{\frac{1}{2}}$ (which is equivalent to $\|\cdot\|_{W^{1,2}}$).

Lemma 5.7. *The Sobolev spaces $W^{1,p}(a,b)$ are Banach spaces, which are separable if $p \neq \infty$. The space $H^1(a,b)$ is a separable Hilbert space.*

Proof. The fact that the $W^{1,p}$ are Banach spaces, or that $H^1$ is a Hilbert space, is an exercise. Recall that $L^p(a,b)$ is separable (Remark 3.9). Hence, the product space $L^p(a,b) \times L^p(a,b)$ is separable, and also every subspace of this product space is separable. Now consider the linear mapping

$$T : W^{1,p}(a,b) \to L^p(a,b) \times L^p(a,b), \quad u \mapsto (u,u'),$$

which is bounded and even isometric. Hence, $W^{1,p}$ is isometrically isomomorphic to a subspace of $L^p \times L^p$ which is separable. Hence $W^{1,p}$ is separable. □

Lemma 5.8. *Let $u \in W^{1,p}(a,b)$ be such that $u' = 0$. Then $u$ is constant.*

Proof. Choose $\psi \in \mathcal{D}(a,b)$ such that $\int_a^b \psi = 1$. Then, for every $\varphi \in \mathcal{D}(a,b)$, the function $\varphi - (\int_a^b \varphi)\psi$ is the derivative of a test function since $\int_a^b (\varphi - (\int_a^b \varphi)\psi) = 0$. Hence, by definition,

$$0 = \int_a^b u(\varphi - (\int_a^b \varphi)\psi),$$

or, with $c = \int_a^b u\psi = \text{const}$,

$$\int_a^b (u-c)\varphi = 0 \quad \forall \varphi \in \mathcal{D}(a,b).$$

By Lemma 5.5, $u = c$ almost everywhere. □

Lemma 5.9. *Let $-\infty < a < b < \infty$ and let $t_0 \in [a,b]$. Let $g \in L^p(a,b)$ and define*

$$u(t) := \int_{t_0}^t g(s)\, ds, \quad t \in [a,b].$$

*Then $u \in W^{1,p}(a,b)$ and $u' = g$.*

PROOF. Let $\varphi \in \mathcal{D}(a,b)$. Then, by Fubini's theorem,

$$\begin{aligned}
\int_a^b u\varphi' &= \int_a^b \int_{t_0}^t g(s)\, ds\varphi'(t)\, dt \\
&= \int_a^{t_0} \int_{t_0}^t g(s)\, ds\varphi'(t)\, dt + \int_{t_0}^b \int_{t_0}^t g(s)\, ds\varphi'(t)\, dt \\
&= -\int_a^{t_0} \int_a^s \varphi'(t)\, dtg(s)\, ds + \int_{t_0}^b \int_s^b \varphi'(t)\, dtg(s)\, ds \\
&= -\int_a^{t_0} \varphi(s)g(s)\, ds - \int_{t_0}^b \varphi(s)g(s)\, ds \\
&= -\int_a^b g\varphi.
\end{aligned}$$

□

THEOREM 5.10. *Let $u \in W^{1,p}(a,b)$ (bounded or unbounded interval). Then there exists $\tilde{u} \in C((a,b))$ which is continuous up to the boundary of $(a,b)$, which coincides with $u$ almost everywhere and such that for every $s$, $t \in (a,b)$*

$$\tilde{u}(t) - \tilde{u}(s) = \int_s^t u'(r)\, dr.$$

PROOF. Fix $t_0 \in (a,b)$ and define $v(t) := \int_{t_0}^t u'(s)\, ds$ ($t \in \overline{(a,b)}$). Clearly, the function $v$ is continuous. By Lemma 5.9, $v \in W^{1,p}(c,d)$ for every bounded interval $(c,d) \subset (a,b)$, and $v' = u'$. By Lemma 5.8, $u - v = C$ for some constant $C$ which clearly does not depend on the choice of the interval $(c,d)$. This proves that $u$ coincides almost everywhere with the continuous function $\tilde{u} = v + C$. By Lemma 5.9,

$$\tilde{u}(t) - \tilde{u}(s) = v(t) - v(s) = \int_s^t u'(r)\, dr.$$

□

REMARK 5.11. By Theorem 5.10, we will identify every function $u \in W^{1,p}(a,b)$ with its continuous representant, and we say that every function in $W^{1,p}(a,b)$ is continuous.

LEMMA 5.12 (Extension lemma). *Let $u \in W^{1,p}(a,b)$. Then there exists $\tilde{u} \in W^{1,p}(\mathbb{R})$ such that $\tilde{u} = u$ on $(a,b)$.*

PROOF. Assume first that $a$ and $b$ are finite and define

$$g(t) := \begin{cases} u'(t) & \text{if } t \in [a,b], \\ u(a) & \text{if } t \in [a-1,a), \\ -u(b) & \text{if } t \in (b,b+1], \\ 0 & \text{else.} \end{cases}$$

Then $g \in L^p(\mathbb{R})$. Let $\tilde{u}(t) := \int_{-\infty}^t g(s)\, ds$, so that $\tilde{u} = u$ on $(a,b)$. By Lemma 5.9, $\tilde{u} \in W^{1,p}(c,d)$ for every bounded interval $(c,d) \in \mathbb{R}$. However, $\tilde{u} = 0$ outside $(a-1, b+1)$ which implies that $\tilde{u} \in W^{1,p}(\mathbb{R})$.

The case of $a = -\infty$ or $b = \infty$ is treated similarly. □

LEMMA 5.13. *For every $1 \le p < \infty$, the space $\mathcal{D}(\mathbb{R})$ is dense in $W^{1,p}(\mathbb{R})$.*

PROOF. Let $u \in W^{1,p}(\mathbb{R})$.

*Regularization:* Choose a positive test function $\varphi \in \mathcal{D}(\mathbb{R})$ such that $\int_{\mathbb{R}} \varphi = 1$ and put $\varphi_n(x) = n\varphi(nx)$. Then $u_n := u * \varphi_n \in C^\infty \cap L^p(\mathbb{R})$, $u_n' = u' * \varphi_n \in L^p(\mathbb{R})$ and

$$\lim_{n\to\infty} \|u - u_n\|_p = 0 \text{ and}$$
$$\lim_{n\to\infty} \|u' - u_n'\|_p = 0,$$

so that $\lim_{n\to\infty} \|u - u_n\|_{W^{1,p}} = 0$. This proves that $W^{1,p}(\mathbb{R}) \cap C^\infty(\mathbb{R})$ is dense in $W^{1,p}(\mathbb{R})$.

*Truncation:* Choose a sequence $(\psi_n) \subset \mathcal{D}(\mathbb{R})$ such that $0 \le \psi_n \le 1$, $\psi_n = 1$ on $[-n,n]$ and $\|\psi_n'\|_\infty \le C$ for all $n \in \mathbb{N}$. Let $\varepsilon > 0$. Choose $v \in C^\infty \cap W^{1,p}(\mathbb{R})$ such that $\|u - v\|_{W^{1,p}} \le \varepsilon$ (regularization step). For every $n \in \mathbb{N}$, one has $v\psi_n \in \mathcal{D}(\mathbb{R})$ and it is easy to check that for all $n$ large enough, $\|v - v\psi_n\|_{W^{1,p}} \le \varepsilon$. The claim is proved. □

COROLLARY 5.14. *For every $u \in W^{1,p}(a,b)$ (bounded or unbounded interval, $1 \le p < \infty$) and every $\varepsilon > 0$, there exists $v \in \mathcal{D}(\mathbb{R})$ such that $\|u - v|_{(a,b)}\|_{W^{1,p}} \le \varepsilon$.*

PROOF. Given $u \in W^{1,p}(a,b)$, we first choose an extension $\tilde{u} \in W^{1,p}(\mathbb{R})$ (extension lemma 5.12) and then a test function $v \in \mathcal{D}(\mathbb{R})$ such that $\|\tilde{u} - v\|_{W^{1,p}(\mathbb{R})} \le \varepsilon$ (Lemma 5.13). Then $\|\tilde{u} - v\|_{W^{1,p}(a,b)} = \|u - v\|_{W^{1,p}(a,b)} \le \varepsilon$. □

COROLLARY 5.15. *Every function $u \in W^{1,p}(a,b)$ is continuous and bounded and there exists a constant $C \ge 0$ such that*

$$\|u\|_\infty \le C\, \|u\|_{W^{1,p}} \quad \forall u \in W^{1,p}(a,b).$$

PROOF. If $p = \infty$, there is nothing to prove. We first prove the claim for the case $(a,b) = \mathbb{R}$.

So let $1 \le p < \infty$ and let $v \in \mathcal{D}(\mathbb{R})$. Then $G(v) := |v|^{p-1}v \in C_c^1(\mathbb{R})$ and $G(v)' = p|v|^{p-1}v'$. By Hölder's inequality,

$$|G(v)(x)| = p\, \Big| \int_{-\infty}^x |v|^{p-1} v' \Big| \le p\, \|v\|_p^{p-1}\, \|v'\|_p,$$

so that by Young's inequality ($ab \le \frac{1}{p}a^p + \frac{1}{p'}b^{p'}$)

$$\|v\|_\infty = \|G(v)\|_\infty^{1/p} \le C \|v\|_{W^{1,p}}.$$

Since $\mathcal{D}(\mathbb{R})$ is dense in $W^{1,p}(\mathbb{R})$ by Lemma 5.13, the claim for $(a,b) = \mathbb{R}$ follows by an approximation argument.

The case $(a,b) \neq \mathbb{R}$ is an exercise. □

THEOREM 5.16 (Product rule, partial integration). *Let $u$, $v \in W^{1,p}(a,b)$ ($1 \leq p \leq \infty$). Then:*

(i) *(Product rule). The product $uv$ belongs to $W^{1,p}(a,b)$ and*

$$(uv)' = u'v + uv'.$$

(ii) *(Partial integration). If $-\infty < a < b < \infty$, then*

$$\int_a^b u'v = u(b)v(b) - u(a)v(a) - \int_a^b uv'.$$

PROOF. Since every function in $W^{1,p}(a,b)$ is bounded, we find that $uv$, $u'v + uv' \in L^p(a,b)$. Choose sequences $(u_n)$, $(v_n) \subset \mathcal{D}(\mathbb{R})$ such that $\lim_{n\to\infty} u_n|_{(a,b)} = u$ and $\lim_{n\to\infty} v_n|_{(a,b)} = v$ in $W^{1,p}(a,b)$ (Corollary 5.14). By Corollary 5.15, this implies also $\lim_{n\to\infty} \|u_n|_{(a,b)} - u\|_\infty = \lim_{n\to\infty} \|v_n|_{(a,b)} - v\|_\infty = 0$. The classical product rule implies

$$(u_n v_n)' = u_n' v_n + u_n v_n' \text{ for every } n \in \mathbb{N},$$

and the classical rule of partial integration implies

$$\int_a^b u_n' v_n = u_n(b)v_n(b) - u_n(a)v_n(a) - \int_a^b u_n v_n' \text{ for every } n \in \mathbb{N}.$$

The claim follows upon letting $n$ tend to $\infty$. □

DEFINITION 5.17. For every $1 \leq p \leq \infty$ and every $k \geq 2$ we define inductively the *Sobolev spaces*

$$W^{k,p}(a,b) := \{u \in W^{1,p}(a,b) : u' \in W^{k-1,p}(a,b)\},$$

which are Banach spaces for the norms

$$\|u\|_{W^{k,p}} := \sum_{j=0}^{k} \|u^{(j)}\|_p.$$

We denote $H^k(a,b) := W^{k,2}(a,b)$ which is a Hilbert space for the scalar product

$$(u,v)_{H^k} := \sum_{j=0}^{k} (u^{(j)}, v^{(j)})_{L^2}.$$

Finally, we define

$$W_0^{k,p}(a,b) := \overline{\mathcal{D}(a,b)}^{\|\cdot\|_{W^{k,p}}},$$

i.e. $W_0^{k,p}(a,b)$ is the closure of the test functions in $W^{k,p}(a,b)$, and we put $H_0^k(a,b) := W_0^{k,2}(a,b)$.

THEOREM 5.18. *Let $-\infty < a < b < \infty$. A function $u \in W_0^{1,p}(a,b)$ if and only if $u \in W^{1,p}(a,b)$ and $u(a) = u(b) = 0$.*

THEOREM 5.19. *Let $-\infty < a < b < \infty$. For every $f \in L^2(a,b)$ there exists a unique function $u \in H_0^1(a,b) \cap H^2(a,b)$ such that*

$$\begin{cases} u - u'' = f & \text{and} \\ u(a) = u(b) = 0 & . \end{cases} \tag{5.3}$$

PROOF. We first note that if $u \in H_0^1(a,b) \cap H^2(a,b)$ is a solution, then, by partial integration (Theorem 5.16), for every $v \in H_0^1(a,b)$

$$\int_a^b (uv + u'v') = (u,v)_{H_0^1} = \int_a^b fv. \tag{5.4}$$

By the Cauchy-Schwarz inequality, the linear functional $\varphi \in H_0^1(a,b)'$ defined by $\varphi(v) = \int_a^b fv$ is bounded:

$$|\varphi(v)| \leq \|f\|_2 \, \|v\|_2 \leq \|f\|_2 \, \|v\|_{H_0^1}.$$

By the theorem of Riesz-Fréchet, there exists a unique $u \in H_0^1(a,b)$ such that (5.4) holds true for all $v \in H_0^1(a,b)$. This proves uniqueness of a solution of (5.3), and if we prove that in addition $u \in H^2(a,b)$, then we prove existence, too. However, (5.4) holds in particular for all $v \in \mathcal{D}(a,b)$, i.e.

$$\int_a^b u'v' = -\int_a^b (u-f)v \quad \forall v \in \mathcal{D}(a,b)$$

and $u - f \in L^2(a,b)$ by assumption. Hence, by definition, $u' \in H^1(a,b)$, i.e. $u \in H^2(a,b)$ and $u'' = u - f$. Using also Theorem 5.18, the claim is proved. □

**5.3. Sobolev spaces in several dimensions.** In order to motivate Sobolev spaces in several space dimensions, we have to recall the partial integration rule in this case.

THEOREM 5.20 (Gauß). *Let $\Omega \subset \mathbb{R}^N$ be open and bounded such that $\partial\Omega \in C^1$. Then there exists a measure $\sigma$ on $\partial\Omega$ such that for every $u$, $v \in C^1(\bar{\Omega})$ and every $1 \leq i \leq N$*

$$\int_\Omega u \frac{\partial v}{\partial x_i} = \int_{\partial\Omega} uvn_i \, d\sigma - \int_\Omega \frac{\partial u}{\partial x_i} v,$$

*where $n(x) = (n_i(x))_{1 \leq i \leq N}$ denotes the outer normal vector at a point $x \in \partial\Omega$.*

In particular, if $u \in C^1(\bar{\Omega})$ and $\varphi \in \mathcal{D}(\Omega)$, then

$$\int_\Omega u \frac{\partial \varphi}{\partial x_i} = -\int_\Omega \frac{\partial u}{\partial x_i} \varphi.$$

DEFINITION 5.21 (Sobolev spaces). Let $\Omega \subset \mathbb{R}^N$ be any open set and $1 \leq p \leq \infty$. We define

$$W^{1,p}(\Omega) := \{u \in L^p(\Omega) : \forall 1 \leq i \leq N \exists g_i \in L^p(\Omega) \quad \forall \varphi \in \mathcal{D}(\Omega) : \int_\Omega u \frac{\partial \varphi}{\partial x_i} = -\int_\Omega g_i \varphi\}.$$

The space $W^{1,p}(\Omega)$ is called (first) *Sobolev space*. If $p = 2$, then we also write $H^1(\Omega) := W^{1,2}(\Omega)$.

Let $u \in W^{1,p}(\Omega)$. By Lemma 5.5, the functions $g_i$ are uniquely determined. We write $\frac{\partial u}{\partial x_i} := g_i$ and call $\frac{\partial u}{\partial x_i}$ the *partial derivative* of $u$ with respect to $x_i$. As in the one-dimensional case, the following holds true.

LEMMA 5.22. *The Sobolev spaces $W^{1,p}(\Omega)$ are Banach spaces for the norms*

$$\|u\|_{W^{1,p}} := \|u\|_p + \sum_{i=1}^N \|\frac{\partial u}{\partial x_i}\|_p \quad (1 \leq p \leq \infty),$$

*and $H^1(\Omega)$ is a Hilbert space for the scalar product*

$$(u, v)_{H^1} := (u, v)_{L^2} + \sum_{i=1}^N (\frac{\partial u}{\partial x_i}, \frac{\partial v}{\partial x_i})_{L^2}.$$

PROOF. Exercise. □

Not all properties of Sobolev spaces on intervals carry over to Sobolev spaces on open sets $\Omega \subset \mathbb{R}^N$. For example, it is *not* true that every function $u \in W^{1,p}(\Omega)$ is continuous (without any further restrictions on $p$ and $\Omega$)!

DEFINITION 5.23. For every open $\Omega \subset \mathbb{R}^N$, $1 \leq p \leq \infty$ and every $k \geq 2$ we define inductively the *Sobolev spaces*

$$W^{k,p}(\Omega) := \{u \in W^{1,p}(\Omega) : \forall 1 \leq i \leq N : \frac{\partial u}{\partial x_i} \in W^{k-1,p}(\Omega)\},$$

which are Banach spaces for the norms

$$\|u\|_{W^{k,p}} := \|u\|_p + \sum_{i=0}^k \|\frac{\partial u}{\partial x_i}\|_{W^{k-1,p}}.$$

We denote $H^k(\Omega) := W^{k,2}(\Omega)$ which is a Hilbert space for the scalar product

$$(u, v)_{H^k} := (u, v)_{L^2} + \sum_{i=0}^k (\frac{\partial u}{\partial x_i}, \frac{\partial v}{\partial x_i})_{H^{k-1}}.$$

Finally, we define

$$W_0^{k,p}(\Omega) := \overline{\mathcal{D}(\Omega)}^{\|\cdot\|_{W^{k,p}}},$$

i.e. $W_0^{k,p}(\Omega)$ is the closure of the test functions in $W^{k,p}(\Omega)$, and we put $H_0^k(\Omega) := W_0^{k,2}(\Omega)$.

## 6. * Elliptic partial differential equations

Let $\Omega \subset \mathbb{R}^N$ be an open, bounded set, $f \in L^2(\Omega)$, and consider the elliptic partial differential equation

$$\begin{cases} u - \Delta u = f & \text{in } \Omega, \\ u = 0 & \text{in } \partial\Omega, \end{cases} \tag{6.1}$$

where

$$\Delta u(x) := \sum_{i=1}^N \frac{\partial^2}{\partial x_i^2} u(x)$$

stands for the *Laplace operator*.

*If* $u \in H_0^1(\Omega) \cap H^2(\Omega)$ is a solution of (6.1), then, by definition of the Sobolev spaces, for every $v \in \mathcal{D}(a,b)$

$$\begin{aligned} (u,v)_{H_0^1} &= \int_\Omega \big(uv + \sum_{i=1}^N \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}\big) \\ &= \int_\Omega \big(uv - \sum_{i=1}^N \frac{\partial^2 u}{\partial x_i^2} v\big) \\ &= \int_\Omega (u - \Delta u)v \\ &= \int_\Omega fv. \end{aligned}$$

By density of the test functions in $H_0^1(\Omega)$, this equality holds actually for all $v \in H_0^1(\Omega)$. This may justify the following definition of a weak solution.

Definition 6.1. A function $u \in H_0^1(\Omega)$ is called a *weak solution* of (6.1) if for every $v \in H_0^1(\Omega)$

$$(u,v)_{H_0^1} = \int_\Omega uv + \int_\Omega \nabla u \nabla v = \int_\Omega fv, \tag{6.2}$$

where $\nabla u = (\frac{\partial}{\partial x_i} u)_{1 \le i \le N}$.

Theorem 6.2. *Let* $\Omega \subset \mathbb{R}^N$ *be an open, bounded set. Then, for every* $f \in L^2(\Omega)$ *there exists a unique weak solution* $u \in H_0^1(\Omega)$ *of the problem* (6.1).

Proof. By the Cauchy-Schwarz inequality, the linear functional $\varphi \in H_0^1(\Omega)'$ defined by $\varphi(v) = \int_\Omega fv$ is bounded:

$$|\varphi(v)| \le \|f\|_2 \, \|v\|_2 \le \|f\|_2 \, \|v\|_{H_0^1}.$$

By the theorem of Riesz-Fréchet, there exists a unique $u \in H_0^1(\Omega)$ such that (6.2) holds true for all $v \in H_0^1(a,b)$. The claim is proved. □

CHAPTER 5

# Dual spaces

## 1. The theorem of Hahn-Banach

Given a normed space $X$, we denote by $X' := \mathcal{L}(X, \mathbb{K})$ the space of all bounded linear functionals on $X$. Recall that $X'$ is always a Banach space by Corollary 3.9 of Chapter 3.

However, *a priori* it is not clear whether there exists any bounded linear functional on a normed space $X$ (apart from the zero functional). This fundamental question and the analysis of dual spaces (analysis of functionals) shall be developed in this chapter.

The existence of nontrivial bounded functionals is guaranteed by the Hahn-Banach theorem which actually admits several versions. However, before stating the first version, we need the following definition.

DEFINITION 1.1. Let $X$ be a real or complex vector space. A function $p : X \to \mathbb{R}$ is called *sublinear* if

(i) $p(\lambda x) = \lambda p(x)$ for every $\lambda > 0$, $x \in X$, and
(ii) $p(x + y) \leq p(x) + p(y)$ for every $x$, $y \in X$.

EXAMPLE 1.2. On a normed space $X$, the norm $\| \cdot \|$ is sublinear. Every linear $p : X \to \mathbb{R}$ is sublinear.

THEOREM 1.3 (Hahn-Banach; version of linear algebra, real case). *Let $X$ be a real vector space, $U \subset X$ a linear subspace, and $p : X \to \mathbb{R}$ sublinear. Let $\varphi : U \to \mathbb{R}$ be linear such that*

$$\varphi(x) \leq p(x) \text{ for all } x \in U.$$

*Then there exists a linear $\tilde{\varphi} : X \to \mathbb{R}$ such that $\tilde{\varphi}(x) = \varphi(x)$ for every $x \in U$ (i.e. $\tilde{\varphi}$ is an extension of $\varphi$) and*

$$\tilde{\varphi}(x) \leq p(x) \text{ for all } x \in X. \tag{1.1}$$

The following lemma asserts that this version of Hahn-Banach is true in the special case when $X/U$ has dimension 1. It is an essential step in the proof of Theorem 1.3.

LEMMA 1.4. *Take the assumptions of Theorem 1.3 and assume in addition that* $\dim X/U = 1$. *Then the assertion of Theorem 1.3 is true.*

Proof. If $\dim X/U = 1$, then there exists $x_0 \in X \setminus U$ such that every $x \in X$ can be uniquely written in the form $x = u + \lambda x_0$ with $u \in U$ and $\lambda \in \mathbb{R}$. So we define $\tilde{\varphi} : X \to \mathbb{R}$ by

$$\tilde{\varphi}(x) := \tilde{\varphi}(u + \lambda x_0) := \varphi(u) + \lambda r,$$

where $r \in \mathbb{R}$ is a parameter which has to be chosen such that (1.1) holds, i.e. such that for every $u \in U$, $\lambda \in \mathbb{R}$,

$$\varphi(u) + \lambda r \le p(u + \lambda x_0). \tag{1.2}$$

If $\lambda = 0$, then this condition clearly holds for every $u \in U$ by the assumption on $\varphi$. If $\lambda > 0$, then (1.2) holds for every $u \in U$ if and only if

$$\begin{aligned} & \lambda r \le p(u + \lambda x_0) - \varphi(u) \text{ for every } u \in U \\ \Leftrightarrow \quad & r \le p(\frac{u}{\lambda} + x_0) - \varphi(\frac{u}{\lambda}) \text{ for every } u \in U \\ \Leftrightarrow \quad & r \le \inf_{v \in U} p(v + x_0) - \varphi(v). \end{aligned}$$

Similarly, if $\lambda < 0$, then (1.2) holds for every $u \in U$ if and only if

$$\begin{aligned} & \lambda r \le p(u + \lambda x_0) - \varphi(u) \text{ for every } u \in U \\ \Leftrightarrow \quad & -r \le p(\frac{u}{-\lambda} - x_0) - \varphi(\frac{u}{-\lambda}) \text{ for every } u \in U \\ \Leftrightarrow \quad & r \ge \sup_{w \in U} \varphi(w) - p(w - x_0). \end{aligned}$$

So it is possible to find an appropriate $r \in \mathbb{R}$ in the definition of $\tilde{\varphi}$ if and only if

$$\varphi(w) - p(w - x_0) \le p(v + x_0) - \varphi(v) \text{ for all } v, w \in U,$$

or, equivalently, if

$$\varphi(w) + \varphi(v) \le p(v + x_0) + p(w - x_0) \text{ for all } v, w \in U.$$

However, by the assumptions on $\varphi$ and $p$, for every $v$, $w \in U$,

$$\varphi(w) + \varphi(v) = \varphi(w + v) \le p(w + v) = p(v + x_0 + w - x_0) \le p(v + x_0) + p(w - x_0).$$

□

For the second step in the proof of Theorem 1.3, we need the Lemma of Zorn.

Lemma 1.5 (Zorn). *Let $(M, \le)$ be a ordered set. Assume that every totally ordered subset $T \subset M$ (i.e. for every $x$, $y \in T$ one either has $x \le y$ or $y \le x$) admits an upper bound. Then for every $x \in M$ there exists a maximal element $m \ge x$ (i.e. an element $m$ such that $m \le \tilde{m}$ implies $m = \tilde{m}$ for every $\tilde{m} \in M$).*

Proof of Theorem 1.3. Define the following set

$$\begin{aligned} M := \{(V, \varphi_V) \; : \; & V \subset X \text{ linear subspace}, U \subset V, \varphi_V : V \to \mathbb{R} \text{ linear, s.t.} \\ & \varphi(x) = \varphi_V(x) \, (x \in U) \text{ and } \varphi_V(x) \le p(x) \, (x \in V)\}, \end{aligned}$$

and equip it with the order relation $\leq$ defined by

$$(V_1, \varphi_{V_1}) \leq (V_2, \varphi_{V_2}) :\Leftrightarrow V_1 \subset V_2 \text{ and } \varphi_{V_1}(x) = \varphi_{V_2}(x) \text{ for all } x \in V_1.$$

Then $(M, \leq)$ is an ordered set. Let $T = ((V_i, \varphi_{V_i}))_{i \in I} \subset M$ be a totally ordered subset. Then the element $(V, \varphi_V) \in M$ defined by

$$V := \bigcup_{i \in I} V_i \text{ and } \varphi_V(x) = \varphi_{V_i}(x) \text{ for } x \in V_i$$

is an upper bound of $T$. By the Lemma of Zorn, the set $M$ admits a maximal element $(X_0, \varphi_{X_0})$. Assume that $X_0 \neq X$. Then, by Lemma 1.4, we could construct an element which is strictly larger than $(X_0, \varphi_{X_0})$, a contradiction to the maximality of $(X_0, \varphi_{X_0})$. Hence, $X = X_0$, and $\tilde{\varphi} := \varphi_{X_0}$ is an element we are looking for. □

The complex version of the Hahn-Banach theorem reads as follows.

THEOREM 1.6 (Hahn-Banach; version of linear algebra, complex case). *Let $X$ be a complex vector space, $U \subset X$ a linear subspace, and $p : X \to \mathbb{R}$ sublinear. Let $\varphi : U \to \mathbb{C}$ be linear such that*

$$\operatorname{Re} \varphi(x) \leq p(x) \textit{ for all } x \in U.$$

*Then there exists a linear $\tilde{\varphi} : X \to \mathbb{C}$ such that $\tilde{\varphi}(x) = \varphi(x)$ for every $x \in U$ (i.e. $\tilde{\varphi}$ is an extension of $\varphi$) and*

$$\operatorname{Re} \tilde{\varphi}(x) \leq p(x) \textit{ for all } x \in X. \tag{1.3}$$

PROOF. We may consider $X$ also as a real vector space. Note that $\psi(x) := \operatorname{Re} \varphi(x)$ is an $\mathbb{R}$-linear functional on $X$. By Theorem 1.3, there exists an extension $\tilde{\psi} : X \to \mathbb{R}$ of $\psi$ satisfying

$$\tilde{\psi}(x) \leq p(x) \text{ for every } x \in X.$$

Let

$$\tilde{\varphi}(x) := \tilde{\psi}(x) - i\tilde{\psi}(ix), \quad x \in X.$$

It is an exercise to show that $\tilde{\varphi}$ is $\mathbb{C}$-linear, that $\varphi(x) = \tilde{\varphi}(x)$ for every $x \in U$ and it is clear from the definition that $\operatorname{Re} \tilde{\varphi}(x) = \tilde{\psi}(x)$. Thus, $\tilde{\varphi}$ is a possible element we are looking for. □

THEOREM 1.7 (Hahn-Banach; extension of bounded functionals). *Let $X$ be a normed space and $U \subset X$ a linear subspace. Then for every bounded linear $u' : U \to \mathbb{K}$ there exists a bounded linear extension $x' : X \to \mathbb{K}$ (i.e. $x'|_U = u'$) such that $\|x'\| = \|u'\|$.*

PROOF. We first assume that $X$ is a real normed space. The function $p : X \to \mathbb{R}$ defined by $p(x) := \|u'\| \, \|x\|$ is sublinear and

$$u'(x) \leq p(x) \text{ for every } x \in U.$$

By the first Hahn-Banach theorem (Theorem 1.3), there exists a linear $x' : X \to \mathbb{R}$ extending $u'$ such that

$$x'(x) \le p(x) = \|u'\| \, \|x\| \text{ for every } x \in X.$$

Replacing $x$ by $-x$, this implies that

$$|x'(x)| \le \|u'\| \, \|x\| \text{ for every } x \in X.$$

Hence, $x'$ is bounded and $\|x'\| \le \|u'\|$. On the other hand, one trivially has

$$\|x'\| = \sup_{\substack{x \in X \\ \|x\| \le 1}} |x'(x)| \ge \sup_{\substack{x \in U \\ \|x\| \le 1}} |x'(x)| = \sup_{\substack{x \in U \\ \|x\| \le 1}} |u'(x)| = \|u'\|.$$

If $X$ is a complex normed space, then the second Hahn-Banach theorem (Theorem 1.6) implies that there exists a linear $x' : X \to \mathbb{C}$ such that

$$\operatorname{Re} x'(x) \le p(x) = \|u'\| \, \|x\| \text{ for every } x \in X.$$

In particular,

$$|x'(x)| = \sup_{\theta \in [0, 2\pi]} \operatorname{Re} x'(e^{i\theta} x) \le \|u'\| \, \|x\| \text{ for every } x \in X,$$

so that again $x'$ is bounded and $\|x'\| \le \|u'\|$. The inequality $\|x'\| \ge \|u'\|$ follows as above. □

Corollary 1.8. *If $X$ is a normed space, then for every $x \in X \setminus \{0\}$ there exists $x' \in X'$ such that*

$$\|x'\| = 1 \text{ and } x'(x) = \|x\|.$$

*In particular, $X'$* separates the points *of $X$, i.e. for every $x_1$, $x_2 \in X$, $x_1 \ne x_2$, there exists $x' \in X'$ such that $x'(x_1) \ne x'(x_2)$.*

Proof. By the Hahn-Banach theorem (Theorem 1.7), there exists an extension $x' \in X'$ of the functional $u' : \operatorname{span}\{x\} \to \mathbb{K}$ defined by $u'(\lambda x) = \lambda \|x\|$ such that $\|x'\| = \|u'\| = 1$.

For the proof of the second assertion, set $x := x_1 - x_2$. □

Corollary 1.9. *If $X$ is a normed space, then for every $x \in X$*

$$\|x\| = \sup_{\substack{x' \in X' \\ \|x'\| \le 1}} |x'(x)|. \tag{1.4}$$

Proof. For every $x' \in X'$ with $\|x'\| \le 1$ one has

$$|x'(x)| \le \|x'\| \, \|x\| \le \|x\|,$$

which proves one of the required inequalities. The other inequality follows from Corollary 1.8. □

Remark 1.10. The equality (1.4) should be compared to the definition of the norm of an element $x' \in X'$:

$$\|x'\| = \sup_{\substack{x \in X \\ \|x\| \leq 1}} |x'(x)|.$$

From now on, it will be convenient to use the following notation. Given a normed space $X$ and elements $x \in X$, $x' \in X'$, we write

$$\langle x', x \rangle := \langle x', x \rangle_{X' \times X} := x'(x).$$

For the bracket $\langle \cdot, \cdot \rangle$, we note the following properties. The function

$$\begin{aligned} \langle \cdot, \cdot \rangle : X' \times X &\to \mathbb{K}, \\ (x', x) &\mapsto \langle x', x \rangle = x'(x) \end{aligned}$$

is bilinear and for every $x' \in X'$, $x \in X$,

$$|\langle x', x \rangle| \leq \|x'\| \, \|x\|.$$

The bracket $\langle \cdot, \cdot \rangle$ thus appeals to the notion of the scalar product on inner product spaces, and the last inequality appeals to the Cauchy-Schwarz inequality, but note, however, that the bracket is *not* a scalar product since it is defined on a pair of two different spaces. Moreover, even if $X = H$ is a complex Hilbert space, then the bracket differs from the scalar product in that it is bilinear instead of sesquilinear.

Corollary 1.11. *Let $X$ be a normed space, $U \subset X$ a closed linear subspace and $x \in X \setminus U$. Then there exists $x' \in X'$ such that*

$$x'(x) \neq 0 \text{ and } x'(u) = 0 \text{ for every } u \in U.$$

Proof. Let $\pi : X \to X/U$ be the quotient map ($\pi(x) = x + U$). Since $x \notin U$, we have $\pi(x) \neq 0$. By Corollary 1.8, there exists $\varphi \in (X/U)'$ such that $\varphi(\pi(x)) \neq 0$. Then $x' := \varphi \circ \pi \in X'$ is a functional we are looking for. □

Definition 1.12. A linear subspace $U$ of a normed space $X$ is called *complemented* if there exists a projection $P \in \mathcal{L}(X)$ such that $\operatorname{Rg} P = U$.

Remark 1.13. If $P$ is a projection (i.e. $P^2 = P$), then $Q = I - P$ is also a projection and $\operatorname{Rg} P = \ker Q$. Hence, if $P$ is a bounded projection, then $\operatorname{Rg} P$ is necessarily closed. Thus, a necessary condition for $U$ to be complemented is that $U$ is closed.

Corollary 1.14. *Every finite dimensional subspace of a normed space is complemented.*

Proof. Let $U$ be a finite dimensional subspace of a normed space $X$. Let $(b_i)_{1 \leq i \leq N}$ be a basis of $U$. By Corollary 1.11, there exist functionals $x_i' \in X'$ such that

$$\langle x_i', b_j \rangle = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Let $P : X \to X$ be defined by

$$Px := \sum_{i=1}^{N} \langle x_i', x\rangle b_i, \quad x \in X.$$

Then $Pb_i = b_i$ for every $1 \le i \le N$, and thus $P^2 = P$, i.e. $P$ is a projection. Moreover, $\operatorname{Rg} P = U$ by construction. By the estimate

$$\begin{aligned} \|Px\| &\le \sum_{i=1}^{N} |\langle x_i', x\rangle| \, \|b_i\| \\ &\le \big( \sum_{i=1}^{N} \|x_i'\| \, \|b_i\| \big) \, \|x\|, \end{aligned}$$

the projection $P$ is bounded. □

The following lemma which does not depend on the Hahn-Banach theorem is stated for completeness.

LEMMA 1.15. *In a Hilbert space every closed linear subspace is complemented.*

PROOF. Take the orthogonal projection onto the closed subspace as a possible projection. □

COROLLARY 1.16. *If $X$ is a normed space such that $X'$ is separable, then $X$ is separable, too.*

PROOF. Let $D' = \{x_n' : n \in \mathbb{N}\}$ be a dense subset of the unit sphere of $X'$. For every $n \in \mathbb{N}$ we choose an element $x_n \in X$ such that $\|x_n\| \le 1$ and $|\langle x_n', x_n\rangle| \ge \frac{1}{2}$. We claim that $D := \operatorname{span}\{x_n : n \in \mathbb{N}\}$ is dense in $X$. If this was not true, i.e. if $\bar{D} \ne X$, then, by Corollary 1.11, we find an element $x' \in X' \setminus \{0\}$ such that $x'(x_n) = 0$ for every $n \in \mathbb{N}$. We may without loss of generality assume that $\|x'\| = 1$. Since $D'$ is dense in the unit sphere of $X'$, we find $n_0 \in \mathbb{N}$ such that $\|x' - x_{n_0}'\| \le \frac{1}{4}$. But then

$$\frac{1}{2} \le |\langle x_{n_0}', x_{n_0}\rangle| = |\langle x_{n_0}' - x', x_{n_0}\rangle| \le \|x_{n_0}' - x'\| \, \|x_{n_0}\| \le \frac{1}{4},$$

which is a contradiction. Hence, $\bar{D} = X$ and $X$ is separable by Lemma 2.3 of Chapter 4. □

## 2. Weak convergence, reflexivity

Given a normed space $X$, we call $X'' := (X')' = \mathcal{L}(X', \mathbb{K})$ the *bidual* of $X$.

LEMMA 2.1. *Let $X$ be a normed space. Then the mapping*

$$\begin{aligned} J : X &\to X'', \\ x &\mapsto (x' \mapsto \langle x', x\rangle), \end{aligned}$$

*is well defined and isometric.*

PROOF. The linearity of $x' \mapsto \langle x', x \rangle$ is clear, and from the inequality

$$|Jx(x')| = |\langle x', x \rangle| \leq \|x'\| \, \|x\|,$$

follows that $Jx \in X''$ (i.e. $J$ is well defined) and $\|Jx\| \leq \|x\|$. The fact that $J$ is isometric follows from Corollary 1.8. □

DEFINITION 2.2. A normed space $X$ is called *reflexive* if the isometry $J$ from Lemma 2.1 is surjective, i.e. if $JX = X''$. In other words: a normed space $X$ is reflexive if for every $x'' \in X''$ there exists $x \in X$ such that

$$\langle x'', x' \rangle = \langle x', x \rangle \text{ for all } x' \in X'.$$

REMARK 2.3. If a normed space is reflexive then $X$ and $X''$ are isometrically isomorphic (via the operator $J$). Since $X''$ is always complete, a reflexive space is necessarily a Banach space.

Note that it can happen that $X$ and $X''$ are isomorphic without $X$ being reflexive (the example of such a Banach space is however quite involved). We point out that reflexivity means that the special operator $J$ is an isomorphism.

LEMMA 2.4. *Every Hilbert space is reflexive.*

PROOF. By the Theorem of Riesz-Fréchet, we may identify $H$ with its dual $H'$ and thus also $H$ with its bidual $H''$. The identification is done via the scalar product. It is an exercise to show that this identification of $H$ with $H''$ coincides with the mapping $J$ from Lemma 2.1. □

REMARK 2.5. It should be noted that for complex Hilbert spaces, the identification of $H$ with its dual $H'$ is only antilinear, but after the second identification ($H'$ with $H''$) it turns out that the identification of $H$ with $H''$ is linear.

LEMMA 2.6. *Every finite dimensional Banach space is reflexive.*

PROOF. It suffices to remark that if $X$ is finite dimensional, then

$$\dim X = \dim X' = \dim X'' < \infty.$$

Surjectivity of the mapping $J$ (which is always injective) thus follows from linear algebra. □

THEOREM 2.7. *The space $L^p(\Omega)$ is reflexive if $1 < p < \infty$ ($(\Omega, \mathcal{A}, \mu)$ being an arbitrary measure space).*

We will actually only prove the following special case.

THEOREM 2.8. *The spaces $l^p$ are reflexive if $1 < p < \infty$.*

The proof of Theorem 2.8 is based on the following lemma.

LEMMA 2.9. *Let $1 \leq p < \infty$ and let $q := \frac{p}{p-1}$ be the conjugate exponent so that $\frac{1}{p} + \frac{1}{q} = 1$. Then the operator*

$$\begin{aligned} T : l^q &\to (l^p)', \\ (a_n) &\mapsto ((x_n) \mapsto \sum_n a_n x_n), \end{aligned}$$

*is an isometric isomorphism, i.e. $(l^p)' = l^q$.*

PROOF. Linearity of $T$ is obvious. Assume first $p > 1$, so that $q < \infty$. Note that for every $a := (a_n) \in l^q \setminus \{0\}$ the sequence $(x_n) := (c\bar{a_n}|a_n|^{q-2})$ $(c = \|a\|_q^{-q/p})$ belongs to $l^p$ and

$$\|x\|_p^p = \|a\|_q^{-q} \sum_n |a_n|^{(q-1)p} = 1.$$

This particular $x \in l^p$ shows that

$$\|Ta\|_{(l^p)'} \geq \sum_n a_n x_n = \|a\|_q^{-q/p} \sum_n |a_n|^q = \|a\|_q^{q(p-1)/p} = \|a\|_q.$$

On the other hand, by Hölder's inequality,

$$\|Ta\|_{(l^p)'} = \sup_{\|x\|_p \leq 1} |\sum_n a_n x_n| \leq \|a\|_q,$$

so that $T$ is isometric in the case $p \in (1, \infty)$. The case $p = 1$ is very similar and will be omitted.

In order to show that $T$ is surjective, let $\varphi \in (l^p)'$. Denote by $e_n$ the $n$-th unit vector in $l^p$, and let $a_n := \varphi(e_n)$. If $p = 1$, then $(a_n) \in l^\infty = l^q$ by the trivial estimate

$$|a_n| = |\varphi(e_n)| \leq \|\varphi\| \, \|e_n\|_1 = \|\varphi\|.$$

If $p > 1$, then we may argue as follows. For every $N \in \mathbb{N}$,

$$\begin{aligned} \sum_{n=1}^N |a_n|^q &= \sum_{n=1}^N a_n \, \bar{a_n} \, |a_n|^{q-2} \\ &= \varphi(\sum_{n=1}^N \bar{a_n} \, |a_n|^{q-2} \, e_n) \\ &\leq \|\varphi\| \, \big( \sum_{n=1}^N |a_n|^{(q-1)p} \big)^{\frac{1}{p}} \\ &= \|\varphi\| \, \big( \sum_{n=1}^N |a_n|^q \big)^{\frac{1}{p}}, \end{aligned}$$

which is equivalent to

$$\big( \sum_{n=1}^N |a_n|^q \big)^{1-\frac{1}{p}} = \big( \sum_{n=1}^N |a_n|^q \big)^{\frac{1}{q}} \leq \|\varphi\|.$$

Since the right-hand side of this inequality does not depend on $N \in \mathbb{N}$, we obtain that $a := (a_n) \in l^q$ and $\|a\|_q \le \|\varphi\|$.

Next, observe that for every $x \in l^p$ one has

$$x = \sum_n x_n e_n = \lim_{N\to\infty} \sum_{n=1}^N x_n e_n,$$

the series converging in $l^p$ (here we need the restriction $p < \infty$!). Hence, for every $x \in l^p$, by the boundedness of $\varphi$,

$$\begin{aligned} \varphi(x) &= \lim_{N\to\infty} \varphi(\sum_{n=1}^N x_n e_n) \\ &= \lim_{N\to\infty} \sum_{n=1}^N x_n a_n \\ &= \sum_n x_n a_n \\ &= Ta(x). \end{aligned}$$

Hence, $T$ is surjective. □

Proof of Theorem 2.8. By Lemma 2.9, we may identify $(l^p)'$ with $l^q$ and, if $1 < p < \infty$ (!), also $(l^p)'' = (l^q)'$ with $l^p$. One just has to notice that this identification of $l^p$ with $(l^p)'' = l^p$ (the identity map on $l^p$) coincides with the operator $J$ from Lemma 2.1, so that $l^p$ is reflexive if $1 < p < \infty$. □

Lemma 2.10. *The spaces $l^1$, $L^1(\Omega)$ ($\Omega \subset \mathbb{R}^N$) and $C([0,1])$ are not reflexive.*

Proof. For every $t \in [0,1]$, let $\delta_t \in C([0,1])'$ be defined by

$$\langle \delta_t, f\rangle := f(t), \quad f \in C([0,1]).$$

Then $\|\delta_t\| = 1$ and whenever $t \neq s$, then

$$\|\delta_t - \delta_s\| = 2.$$

In particular, the uncountably many balls $B(\delta_t, \frac{1}{2})$ ($t \in [0,1]$) are mutually disjoint so that $C([0,1])'$ is not separable.

Now, if $C([0,1])$ were reflexive, then $C([0,1])'' = C([0,1])$ would be separable (since $C([0,1])$ is separable), and then, by Corollary 1.16, $C([0,1])'$ would be separable; a contradiction to what has been said before. This proves that $C([0,1])$ is not reflexive.

The cases of $l^1$ and $L^1(\Omega)$ are proved similarly. They are separable Banach spaces with nonseparable dual. □

Theorem 2.11. *Every closed subspace of a reflexive Banach space is reflexive.*

Proof. Let $X$ be a reflexive Banach space, and let $U \subset X$ be a closed subspace. Let $u'' \in U''$. Then the mapping $x'' : X' \to \mathbb{K}$ defined by

$$\langle x'', x' \rangle = \langle u'', x'|_U \rangle, \quad x' \in X',$$

is linear and bounded, i.e. $x'' \in X''$. By reflexivity of $X$, there exists $x \in X$ such that

$$\langle x', x \rangle = \langle u'', x'|_U \rangle, \quad x' \in X'. \tag{2.1}$$

Assume that $x \notin U$. Then, by Corollary 1.9, there exists $x' \in X'$ such that $x'|_U = 0$ and $\langle x', x \rangle \neq 0$; a contradiction to the last equality. Hence, $x \in U$. We need to show that

$$\langle u'', u' \rangle = \langle u', x \rangle, \forall u' \in U'. \tag{2.2}$$

However, if $u' \in U'$, then, by Hahn-Banach we can choose an extension $x' \in X'$, i.e. $x'|_U = u'$. The equation (2.2) thus follows from (2.1). □

Corollary 2.12. *The Sobolev spaces $W^{k,p}(\Omega)$ ($\Omega \subset \mathbb{R}^N$ open) are reflexive if $1 < p < \infty$, $k \in \mathbb{N}$.*

Proof. For example, for $k = 1$, the operator

$$\begin{aligned} T : W^{1,p}(\Omega) &\to L^p(\Omega)^{1+N}, \\ u &\mapsto (u, \frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_N}), \end{aligned}$$

is isometric, so that we may consider $W^{1,p}(\Omega)$ as a closed subspace of $L^p(\Omega)^{1+N}$ which is reflexive by Theorem 2.7. The claim follows from Theorem 2.11. □

Corollary 2.13. *A Banach space is reflexive if and only if its dual is reflexive.*

Proof. Assume that the Banach space $X$ is reflexive. Let $x''' \in X'''$ (the tridual!). Then the mapping $x' : X \to \mathbb{K}$ defined by

$$\langle x', x \rangle := \langle x''', J_X(x) \rangle, \quad x \in X,$$

is linear and bounded, i.e. $x' \in X'$ (here $J_X$ denotes the isometry $X \to X''$). Let $x'' \in X''$ be arbitrary. Since $X$ is reflexive, there exists $x \in X$ such that $J_X x = x''$. Hence,

$$\langle x''', x'' \rangle = \langle x''', J_X x \rangle = \langle x', x \rangle = \langle x'', x' \rangle,$$

which proves that $J_{X'} x' = x'''$, i.e. the isometry $J_{X'} : X' \to X'''$ is surjective. Hence, $X'$ is reflexive.

On the other hand, assume that $X'$ is reflexive. Then $X''$ is reflexive by the preceeding argument, and therefore $X$ (considered as a closed subspace of $X''$ via the isometry $J$) is reflexive by Theorem 2.11. □

DEFINITION 2.14. Let $X$ be a normed space. We say that a sequence $(x_n) \subset X$ *converges weakly* to some $x \in X$ if

$$\lim_{n\to\infty} \langle x', x_n\rangle = \langle x', x\rangle \text{ for every } x' \in X'.$$

Notations: if $(x_n)$ converges weakly to $x$, then we write $x_n \rightharpoonup x$, $w-\lim_{n\to\infty} x_n = x$, $x_n \to x$ in $\sigma(X, X')$, or $x_n \to x$ weakly.

THEOREM 2.15. *In a reflexive Banach space every bounded sequence admits a weakly convergent subsequence.*

PROOF. Let $(x_n)$ be a bounded sequence in a reflexive Banach space $X$. We first assume that $X$ is separable. Then $X''$ is separable by reflexivity, and $X'$ is separable by Corollary 1.16. Let $(x'_m) \subset X'$ be a dense sequence.

Since $(\langle x'_1, x_n\rangle)$ is bounded by the boundedness of $(x_n)$, there exists a subsequence $(x_{\varphi_1(n)})$ of $(x_n)$ ($\varphi_1 : \mathbb{N} \to \mathbb{N}$ is increasing, unbounded) such that

$$\lim_{n\to\infty} \langle x'_1, x_{\varphi_1(n)}\rangle \text{ exists.}$$

Similarly, there exists a subsequence $(x_{\varphi_2(n)})$ of $(x_{\varphi_1(n)})$ such that

$$\lim_{n\to\infty} \langle x'_2, x_{\varphi_2(n)}\rangle \text{ exists.}$$

Note that for this subsequence, we also have that

$$\lim_{n\to\infty} \langle x'_1, x_{\varphi_2(n)}\rangle \text{ exists.}$$

Iterating this argument, we find a subsequence $(x_{\varphi_3(n)})$ of $(x_{\varphi_2(n)})$ and finally for every $m \in \mathbb{N}$, $m \geq 2$, a subsequence $(x_{\varphi_m(n)})$ of $(x_{\varphi_{m-1}(n)})$ such that

$$\lim_{n\to\infty} \langle x'_j, x_{\varphi_m(n)}\rangle \text{ exists for every } 1 \leq j \leq m.$$

Let $(y_n) := (x_{\varphi_n(n)})$ be the 'diagonal sequence'. Then $(y_n)$ is a subsequence of $(x_n)$ and

$$\lim_{n\to\infty} \langle x'_m, y_n\rangle \text{ exists for every } m \in \mathbb{N}.$$

By Lemma 5.4 of Chapter 4, there exists $x'' \in X''$ such that

$$\lim_{n\to\infty} \langle x', y_n\rangle = \langle x', x''\rangle \text{ for every } x' \in X'.$$

Since $X$ is reflexive, there exists $x \in X$ such that $Jx = x''$. For this $x$, we have by definition of $J$

$$\lim_{n\to\infty} \langle x', y_n\rangle = \langle x', x\rangle \text{ exists for every } x' \in X',$$

i.e. $(y_n)$ converges weakly to $x$.

If $X$ is not separable as we first assumed, then one may replace $X$ by $\tilde{X} := \overline{\operatorname{span}}\,\{x_n : n \in \mathbb{N}\}$ which is separable. By the above, there exists $x \in \tilde{X}$ and a subsequence of $(x_n)$ (which we denote again by $(x_n)$) such that for every $\tilde{x}' \in \tilde{X}'$,

$$\lim_{n\to\infty} \langle \tilde{x}', x_n\rangle = \langle \tilde{x}', x\rangle,$$

i.e. $(x_n)$ converges weakly in $\tilde{X}$. If $x' \in X'$, then $x'|_{\tilde{X}} \in \tilde{X}'$, and it follows easily that the sequence $(x_n)$ also converges weakly in $X$ to the element $x$. $\square$

## 3. * Minimization of convex functionals

THEOREM 3.1 (Hahn-Banach; separation of convex sets). *Let $X$ be a Banach space, $K \subset X$ a closed, nonempty, convex subset, and $x_0 \in X \setminus K$. Then there exists $x' \in X'$ and $\varepsilon > 0$ such that*

$$\operatorname{Re}\langle x', x\rangle + \varepsilon \leq \operatorname{Re}\langle x', x_0\rangle, \quad x \in K.$$

LEMMA 3.2. *Let $K$ be an open, nonempty, convex subset of a Banach space $X$ such that $0 \in K$. Define the* Minkowski functional $p : X \to \mathbb{R}$ *by*

$$p(x) = \inf\{\lambda > 0 : \frac{x}{\lambda} \in K\}.$$

*Then $p$ is sublinear, there exists $M \geq 0$ such that*

$$p(x) \leq M\,\|x\|, \quad x \in X,$$

*and $K = \{x \in X : p(x) < 1\}$.*

PROOF. Since $B(0, r) \subset K$ for some $r > 0$, we find that

$$p(x) \leq \frac{1}{r}\,\|x\| \text{ for every } x \in X.$$

The property $p(\alpha x) = \alpha p(x)$ for every $\alpha > 0$ and every $x \in X$ is obvious.

Next, if $p(x) < 1$, then there exists $\lambda \in (0, 1)$ such that $x/\lambda \in K$. Hence, by convexity, $x = \lambda\frac{x}{\lambda} = \lambda\frac{x}{\lambda} + (1 - \lambda)0 \in K$. On the other hand, if $x \in K$, then $(1 + \varepsilon)x \in K$, since $K$ is open. Hence, $p(x) \leq (1 + \varepsilon)^{-1} < 1$, so that $K = \{x \in X : p(x) < 1\}$.

Let finally $x$, $y \in X$. Then for every $\varepsilon > 0$, $x/(p(x)+\varepsilon) \in K$ and $y/(p(y)+\varepsilon) \in K$. In particular, for every $t \in [0, 1]$,

$$\frac{t}{p(x) + \varepsilon}\,x + \frac{1 - t}{p(y) + \varepsilon}\,y \in K.$$

Setting $t = (p(x) + \varepsilon)/(p(x) + p(y) + 2\varepsilon)$, one finds that

$$\frac{x + y}{p(x) + p(y) + 2\varepsilon} \in K,$$

so that $p(x + y) \leq p(x) + p(y) + 2\varepsilon$. Since $\varepsilon > 0$ was arbitrary, we find $p(x + y) \leq p(x) + p(y)$. The claim is proved. $\square$

PROOF OF THEOREM 3.1. We prove the theorem for the case when $X$ is a real Banach space. The complex case is proved similarly.

We may without loss of generality assume that $0 \in K$; it suffices to translate $K$ and $x_0$ for this. Since $x_0 \notin K$ and since $K$ is closed, we find that $d := \operatorname{dist}(x_0, K) > 0$. Put

$$K_d := \{x \in X : \operatorname{dist}(x, K) < d/2\},$$

so that $K_d$ is an open, convex subset such that $0 \in K_d$. Let $p$ be the corresponding Minkowski functional (see Lemma 3.2).

Define on the one-dimensional subspace $U := \{\lambda x_0 : \lambda \in \mathbb{R}\}$ the functional $u' : U \to \mathbb{R}$ by $\langle u', \lambda x_0 \rangle = \lambda$. Then

$$\langle u', u \rangle \leq p(u), \quad u \in U.$$

By the Hahn-Banach theorem 1.3, there exists a linear extension $x' : X \to \mathbb{R}$ such that

$$\langle x', x \rangle \leq p(x), \quad x \in X. \tag{3.1}$$

In particular, by Lemma 3.2,

$$|\langle x', x \rangle| \leq M \, \|x\|,$$

so that $x' \in X'$ and $\|x'\| \leq M$. By construction, $\langle x', x_0 \rangle = 1$. Moreover, by (3.1) and Lemma 3.2, $\langle x', x \rangle < 1$ for every $x \in K \subset K_d$, so that

$$\langle x', x \rangle \leq \langle x', x_0 \rangle \, (= 1), \quad x \in K_d.$$

Replacing the above argument with $(1 - \varepsilon')x_0$ instead of $x_0$ (where $\varepsilon' > 0$ is chosen so small that $(1 - \varepsilon')x_0 \notin K_d$), we find that

$$\langle x', x \rangle + \varepsilon' \langle x', x_0 \rangle \leq \langle x', x_0 \rangle, \quad x \in K \subset K_d,$$

and putting $\varepsilon := \varepsilon' = \varepsilon' \langle x', x_0 \rangle > 0$ yields the claim. □

COROLLARY 3.3. *Let $X$ be a Banach space and $K \subset X$ a closed, convex subset (closed for the norm topology). If $(x_n) \subset K$ converges weakly to some $x \in X$, then $x \in K$.*

PROOF. Assume the contrary, i.e. $x \notin K$. By the Hahn-Banach theorem (Theorem 3.1), there exist $x' \in X'$ and $\varepsilon > 0$ such that

$$\operatorname{Re} \langle x', x_n \rangle + \varepsilon \leq \operatorname{Re} \langle x', x \rangle \text{ for every } n \in \mathbb{N},$$

a contradiction to the assumption that $x_n \rightharpoonup x$. □

A function $f : K \to \mathbb{R}$ on a convex subset $K$ of a Banach space $X$ is called *convex* if for every $x$, $y \in K$, and every $t \in [0, 1]$,

$$f(tx + (1 - t)y) \leq t \, f(x) + (1 - t) \, f(y). \tag{3.2}$$

COROLLARY 3.4. *Let $X$ be a Banach space, $K \subset X$ a closed, convex subset, and $f : K \to \mathbb{R}$ a continuous, convex function. If $(x_n) \subset K$ converges weakly to $x \in K$, then*

$$f(x) \leq \liminf_{n \to \infty} f(x_n).$$

PROOF. For every $l \in \mathbb{R}$, the set $K_l := \{x \in K : f(x) \leq l\}$ is closed (by continuity of $f$) and convex (by convexity of $f$). After extracting a subsequence, if necessary, we may assume that $l := \liminf_{n\to\infty} f(x_n) = \lim_{n\to\infty} f(x_n)$. Then for every $\varepsilon > 0$ the sequence $(x_n)$ is eventually in $K_{l+\varepsilon}$, i.e. except for finitely many $x_n$, the sequence $(x_n)$ lies in $K_{l+\varepsilon}$. Hence, by Corollary 3.3, $x \in K_{l+\varepsilon}$, which means that $f(x) \leq l + \varepsilon$. Since $\varepsilon > 0$ was arbitrary, the claim follows. □

**Theorem 3.5.** *Let $X$ be a reflexive Banach space, $K \subset X$ a closed, convex, nonempty subset, and $f : K \to \mathbb{R}$ a continuous, convex function such that*

$$\lim_{\substack{\|x\|\to\infty \\ x\in K}} f(x) = +\infty \text{ (coercivity).}$$

*Then there exists $x_0 \in K$ such that*

$$f(x_0) = \inf\{f(x) : x \in K\} > -\infty.$$

Proof. Let $(x_n) \subset K$ be such that $\lim_{n\to\infty} f(x_n) = \inf\{f(x) : x \in K\}$. By the coercivity assumption on $f$, the sequence $(x_n)$ is bounded. Since $X$ is reflexive, there exists a weakly convergent subsequence (Theorem 2.15); we denote by $x_0$ the limit. By Corollary 3.3, $x_0 \in K$. By Corollary 3.4,

$$f(x_0) \leq \lim_{n\to\infty} f(x_n) = \inf\{f(x) : x \in K\}.$$

The claim is proved. □

Remark 3.6. Theorem 3.5 remains true if $f$ is only lower semicontinuous, i.e. if

$$\liminf_{n\to\infty} f(x_n) \geq f(x)$$

for every convergent $(x_n) \subset K$ with $x = \lim x_n$. In fact, already Corollary 3.4 remains true if $f$ is only lower semicontinuous (and then Corollary 3.4 says that lower semicontinuity of a convex function in the norm topology implies lower semicontinuity in the weak topology). It suffices for example to remark that the sets $K_l := \{f \leq l\}$ ($l \in \mathbb{R}$) are closed as soon as $f$ is lower semicontinuous.

## 4. * The von Neumann minimax theorem

In the following theorem, we call a function $f : K \to \mathbb{R}$ on a convex subset $K$ of a Banach space $X$ *concave* if $-f$ is convex, or, equivalently, if for every $x$, $y \in K$ and every $t \in [0, 1]$,

$$f(tx + (1-t)y) \geq t\, f(x) + (1-t)\, f(y). \tag{4.1}$$

A function $f : K \to \mathbb{R}$ is called *strictly convex* (resp. *strictly concave*) if for every $x$, $y \in K$, $x \neq y$, $f(x) = f(y)$ the inequality in (3.2) (resp. (4.1)) is strict for $t \in (0, 1)$.

**Theorem 4.1** (von Neumann). *Let $K$ and $L$ be two closed, bounded, nonempty, convex subsets of reflexive Banach spaces $X$ and $Y$, respectively. Let $f : K \times L \to \mathbb{R}$ be a continuous function such that*

$$x \mapsto f(x, y) \text{ is strictly convex for every } y \in L, \text{ and}$$
$$y \mapsto f(x, y) \text{ is concave for every } x \in K.$$

*Then there exists $(\bar{x}, \bar{y}) \in K \times L$ such that*

$$f(\bar{x}, y) \leq f(\bar{x}, \bar{y}) \leq f(x, \bar{y}) \text{ for every } x \in K,\ y \in L. \tag{4.2}$$

REMARK 4.2. A point $(\bar{x},\bar{y}) \in K\times L$ satisfying (4.2) is called a *saddle point* of $f$.

A saddle point is a point of *equilibrium* in a two-person zero-sum game in the following sense: If the player controlling the strategy $x$ modifies his strategy when the second player plays $\bar{y}$, he increases his loss; hence, it is his interest to play $\bar{x}$. Similarly, if the player controlling the strategy $y$ modifies his strategy when the first player plays $\bar{x}$, he diminishes his gain; thus it is in his interest to play $\bar{y}$. This property of equilibrium of saddle points justifies their use as a (reasonable) solution in a two-person zero-sum game ([**1**]).

PROOF. Define the function $F : L \to \mathbb{R}$ by $F(y) := \inf_{x\in K} f(x,y)$ $(y \in L)$. By Theorem 3.5, for every $y \in L$ there exists $x \in K$ such that $F(y) = f(x,y)$. By strict convexity, this element $x$ is uniquely determined. We denote $x := \Phi(y)$ and thus obtain

$$F(y) = \inf_{x\in K} f(x,y) = f(\Phi(y),y), \quad y \in L. \tag{4.3}$$

By concavity of the function $y \mapsto f(x,y)$ and by the definition of $F$, for every $y_1$, $y_2 \in L$ and every $t \in [0,1]$,

$$\begin{aligned} F(ty_1 + (1-t)y_2) &= f(\Phi(ty_1 + (1-t)y_2), ty_1 + (1-t)y_2) \\ &\geq t\,f(\Phi(ty_1 + (1-t)y_2), y_1) + (1-t)\,f(\Phi(ty_1 + (1-t)y_2), y_2) \\ &\geq t\,F(y_1) + (1-t)\,F(y_2), \end{aligned}$$

so that $F$ is concave. Moreover, $F$ is upper semicontinuous: let $(y_n) \subset L$ be convergent to $y \in L$. For every $x \in K$ and every $n \in \mathbb{N}$ one has $F(y_n) \leq f(x,y_n)$, and taking the limes superior on both sides, we obtain, by continuity of $f$,

$$\limsup_{n\to\infty} F(y_n) \leq \limsup_{n\to\infty} f(x,y_n) = f(x,y).$$

Since $x \in K$ was arbitrary, this inequality implies $\limsup_{n\to\infty} F(y_n) \leq F(y)$, i.e. $F$ is upper semicontinuous.

By Theorem 3.5 (applied to $-F$; use also Remark 3.6), there exists $\bar{y} \in L$ such that

$$f(\Phi(\bar{y}),\bar{y}) = F(\bar{y}) = \sup_{y\in L} F(y).$$

We put $\bar{x} = \Phi(\bar{y})$ and show that $(\bar{x},\bar{y})$ is a saddle point. Clearly, for every $x \in K$,

$$f(\bar{x},\bar{y}) \leq f(x,\bar{y}). \tag{4.4}$$

Therefore it remains to show that for every $y \in L$,

$$f(\bar{x},\bar{y}) \geq f(\bar{x},y). \tag{4.5}$$

Let $y \in L$ be arbitrary and put $y_n := (1-\frac{1}{n})\bar{y} + \frac{1}{n}y$ and $x_n = \Phi(y_n)$. Then, by concavity,

$$\begin{aligned} F(\bar{y}) \quad &\geq \quad F(y_n) \quad = \quad f(x_n, y_n) \\ &\geq \quad (1-\frac{1}{n})f(x_n,\bar{y}) + \frac{1}{n}f(x_n,y) \\ &\geq \quad (1-\frac{1}{n})F(\bar{y}) + \frac{1}{n}f(x_n,y), \end{aligned}$$

or

$$F(\bar{y}) \geq f(x_n, y) \text{ for every } n \in \mathbb{N}.$$

Since $K$ is bounded and closed, the sequence $(x_n) \subset K$ has a weakly convergent subsequence which converges to some element $x_0 \in K$ (Theorem 2.15 and Corollary 3.3). By the preceeding inequality and Corollary 3.4,

$$F(\bar{y}) \geq f(x_0, y).$$

This is just the remaining inequality (4.5) if we can prove that $x_0 = \bar{x}$. By concavity, for every $x \in K$ and every $n \in \mathbb{N}$,

$$\begin{aligned} f(x, y_n) \quad &\geq \quad f(x_n, y_n) \\ &\geq \quad (1-\frac{1}{n})f(x_n,\bar{y}) + \frac{1}{n}f(x_n,y) \\ &\geq \quad (1-\frac{1}{n})f(x_n,\bar{y}) + \frac{1}{n}F(y). \end{aligned}$$

Letting $n \to \infty$ in this inequality and using Corollary 3.4 again, we obtain that for every $x \in K$,

$$f(x, \bar{y}) \geq f(x_0, \bar{y}).$$

Hence, $x_0 = \Phi(\bar{y}) = \bar{x}$ and the theorem is proved. □

CHAPTER 6

# Uniform boundedness, bounded inverse and closed graph

This chapter is devoted to the other fundamental theorems in functional analysis; other than the Hahn-Banach theorem which has been discussed in the previous chapter. These fundamental results are

- the uniform boundedness principle or the Banach-Steinhaus theorem,
- the bounded inverse theorem (and the related open mapping theorem), and
- the closed graph theorem.

All these fundamental results rely on an abstract lemma for metric spaces.

## 1. The lemma of Baire

LEMMA 1.1 (Baire). *Let $(M, d)$ be a complete metric space, and let $(O_n)$ be a sequence of open and dense subsets of $M$. Then $\bigcap_n O_n$ is dense in $M$.*

PROOF. We can assume that $M$ is not empty since the statement is trivial otherwise. Let $x_0 \in M$ and $\varepsilon > 0$ be arbitrary. We have to prove that $\bigcap_n O_n \cap B(x_0, \varepsilon)$ is not empty.

Since $O_1$ is dense and open in $M$, the intersection $B(x_0, \varepsilon) \cap O_1$ is open and nonempty. Hence, there exists $\varepsilon_1 > 0$ (w.l.o.g. $\varepsilon_1 \leq \varepsilon/2$) and $x_1 \in B(x_0, \varepsilon) \cap O_1$ such that

$$B(x_1, \varepsilon_1) \subset B(x_0, \varepsilon) \cap O_1.$$

Choosing $\varepsilon_1$ a little bit smaller, if necessary, we can even assume that

$$\overline{B(x_1, \varepsilon_1)} \subset B(x_0, \varepsilon) \cap O_1.$$

Since $O_2$ is dense and open in $M$, the intersection $B(x_1, \varepsilon_1) \cap O_2$ is open and nonempty. Hence, there exists $\varepsilon_2 > 0$ (w.l.o.g. $\varepsilon_2 \leq \varepsilon_1/2$) and $x_2 \in B(x_1, \varepsilon_1) \cap O_2$ such that

$$\overline{B(x_2, \varepsilon_2)} \subset B(x_1, \varepsilon_1) \cap O_2 \subset B(x_0, \varepsilon) \cap O_1 \cap O_2.$$

Proceeding inductively, we can construct sequences $(\varepsilon_n) \subset (0, \infty)$ and $(x_n) \subset M$ such that

(i) $\varepsilon_n \leq \varepsilon_{n-1}/2$ and
(ii) for every $n \in \mathbb{N}$

$$\overline{B(x_n, \varepsilon_n)} \subset B(x_{n-1}, \varepsilon_{n-1}) \cap O_n \subset B(x_0, \varepsilon) \cap \bigcap_{j=1}^{n} O_j.$$

In particular, $x_m \in B(x_n, \varepsilon_n)$ for every $m \geq n$ and $\lim_{n\to\infty} \varepsilon_n = 0$. Hence, the sequence $(x_n)$ is a Cauchy sequence in $M$. Since $M$ is complete, there exists $x := \lim_{n\to\infty} x_n \in M$. By the above,

$$x \in \overline{B(x_n, \varepsilon_n)} \text{ for every } n \in \mathbb{N},$$

or

$$x \in \bigcap_n \overline{B(x_n, \varepsilon_n)} \subset B(x_0, \varepsilon) \cap \bigcap_n O_n.$$

The claim is proved. □

Lemma 1.2 (Baire). *Let $(M, d)$ be a complete, nonempty, metric space, and let $(A_n)$ be a sequence of closed subsets in $M$ such that $M = \bigcup_n A_n$. Then there exists $n_0 \in \mathbb{N}$ such that $A_{n_0}$ has nonempty interior.*

Proof. Assume the contrary, i.e. that every $A_n$ has empty interior. In this case, the sets $O_n := M \setminus A_n$ are open and dense. By assumption,

$$\emptyset = M \setminus \bigcup_n A_n = \bigcap_n O_n,$$

a contradiction to Lemma 1.1 and the assumption that $M$ is nonempty. □

Remark 1.3. The assumption in Lemma 1.1 or Lemma 1.2 that $M$ is complete is necessary in general. For example,

$$\mathbb{Q} = \bigcup_{x\in\mathbb{Q}} \{x\},$$

and this union is countable. Each one point set $\{x\}$ is closed but in this example, none of these sets has nonempty interior.

Remark 1.4. As a corollary to the lemma of Baire one obtains for example that there exists a continuous function $f \in C([0, 1])$ which is nowhere differentiable. In fact, the set of such functions is dense in $C([0, 1])$; see [**5**].

## 2. The uniform boundedness principle

Theorem 2.1 (Uniform boundedness principle; Banach-Steinhaus). *Let $X$, $Y$ be Banach spaces and let $(T_i)_{i\in I} \subset \mathcal{L}(X, Y)$ be a family of bounded linear operators such that*

$$\sup_{i\in I} \|T_i x\| < \infty \textit{ for every } x \in X.$$

*Then*

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

Remark 2.2. Theorem 2.1 is in general not true if $X$ is only a normed space. For example, let $X = c_{00} (= Y)$ be the space of all finite sequences equipped with the supremum norm (or any other reasonable norm). Let

$$T_n x = T_n(x_m) = (a_{nm} x_m)$$

with

$$a_{nm} = \begin{cases} m & \text{if } m \leq n, \\ 0 & \text{if } m > n. \end{cases}$$

Then $\sup_n \|T_n x\|$ is finite for every $x \in X$, but $\|T_n\| = n$ is unbounded.

REMARK 2.3. The fact that in Theorem 2.1 we suppose also $Y$ to be a Banach space is not important. In fact, if $Y$ is not complete, then we may embed $Y$ into its completion $\tilde{Y}$ and consider every operator $T_i \in \mathcal{L}(X,Y)$ also as an operator in $\mathcal{L}(X,\tilde{Y})$.

PROOF OF THEOREM 2.1. Let $A_n := \{x \in X : \sup_{i\in I} \|T_i x\| \leq n\}$. Since arbitrary intersections of closed sets are closed, and by the boundedness of the $T_i$, the sets $A_n$ are closed for every $n \in \mathbb{N}$. By assumption, $X = \bigcup_n A_n$.

Hence, by the lemma of Baire (Lemma 1.2), there exists $n_0 \in \mathbb{N}$ such that $A_{n_0}$ has nonempty interior, i.e. there exist $n_0 \in \mathbb{N}$, $x_0 \in X$ and $\varepsilon > 0$ such that

$$\sup_{i\in I} \|T_i x\| \leq n_0 \text{ for every } x \in B(x_0, \varepsilon),$$

or, in other words, there exists $n_0 \in \mathbb{N}$, $x_0 \in X$ and $\varepsilon > 0$ such that

$$\|T_i(x_0 + \varepsilon x)\| \leq n_0 \text{ for every } x \in B(0,1),\ i \in I.$$

This implies, by the triangle inequality,

$$\varepsilon \|T_i x\| \leq n_0 + \|T_i x_0\| \leq 2n_0 \text{ for every } x \in B(0,1),\ i \in I.$$

The claim is proved. □

COROLLARY 2.4. *Let $X$, $Y$ be Banach spaces and let $(T_n) \subset \mathcal{L}(X,Y)$ be a* strongly convergent *sequence of bounded linear operators, i.e.*

$$Tx := \lim_{n\to\infty} T_n x \text{ exists for every } x \in X.$$

*Then* $\sup_{n\in\mathbb{N}} \|T_n\| =: M < \infty$ *and* $T \in \mathcal{L}(X,Y)$.

PROOF. Linearity of $T$ is clear. Since $(T_n)$ is strongly convergent, the sequence $(T_n x)$ is bounded for every $x \in X$. By the uniform bounded principle (Theorem 2.1), $\sup_{n\in\mathbb{N}} \|T_n\| =: M < \infty$. As a consequence, for every $x \in X$,

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

so that $T$ is bounded. □

COROLLARY 2.5. *Every weakly convergent sequence in a Banach space is bounded.*

PROOF. Let $X$ be a Banach space and $(x_n) \subset X$ be weakly convergent. Considering the $x_n$ as elements in $X'' = \mathcal{L}(X', \mathbb{K})$ by the embedding $J : X \to X''$, the claim follows from Corollary 2.4. □

## 3. Open mapping theorem, bounded inverse theorem

THEOREM 3.1 (Open mapping theorem). *Let $X$, $Y$ be two Banach spaces and let $T \in \mathcal{L}(X, Y)$ be surjective. Then there exists $r > 0$ such that*

$$TB_X(0,1) \supset B_Y(0,r). \tag{3.1}$$

PROOF. *First step:* We show that there exists $r > 0$ such that

$$B(0,2r) \subset \overline{TB(0,1)}. \tag{3.2}$$

For this, we remark first that by surjectivity,

$$Y = TX = \bigcup_n TB(0,n) = \bigcup_n \overline{TB(0,n)}.$$

By the Lemma of Baire, there exists $n_0$ such that $\overline{TB(0,n_0)}$ has nonempty interior, i.e. there exist $x \in \overline{TB(0,n_0)}$ and $\varepsilon > 0$ such that

$$B(x,\varepsilon) \subset \overline{TB(0,n_0)}.$$

By symmetry,

$$B(-x,\varepsilon) \subset \overline{TB(0,n_0)},$$

and adding both 'inequalities' together, we obtain

$$B(0,\varepsilon) \subset \overline{TB(0,n_0)},$$

which implies the required inclusion (3.2) if we put $r = \frac{\varepsilon}{2n_0}$.

*Second step:* We prove (3.1). Let $y \in B(0,r)$, where $r > 0$ is as in (3.2) from the first step. Then, by (3.2), for every $\varepsilon > 0$ there exists $x \in B(0,\frac{1}{2})$ such that $\|y - Tx\| < \varepsilon$. In particular, if we choose $\varepsilon = \frac{r}{2}$, then there exists $x_1 \in B(0,\frac{1}{2})$ such that $\|y - Tx_1\| < \frac{r}{2}$.

Similarly, since $y - Tx_1 \in B(0,\frac{r}{2})$, there exists $x_2 \in B(0,\frac{1}{4})$ such that $\|(y - Tx_1) - Tx_2\| \leq \frac{r}{4}$. Iterating this construction, we find a sequence $(x_n)$ such that $x_n \in B(0,2^{-n})$ and such that $\|y - \sum_{j=1}^n Tx_n\| \leq 2^{-n}r$. Since $X$ is complete and since $\sum_n x_n$ is absolutely convergent with $\sum_n \|x_n\| < 1$, the limit $x = \sum_n x_n$ exists and $x \in B(0,1)$. By the preceeding estimates, $\|y - Tx\| = 0$ or $Tx = y$. Thus we have proved (3.1). □

REMARK 3.2. It is not difficult to prove that if an operator $T \in \mathcal{L}(X,Y)$ satisfies (3.1), then $TO$ is open for every open $O \subset X$. A function which maps open sets into open sets is called *open*; whence the name of the open mapping theorem.

COROLLARY 3.3 (Bounded inverse theorem). *Let $X$, $Y$ be two Banach spaces and let $T \in \mathcal{L}(X,Y)$ be bijective. Then $T^{-1} \in \mathcal{L}(Y,X)$.*

PROOF. Linearity of $T^{-1}$ is clear. By the open mapping theorem (Theorem 3.1), we have

$$T^{-1}B_Y(0,1) \subset B_X(0,\frac{1}{r})$$

for some $r > 0$. Hence, $T^{-1}$ is bounded. □

COROLLARY 3.4. *Let* $\|\cdot\|_1$ *and* $\|\cdot\|_2$ *be two norms on a vector space* $X$ *such that* $(X, \|\cdot\|_1)$ *and* $(X, \|\cdot\|_2)$ *are complete. If there exists a constant* $C > 0$ *such that*

$$\|x\|_2 \leq C\,\|x\|_1 \text{ for every } x \in X,$$

*then the two norms are equivalent.*

PROOF. It suffices to consider the identity $I : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$. It is bounded by assumption, and clearly it is bijective. By the bounded inverse theorem (Corollary 3.3), the inverse $I^{-1} : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$ is bounded, i.e. there exists $c > 0$ such that

$$\|x\|_1 \leq c\,\|x\|_2 \text{ for every } x \in X.$$

□

## 4. Closed graph theorem

DEFINITION 4.1. Let $X$, $Y$ be two Banach spaces, and let $D(T) \subset X$ be a linear subspace. A linear operator $T : D(T) \to Y$ is called a *closed operator* if the graph

$$\operatorname{Graph} T := \{(x, Tx) : x \in D(T)\}$$

is closed in $X \times Y$.

LEMMA 4.2. *A linear operator* $T : X \supset D(T) \to Y$ *is closed if and only if*

$$\left.\begin{array}{l} D(T) \ni x_n \to x \text{ in } X \text{ and} \\ Tx_n \to y \text{ in } Y \end{array}\right\} \Rightarrow \quad x \in D(T) \text{ and } Tx = y. \tag{4.1}$$

PROOF. Exercise. □

LEMMA 4.3. *Every bounded linear operator* $T \in \mathcal{L}(X,Y)$ *(*$X$, $Y$ *Banach spaces) is closed.*

PROOF. This is an immediate consequence of Lemma 4.2. □

LEMMA 4.4. *A linear operator* $T : X \supset D(T) \to Y$ *is closed if and only if the space* $D(T)$ *equipped with the* graph norm

$$\|x\|_{D(T)} := \|x\|_X + \|Tx\|_Y, \quad x \in X,$$

*is complete.*

PROOF. $\Rightarrow$ Assume that $T$ is closed. Let $(x_n)$ be a Cauchy sequence in $(D(T), \|\cdot\|_{D(T)})$. Then $(x_n)$ is a Cauchy sequence in $X$ and $(Tx_n)$ is a Cauchy sequence in $Y$. Since $X$ and $Y$ are complete, there exist $x \in X$ and $y \in Y$ such that $x_n \to x$ and $Tx_n \to y$. Since $T$ is closed, and by Lemma 4.2, this implies $x \in D(T)$ and $Tx = y$. Moreover,

$$\|x_n - x\|_{D(T)} = \|x_n - x\|_X + \|Tx_n - Tx\|_Y \to 0,$$

so that $(x_n)$ converges in $(D(T), \|\cdot\|_{D(T)})$. Hence, $D(T)$ euipped with the graph norm is complete.

$\Leftarrow$ Assume that $(D(T), \|\cdot\|_{D(T)})$ is complete. Assume that $D(T) \ni x_n \to x \in X$ and $Tx_n \to y \in Y$. Then $(x_n)$ and $(Tx_n)$ are Cauchy sequences in $X$ and $Y$, respectively. By the definition of $\|\cdot\|_{D(T)}$, this implies that $(x_n)$ is a Cauchy sequence in $(D(T), \|\cdot\|_{D(T)})$. By completeness, there exists $\bar{x} \in D(T)$ such that $x_n \to \bar{x}$ in $D(T)$ (with respect to the graph norm). Since convergence of $(x_n)$ in $D(T)$ implies the convergence of $(x_n)$ in $X$, and since $(x_n)$ converges to $x$ in $X$, we find $x = \bar{x} \in D(T)$ by the uniqueness of the limit. Moreover, since $T$ is always bounded from $D(T)$ (when equipped with the graph norm) into $Y$, we have $Tx = \lim_{n\to\infty} Tx_n = y$. Hence, by Lemma 4.2, $T$ is closed. □

Example 4.5. Let $X = Y = C([0,1])$ be equipped with the supremum norm, and let $D(T) := C^1([0,1]) \subset X$. Let $Tf := f'$ for $f \in D(T)$. Then $T$ is a closed operator. In fact, the graph norm $\|\cdot\|_{D(T)}$ coincides with the canonical norm on $C^1([0,1])$, i.e.

$$\|f\|_{C^1} := \|f\|_\infty + \|f'\|_\infty,$$

and $(C^1([0,1]), \|\cdot\|_{C^1})$ is complete.

Theorem 4.6 (Closed graph theorem). *Let $X$, $Y$ be two Banach spaces and let $T : X \to Y$ be a closed operator. Then $T$ is bounded.*

Remark 4.7. The important assumption in Theorem 4.6, besides the assumption that $T$ is closed, is the assumption that $D(T) = X$! The Example 4.5 shows that closed operators need not be bounded in general; when considered from $(D(T), \|\cdot\|_X)$ with values in $Y$. Note that in Example 4.5, $D(T)$ is not complete when equipped with the norm coming from $X$.

Proof of Theorem 4.6. By assumption $(X, \|\cdot\|_X)$ is a Banach space, and by closedness of $T$ and Lemma 4.4, also $(X, \|\cdot\|_{D(T)})$ is a Banach space, where $\|\cdot\|_{D(T)}$ denotes the graph norm. Moreover, trivially,

$$\|x\|_X \le \|x\|_{D(T)} \text{ for every } x \in X.$$

By Corollary 3.4, the two norms $\|\cdot\|_X$ and $\|\cdot\|_{D(T)}$ are equivalent, i.e. there exists a constant $C \ge 0$ such that

$$\|x\|_X + \|Tx\|_Y \le C\,\|x\|_X \text{ for every } x \in X.$$

As a consequence, $T$ is bounded. □

Example 4.8 (Sobolev embedding). Let $-\infty < a < b < \infty$. Then the embedding

$$\begin{aligned} J : W^{1,p}(a,b) &\to C([a,b]), \\ u &\mapsto u \end{aligned}$$

is well defined and bounded, i.e. there exists a constant $C \ge 0$ such that

$$\|u\|_\infty \le C\,\|u\|_{W^{1,p}} \text{ for every } u \in W^{1,p}(a,b).$$

Recall that this embedding is well defined since every function $u \in W^{1,p}(a,b)$ is continuous on $[a,b]$ by Theorem 5.10 of Chapter 4.

In order to see that $J$ is also bounded, we apply the closed graph theorem together with the characterization in Lemma 4.2: let $(u_n) \subset W^{1,p}(a,b)$ be such that $u = \lim_{n\to\infty} u_n$ exists in $W^{1,p}(a,b)$ and such that $v = \lim_{n\to\infty} u_n$ exists in $C([a,b])$. The convergence in $W^{1,p} \subset L^p$ implies that $u_n \to u$ almost everywhere if we extract a subsequence. The convergence in $C$ implies that $u_n \to v$ everywhere. Hence $u = v$ almost everyhwere, and since both functions are continuous, we obtain $u = v$. Hence, the embedding is closed. By the closed graph theorem, the embedding $W^{1,p} \to C$ is bounded.

Exercise 4.9. Let $T : X \supset D(T) \to Y$ be a closed, injective operator. Define

$$\begin{aligned} D(T^{-1}) &:= \operatorname{Rg} T = \{Tx : x \in D(T)\} \subset Y, \\ T^{-1}y &:= x \text{ where } x \in D(T) \text{ is the unique element such that } Tx = y. \end{aligned}$$

Then $T^{-1}$ is a closed operator.

If in addition $T$ is surjective, then $T^{-1} : Y \to X$ is bounded.

## 5. * Vector-valued analytic functions

Let $X$ be a complex Banach space and let $\Omega \subset \mathbb{C}$ be an open subset.

Definition 5.1. We say that a function $f : \Omega \to X$ is *analytic* (or: *holomorphic*) if

$$f'(z_0) := \lim_{z\to z_0} \frac{f(z) - f(z_0)}{z - z_0} \text{ exists for every } z_0 \in \Omega.$$

We say that $f : \Omega \to X$ is *weakly analytic* (or: *weakly holomorphic*) if $x' \circ f : \Omega \to \mathbb{C}$ is analytic for every $x' \in X'$.

Theorem 5.2. *A function $f : \Omega \to X$ is analytic if and only if it is weakly analytic.*

Proof. Cleary, if $f$ is analytic, then $f$ is weakly analytic. So we only have to prove the other direction.

By considering $X$ as a closed subspace of $X''$ (via the embedding $J$), and by replacing then $X$ by $X''$ (so that the function $f$ becomes $X''$-valued), we can assume that $X$ is a dual space. But doing this, we no longer assume that $f$ is weakly analytic. The assertion which we have to prove is then the following:

Let $X$ be a complex Banach space, and let $X'$ be its dual. Let $f : \Omega \to X'$ be such that $\langle f, x\rangle : \Omega \to \mathbb{C}$ is analytic for every $x \in X$. Then $f$ is analytic.

In fact, it suffices to prove that for fixed $z_0 \in \Omega$ there exists $M \geq 0$ such that for every $y$, $z \in \Omega \setminus \{z_0\}$ 'close' to $z_0$,

$$\left\| \frac{f(z) - f(z_0)}{z - z_0} - \frac{f(y) - f(z_0)}{y - z_0} \right\| \leq M\, |z - y|. \tag{5.1}$$

Let $K := \overline{B(z_0, r)} \setminus \{z_0\}$, where $r > 0$ is chosen so small that $K \subset \Omega$. Let

$$\tilde{K} = (K \times K) \setminus \{(z, z) : z \in K\}$$

be the cartesian product of $K$ and $K$ from which we take out the 'diagonal'.

By assumption, for every $x \in X$, the function $\langle f, x\rangle$ is analytic. Hence, for every $x \in X$ we have

$$\sup_{(y,z)\in\tilde{K}} \left| \left\langle \frac{\frac{f(z)-f(z_0)}{z-z_0} - \frac{f(y)-f(z_0)}{y-z_0}}{y-z}, x \right\rangle \right| < \infty.$$

By the uniform boundedness principle, this implies

$$\sup_{(y,z)\in\tilde{K}} \left\| \frac{\frac{f(z)-f(z_0)}{z-z_0} - \frac{f(y)-f(z_0)}{y-z_0}}{y-z} \right\| =: M < \infty,$$

which actually implies (5.1) for every $y$, $z \in K$. □

By Theorem 5.2, many important properties of 'classical' analytic functions $\Omega \to \mathbb{C}$ carry over to vector-valued analytic functions $\Omega \to X$. For example:

- Every analytic function $f : \Omega \to X$ is infinitely many times differentiable.
- Every analytic function $f : \Omega \to X$ can be locally developed into a power series of the form $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ with $a_n \in X$. In fact: $a_n = \frac{1}{n!} f^{(n)}(z_0)$.
- Cauchy's integral formula $f(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(y)}{z-y}\, dy$ holds true for appropriate paths $\gamma$. Note, however, that we have not yet defined integrals of vector-valued functions.

An important example of a vector-valued analytic function will be the resolvent of an operator $T \in \mathcal{L}(X)$; see the following Chapter 7.

CHAPTER 7

# Compact operators and spectral theory

## 1. Compact operators

DEFINITION 1.1 (Compact operator). Let $X$, $Y$ be Banach spaces. A linear operator $T : X \to Y$ is called a *compact operator* if $TB(0,1)$ is relatively compact in $Y$. The set of all compact linear operators from $X$ into $Y$ is denoted by $\mathcal{K}(X,Y)$. We denote $\mathcal{K}(X) := \mathcal{K}(X,X)$.

REMARK 1.2. A linear operator $T : X \to Y$ is compact if and only if for every sequence $(x_n) \subset B(0,1)$ there exists a subsequence (again denoted by $(x_n)$) such that $(Tx_n)$ is convergent (or Cauchy).

Since relatively compact subsets of Banach spaces are necessarily bounded, every compact operator is bounded.

LEMMA 1.3. *Let $X$, $Y$, $Z$ be Banach spaces. Then:*

(a) *The set $\mathcal{K}(X,Y)$ is a closed linear subspace of $\mathcal{L}(X,Y)$.*
(b) *If $T \in \mathcal{K}(X,Y)$ and $S \in \mathcal{L}(Y,Z)$, then $ST \in \mathcal{K}(X,Z)$.*
(c) *If $T \in \mathcal{L}(X,Y)$ and $S \in \mathcal{K}(Y,Z)$, then $ST \in \mathcal{K}(X,Z)$.*
(d) *The set $\mathcal{K}(X)$ is a two-sided ideal in $\mathcal{L}(X)$.*

PROOF. (a) If $T$, $S \in \mathcal{K}(X,Y)$, $\lambda \in \mathbb{K}$, then clearly $\lambda T \in \mathcal{K}(X,Y)$. Moreoever, if $(x_n) \subset B(0,1)$ is any sequence, then we can choose a subsequence (again denoted by $(x_n)$) such that $(Tx_n)$ converges. From this subsequence, we extract another subsequence (again denoted by $(x_n)$) such that $(Sx_n)$ converges. Then $(Tx_n + Sx_n)$ converges, and therefore $T + S \in \mathcal{K}(X,Y)$. Hence, $\mathcal{K}(X,Y)$ is a linear subspace of $\mathcal{L}(X,Y)$.

In order to see that $\mathcal{K}(X,Y)$ is closed in $\mathcal{L}(X,Y)$, let $(T_n) \subset \mathcal{K}(X,Y)$ be convergent to some element in $T \in \mathcal{L}(X,Y)$. Let $(x_j) \subset B(0,1)$ be any sequence. A diagonal sequence argument implies that we can choose a subsequence (again denoted by $(x_j)$) such that

$$\lim_{j\to\infty} T_n x_j \text{ exists for every } n \in \mathbb{N}.$$

Let $\varepsilon > 0$ be arbitrary, and choose $n \in \mathbb{N}$ so large such that $\|T - T_n\| < \varepsilon$. Choose $j_0 \in \mathbb{N}$ so large that $\|T_n x_j - T_n x_k\| < \varepsilon$ for every $j$, $k \geq j_0$. Then, for every $j$, $k \geq j_0$,

$$\|Tx_j - Tx_k\| \leq \|Tx_j - T_n x_j\| + \|T_n x_j - T_n x_k\| + \|T_n x_k - Tx_k\| < 3\varepsilon.$$

Hence, $(Tx_j)$ is a Cauchy sequence. Since $Y$ is complete, $(Tx_j)$ is convergent. As a consequence, for every sequence $(x_j) \subset B(0,1)$ we have extracted a subsequence (again denoted by $(x_j)$) such that $(Tx_j)$ converges. This means that $T \in \mathcal{K}(X,Y)$. Hence, $\mathcal{K}(X,Y)$ is closed in $\mathcal{L}(X,Y)$.

(b), (c) Let $T \in \mathcal{L}(X,Y)$ and $S \in \mathcal{L}(Y,Z)$. If $T$ is compact, then $TB(0,1)$ is relatively compact, and since $S$ is continuous, $STB(0,1)$ is relatively compact in $Z$ by Lemma 4.4 of chapter 2. Hence, $ST \in \mathcal{K}(X,Z)$. If on the other hand $T$ is only bounded and $S$ is compact, then $TB(0,1)$ is bounded in $Y$, and therefore $STB(0,1)$ is relatively compact in $Z$, i.e. $ST \in \mathcal{K}(X,Z)$.

(d) This is an immediate consequence of (b) and (c). □

Lemma 1.4. *Let $X$, $Y$ be Banach spaces. Then:*

(a) *If $T \in \mathcal{L}(X,Y)$ has* finite rank*, i.e. if $\dim \operatorname{Rg} T < \infty$, then $T \in \mathcal{K}(X,Y)$.*
(b) *If $(T_n) \in \mathcal{K}(X,Y)$ is a uniformly convergent sequence of finite rank operators, then $T := \lim_{n\to\infty} T_n \in \mathcal{K}(X,Y)$.*

Proof. Assertion (a) follows from the Theorem of Heine-Borel, while (b) is a consequence of Lemma 1.3. □

Lemma 1.5. *A Banach space $X$ is finite dimensional if and only if the identity operator $I \in \mathcal{L}(X)$ is compact.*

Proof. This is an immediate consequence of Theorem 1.20 of Chapter 3 which itself was a consequence of the Lemma of Riesz (Lemma 1.19). □

A difficult problem is in general to decide which operators are compact. By the very definition of compact operators, it is thus important to know which subsets of (infinite dimensional) Banach spaces are relatively compact. Boundedness of the subset alone does not suffice as the Lemma of Riesz shows (see also the preceeding lemma). In the case when the underlying Banach space is $C(K)$ ($K$ a compact metric space) there is a satisfactory characterization of relatively compact subsets.

A subset $M \subset C(K)$ is called *equicontinuous in some point* $x \in K$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\forall y \in K\, \forall f \in M : \quad d(x,y) < \delta \quad \Rightarrow \quad |f(x) - f(y)| < \varepsilon. \tag{1.1}$$

The same subset $M$ is called *equicontinuous* if

$$\forall \varepsilon\, \exists \delta > 0\, \forall x,\, y \in K\, \forall f \in M : \quad d(x,y) < \delta \quad \Rightarrow \quad |f(x) - f(y)| < \varepsilon.$$

It is not too difficult to show that $M \subset C(K)$ is equicontinuous if and only if it is equicontinuous in every point $x \in K$. Moreover, if $M$ is equicontinuous, then the closure $\bar{M}$ is also equicontinuous.

Theorem 1.6 (Arzela-Ascoli). *Let $K$ be a compact metric space. A subset $M$ of $C(K)$ is relatively compact if and only if it is bounded and equicontinuous.*

PROOF. We first note that every compact metric space $K$ is separable. In fact, for every $n \in \mathbb{N}$ the open covering $(B(x, \frac{1}{n}))_{x \in K}$ of $K$ admits a finite subcovering $(B(x_j^n, \frac{1}{n}))_{1 \leq j \leq m_n}$. Then the set

$$D := \{x_j^n : n \in \mathbb{N},\, 1 \leq j \leq m_n\}$$

is countable and dense in $K$.

Assume now that $K$ is a compact metric space, and let $M \subset C(K)$ be bounded and equicontinuous. Let $D \subset K$ be dense and countable.

Let $(f_n) \subset M$. Since $M$ is bounded, a diagonal sequence argument implies that there exists a subsequence (again denoted by $(f_n)$) such that

$$\lim_{n \to \infty} f_n(x) \text{ exists for every } x \in D.$$

We show that this subsequence converges already uniformly on $K$. Actually, we only show that this subsequence $(f_n)$ is a Cauchy sequence (this suffices since $C(K)$ is complete). Let $\varepsilon > 0$ be arbitrary. By equicontinuity of $(f_n)$, there exists $\delta > 0$ such that (1.1) holds. By compactness of $K$ and the density of $D$ in $K$ there exist $(x_j)_{1 \leq j \leq k} \subset D$ such that $K = \bigcup_{1 \leq j \leq k} B(x_j, \delta)$. Since the family $(x_j)$ is finite, there exists $n_0 \in \mathbb{N}$ such that

$$|f_n(x_j) - f_m(x_j)| < \varepsilon \text{ for every } n,\, m \geq n_0,\, 1 \leq j \leq k.$$

Let $x \in K$ be arbitrary. Then there exists $j \in \{1, \ldots, k\}$ such that $x \in B(x_j, \delta)$. Hence, for every $n,\, m \geq n_0$,

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f_n(x_j)| + |f_n(x_j) - f_m(x_j)| + |f_m(x_j) - f_m(x)| < 3\varepsilon.$$

Since the $n_0$ does not depend on $x \in K$, we have thus proved

$$\|f_n - f_m\|_\infty \leq 3\varepsilon \text{ for every } n, m \geq n_0.$$

Hence, every sequence in $M$ admits a convergent subsequence. Therefore, $M$ is relatively compact.

The other direction is left as an exercise. □

EXAMPLE 1.7 (Sobolev embedding). Consider the embedding $J : W^{1,p}(a,b) \to C([a,b])$ from Example 4.8 of Chapter 6. The closed graph theorem showed that $J$ is bounded, i.e. there exists $C \geq 0$ such that

$$\|u\|_\infty \leq C\, \|u\|_{W^{1,p}}, \quad u \in W^{1,p}(a,b).$$

We can show in addition that the embedding is compact if $p > 1$. Let

$$M := \{u \in W^{1,p}(a,b) : \|u\|_{W^{1,p}} < 1\} = JB(0,1) \subset C([a,b])$$

be the image of the unit ball under $J$. By boundedness of $J$, $M$ is bounded in $C([a,b])$. Moreover, by Hölder's inequality (we assume $p > 1$), for every $t$, $s \in [a,b]$ $(t \geq s)$and every $u \in M$,

$$|u(t) - u(s)| = |\int_s^t u'(r)\, dr| \leq \int_s^t |u'(r)|\, dr \leq \|u'\|_p\, (t-s)^{\frac{p-1}{p}} \leq (t-s)^{\frac{p-1}{p}}.$$

This implies that $M$ is equicontinuous if $p > 1$ (choose for every $\varepsilon > 0$ the $\delta$ equal to $\varepsilon^{\frac{p}{p-1}}$ in order to see that (1.1) holds).

By Arzela-Ascoli, $M$ is relatively compact in $C([a,b])$, and therefore the embedding $W^{1,p}(a,b) \hookrightarrow C([a,b])$ is compact if $p > 1$.

Exercise 1.8. (Sobolev embedding) Show that the embedding $W^{1,1}(a,b) \hookrightarrow C([a,b])$ is not compact.

Exercise 1.9 (Multiplication operators in sequence spaces). Let $X = l^p$ $(1 \leq p < \infty)$ or let $X = c_0$. Let $m \in l^\infty$ and define the associated multiplication operator $M \in \mathcal{L}(X)$ by

$$Mx = M(x_n) = (m_n x_n), \quad x \in X.$$

Show that $M$ is compact if and only if $m \in c_0$.

*Hint:* Use Lemma 1.4.

Exercise 1.10 (Kernel operators). Let $\Omega \subset \mathbb{R}^n$ be a compact (!) set. Let $k \in C(\Omega \times \Omega)$, and define the associated *kernel operator* $K \in \mathcal{L}(C(\Omega))$ by

$$Kf(t) = \int_\Omega k(t,s) f(s)\, ds, \quad t \in \Omega,\ f \in C(\Omega).$$

Then $K$ is compact.

Definition 1.11. Let $X, Y$ be two Banach spaces, $T \in \mathcal{L}(X,Y)$. For every $y' \in Y'$, the linear mapping $X \to \mathbb{K}$, $x \mapsto \langle y', Tx\rangle$ is bounded on $X$. This linear mapping is denoted by $T'y' \in X'$. The resulting operator $T' : Y' \to X'$ is called the *adjoint* of $T$. For every $x \in X$ and every $y' \in Y'$,

$$\langle y', Tx\rangle = \langle T'y', x\rangle.$$

Lemma 1.12. *For every $T \in \mathcal{L}(X,Y)$, the adjoint $T' : Y' \to X'$ is bounded and $\|T\| = \|T'\|$.*

Proof. For every $y' \in Y'$,

$$\|T'y'\| = \sup_{\|x\| \leq 1} |\langle T'y', x\rangle| = \sup_{\|x\| \leq 1} |\langle y', Tx\rangle| \leq \|T\|\,\|y'\|,$$

which proves that $T'$ is bounded and that $\|T'\| \leq \|T\|$. On the other hand, by Hahn-Banach (Corollary 1.9 of Chapter 5),

$$\begin{aligned} \|T'\| &= \sup_{\|y'\| \leq 1} \|T'y'\| \\ &= \sup_{\|y'\| \leq 1} \sup_{\|x\| \leq 1} |\langle T'y', x\rangle| \\ &= \sup_{\|x\| \leq 1} \sup_{\|y'\| \leq 1} |\langle y', Tx\rangle| \\ &= \sup_{\|x\| \leq 1} \|Tx\| \\ &= \|T\|, \end{aligned}$$

and the claim is proved. □

THEOREM 1.13 (Schauder). *An operator $T \in \mathcal{L}(X,Y)$ is compact if and only if $T' \in \mathcal{L}(Y',X')$ is compact.*

PROOF. Assume that $T \in \mathcal{K}(X,Y)$, and let $K := \overline{TB_X(0,1)} \subset Y$. Then $K$ is compact. Let $M := B_{Y'}(0,1)$ be considered as a subset of $C(K)$. Then clearly $M$ is bounded, and it is not difficult to see that $M$ is also equicontinuous. By the theorem of Arzela-Ascoli, $M$ is relatively compact in $C(K)$. This means that for every sequence $(y_n') \in B_{Y'}(0,1)$ there exists a convergent subsequence (convergent in $C(K)$!). If we denote this subsequence again by $(y_n')$, then we obtain

$$0 = \lim_{n,m\to\infty} \|y_n' - y_m'\|_{C(K)} \geq \lim_{n,m\to\infty} \sup_{\|x\|\leq 1} |\langle y_n' - y_m', Tx\rangle| = \lim_{n,m\to\infty} \|T'y_n' - T'y_m'\|_{X'},$$

which just means that $T'$ is compact.

Assume on the other hand that $T' \in \mathcal{K}(Y',X')$. By what we have just proved, this implies $T'' \in \mathcal{K}(X'',Y'')$. Hence, if $(x_n) \in B_X(0,1)$ is any sequence, then there exists a subsequence (again denoted by $(x_n)$) such that $(T''x_n)$ is convergent in $Y''$ (note that we have considered $(x_n)$ also as a sequence in $X''$ via the embedding $J$). However, $T''x_n = Tx_n$, and the claim is proved. □

## 2. Spectrum of bounded operators

DEFINITION 2.1. Let $X$ be a Banach space and let $A : X \supset D(A) \to X$ be a linear operator. For every $\lambda \in \mathbb{K}$ we write $\lambda - A := \lambda I - A$, and we write $\lambda \in \varrho(A)$ (the *resolvent set* of $A$) if $\lambda - A : D(A) \to X$ is bijective and the inverse $(\lambda - A)^{-1}$ is bounded on $X$. For every $\lambda \in \varrho(A)$ we denote by $R(\lambda, A) := (\lambda - A)^{-1}$ the *resolvent* of $A$ in $\lambda$.

The set $\sigma(A) := \mathbb{C} \setminus \varrho(A)$ is called the *spectrum* of $A$. Moreover, we define the *point spectrum*, the *approximative point spectrum* and the *residual spectrum*, respectively, by

$$\begin{aligned} \sigma_p(A) &:= \{\lambda \in \mathbb{C} : \exists x \in X \setminus \{0\} \text{ s.t. } Ax = \lambda x\} \\ \sigma_{ap}(A) &:= \{\lambda \in \mathbb{C} : \exists (x_n) \subset X,\ \|x_n\| = 1, \text{ s.t. } (\lambda - A)x_n \to 0\} \text{ and} \\ \sigma_r(A) &:= \{\lambda \in \mathbb{C} : \operatorname{Rg}(\lambda - A) \text{ is not dense in } X\}. \end{aligned}$$

LEMMA 2.2 (Resolvent identity). *For every $\lambda$, $\mu \in \varrho(A)$ one has*

$$R(\lambda, A) - R(\mu, A) = (\mu - \lambda) R(\mu, A) R(\lambda, A).$$

PROOF. For every $\lambda$, $\mu \in \varrho(A)$

$$\mu - \lambda = (\mu - A) - (\lambda - A).$$

Multiplying both sides by $R(\mu, A)$ and $R(\lambda, A)$, one obtains the claim. □

In what follows, we will always consider bounded linear operators $T$ on complex Banach spaces $X$.

LEMMA 2.3. *The set $\varrho(T)$ is open in $\mathbb{C}$ and the function $\varrho(T) \to \mathcal{L}(X)$, $\lambda \mapsto R(\lambda, T)$ is analytic (= differentiable).*

PROOF. Let $\lambda \in \varrho(T)$ and $\mu \in \mathbb{C}$. Then

$$\mu - T = \mu - \lambda + \lambda - T = ((\mu - \lambda)R(\lambda, T) + I)(\lambda - T),$$

and the right-hand side is invertible if $|\mu - \lambda| < 1/\|R(\lambda, T)\|$ by the Neumann series. Hence, $\varrho$ is open in $\mathbb{C}$. The Neumann series precisely yields

$$R(\mu, T) = \sum_{n=0}^{\infty} (-1)^n R(\lambda, T)^{n+1} (\mu - \lambda)^n,$$

so that the function $\lambda \mapsto R(\lambda, A)$ can be locally developped into a power series. As a consequence, this function is analytic. □

REMARK 2.4. One may also employ the resolvent identity in order to prove that the function $\lambda \to R(\lambda, T)$ is analytic; but in this case one should at least prove continuity of $R(\cdot, T)$.

LEMMA 2.5. *One has*

$$\{\lambda \in \mathbb{C} : |\lambda| > \|T\|\} \subset \varrho(T),$$

*and*

$$R(\lambda, T) = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}, \quad |\lambda| > \|T\|.$$

PROOF. Use the identity

$$\lambda - T = \lambda(I - \frac{T}{\lambda})$$

and the Neumann series. □

REMARK 2.6. In fact, $\lambda \in \varrho(T)$ as soon as

$$|\lambda| > \liminf_{n\to\infty} \|T^n\|^{\frac{1}{n}} =: r(T).$$

The number $r(T) \geq 0$ is called the *spectral radius* of $T$.

LEMMA 2.7. *For every $T \in \mathcal{L}(X)$ and every $\lambda \in \varrho(T)$,*

$$\|R(\lambda, T)\| \geq \operatorname{dist}(\lambda, \sigma(T))^{-1}.$$

PROOF. As we have seen in the proof of Lemma 2.3, for $\lambda \in \varrho(T)$ the condition

$$|\mu - \lambda| \, \|R(\lambda, T)\| < 1$$

implies $\mu \in \varrho(T)$. The claim follows. □

LEMMA 2.8. *For every $T \in \mathcal{L}(X)$, $X \neq \{0\}$, the spectrum $\sigma(T)$ is nonempty and compact.*

Proof. The compactness of $\sigma(T)$ follows Lemma 2.3 and 2.5. If $\sigma(T)$ was empty, then, by Lemma 2.3, the function $\lambda \mapsto R(\lambda, T)$ is entire. On the other hand, by Lemma 2.5,

$$\lim_{|\lambda|\to\infty} \|R(\lambda, T)\| = 0.$$

By Liouville's theorem, this implies $R(\lambda, T) \equiv 0$ which is only possible if $X = \{0\}$ is the trivial space. □

Lemma 2.9. *For every $T \in \mathcal{L}(X)$,*

$$\partial\sigma(T) \subset \sigma_{ap}(T).$$

Proof. If $\lambda \in \partial\sigma(T)$, then there exists $(\lambda_n) \subset \varrho(T)$ such that $\lim_{n\to\infty} \lambda_n = \lambda$. By Lemma 2.7, $\lim_{n\to\infty} \|R(\lambda_n, T)\| = \infty$. By the definition of the operator norm, there exists a sequence $(y_n) \subset X$, $\|y_n\| = 1$, such that

$$\lim_{n\to\infty} \|R(\lambda_n, T)y_n\| = \infty.$$

Put $x_n := \frac{R(\lambda_n,T)y_n}{\|R(\lambda_n,T)y_n\|}$, so that $\|x_n\| = 1$. Then

$$\lambda x_n - T x_n = (\lambda - \lambda_n)x_n + \frac{y_n}{\|R(\lambda_n, T)y_n\|} \to 0 \quad (n \to \infty).$$

As a consequence, $\lambda \in \sigma_{ap}(T)$. □

Lemma 2.10. *For every $T \in \mathcal{L}(X)$ one has $\sigma(T) = \sigma(T')$. For every $\lambda \in \varrho(T)$ one has*

$$R(\lambda, T)' = R(\lambda, T').$$

Proof. For every $x' \in X'$ and every $x \in X$ we have

$$\begin{aligned} \langle R(\lambda, T)'(\lambda - T')x', x\rangle &= \langle(\lambda - T')x', R(\lambda, T)x\rangle \\ &= \langle x', (\lambda - T)R(\lambda, T)x\rangle \\ &= \langle x', x\rangle, \end{aligned}$$

and

$$\begin{aligned} \langle(\lambda - T')R(\lambda, T)'x', x\rangle &= \langle R(\lambda, T)'x', (\lambda - T)x\rangle \\ &= \langle x', R(\lambda, T)(\lambda - T)x\rangle \\ &= \langle x', x\rangle. \end{aligned}$$

Hence, $R(\lambda, T)'(\lambda - T') = (\lambda - T')R(\lambda, T)' = I$, or $R(\lambda, T)' = R(\lambda, T')$. □

## 3. Spectrum of compact operators, Fredholm alternative

Let $X$ be a Banach space and $X'$ its dual. For every subset $M \subset X$ we define the *annihilator*

$$M^\perp := \{x' \in X' : \langle x', x\rangle = 0 \, \forall x \in M\}.$$

For every subset $M' \subset X'$, we define the *preannihilator*

$$M'_\perp := \{x \in X : \langle x', x\rangle = 0 \, \forall x' \in M'\}.$$

It is easy to show that $M^\perp$ and $M'_\perp$ are closed linear subspaces of $X'$ and $X$, respectively.

LEMMA 3.1. *Let $X$ be a Banach space and let $T \in \mathcal{L}(X)$. Then:*
(a) $(\operatorname{Rg} T)^\perp = \ker T'$.
(b) $\overline{\operatorname{Rg} T} = (\ker T')_\perp$.
(c) $(\ker T)^\perp \supset \overline{\operatorname{Rg} T'}$
(d) $\ker T = (\operatorname{Rg} T')_\perp$.

PROOF. In order to prove (a), we observe

$$\begin{aligned} x' \in (\operatorname{Rg} T)^\perp \quad &\Leftrightarrow \quad \forall x \in X : \langle x', Tx \rangle = 0 \\ &\Leftrightarrow \quad \forall x \in X : \langle T'x', x \rangle = 0 \\ &\Leftrightarrow \quad T'x' = 0 \\ &\Leftrightarrow \quad x' \in \ker T'. \end{aligned}$$

(b) If $x \in \operatorname{Rg} T$, $x = Ty$, and if $x' \in \ker T'$, then

$$\langle x', x \rangle = \langle x', Ty \rangle = \langle T'x', y \rangle = 0.$$

Hence, $\operatorname{Rg} T \subset (\ker T')_\perp$, and since the latter space is closed, we obtain $\overline{\operatorname{Rg} T} \subset (\ker T')_\perp$. Assume that the inclusion is strict. Then there exists $x_0 \in (\ker T')_\perp$ which does not belong to $\overline{\operatorname{Rg} T}$. By Hahn-Banach (Theorem 3.1 of Chapter 5), there exist $x' \in X'$ and $\varepsilon > 0$ such that

$$\operatorname{Re} \langle x', x \rangle + \varepsilon \leq \operatorname{Re} \langle x', x_0 \rangle, \quad x \in \overline{\operatorname{Rg} T}. \tag{3.1}$$

Since $\overline{\operatorname{Rg} T}$ is a subspace of $X$, in particular $x \in \overline{\operatorname{Rg} T}$ implies $\lambda x \in \overline{\operatorname{Rg} T}$ for every $\lambda \in \mathbb{K}$, we deduce from this inequality that $\langle x', x \rangle = 0$ for every $x \in \overline{\operatorname{Rg} T}$. Hence, by (a), $x' \in \ker T'$. But then $\langle x', x_0 \rangle = 0$, too, and this is a contradiction to (3.1). Hence, we have proved (b).

(c) If $x' \in \operatorname{Rg} T'$, $x' = T'y'$, and if $x \in \ker T$, then

$$\langle x', x \rangle = \langle T'y', x \rangle = \langle y', Tx \rangle = 0.$$

This implies $\operatorname{Rg} T' \subset (\ker T)^\perp$, and since the latter space is closed, we obtain (c).

(d) Similarly as in (a), we observe

$$\begin{aligned} x \in \ker T \quad &\Leftrightarrow \quad Tx = 0 \\ &\Leftrightarrow \quad \forall x' \in X' : \langle x', Tx \rangle = 0 \\ &\Leftrightarrow \quad \forall x' \in X' : \langle T'x', x \rangle = 0 \\ &\Leftrightarrow \quad x \in (\operatorname{Rg} T')_\perp. \end{aligned}$$

□

THEOREM 3.2 (Fredholm alternative). *Let $X$ be a Banach space, $T \in \mathcal{K}(X)$. Then:*
(a) $\ker(I - T)$ *is finite dimensional.*
(b) $\operatorname{Rg}(I - T)$ *is closed and* $\operatorname{Rg}(I - T) = \ker(I - T')_\perp$.

(c) $\ker(I - T) = \{0\}$ *if and only if* $\operatorname{Rg}(I - T) = X$.
(d) $\dim \ker(I - T) = \dim \ker(I - T') = \dim (X/\operatorname{Rg}(I - T))$.

REMARK 3.3. Fredholm's alternative says the following for the equation

$$x - Tx = y, \tag{3.2}$$

where $y \in X$ is given and $T \in \mathcal{K}(X)$. Either for every $y$ there exists a solution $x$ of this equation, and in this case the solution $x$ is unique, or the homogeneous equation

$$x - Tx = 0$$

has a a finite number of linearly independent solutions $(x_i)_{1 \le i \le n}$ and the equation (3.2) has a solution if and only if $y$ satisfies $n$ equations of orthogonality $\langle x_i', y\rangle = 0$, where the $x_i' \in \ker(I - T')$ are linearly independent.

REMARK 3.4. If $T \in \mathcal{K}(X)$, then, by property (c), $I - T$ is injective if and only if $I - T$ is surjective. In finite dimensions, this property of linear mappings is well-known. This property of operators of the form $I - T$ with $T$ compact is however not shared by arbitrary bounded operators on infinite-dimensional Banach spaces. For example, the left-shift $L$ on $l^p(\mathbb{N})$ defined by $Lx = L(x_n) := (x_{n+1})$ is surjective but not injective.

REMARK 3.5. An operator $S \in \mathcal{L}(X, Y)$ such that $\ker S$ is finite dimensional and such that $\operatorname{Rg} S$ is closed and has finite codimension (i.e. $\dim (X/\operatorname{Rg} S) < \infty$) is called a *Fredholm operator*, and

$$\operatorname{ind} S := \dim \ker S - \dim (X/\operatorname{Rg} S)$$

is called the *index* of $S$. By Theorem 3.2, $S = I - T \in \mathcal{L}(X)$ is a Fredholm operator of index 0 if $T \in \mathcal{K}(X)$.

PROOF OF THEOREM 3.2. (a) On $\ker(I - T)$ we have $T = I$, and since $T$ is compact, $\ker(I - T)$ must be finite dimensional.

(b) Let $(x_n) \subset X$ be such that $u_n := x_n - Tx_n \to u \in X$. We have to show that $u \in \operatorname{Rg}(I - T)$. Since $\ker(I - T)$ is finite dimensional, for every $n \in \mathbb{N}$ there exists $y_n \in \ker(I - T)$ such that

$$\operatorname{dist}(x_n, \ker(I - T)) = \|x_n - y_n\|.$$

We show that the sequence $(x_n - y_n)$ is bounded. Otherwise, after extracting a subsequence, we may assume that $\lim_{n\to\infty} \|x_n - y_n\| = \infty$. Putting $w_n := \frac{x_n - y_n}{\|x_n - y_n\|}$, we find that $w_n - Tw_n = u_n/\|x_n - y_n\| \to 0$. After extracting a subsequence, we may assume that $Tw_n \to z$ ($T$ is compact). But then $w_n \to z$, too, and therefore $z \in \ker(I - T)$. On the other hand,

$$\operatorname{dist}(w_n, \ker(I - T)) = \frac{\operatorname{dist}(x_n, \ker(I - T))}{\|x_n - y_n\|} = 1,$$

a contradiction. Hence, the sequence $(x_n - y_n)$ is bounded.

But then, by compactness of $T$, we can extract a subsequence (again denoted by $(x_n - y_n)$) such that $T(x_n - y_n) \to v$. Hence,

$$x_n - y_n = u_n + T(x_n - y_n) \to u + v.$$

We deduce that $T(u+v) = v$, or $u = (u+v) - T(u+v)$, so that $u \in \operatorname{Rg}(I-T)$. Hence, $\operatorname{Rg}(I-T)$ is closed.

Since the equality $\overline{\operatorname{Rg}(I-T)} = \ker(I-T')_\perp$ always holds true (Lemma 3.1), we have thus proved (b).

(c) Assume first that $I-T$ is injective, i.e. $\ker(I-T) = \{0\}$. Assume that $X_1 := \operatorname{Rg}(I-T) \neq X$, i.e. $X_1$ is a closed (by (b)) proper subspace of $X$. Then $T|_{X_1} \in \mathcal{K}(X_1)$, so that, by (b) again, $X_2 = (I-T)X_1$ is a closed subspace of $X_1$. Since $I-T$ is injective, $X_2 \neq X_1$. Iterating this argument and putting $X_n = (I-T)^n X$, we obtain a decreasing sequence $(X_n)$ of closed subspaces of $X$ such that $X_{n+1} \neq X_n$. By the Lemma of Riesz, for every $n \geq 1$ there exists $x_n \in X_n$ such that $\|x_n\| = 1$ and $\operatorname{dist}(x_n, X_{n+1}) \geq \frac{1}{2}$. For every $n > m$ we have

$$Tx_n - Tx_m = -(x_n - Tx_n) + (x_m - Tx_m) + x_n - x_m$$

and

$$-(x_n - Tx_n) + (x_m - Tx_m) + x_n \in X_{m+1}.$$

Hence, $\|Tx_n - Tx_m\| \geq \frac{1}{2}$ whenever $n \neq m$, a contradiction to the assumption that $T$ is compact. Hence, $\operatorname{Rg}(I-T) = X$.

Assume now on the other hand that $\operatorname{Rg}(I-T) = X$. Then, by Lemma 3.1, $\ker(I-T') = \{0\}$. Since $T'$ is compact by Schauder's theorem, this implies $\operatorname{Rg}(I-T') = X'$ by the preceeding step. By Lemma 3.1, $\ker(I-T) = \{0\}$.

(d) For every closed subspace $U$ of $X$ the dual $(X/U)'$ is isomorphic to $U^\perp$. In particular, for $U = \operatorname{Rg}(I-T)$ one obtains (using Lemma 3.1)

$$\ker(I-T') = (\operatorname{Rg}(I-T))^\perp \cong (X/\operatorname{Rg}(I-T))' \cong X/\operatorname{Rg}(I-T).$$

The last isomorphy holds since we know by the first isomorphy that $(X/\operatorname{Rg}(I-T))'$ is finite dimensional. In particular,

$$\dim \ker(I-T') = \dim X/\operatorname{Rg}(I-T),$$

so that we have proved the second inequality.

It remains to prove that

$$\dim X/\operatorname{Rg}(I-T) = \dim \ker(I-T).$$

Since $Tx = x$ for every $x \in \ker(I-T)$, we see that $T$ leaves $\ker(I-T)$ invariant. In particular, the operator

$$\begin{aligned} \tilde{T} : X/\ker(I-T) &\to X/\ker(I-T), \\ x + \ker(I-T) &\mapsto Tx + \ker(I-T), \end{aligned}$$

is well-defined and one easily checks that $\tilde{T}$ is compact since $T$ is compact. By construction, $\ker(I-\tilde{T}) = \{0\}$ so that, by (c), $\operatorname{Rg}(I-\tilde{T}) = X/\ker(I-T)$.

This means that for every $y \in X$ there exists $x \in X$ and $x_0 \in \ker(I-T)$ such that

$$(I-T)x = y - x_0,$$

or

$$y = (I-T)x + x_0 =: x_1 + x_0.$$

In particular, every $y \in X$ can be written as a sum $x_1 + x_0$ of an element $x_1 \in \mathrm{Rg}\,(I-T)$ and an element $x_0 \in \ker(I-T)$. Hence,

$$\dim \ker(I-T') = \dim X/\mathrm{Rg}\,(I-T) \leq \dim \ker(I-T).$$

Replacing $T$ by $T'$ (which is compact by Schauder's theorem), we obtain

$$\dim \ker(I-T'') \leq \dim \ker(I-T') \leq \dim \ker(I-T).$$

On the other hand, since $I-T''$ extends $I-T$, one trivially has

$$\dim \ker(I-T) \leq \dim \ker(I-T'').$$

The claim is proved □

THEOREM 3.6 (Spectrum of a compact operator). *Let $X$ be an infinite dimensional Banach space and let $T \in \mathcal{K}(X)$. Then:*

(a) $0 \in \sigma(T)$.
(b) $\sigma(T) \setminus \{0\} = \sigma_p(T) \setminus \{0\}$.
(c) *Either $\sigma(T)$ is finite or $\sigma(T) \setminus \{0\} = \{\lambda_n : n \in \mathbb{N}\}$ for some sequence $(\lambda_n) \subset \mathbb{C}$ such that $\lim_{n\to\infty} \lambda_n = 0$.*

PROOF. (a) If $0 \in \varrho(T)$, then $T^{-1}$ exists and is bounded. Hence, $I = TT^{-1}$ is compact; a contradiction to the assumption that $X$ is infinite dimensional.

(b) Let $\lambda \in \sigma(T) \setminus \{0\}$. If $\lambda \notin \sigma_p(T)$, then $\ker(\lambda - T) = \{0\}$. By the Fredholm alternative, this implies $\mathrm{Rg}\,(\lambda - T) = X$ so that $\lambda - T$ is bijective; a contradiction to the assumption $\lambda \in \sigma(T)$.

(c) It suffices to prove that $\sigma(T) \cap \{\lambda \in \mathbb{C} : |\lambda| \geq R\}$ is finite for every $R > 0$. If this was not the case, then we find a sequence $(\lambda_n) \subset \sigma(T) \setminus \{0\}$ such that $\lambda_n \neq \lambda_m$ for $n \neq m$ and $|\lambda_n| \geq R > 0$. By (b), for every $n \in \mathbb{N}$ there exists $x_n \in X \setminus \{0\}$ such that $\lambda_n x_n - Tx_n = 0$. Note that the family $(x_n)$ are linearly independent. Otherwise, we find a smallest $n \in \mathbb{N}$ such that the family $(x_i)_{1\leq i\leq n}$ is linearly independent, but $x_{n+1} = \sum_{i=1}^n \alpha_i x_i$ for some scalars $\alpha_i$. Then

$$\sum_{i=1}^n \alpha_i \lambda_{n+1} x_i = \lambda_{n+1} x_{n+1} = Tx_{n+1} = \sum_{i=1}^n \alpha_i \lambda_i x_i,$$

and this implies $\alpha_i(\lambda_{n+1} - \lambda_i) = 0$ for every $1 \leq i \leq n$. Since $\lambda_{n+1} \neq \lambda_i$ for $1 \leq i \leq n$, we obtain $\alpha_i = 0$; a contradiction to $x_{n+1} \neq 0$. Let $X_n := \mathrm{span}\{x_i : 1 \leq i \leq n\}$. Then $(X_n)$ is an increasing sequence of closed subspaces of $X$ such that $X_n \neq X_{n+1}$ (the latter by linear independence of the vectors $x_n$). By the

Lemma of Riesz, for every $n \geq 2$ there exists $y_n \in X_n$ such that $\|y_n\| = 1$ and $\operatorname{dist}(y_n, X_{n-1}) \geq \frac{1}{2}$. Then, for every $n > m \geq 2$,

$$\begin{aligned} \|Ty_n - Ty_m\| &= \| -(\lambda_n y_n - Ty_n) + (\lambda_m y_m - Ty_m) + \lambda_n y_n - \lambda_m y_m \| \\ &\geq \operatorname{dist}(\lambda_n y_n, X_{n-1}) \\ &\geq \frac{\lambda_n}{2} \quad \geq \quad \frac{R}{2}. \end{aligned}$$

This is a contradiction to the compactness of $T$, and hence (c) is proved. □

## 4. Spectral theorem for self-adjoint compact operators

DEFINITION 4.1. Let $H$, $K$ be two Hilbert spaces, $T \in \mathcal{L}(H, K)$. For every $y \in K$ the mapping $H \to \mathbb{K}$, $x \mapsto (Tx, y)_K$ is a bounded linear functional on $H$ which admits a unique representation by $T^* y \in H$ such that

$$(Tx, y)_K = (x, T^* y)_H, \quad x \in H.$$

The resulting linear operator $T^* : K \to H$ is called the *(Hilbert space) adjoint* of $T$.

DEFINITION 4.2. Let $H$ be a complex Hilbert space. An operator $T \in \mathcal{L}(H)$ is called *self-adjoint* if $T = T^*$, or, equivalently, if for every $x$, $y \in H$,

$$(Tx, y) = (x, Ty).$$

REMARK 4.3. Let $\mathcal{A}$ be a complex Banach algebra. A mapping $* : \mathcal{A} \to \mathcal{A}$ is called an *involution* if for every $a$, $b \in \mathcal{A}$, $\lambda \in \mathbb{C}$,

$$(a+b)^* = a^* + b^*,\ (ab)^* = b^* a^*,\ (\lambda a)^* = \bar{\lambda} a^*,\ (a^*)^* = a.$$

If a complex Banach algebra $\mathcal{A}$ admits an involution $*$ such that for every $a \in \mathcal{A}$,

$$\|a^* a\| = \|a\|^2,$$

then $\mathcal{A}$ is called a *$C^*$-algebra*.

If $H$ is a Hilbert space, then $\mathcal{L}(H)$ is a $C^*$-algebra for the involution $T \mapsto T^*$, where $T^*$ is the (Hilbert space) adjoint of $T$. Note that

$$\begin{aligned} \|T^* T\| &= \sup_{\|x\| \leq 1} \|T^* T x\| \\ &= \sup_{\|x\| \leq 1} \sup_{\|y\| \leq 1} \big|(T^* T x, y)\big| \\ &= \sup_{\|x\| \leq 1} \sup_{\|y\| \leq 1} \big|(Tx, Ty)\big| \\ &\geq \sup_{\|x\| \leq 1} \big|(Tx, Tx)\big| \\ &= \sup_{\|x\| \leq 1} \|Tx\|^2 \\ &= \|T\|^2, \end{aligned}$$

while the inequality $\|T^* T\| \leq \|T^*\| \, \|T\| = \|T\|^2$ is trivial.

The simplest $C^*$-algebra is $\mathbb{C}$ (the involution being the complex conjugation). In every $C^*$-algebra $\mathcal{A}$ one can define that an element $a$ is self-adjoint if $a = a^*$. The self-adjoint elements of $\mathcal{A} = \mathbb{C}$ are the real numbers. The self-adjoint elements of $\mathbb{C}^{N\times N}$ are the hermitian matrices, i.e. the matrices $A$ for which $A = \overline{A^t}$.

THEOREM 4.4 (Hellinger-Toeplitz). *Let $T : H \to H$ be linear and* symmetric*, i.e.*

$$(Tx, y) = (x, Ty) \text{ for every } x,\ y \in H.$$

*Then $T$ is bounded.*

PROOF. Let $(x_n) \subset H$ be convergent to $x \in H$ and such that $(Tx_n)$ converges to $y \in H$. Then, for every $z \in H$,

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

Hence, $Tx = y$. This means that $T$ is closed, and by the closed graph theorem, $T$ is bounded. □

LEMMA 4.5. *Let $T \in \mathcal{L}(H)$ be a self-adjoint operator on a complex Hilbert space. Then*

$$\sigma(T) \subset \overline{W(T)} \subset \mathbb{R}. \tag{4.1}$$

*where*

$$W(T) := \{(Tx, x) : \|x\| = 1\}. \tag{4.2}$$

PROOF. Since $(Tx, x) = (x, Tx) = \overline{(Tx, x)}$ by symmetry, we obtain $W(T) \subset \mathbb{R}$.

Let $\lambda \in \mathbb{C}$ be such that $d := \operatorname{dist}(\lambda, W(T)) > 0$. Then, for every $x \in H$ such that $\|x\| = 1$,

$$d = d\,\|x\| \le |\lambda - (Tx, x)| = |((\lambda - T)x, x)| \le \|(\lambda - T)x\|.$$

This estimate implies that $\lambda - T$ is injective and that $\operatorname{Rg}(\lambda - T)$ is closed. If $\operatorname{Rg}(\lambda - T) \ne H$, then there exists $x_0 \in (\operatorname{Rg}(\lambda - T))^\perp$ such that $\|x_0\| = 1$. For this $x_0$ we have

$$0 = ((\lambda - T)x_0, x_0) = \lambda - (Tx_0, x_0) \ge d > 0,$$

a contradiction. Hence, $\lambda - T$ is invertible, or $\lambda \in \varrho(T)$. Thus we have proved also the first inclusion in (4.1). □

LEMMA 4.6. *Let $T \in \mathcal{L}(H)$ be a self-adjoint operator on a complex Hilbert space. Then*

$$\sup W(T) \in \sigma(T) \text{ and } \inf W(T) \in \sigma(T),$$

*where $W(T)$ is defined as in (4.2).*

Proof. Let $\lambda := \sup W(T)$. By definition of $W(T)$, the form $a(x,y) := \lambda(x,y)-(Tx,y)$ is sesquilinear in the case of a complex Hilbert space or bilinear and symmetric in the case of a real Hilbert space. Moreover, this form is positive semidefinite, i.e. $a(x,x) \geq 0$ for every $x \in H$.

By the Cauchy-Schwarz inequality applied to the form $a(x,y)$, for every $x$, $y \in H$,

$$|(\lambda x - Tx, y)| \leq (\lambda x - Tx, x)^{\frac{1}{2}} (\lambda y - Ty, y)^{\frac{1}{2}}.$$

This inequality implies that there exists a constant $C \geq 0$ such that for every $x \in H$,

$$\|\lambda x - Tx\| \leq C (\lambda x - Tx, x)^{\frac{1}{2}}.$$

Let $(x_n) \subset H$, $\|x_n\| = 1$ be such that $(Tx_n, x_n) \to \lambda$. Then the preceeding inequality implies that $\lim_{n\to\infty} \|\lambda x_n - Tx_n\| = 0$. Hence, $\lambda \in \sigma_{ap}(T) \subset \sigma(T)$.

The proof that $\inf W(T) \in \sigma(T)$ is similar. □

Lemma 4.7. *Let $T \in \mathcal{L}(H)$ be a self-adjoint operator on a complex Hilbert space, and let $x$, $y \in H$ be two eigenvectors corresponding to two distinct eigenvalues $\lambda$, $\mu \in \sigma_p(T)$. Then $(x,y) = 0$.*

Proof. Since $T$ is self-adjoint and $\lambda$, $\mu \in \mathbb{R}$ (Lemma 4.5),

$$\lambda(x,y) = (\lambda x, y) = (Tx, y) = (x, Ty) = (x, \mu y) = \mu(x,y).$$

This equality can only hold if $(x,y) = 0$. □

Theorem 4.8 (Spectral theorem for self-adjoint compact operators). *Let $H$ be a complex Hilbert space, and let $T \in \mathcal{K}(H)$ be an injective, self-adjoint, compact operator. Then there exists an orthonormal basis $(e_n) \subset H$ and a sequence $(\lambda_n)$ such that $\lim_{n\to\infty} \lambda_n = 0$ such that*

$$Te_n = \lambda_n e_n,$$

*i.e. $(e_n)$ is an orthonormal basis consisting only of eigenvectors of $T$.*

Remark 4.9. If $T \in \mathcal{K}(H)$ is self-adjoint, compact, but not injective, then $H = \ker T \oplus (\ker T)^\perp$, and there exists a orthonormal basis $(e_n) \subset (\ker T)^\perp$ consisting only of eigenvectors of $T$. This follows from applying Theorem 4.8 to the restriction of $T$ to $(\ker T)^\perp$ (which is invariant under $T$!).

Proof of Theorem 4.8. Let $(\mu_n)$ be the sequence of *all* nonzero eigenvalues of $T$ ($\mu_n \neq \mu_m$ if $n \neq m$) and let $d_n = \dim \ker(\mu_n - T)$ be the multiplicity of $\mu_n$. Let $(\lambda_n)$ be the sequence of all nonzero eigenvalues of $T$ counted with their finite multiplicity $d_n$, i.e. $\lambda_1 = \mu_1, \dots, \lambda_{d_1} = \mu_1, \lambda_{d_1+1} = \mu_2, \dots, \lambda_{d_1+d_2} = \mu_2$, etc.

Let $(f_i^n)_{1\leq i\leq d_n}$ be an orthonormal basis of $\ker(\mu_n - T)$, and let $(e_n)$ be the family which is obtained by taking successively the union over all $f_i^n$, i.e. $e_1 = f_1^1, \dots, e_{d_1} = f_{d_1}^1, e_{d_1+1} = f_1^2, \dots, e_{d_1+d_2} = f_{d_2}^2$, etc.

The family $(e_n)$ is orthonormal by Lemma 4.7. Moreover, by construction, $Te_n = \lambda_n e_n$ for every $n \in \mathbb{N}$. It remains only to show that $\operatorname{span}\{e_n : n \in \mathbb{N}\} =: H_0$ is dense in $H$.

Let $H_1 := H_0^\perp$. For every $x \in H_1$ and every $n \in \mathbb{N}$, since $T$ is self-adjoint,

$$(Tx, e_n) = (x, Te_n) = (x, \lambda_n e_n) = \bar{\lambda_n}(x, e_n) = 0.$$

Hence, $TH_1 \subset H_1$, or, equivalently, $T_1 := T|_{H_1} \in \mathcal{K}(H_1)$. Since *all* eigenvectors of $T$ (for nonzero eigenvalues) are contained in $H_0$, $T_1$ does not have any nonzero eigenvalue. In other words, $\sigma(T_1) = \{0\}$. By Lemma 4.6, this implies $(T_1x, x) = 0$ for every $x \in H_1$. But then, for every $x, y \in H_1$,

$$2\,\mathrm{Re}\,(T_1x, y) = (T_1(x + y), x + y) - (T_1x, x) - (T_1y, y) = 0,$$

so that $T_1 = 0$. But this means that $\ker T$ contains $H_1$. Since $T$ is injective, we obtain $H_1 = \{0\}$, or, in other words, $H_0$ is dense in $H$. The claim is proved. □

## 5. * Elliptic partial differential equations

Let $\Omega \subset \mathbb{R}^N$ be open and bounded, $\lambda \in \mathbb{C}$, and consider the elliptic partial differential equation

$$\begin{cases} \lambda u - \Delta u = f & \text{in } \Omega, \\ u = 0 & \text{in } \partial\Omega, \end{cases} \tag{5.1}$$

where $\Delta$ stands for the Laplace operator and $f \in L^2(\Omega)$.

Recall from Chapter 4 that a function $u \in H_0^1(\Omega)$ is a *weak solution* of (5.1) if for every $\varphi \in H_0^1(\Omega)$ one has

$$\lambda \int_\Omega u\bar{\varphi} + \int_\Omega \nabla u \overline{\nabla\varphi} = \int_\Omega f\bar{\varphi}.$$

Let $H := L^2(\Omega)$ and define

$$\begin{aligned} D(A) \;&:=\; \{u \in H_0^1(\Omega) : \exists f \in L^2(\Omega)\, \forall \varphi \in H_0^1(\Omega) : \\ &\qquad\qquad \int_\Omega \nabla u \overline{\nabla\varphi} = -\int_\Omega f\bar{\varphi}\} \\ Au \;&:=\; f, \end{aligned} \tag{5.2}$$

so that $A : D(A) \to L^2(\Omega)$ is a linear operator on $L^2(\Omega)$. By definition, $u \in D(A)$ and $-Au = f$ if and only if $u$ is a weak solution of (5.1) for $\lambda = 0$. Moreover, a function $u \in H_0^1(\Omega)$ is a weak solution of (5.1) if and only if

$$u \in D(A) \text{ and } \lambda u - Au = f. \tag{5.3}$$

In this sense, we may say that $A$ is the *realization* of the Laplace operator with Dirichlet boundary conditions. The problem (5.3) is a *functional analytic* reformulation of (5.1). Instead of solving a partial differential equation we now have to solve an algebraic equation. Clearly, the operator $A$ is linear.

THEOREM 5.1. *There exists an orthonormal basis* $(e_n)$ *of* $L^2(\Omega)$ *and a sequence* $(\lambda_n) \subset \mathbb{R}_-$ *such that* $\lim_{n\to\infty} \lambda_n = -\infty$ *and for every* $n \in \mathbb{N}$

$$e_n \in D(A) \text{ and } \lambda_n e_n - Ae_n = 0.$$

*Moreover,* $\sigma(A) = \sigma_p(A) = \{\lambda_n : n \in \mathbb{N}\}$.

Remark 5.2. Theorem 5.1 gives also a description of the *spectrum* of the Dirichlet-Laplace operator $A$. Every spectral value is an eigenvalue. Every eigenspace is finite dimensional and there exists an orthonormal basis consisting only of eigenvectors. For every $\lambda \notin \sigma(A)$ and every $f \in L^2(\Omega)$ there exists a unique weak solution $u \in H_0^1(\Omega)$ of (5.1).

Theorem 5.1 also implies that the Dirichlet-Laplace operator is unitarily equivalent to a multiplication operator on an $l^2$ space, i.e. the Dirichlet-Laplace operator is *diagonalizable*.

In order to prove Theorem 5.1, we need the following theorem which will not be proved here. We only remark that in the case when $\Omega \subset \mathbb{R}$ is a bounded interval we have given a proof in Example 1.7. For a proof for general $\Omega$, see [**3**].

Theorem 5.3 (Rellich-Kondrachov). *Let $\Omega \subset \mathbb{R}^N$ be open and bounded. Then the embedding*

$$H_0^1(\Omega) \to L^2(\Omega), \quad u \mapsto u,$$

*is compact.*

Proof of Theorem 5.1. Let $u$, $v \in D(A)$. Then,

$$\begin{aligned} (Au, v)_{L^2} &= \int_\Omega Au\bar{v} &&= -\int_\Omega \nabla u \overline{\nabla v} \\ &= \overline{-\int_\Omega \nabla v \overline{\nabla u}} &&= \overline{\int_\Omega Av\bar{u}} \\ &= \overline{(Av, u)_{L^2}} &&= (u, Av). \end{aligned}$$

This equality means that $A$ is *symmetric*.

By Theorem 6.2 of Chapter 4, for every $f \in L^2(\Omega)$ there exists a unique weak solution $u \in H_0^1(\Omega)$ of (5.1) with $\lambda = 1$. This means that $I - A : D(A) \to H$ is bijective. Let $J := (I - A)^{-1} : H \to D(A) \subset H$ be the inverse. For every $u$, $v \in H$, $u = u_1 - Au_1$, $v = v_1 - Av_1$, by the symmetry of $A$,

$$(Ju, v) = (u_1, v_1 - Av_1) = (u_1 - Au_1, v_1) = (u, v_1) = (u, Jv).$$

Hence, $J$ is symmetric. By the Theorem of Hellinger-Toeplitz (Theorem 4.4), $J : H \to H$ is bounded, and thus also self-adjoint. Since $J$ is also a linear operator from $H$ into $H_0^1(\Omega)$, and since $J$ is closed when considered as such an operator, we obtain in fact that $J : H \to H_0^1(\Omega)$ is bounded by the closed graph theorem. Since the embedding $H_0^1(\Omega) \to L^2(\Omega)$ is compact by the Rellich-Kondrachov theorem, we obtain that $J \in \mathcal{K}(H)$.

By the spectral theorem for self-adjoint compact operators, there exists an orthonormal basis $(e_n)$ of $H = L^2(\Omega)$ and a sequence $(\mu_n) \subset \mathbb{R}$ such that $\lim_{n\to\infty} \mu_n = 0$ and

$$\mu_n e_n = Je_n \text{ for every } n \in \mathbb{N}.$$

Since $\operatorname{Rg} J = D(A)$, we obtain also that $e_n \in D(A)$ for every $n \in \mathbb{N}$. Multiplying the above equation by $I - A$, we obtain

$$\lambda_n e_n - Ae_n = 0 \text{ for every } n \in \mathbb{N},$$

with $\lambda_n := \frac{\mu_n - 1}{\mu_n} \in \mathbb{R}$. Since, by Theorem 6.2 of Chapter 4, $\lambda - A$ is invertible for every $\lambda > 0$, we obtain $\lambda_n \in \mathbb{R}_-$. Clearly, the sequence $(\lambda_n)$ is unbounded since $\mu_n \to 0$.

Now let $\lambda \notin \{\lambda_n : n \in \mathbb{N}\}$, and let $f \in L^2(\Omega)$. If $\lambda = 1$ (or even $\lambda > 0$), then we have seen above that the operator $\lambda - A : D(A) \to H$ is bijective. So we can assume that $\lambda \neq 1$. Then $\frac{1}{1-\lambda} \in \varrho(J)$ and we can define $u := R(1, A)R(\frac{1}{1-\lambda}, J)\frac{f}{\lambda - 1}$. Clearly, $u \in D(A)$, and an easy calculation shows that $\lambda u - Au = f$. Moreover, every solution of $\lambda u - Au = f$ is of the form above, and thus $\lambda - A$ is bijective.

The claim is proved. □

Corollary 5.4. *The operator $A$ is closed and*

$$D(A) = \{u \in L^2(\Omega) : (\lambda_n(u, e_n)) \in l^2\}.$$

Proof. If an operator $A : X \supset D(A) \to X$ on a Banach space $X$ has nonempty resolvent set, then $A$ is necessarily closed. In fact, $(\lambda - A)^{-1}$ is bounded for some $\lambda \in \varrho(A) \neq \emptyset$; in particular, $(\lambda - A)^{-1}$ is closed, and thus $\lambda - A$ is closed.

Note that the Dirichlet-Laplace operator $A$ defined above has nonempty resolvent set by Theorem 5.1, and thus $A$ is closed.

The remaining claim follows easily from the fact that, by Theorem 5.1, $A$ is unitarily equivalent to the (unbounded) multiplication operator

$$\begin{aligned} D(M) &:= \{(x_n) \in l^2 : (\lambda_n x_n) \in l^2\}, \\ M(x_n) &:= (\lambda_n x_n), \end{aligned}$$

where the unitary operator is given by

$$\begin{aligned} U : L^2(\Omega) &\to l^2, \\ u &\mapsto ((u, e_n)), \end{aligned}$$

i.e. $A = U^{-1}MU$. □

## 6. * The heat equation

Let $\Omega \subset \mathbb{R}^N$ be open and bounded, and consider the heat equation

$$\text{(6.1)} \qquad \begin{cases} u_t - \Delta u = 0 & \text{in } \mathbb{R}_+ \times \Omega, \\ u = 0 & \text{in } \mathbb{R}_+ \times \partial\Omega, \\ u(0, x) = u_0(x) & \text{in } \Omega, \end{cases}$$

where $\Delta$ denotes the Laplace operator, and $u_0 \in L^2(\Omega)$.

We call a function $u \in C(\mathbb{R}_+; L^2(\Omega))$ a *mild solution* of (6.1) if $u(0) = u_0$ and if for every $\varphi \in D(A)$ the function $t \mapsto (u, \varphi)_{L^2}$ is continuously differentiable and if

$$\frac{d}{dt}(u, \varphi)_{L^2} = (u, A\varphi)_{L^2}.$$

Here, $A$ is the realization of the Dirichlet-Laplace operator on $L^2(\Omega)$ defined in (5.2).

THEOREM 6.1. *For every $u_0 \in L^2(\Omega)$ there exists a unique mild solution $u$ of* (6.1).

PROOF. Let $A$ be the realization of the Dirichlet-Laplace operator as defined in the previous section. By Theorem 5.1, there exists an orthonormal basis $(e_n)$ and an unbounded sequence $(\lambda_n) \subset \mathbb{R}_-$ such that for every $n \in \mathbb{N}$ one has $\lambda_n e_n = A e_n$.

Assume that $u$ is a mild solution of the heat equation (6.1). Then, for every $n \in \mathbb{N}$,

$$\frac{d}{dt}(u(t), e_n)_{L^2} = (u(t), Ae_n)_{L^2} = \lambda_n (u(t), e_n)_{L^2}.$$

This implies

$$(u(t), e_n)_{L^2} = e^{\lambda_n t}(u_0, e_n)_{L^2}, \quad t \geq 0.$$

Hence, since $(e_n)$ is an orthonormal basis,

$$u(t) = \sum_{n\in\mathbb{N}} e^{\lambda_n t}(u_0, e_n)_{L^2}\, e_n, \quad t \geq 0. \tag{6.2}$$

This proves uniqueness of mild solutions.

On the other hand, let $u_0 \in L^2(\Omega)$ and define $u(t)$ as in (6.2). Since $|e^{\lambda_n t}| \leq 1$ for every $t \geq 0$ and since $t \mapsto e^{\lambda_n t}$ is continuous, $u(t) \in L^2(\Omega)$ for every $t \geq 0$, and the function $t \mapsto u(t)$, $\mathbb{R}_+ \to L^2(\Omega)$ is continuous. Moreover, $u(0) = u_0$.

Let $\varphi \in D(A)$. By Corollary 5.4, $(\lambda_n(\varphi, e_n)) \in l^2$. As a consequence, $t \mapsto (u, \varphi)_{L^2}$ is continuously differentiable and, by the symmetry of $A$,

$$\begin{aligned} \frac{d}{dt}(u, \varphi)_{L^2} &= \sum_{n\in\mathbb{N}} \lambda_n e^{\lambda_n t}(u_0, e_n)_{L^2}\,(e_n, \varphi)_{L^2} \\ &= \sum_{n\in\mathbb{N}} e^{\lambda_n t}(u_0, e_n)_{L^2}\,(Ae_n, \varphi)_{L^2} \\ &= \sum_{n\in\mathbb{N}} e^{\lambda_n t}(u_0, e_n)_{L^2}\,(e_n, A\varphi)_{L^2} \\ &= (u, A\varphi)_{L^2}, \quad t \geq 0. \end{aligned}$$

This proves existence of mild solutions. □

REMARK 6.2. The concrete form (6.2) of the solution $u$ of the heat equation (6.1) allows us to prove that in fact

$$u \in C^\infty((0, \infty); L^2(\Omega)),$$

or even

$$u \in C^\infty((0,\infty); D(A^k)) \text{ for every } k \geq 1,$$

where $D(A^k)$ is the domain of $A^k$ equipped with the graph norm. The heat equation thus has a regularizing effect in space and time; even if $u_0$ belongs 'only' to $L^2(\Omega)$, then $u(t)$ belongs already to $D(A^k)$ for every $k \geq 1$. Moreover, the solution is $C^\infty$ with values in $D(A^k)$ for every $k \geq 1$.

## 7. * The wave equation

Let $\Omega \subset \mathbb{R}^N$ be open and bounded, and consider the wave equation

$$\begin{cases} u_{tt} - \Delta u = 0 & \text{in } \mathbb{R}_+ \times \Omega, \\ u = 0 & \text{in } \mathbb{R}_+ \times \partial\Omega, \\ u(0,x) = u_0(x) & \text{in } \Omega, \\ u_t(0,x) = u_1(x) & \text{in } \Omega, \end{cases} \tag{7.1}$$

where $\Delta$ denotes the Laplace operator, $u_0 \in H_0^1(\Omega)$, and $u_1 \in L^2(\Omega)$.

We call a function $u \in C(\mathbb{R}_+; H_0^1(\Omega)) \cap C^1(\mathbb{R}_+; L^2(\Omega))$ a *mild solution* of (7.1) if $u(0) = u_0$, $u_t(0) = u_1$, if for every $\varphi \in H_0^1(\Omega)$ the function $t \mapsto (u,\varphi)_{L^2}$ is twice continuously differentiable and if

$$\frac{d^2}{dt^2}(u(t),\varphi)_{L^2} + \int_\Omega \nabla u(t) \overline{\nabla\varphi} = 0.$$

THEOREM 7.1. *For every $u_0 \in H_0^1(\Omega)$ and every $u_1 \in L^2(\Omega)$ there exists a unique mild solution of* (7.1).

For the proof of Theorem 7.1, we need the following result which we shall not prove here; compare with Corollary 5.4.

LEMMA 7.2. *Let $A$ be the Dirichlet-Laplace operator as defined in* (5.2)*, and let $(e_n)$ and $(\lambda_n)$ be as in Theorem 5.1. Then*

$$H_0^1(\Omega) = \{u \in L^2(\Omega) : (\sqrt{-\lambda_n}(u,e_n)) \in l^2\}.$$

PROOF OF THEOREM 7.1. Let $A$ be the realization of the Dirichlet-Laplace operator as defined in Section 5. By Theorem 5.1, there exists an orthonormal basis $(e_n)$ and an unbounded sequence $(\lambda_n) \subset \mathbb{R}_-$ such that for every $n \in \mathbb{N}$ one has $\lambda_n e_n = A e_n$.

Assume that $u$ is a mild solution of the wave equation (7.1). Then, for every $n \in \mathbb{N}$,

$$\frac{d^2}{dt^2}(u(t),e_n)_{L^2} = (u(t),Ae_n)_{L^2} = \lambda_n (u(t),e_n)_{L^2}.$$

Setting $\alpha_n := \sqrt{-\lambda_n}$, this implies

$$(u(t),e_n)_{L^2} = \cos(\alpha_n t)(u_0,e_n)_{L^2} + \frac{1}{\alpha_n}\sin(\alpha_n t)(u_1,e_n)_{L^2}, \quad t \geq 0.$$

Hence, since $(e_n)$ is an orthonormal basis,

$$u(t) = \sum_{n\in\mathbb{N}} \cos(\alpha_n t)(u_0, e_n)_{L^2}\, e_n + \sum_{n\in\mathbb{N}} \frac{1}{\alpha_n} \sin(\alpha_n t)(u_1, e_n)_{L^2}\, e_n, \quad t \ge 0. \tag{7.2}$$

This proves uniqueness of mild solutions.

On the other hand, let $u_0 \in H_0^1(\Omega)$ and $u_1 \in L^2(\Omega)$, and define $u(t)$ as in (7.2). Since $|\cos(\alpha_n t)| \le 1$ and $|\sin(\alpha_n t)| \le 1$ for every $t \ge 0$ and since cos and sin are continuous, by Lemma 7.2, $u(t) \in H_0^1(\Omega)$ for every $t \ge 0$, and the function $t \mapsto u(t)$, $\mathbb{R}_+ \to H_0^1(\Omega)$ is continuous. Moreover, $u(0) = u_0$. By the same reasons, $t \mapsto u(t)$, $\mathbb{R}_+ \to L^2(\Omega)$ is continuously differentiable and $u_t(0) = u_1$.

Let $\varphi \in H_0^1(\Omega)$. By Lemma 7.2, $(\alpha_n(\varphi, e_n)) \in l^2$. As a consequence, $t \mapsto (u, \varphi)$ is twice continuously differentiable and, by the symmetry of $A$,

$$\begin{aligned}
\frac{d^2}{dt^2}(u,\varphi) &= -\sum_{n\in\mathbb{N}} \lambda_n \cos(\alpha_n t)(u_0, e_n)_{L^2}\,(e_n,\varphi)_{L^2} - \\
&\qquad - \sum_{n\in\mathbb{N}} \alpha_n \sin(\alpha_n t)(u_1, e_n)_{L^2}\,(e_n,\varphi)_{L^2} \\
&= -\sum_{n\in\mathbb{N}} \cos(\alpha_n t)(u_0, e_n)_{L^2}\,(Ae_n,\varphi)_{L^2} - \\
&\qquad - \sum_{n\in\mathbb{N}} \frac{1}{\alpha_n} \sin(\alpha_n t)(u_1, e_n)_{L^2}\,(Ae_n,\varphi)_{L^2} \\
&= -\sum_{n\in\mathbb{N}} \cos(\alpha_n t)(u_0, e_n)_{L^2} \int_\Omega \nabla e_n \nabla\varphi - \\
&\qquad - \sum_{n\in\mathbb{N}} \frac{1}{\alpha_n} \sin(\alpha_n t)(u_1, e_n)_{L^2} \int_\Omega \nabla e_n \nabla\varphi \\
&= -\int_\Omega \nabla u \nabla \varphi, \quad t \ge 0.
\end{aligned}$$

This proves existence of mild solutions. □

Remark 7.3. The concrete form (7.2) of the solution $u$ of the wave equation (7.1) shows that it can be uniquely extended to a solution $u$ defined on $\mathbb{R}$. However, for the wave equation (7.1) there is no regularizing effect as for the heat equation (6.1).

## 8. * The Schrödinger equation

Let $\Omega \subset \mathbb{R}^N$ be open and bounded, and consider the Schrödinger equation

$$\begin{cases} u_t - i\Delta u = 0 & \text{in } \mathbb{R}_+ \times \Omega, \\ u = 0 & \text{in } \mathbb{R}_+ \times \partial\Omega, \\ u(0,x) = u_0(x) & \text{in } \Omega, \end{cases} \tag{8.1}$$

where $\Delta$ denotes the Laplace operator, $i = \sqrt{-1}$ is the complex unity, and $u_0 \in L^2(\Omega)$.

We call a function $u \in C(\mathbb{R}_+; L^2(\Omega))$ a *mild solution* of (8.1) if $u(0) = u_0$ and if for every $\varphi \in D(A)$ the function $t \mapsto (u, \varphi)_{L^2}$ is continuously differentiable and if

$$\frac{d}{dt}(u, \varphi)_{L^2} = i\,(u, A\varphi)_{L^2}, \quad t \geq 0.$$

Here, $A$ is the realization of the Dirichlet-Laplace operator on $L^2(\Omega)$ defined in (5.2).

THEOREM 8.1. *For every $u_0 \in L^2(\Omega)$ there exists a unique mild solution $u$ of* (8.1).

PROOF. Let $A$ be the realization of the Dirichlet-Laplace operator as defined in (5.2). By Theorem 5.1, there exists an orthonormal basis $(e_n)$ and an unbounded sequence $(\lambda_n) \subset \mathbb{R}_-$ such that for every $n \in \mathbb{N}$ one has $\lambda_n e_n = A e_n$.

Assume that $u$ is a mild solution of the Schrödinger equation (8.1). Then, for every $n \in \mathbb{N}$,

$$\frac{d}{dt}(u(t), e_n)_{L^2} = i\,(u(t), Ae_n)_{L^2} = i\lambda_n\,(u(t), e_n)_{L^2}.$$

This implies

$$(u(t), e_n)_{L^2} = e^{i\lambda_n t}(u_0, e_n)_{L^2}, \quad t \geq 0.$$

Hence, since $(e_n)$ is an orthonormal basis,

$$u(t) = \sum_{n\in\mathbb{N}} e^{i\lambda_n t}(u_0, e_n)_{L^2}\, e_n, \quad t \geq 0. \tag{8.2}$$

This proves uniqueness of mild solutions.

On the other hand, let $u_0 \in L^2(\Omega)$ and define $u(t)$ as in (8.2). Since $|e^{i\lambda_n t}| \leq 1$ for every $t \geq 0$ and since $t \mapsto e^{i\lambda_n t}$ is continuous, $u(t) \in L^2(\Omega)$ for every $t \geq 0$, and the function $t \mapsto u(t)$, $\mathbb{R}_+ \to L^2(\Omega)$ is continuous. Moreover, $u(0) = u_0$.

Let $\varphi \in D(A)$. By Corollary 5.4, $(\lambda_n(\varphi, e_n)) \in l^2$. As a consequence, $t \mapsto (u, \varphi)_{L^2}$ is continuously differentiable and, by the symmetry of $A$,

$$\begin{aligned}
\frac{d}{dt}(u, \varphi)_{L^2} &= \sum_{n\in\mathbb{N}} i\lambda_n\, e^{i\lambda_n t}(u_0, e_n)_{L^2}\,(e_n, \varphi)_{L^2} \\
&= i \sum_{n\in\mathbb{N}} e^{i\lambda_n t}(u_0, e_n)_{L^2}\,(Ae_n, \varphi)_{L^2} \\
&= i \sum_{n\in\mathbb{N}} e^{i\lambda_n t}(u_0, e_n)_{L^2}\,(e_n, A\varphi)_{L^2} \\
&= i\,(u, A\varphi)_{L^2}, \quad t \geq 0.
\end{aligned}$$

This proves existence of mild solutions. □

REMARK 8.2. The concrete form (8.2) of the solution $u$ of the Schrödinger equation (8.1) shows that it can be uniquely extended to a solution $u$ defined on $\mathbb{R}$. However, similarly as for the wave equation (7.1), there is no regularizing effect for the Schrödinger equation (8.1).

## 9. * Spectral theorem for unbounded self-adjoint operators

In the preceeding two sections, we have actually proved more than just solvability of an elliptic and a hyperbolic partial differential equation. We have proved that the Dirichlet-Laplace operator is self-adjoint, that it has a compact resolvent, and that therefore it is diagonalisable. In this last section, we discuss the spectral theorem for unbounded self-adjoint operators with compact resolvent.

DEFINITION 9.1. Let $H$ be a complex Hilbert space, and let $A : H \supset D(A) \to H$ be a densely defined (i.e. $D(A)$ is dense in $H$) and linear operator. We define

$$\begin{aligned} D(A^*) &:= \{x \in H : \exists y \in H \, \forall z \in D(A) : \, (Az, x)_H = (z, y)_H\}, \\ A^* x &:= y. \end{aligned}$$

The operator $(A^*, D(A^*))$ is called the *(Hilbert space) adjoint* of $A$. For every $x \in D(A)$, $y \in D(A^*)$ one has

$$(Ax, y) = (x, A^* y).$$

REMARK 9.2. The adjoint $A^*$ is well-defined in the sense that the element $y \in H$ is uniquely determined (use that $D(A)$ is dense in $H$).

LEMMA 9.3. *Let $A : D(A) \to H$ be a densely defined, linear operator. Then $A^* : D(A^*) \to H$ is closed.*

PROOF. Let $(x_n) \subset D(A^*)$ be convergent to some $x \in H$ and such that $(A^* x_n)$ converges to $y \in H$. Then, for every $z \in D(A)$,

$$\begin{aligned} (z, y) &= \lim_{n \to \infty} (z, A^* x_n) \\ &= \lim_{n \to \infty} (Az, x_n) \\ &= (Az, x). \end{aligned}$$

By definition of $A^*$ this implies $x \in D(A^*)$ and $A^* x = y$. Hence, $A^*$ is closed. □

DEFINITION 9.4. Let $H$ be a complex Hilbert space, and let $A : H \supset D(A) \to H$ be a densely defined, linear operator. We say that $A$ is *symmetric* if for every $x$, $y \in D(A)$,

$$(Ax, y) = (x, Ay).$$

We say that $A$ is *self-adjoint* if $A = A^*$.

Remark 9.5. Saying that $A$ is self-adjoint, i.e. that $A = A^*$, means that $D(A) = D(A^*)$ *and* $A = A^*$. By Lemma 9.3, every self-adjoint operator is necessarily closed. Note, however, that a symmetric closed linear operator $A$ need in general not be self-adjoint! However, if $D(A) = H$, then symmetric implies self-adjoint by the Theorem of Hellinger-Toeplitz (Theorem 4.4).

Remark 9.6. If $A : H \to H$ ($D(A) = H$!) is self-adjoint in the sense of Definition 9.1, then $A$ is self-adjoint in the sense of Definition 4.2 and vice versa.

Lemma 9.7. *Let $A : D(A) \to H$ be densely defined and symmetric. Then the following are equivalent:*

(i) *$A$ is self-adjoint.*
(ii) *$A$ is closed and* $\ker(A^* \pm i) = \{0\}$.
(iii) $\operatorname{Rg}(A \pm i) = H$.

Proof. (i)⇒(ii). By Lemma 9.3, $A^*$ is closed, and therefore $A$ is closed. Let $x \in H$ be such that $(A^* - i)x = 0$. Since $A^*$ is symmetric,

$$i\|x\|^2 = (ix, x) = (A^*x, x) = (x, A^*x) = -i\|x\|^2.$$

Hence, $x = 0$. Similarly, one proves $\ker(A^* + i) = \{0\}$.

(ii)⇒(iii). Similarly as in Lemma 3.1 one proves that

$$\ker(A^* - i) = (\operatorname{Rg}(A + i))^\perp,$$

where $\perp$ now means the Hilbert space orthogonal. Hence, if $\ker(A^* - i) = \{0\}$, then $\operatorname{Rg}(A + i)$ is dense in $H$. We prove that $\operatorname{Rg}(A + i)$ is also closed. Since $A$ is symmetric, we have $(Ax, x) \in \mathbb{R}$ for every $x \in D(A)$. Hence, for every $x \in D(A)$,

$$\begin{aligned} \|(A + i)x\| &= \|Ax\|^2 + \|x\|^2 + 2\mathrm{Re}\,(Ax, ix) \\ &= \|Ax\|^2 + \|x\|^2 \quad \geq \quad \|x\|^2. \end{aligned}$$

Let $(x_n) \subset D(A)$ be such that $\lim_{n\to\infty}(A + i)x_n = y \in H$ exists. By the preceeding inequality, this implies that $(x_n)$ is a Cauchy sequence in $H$. Hence, $x := \lim_{n\to\infty} x_n \in H$ exists. Since $A + i$ is closed, we obtain $x \in D(A)$ and $(A + i)x = y$. Hence, $\operatorname{Rg}(A + i)$ is closed. Similarly, $\operatorname{Rg}(A - i)$ is closed.

(iii)⇒(i). Since $A$ is symmetric, $D(A) \subset D(A^*)$ and $Ax = A^*x$ for every $x \in D(A)$. It remains to show that $D(A^*) \subset D(A)$. Let $y \in D(A^*)$. Since $\operatorname{Rg}(A + i) = H$, there exists $x \in D(A)$ such that $(A^* - i)y = (A + i)x$. By symmetry of $A$, $(A + i)x = (A^* - i)x = (A^* - i)y$. Since $\operatorname{Rg}(A + i) = H$ implies $\ker(A^* - i) = \{0\}$ (compare again with Lemma 3.1), this implies $x = y \in D(A)$. □

Exercise 9.8. The Dirichlet-Laplace operator $A$ defined in (5.2) is self-adjoint.

LEMMA 9.9. *Let $A : D(A) \to H$ be densely defined and closed. Then, for every $\lambda \in \varrho(A)$ one has $\bar\lambda \in \varrho(A^*)$ and*

$$R(\lambda, A)^* = R(\bar\lambda, A^*).$$

PROOF. For every $x \in D(A)$ and every $y \in D(A^*)$ one has

$$\begin{aligned} (x, R(\lambda, A)^*(\bar\lambda - A^*)y) &= (R(\lambda, A)x, (\bar\lambda - A^*)y) \\ &= ((\lambda - A)R(\lambda, A)x, y) \\ &= (x, y) \end{aligned}$$

and

$$\begin{aligned} (x, (\bar\lambda - A^*)R(\lambda, A)^*y) &= ((\lambda - A)x, R(\lambda, A)^*y) \\ &= (R(\lambda, A)(\lambda - A)x, y) \\ &= (x, y) \end{aligned}$$

so that $\bar\lambda - A^*$ is invertible and $R(\bar\lambda, A^*) = R(\lambda, A)^*$. □

THEOREM 9.10 (Spectral mapping theorem). *Let $A : D(A) \to H$ be densely defined, closed. Assume that $\varrho(A)$ is not empty. Then, for every $\lambda \in \varrho(A)$,*

$$(\lambda - \sigma(A))^{-1} = \sigma((\lambda - A)^{-1}) \setminus \{0\}.$$

PROOF. The proof is an exercise. □

THEOREM 9.11 (Spectral theorem for unbounded self-adjoint operators with compact resolvent). *Let $A : D(A) \to H$ be densely defined, self-adjoint, such that $R(\lambda, A) \in \mathcal{K}(H)$ for some $\lambda \in \varrho(A)$ (for example for $\lambda = \pm i$). Then there exists an orthonormal basis $(e_n) \subset H$ and a sequence $(\lambda_n) \subset \mathbb{R}$ such that* $\lim_{n\to\infty} |\lambda_n| = \infty$,

$$e_n \in D(A) \text{ and } Ae_n = \lambda_n e_n \text{ for every } n \in \mathbb{N}.$$

*Moreover, $\sigma(A) = \sigma_p(A) = \{\lambda_n : n \in \mathbb{N}\}$.*

PROOF. Let $\lambda \in \varrho(A)$ be such that $R(\lambda, A) \in \mathcal{K}(H)$. By Theorem 3.6, $\sigma(R(\lambda, A))$ is countable. Hence, by Theorem 9.10, $\sigma(A)$ is countable. In particular, there exists $\mu \in \varrho(A) \cap \mathbb{R}$. Let $J := R(\mu, A)$. By the resolvent identity

$$R(\mu, A) = R(\lambda, A) + (\lambda - \mu)R(\mu, A)R(\lambda, A),$$

so that $R(\mu, A)$ is compact, too. Moreover, since $\mu \in \mathbb{R}$, for every $x, y \in H$,

$$(Jx, y) = (Jx, (\mu - A)Jy) = ((\mu - A)Jx, Jy) = (x, Jy),$$

so that $J$ is self-adjoint. By the spectral theorem for self-adjoint compact operators, there exists an orthonormal basis $(e_n)$ and a sequence $(\mu_n) \subset \mathbb{R} \setminus \{0\}$ such that $\lim_{n\to\infty} \mu_n = 0$ and such that

$$\mu_n e_n = Je_n \text{ for every } n \in \mathbb{N}.$$

This equation implies on the one hand that $e_n \in D(A)$ and on the other hand, when we multiply by $\mu - A$,

$$\lambda_n e_n = A e_n \text{ for every } n \in \mathbb{N},$$

with $\lambda_n = \mu - \frac{1}{\mu_n}$. Clearly, $\lim_{n\to\infty} |\lambda_n| = \infty$, and by the spectral mapping theorem (Theorem 9.10), $\sigma(A) = \sigma_p(A) = \{\lambda_n : n \in \mathbb{N}\}$. The claim is proved.

□

# Bibliography


1. J.-P. Aubin, *Applied Functional Analysis*, Pure Applied Mathematics, John Wiley & Sons, New York, Chichester, Brisbane, Toronto, 1979.
2. A. V. Balakrishnan, *Applied Functional Analysis*, Springer Verlag, New York, 1981.
3. H. Brézis, *Analyse fonctionnelle*, Masson, Paris, 1992.
4. V. Hutson and J. S. Pym, *Applications of Functional Analysis and Operator Theory*, Academic Press, New York, 1980.
5. D. Werner, *Funktionalanalysis*, Springer Verlag, Berlin, Heidelberg, New York, 1997.
6. K. Yosida, *Functional analysis*, Springer Verlag, Berlin, Heidelberg, New York, 1978.
7. E. Zeidler, *Nonlinear Functional Analysis and Its Applications. I*, Springer Verlag, New York, Berlin, Heidelberg, 1990.
8. ———, *Applied Functional Analysis: Applications to Mathematical Physics*, Springer Verlag, New York, Berlin, Heidelberg, 1995.