THE UNIVERSITY OF SYDNEY

# Introduction to Functional Analysis

Daniel Daners

School of Mathematics and Statistics
University of Sydney, NSW 2006
Australia

Semester 1, 2017

# Contents

## Acknowledgement

Thanks to Fan Wu from the 2008 Honours Year for providing an extensive list of misprints.

Chapter I

# Preliminary Material

In functional analysis many different fields of mathematics come together. The objects we look at are *vector spaces* and *linear operators*. Hence you need to some basic linear algebra in general vector spaces. I assume your knowledge of that is sufficient. Second we will need some basic set theory. In particular, many theorems depend on the *axiom of choice*. We briefly discuss that most controversial axiom of set theory and some equivalent statements. In addition to the algebraic structure on a vector space, we will look at topologies on them. Of course, these topologies should be compatible with the algebraic structure. This means that addition and multiplication by scalars should be continuous with respect to the topology. We will only look at one class of such spaces, namely *normed spaces* which are naturally metric spaces. Hence it is essential you know the basics of metric spaces, and we provide a self contained introduction of what we need in the course.

## 1 The Axiom of Choice and Zorn's Lemma

Suppose that $A$ is a set, and that for each $\alpha \in A$ there is a set $X_\alpha$. We call $(X_\alpha)_{\alpha\in A}$ a family of sets indexed by $A$. The set $A$ may be finite, countable or uncountable. We then consider the *Cartesian product* of the sets $X_\alpha$:

$$\prod_{\alpha\in A} X_\alpha$$

consisting of all "collections" $(x_\alpha)_{\alpha\in A}$, where $x_\alpha \in X_\alpha$. More formally, $\prod_{\alpha\in A} X_\alpha$ is the set of functions

$$x\colon A \to \bigcup_{\alpha\in A} X_\alpha$$

such that $x(\alpha) \in X_\alpha$ for all $\alpha \in A$. We write $x_\alpha$ for $x(\alpha)$ and $(x_\alpha)_{\alpha\in A}$ or simply $(x_\alpha)$ for a given such function $x$. Suppose now that $A \neq \emptyset$ and $X_\alpha \neq \emptyset$ for all $\alpha \in A$. Then there is a fundamental question:

Is $\prod_{\alpha\in A} X_\alpha$ nonempty in general?

Here some brief history about the problem, showing how basic and difficult it is:

- Zermelo (1904) (see [14]) observed that it is not obvious from the existing axioms of set theory that there is a procedure to select a single $x_\alpha$ from each $X_\alpha$ in general. As a consequence he introduced what we call the *axiom of choice*, asserting that $\prod_{\alpha\in A} X_\alpha \neq \emptyset$ whenever $A \neq \emptyset$ and $X_\alpha \neq \emptyset$ for all $\alpha \in A$.

  It remained open whether his axiom of choice could be derived from the other axioms of set theory. There was an even more fundamental question on whether the axiom is consistent with the other axioms!

- Gödel (1938) (see [8]) proved that the axiom of choice is consistent with the other axioms of set theory. The open question remaining was whether it is independent of the other axioms.

- P.J. Cohen (1963/1964) (see [4, 5]) finally showed that the axiom of choice is in fact independent of the other axioms of set theory, that is, it cannot be derived from them.

The majority of mathematicians accept the axiom of choice, but there is a minority which does not. Many very basic and important theorems in functional analysis cannot be proved without the axiom of choice.

We accept the axiom of choice.

There are some non-trivial equivalent formulations of the axiom of choice which are useful for our purposes. Given two sets $X$ and $Y$ recall that a *relation* from $X$ to $Y$ is simply a subset of the Cartesian product $X \times Y$. We now explore some special relations, namely *order relations*.

**1.1 Definition (partial ordering)** A relation $\prec$ on a set $X$ is called a *partial ordering* of $X$ if

- $x \prec x$ for all $x \in X$ (reflexivity);
- $x \prec y$ and $y \prec z$ imply $x \prec z$ (transitivity);
- $x \prec y$ and $y \prec x$ imply $x = y$ (anti-symmetry).

We also write $x \succ y$ for $y \prec x$. We call $(X, \prec)$ a *partially ordered set*.

**1.2 Examples** (a) The usual ordering $\leq$ on $\mathbb{R}$ is a partial ordering on $\mathbb{R}$.

(b) Suppose $\mathcal{S}$ is a collection of subsets of a set $X$. Then inclusion is a partial ordering. More precisely, if $S, T \in \mathcal{S}$ then $S \prec T$ if and only if $S \subseteq T$. We say $\mathcal{S}$ is partially ordered by inclusion.

(c) Every subset of a partially ordered set is a partially ordered set by the induced partial order.

There are more expressions appearing in connection with partially ordered sets.

**1.3 Definition** Suppose that $(X, \prec)$ is a partially ordered set. Then

(a) $m \in X$ is called a *maximal element in* $X$ if for all $x \in X$ with $x \succ m$ we have $x \prec m$;

(b) $m \in X$ is called an *upper bound* for $S \subseteq X$ if $x \prec m$ for all $x \in S$;

(c) A subset $C \subseteq X$ is called a *chain in* $X$ if $x \prec y$ or $y \prec x$ for all $x, y \in C$;

(d) If a partially ordered set $(X, \prec)$ is a chain we call it a *totally ordered set*.

(e) If $(X, \prec)$ is partially ordered and $x_0 \in X$ is such that $x_0 \prec x$ for all $x \in X$, then we call $x_0$ a *first element*.

There is a special class of partially ordered sets playing a particularly important role in relation to the axiom of choice as we will see later.

**1.4 Definition (well ordered set)** A partially ordered set $(X, \prec)$ is called a *well ordered set* if every subset has a first element.

**1.5 Examples** (a) $\mathbb{N}$ is a well ordered set, but $\mathbb{Z}$ or $\mathbb{R}$ are not well ordered with the usual order.

(b) $\mathbb{Z}$ and $\mathbb{R}$ are totally ordered with the usual order.

**1.6 Remark** Well ordered sets are always totally ordered. To see this assume $(X, \prec)$ is well ordered. Given $x, y \in X$ we consider the subset $\{x, y\}$ of $X$. By definition of a well ordered set we have either $x \prec y$ or $y \prec x$, which shows that $(X, \prec)$ is totally ordered. The converse is not true as the example of $\mathbb{Z}$ given above shows.

There is another, highly non-obvious but very useful statement appearing in connection with partially ordered sets:

**1.7 Zorn's Lemma** *Suppose that* $(X, \prec)$ *is a partially ordered set such that each chain in* $X$ *has an upper bound. Then* $X$ *has a maximal element.*

There is a non-trivial connection between all the apparently different topics we discussed so far. We state it without proof (see for instance [7]).

**1.8 Theorem** *The following assertions are equivalent*

(i) *The axiom of choice;*

(ii) *Zorn's Lemma;*

(iii) *Every set can be well ordered.*

The axiom of choice may seem "obvious" at the first instance. However, the other two equivalent statements are certainly not. For instance take $X = \mathbb{R}$, which we know is not well ordered with the usual order. If we accept the axiom of choice then it follows from the above theorem that there exists a partial ordering making $\mathbb{R}$ into a well ordered set. This is a typical "existence proof" based on the axiom of choice. It does not give us any hint on *how to find* a partial ordering making $\mathbb{R}$ into a well ordered set. This reflects Zermelo's observation that it is not obvious how to choose precisely one element from each set when given an arbitrary collection of sets. Because of the *non-constructive* nature of the axiom of choice and its equivalent counterparts, there are some mathematicians rejecting the axiom. These mathematicians have the point of view that everything should be "constructible," at least in principle, by some means (see for instance [2]).

# 2 Metric Spaces

Metric spaces are sets in which we can measure distances between points. We expect such a "distance function," called a *metric*, to have some obvious properties, which we postulate in the following definition.

**2.1 Definition (Metric Space)** Suppose $X$ is a set. A map $d: X \times X \to \mathbb{R}$ is called a *metric* on $X$ if the following properties hold:

(i) $d(x, y) \geq 0$ for all $x, y \in x$;

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

(iii) $d(x, y) = d(y, x)$ for all $x, y \in X$.

(iv) $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$ (triangle inequality).

We call $(X, d)$ a *metric space*. If it is clear what metric is being used we simply say $X$ is a metric space.

**2.2 Example** The simplest example of a metric space is $\mathbb{R}$ with $d(x, y) := |x - y|$. The standard metric used in $\mathbb{R}^N$ is the *Euclidean metric* given by

$$d(\boldsymbol{x}, \boldsymbol{y}) = |\boldsymbol{x} - \boldsymbol{y}|_2 := \sqrt{\sum_{i=1}^{N} |x_i - y_i|^2}$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$.

**2.3 Remark** If $(X, d)$ is a metric space, then every subset $Y \subseteq X$ is a metric space with the metric restricted to $Y$. We say the metric on $Y$ is *induced* by the metric on $X$.

**2.4 Definition (Open and Closed Ball)** Let $(X, d)$ be a metric space. For $r > 0$ we call

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

the *open ball* about $x$ with radius $r$. Likewise we call

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

the *closed ball* about $x$ with radius $r$.

Using open balls we now define a "topology" on a metric space.

**2.5 Definition (Open and Closed Set)** Let $(X, d)$ be a metric space. A subset $U \subseteq X$ is called *open* if for every $x \in X$ there exists $r > 0$ such that $B(x, r) \subseteq U$. A set $U$ is called *closed* if its complement $X \setminus U$ is open.

**2.6 Remark** For every $x \in X$ and $r > 0$ the open ball $B(x, r)$ in a metric space is open. To prove this fix $y \in B(x, r)$. We have to show that there exists $\varepsilon > 0$ such that $B(y, \varepsilon) \subseteq B(x, r)$. To do so note that by definition $d(x, y) < r$. Hence we can choose $\varepsilon \in \mathbb{R}$ such that $0 < \varepsilon < r - d(x, y)$. Thus, by property (iv) of a metric, for $z \in B(y, \varepsilon)$ we have $d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + r - d(x, y) = r$. Therefore $z \in B(x, r)$, showing that $B(y, \varepsilon) \subseteq B(x, r)$.

Next we collect some fundamental properties of open sets.

**2.7 Theorem** *Open sets in a metric space $(X, d)$ have the following properties.*

(i) *$X$, $\emptyset$ are open sets;*

(ii) *arbitrary unions of open sets are open;*

(iii) *finite intersections of open sets are open.*

*Proof.* Property (i) is obvious. To prove (ii) let $U_\alpha$, $\alpha \in A$ be an arbitrary family of open sets in $X$. If $x \in \bigcup_{\alpha \in A} U_\alpha$ then $x \in U_\beta$ for some $\beta \in A$. As $U_\beta$ is open there exists $r > 0$ such that $B(x, r) \subseteq U_\beta$. Hence also $B(x, r) \subseteq \bigcup_{\alpha \in A} U_\alpha$, showing that $\bigcup_{\alpha \in A} U_\alpha$ is open. To prove (iii) let $U_i$, $i = 1, \dots, n$ be open sets. If $x \in \bigcap_{i=1}^n U_i$ then $x \in U_i$ for all $i = 1, \dots, n$. As the sets $U_i$ are open there exist $r_i > 0$ such that $B(x, r_i) \subseteq U_i$ for all $i = 1, \dots, n$. If we set $r := \min_{i=1,\dots,n} r_i$ then obviously $r > 0$ and $B(x, r) \subseteq \bigcap_{i=1}^n U_i$, proving (iii). ■

**2.8 Remark** There is a more general concept than that of a metric space, namely that of a "topological space." A collection $\mathcal{T}$ of subsets of a set $X$ is called a *topology* if the following conditions are satisfied

(i) $X, \emptyset \in \mathcal{T}$;

(ii) arbitrary unions of sets in $\mathcal{T}$ are in $\mathcal{T}$;

(iii) finite intersections of sets in $\mathcal{T}$ are in $\mathcal{T}$.

The elements of $\mathcal{T}$ are called *open sets*, and $(X, \mathcal{T})$ a *topological space*. Hence the open sets in a metric space form a topology on $X$.

**2.9 Definition (Neighbourhood)** Suppose that $(X, d)$ is a metric space (or more generally a topological space). We call a set $U$ a *neighbourhood* of $x \in X$ if there exists an open set $V \subseteq U$ with $x \in V$.

Now we define some sets associated with a given subset of a metric space.

**2.10 Definition (Interior, Closure, Boundary)** Suppose that $U$ is a subset of a metric space $(X, d)$ (or more generally a topological space). A point $x \in U$ is called an *interior point* of $U$ if $U$ is a neighbourhood of $x$. We call

(i) $\mathring{U} := \operatorname{Int}(U) := \{x \in U : x \text{ interior point of } U\}$ the *interior* of $U$;

(ii) $\overline{U} := \{x \in X : U \cap V \neq \emptyset$ for every neighbourhood $V$ of $x\}$ the *closure* of $U$;

(iii) $\partial U := \bar{U} \setminus \operatorname{Int}(U)$ the *boundary* of $U$.

**2.11 Remark** A set is open if and only if $\mathring{U} = U$ and closed if and only if $\overline{U} = U$. Moreover, $\partial U = \overline{U} \cap \overline{X \setminus U}$.

Sometimes it is convenient to look at products of a (finite) number of metric spaces. It is possible to define a metric on such a product as well.

**2.12 Proposition** *Suppose that $(X_i, d_i)$, $i = 1, \dots, n$ are metric spaces. Then $X = X_1 \times X_2 \times \cdots \times X_n$ becomes a metric space with the metric $d$ defined by*

$$d(x, y) := \sum_{i=1}^{n} d_i(x_i, y_i)$$

*for all $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ in $X$.*

*Proof.* Obviously, $d(x, y) \geq 0$ and $d(x, y) = d(y, x)$ for all $x, y \in X$. Moreover, as $d_i(x_i, y_i) \geq 0$ we have $d(x, y) = 0$ if and only if $d_i(x_i, y_i) = 0$ for all $i = 1, \dots, n$. As $d_i$ are metrics we get $x_i = y_i$ for all $i = 1, \dots, n$. For the triangle inequality note that

$$\begin{aligned} d(x, y) = \sum_{i=1}^{n} d(x_i, y_i) &\leq \sum_{i=1}^{n} \big(d(x_i, z_i) + d(z_i, y_i)\big) \\ &= \sum_{i=1}^{n} d(x_i, z_i) + \sum_{i=1}^{n} d(z_i, y_i) = d(x, y) + d(z, y) \end{aligned}$$

for all $x, y, z \in X$. ■

**2.13 Definition (Product space)** The space and metric introduced in Proposition 2.12 is called a *product space* and a *product metric*, respectively.

# 3 Limits

Once we have a notion of "closeness" we can discuss the asymptotics of sequences and continuity of functions.

**3.1 Definition (Limit)** Suppose $(x_n)_{n\in\mathbb{N}}$ is a sequence in a metric space $(X, d)$, or more generally a topological space. We say $x_0$ is a *limit* of $(x_n)$ if for every neighbourhood $U$ of $x_0$ there exists $n_0 \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq n_0$. We write

$$x_0 = \lim_{n\to\infty} x_n \qquad \text{or} \qquad x_n \to x \text{ as } n \to \infty.$$

If the sequence has a limit we say it is *convergent*, otherwise we say it is *divergent*.

**3.2 Remark** Let $(x_n)$ be a sequence in a metric space $(X,d)$ and $x_0 \in X$. Then the following statements are equivalent:

(1) $\lim_{n\to\infty} x_n = x_0$;

(2) for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $d(x_n, x_0) < \varepsilon$ for all $n \geq n_0$.

*Proof.* Clearly (1) implies (2) by choosing neighbourhoods of the form $B(x,\varepsilon)$. If (2) holds and $U$ is an arbitrary neighbourhood of $x_0$ we can choose $\varepsilon > 0$ such that $B(x_0,\varepsilon) \subseteq U$. By assumption there exists $n_0 \in \mathbb{N}$ such that $d(x_n, x_0) < \varepsilon$ for all $n \geq n_0$, that is, $x_n \in B(x_0,\varepsilon) \subseteq U$ for all $n \geq n_0$. Therefore, $x_n \to x_0$ as $n \to \infty$. ■

**3.3 Proposition** *A sequence in a metric space $(X,d)$ has at most one limit.*

*Proof.* Suppose that $(x_n)$ is a sequence in $(X,d)$ and that $x$ and $y$ are limits of that sequence. Fix $\varepsilon > 0$ arbitrary. Since $x$ is a limit there exists $n_1 \in \mathbb{N}$ such that $d(x_n, x) < \varepsilon/2$ for all $n > n_1$. Similarly, since $y$ is a limit there exists $n_2 \in \mathbb{N}$ such that $d(x_n, y) < \varepsilon/2$ for all $n > n_2$. Hence $d(x,y) \leq d(x,x_n) + d(x_n,y) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$ for all $n > \max\{n_1, n_2\}$. Since $\varepsilon > 0$ was arbitrary it follows that $d(x,y) = 0$, and so by definition of a metric $x = y$. Thus $(x_n)$ has at most one limit. ■

We can characterise the closure of sets by using sequences.

**3.4 Theorem** *Let $U$ be a subset of the metric space $(X,d)$ then $x \in \overline{U}$ if and only if there exists a sequence $(x_n)$ in $U$ such that $x_n \to x$ as $n \to \infty$.*

*Proof.* Let $U \subseteq X$ and $x \in \overline{U}$. Hence $B(x,\varepsilon) \cap U) \neq \emptyset$ for all $\varepsilon > 0$. For all $n \in \mathbb{N}$ we can therefore choose $x_n \in U$ with $d(x,x_n) < 1/n$. By construction $x_n \to x$ as $n \to \infty$. If $(x_n)$ is a sequence in $U$ converging to $x$ then for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $x_n \in B(x,\varepsilon)$ for all $n \geq n_0$. In particular, $B(x,\varepsilon) \cap U \neq \emptyset$ for all $\varepsilon > 0$, implying that $x \in \overline{U}$ as required. ■

There is another concept closely related to convergence of sequences.

**3.5 Definition (Cauchy Sequence)** Suppose $(x_n)$ is a sequence in the metric space $(X,d)$. We call $(x_n)$ a *Cauchy sequence* if for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon$ for all $m, n \geq n_0$.

Some sequences may not converge, but they accumulate at certain points.

**3.6 Definition (Point of Accumulation)** Suppose that $(x_n)$ is a sequence in a metric space $(X,d)$ or more generally in a topological space. We say that $x_0$ is a *point of accumulation* of $(x_n)$ if for every neighbourhood $U$ of $x_0$ and every $n_0 \in \mathbb{N}$ there exists $n \geq n_0$ such that $x_n \in U$.

**3.7 Remark** Equivalently we may say $x_0$ is an accumulation point of $(x_n)$ if for every $\varepsilon > 0$ and every $n_0 \in \mathbb{N}$ there exists $n \geq n_0$ such that $d(x_n, x_0) < \varepsilon$. Note that it follows from the definition that every neighbourhood of $x_0$ contains infinitely many elements of the sequence $(x_n)$.

**3.8 Proposition** *Suppose that $(X, d)$ is a metric space and $(x_n)$ a sequence in that space. Then $x \in X$ is a point of accumulation of $(x_n)$ if and only if*

$$x \in \bigcap_{k=1}^{\infty} \overline{\{x_j : j \geq k\}}. \tag{3.1}$$

*Proof.* Suppose that $x \in \bigcap_{k=1}^{\infty} \overline{\{x_j : j \geq k\}}$. Then $x \in \overline{\{x_j : j \geq k\}}$ for all $k \in \mathbb{N}$. By Theorem 3.4 we can choose for every $k \in \mathbb{N}$ an element $x_{n_k} \in \{x_j : j \geq k\}$ such that $d(x_{n_k}, x) < 1/k$. By construction $x_{n_k} \to x$ as $k \to \infty$, showing that $x$ is a point of accumulation of $(x_n)$. If $x$ is a point of accumulation of $(x_n)$ then for all $k \in \mathbb{N}$ there exists $n_k \geq k$ such that $d(x_{n_k}, x) < 1/k$. Clearly $x_{n_k} \to x$ as $k \to \infty$, so that $x \in \overline{\{x_{n_j} : j \geq k\}}$ for all $k \in \mathbb{N}$. As $\{x_{n_j} : j \geq k\} \subseteq \{x_j : j \geq k\}$ for all $k \in \mathbb{N}$ we obtain (3.1). ■

In the following theorem we establish a connection between Cauchy sequences and converging sequences.

**3.9 Theorem** *Let $(X, d)$ be a metric space. Then every convergent sequence is a Cauchy sequence. Moreover, if a Cauchy sequence $(x_n)$ has an accumulation point $x_0$, then $(x_n)$ is a convergent sequence with limit $x_0$.*

*Proof.* Suppose that $(x_n)$ is a convergent sequence with limit $x_0$. Then for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $d(x_n, x_0) < \varepsilon/2$ for all $n \geq n_0$. Now

$$d(x_n, x_m) \leq d(x_n, x_0) + d(x_0, x_m) = d(x_n, x_0) + d(x_m, x_0) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for all $n, m \geq n_0$, showing that $(x_n)$ is a Cauchy sequence. Now assume that $(x_n)$ is a Cauchy sequence, and that $x_0 \in X$ is an accumulation point of $(x_n)$. Fix $\varepsilon > 0$ arbitrary. Then by definition of a Cauchy sequence there exists $n_0 \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon/2$ for all $n, m \geq n_0$. Moreover, since $x_0$ is an accumulation point there exists $m_0 \geq n_0$ such that $d(x_{m_0}, x_0) < \varepsilon/2$. Hence

$$d(x_n, x_0) \leq d(x_n, x_{m_0}) + d(x_{m_0}, x_0) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for all $n \geq n_0$. Hence by Remark 3.2 $x_0$ is the limit of $(x_n)$. ■

In a general metric space not all Cauchy sequences have necessarily a limit, hence the following definition.

**3.10 Definition (Complete Metric Space)** A metric space is called *complete* if every Cauchy sequence in that space has a limit.

One property of the real numbers is that the intersection of a nested sequence of closed bounded intervals whose lengths shrinks to zero have a non-empty intersection. This property is in fact equivalent to the “completeness” of the real number system. We now prove a counterpart of that fact for metric spaces. There are no intervals in general metric spaces, so we look at a sequence of nested closed sets whose diameter goes to zero. The diameter of a set $K$ in a metric space $(X, d)$ is defined by

$$\operatorname{diam}(K) := \sup_{x,y \in K} d(x, y).$$

**3.11 Theorem (Cantor's Intersection Theorem)** *Let $(X, d)$ be a metric space. Then the following two assertions are equivalent:*

(i) *$(X, d)$ is complete;*

(ii) *For every sequence of closed sets $K_n \subseteq X$ with $K_{n+1} \subseteq K_n$ for all $n \in \mathbb{N}$*

$$\operatorname{diam}(K_n) := \sup_{x,y \in K_n} d(x, y) \to 0$$

*as $n \to \infty$ we have $\bigcap_{n \in \mathbb{N}} K_n \neq \emptyset$.*

*Proof.* First assume that $X$ is complete and let $K_n$ be as in (ii). For every $n \in \mathbb{N}$ we choose $x_n \in K_n$ and show that $(x_n)$ is a Cauchy sequence. By assumption $K_{n+1} \subseteq K_n$ for all $n \in \mathbb{N}$, implying that $x_m \in K_m \subseteq K_n$ for all $m > n$. Since $x_m, x_n \in K_n$ we have

$$d(x_m, x_n) \leq \sup_{x,y \in K_n} d(x, y) = \operatorname{diam}(K_n)$$

for all $m > n$. Since $\operatorname{diam}(K_n) \to 0$ as $n \to \infty$, given $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $\operatorname{diam}(K_{n_0}) < \varepsilon$. Hence, since $K_m \subseteq K_n \subseteq K_{n_0}$ we have

$$d(x_m, x_n) \leq \operatorname{diam}(K_n) \leq \operatorname{diam}(K_{n_0}) < \varepsilon$$

for all $m > n > n_0$, showing that $(x_n)$ is a Cauchy sequence. By completenes of $S$, the sequence $(x_n)$ converges to some $x \in X$. We know from above that $x_m \in K_n$ for all $m > n$. As $K_n$ is closed $x \in K_n$. Since this is true for all $n \in \mathbb{N}$ we conclude that $x \in \bigcap_{n \in \mathbb{N}} K_n$, so the intersection is non-empty as claimed.

Assume now that (ii) is true and let $(x_n)$ be a Cauchy sequence in $(X, d)$. Hence there exists $n_0 \in X$ such that $d(x_{n_0}, x_n) < 1/2$ for all $n \geq n_0$. Similarly, there exists $n_1 > n_0$ such that $d(x_{n_1}, x_n) < 1/2^2$ for all $n \geq n_1$. Continuing that way we construct a sequence $(n_k)$ in $\mathbb{N}$ such that for every $k \in \mathbb{N}$ we have $n_{k+1} > n_k$ and $d(x_{n_k}, x_n) < 1/2^{k+1}$ for all $n > n_k$. We now set $K_k := \overline{B(x_k, 2^{-k})})$. If $x \in K_{k+1}$, then since $n_{k+1} > n_k$

$$d(x_{n_k}, x) \leq d(x_{n_k}, x_{n_{k+1}}) + d(x_{n_{k+1}}, x) < \frac{1}{2^{k+1}} + \frac{1}{2^{k+1}} = \frac{1}{2^k}.$$

Hence $x \in K_k$, showing that $K_{k+1} \subseteq K_k$ for all $k \in \mathbb{N}$. By assumption (ii) we have $\bigcap_{k \in \mathbb{N}} K_k \neq \emptyset$, so choose $x \in \bigcap_{k \in \mathbb{N}} K_k \neq \emptyset$. Then $x \in K_k$ for all $k \in \mathbb{N}$, so $d(x_{n_k}, x) \leq 1/2^k$ for all $k \in \mathbb{N}$. Hence $x_{n_k} \to x$ as $k \to \infty$. By Theorem 3.9 the Cauchy sequence $(x_n)$ converges, proing (i). ■

We finally look at product spaces defined in Definition 2.13. The rather simple proof of the following proposition is left to the reader.

**3.12 Proposition** *Suppose that $(X_i, d_i)$, $i = 1, \dots, n$ are complete metric spaces. Then the corresponding product space is complete with respect to the product metric.*

# 4 Compactness

We start by introducing some additional concepts, and show that they are all equivalent in a metric space. They are all generalisations of “finiteness” of a set.

**4.1 Definition (Open Cover, Compactness)** Let $(X, d)$ be a metric space. We call a collection of open sets $(U_\alpha)_{\alpha\in A}$ an *open cover* of $X$ if $X \subseteq \bigcup_{\alpha\in A} U_\alpha$. The space $X$ is called *compact* if for every open cover $(U_\alpha)_{\alpha\in A}$ there exist finitely many $\alpha_i \in A$, $i = 1, \dots, m$ such that $(U_{\alpha_i})_{i=1,\dots,m}$ is an open cover of $X$. We talk about a finite sub-cover of $X$.

**4.2 Definition (Sequential Compactness)** We call a metric space $(X, d)$ *sequentially compact* if every sequence in $X$ has an point of accumulation.

**4.3 Definition (Total Boundedness)** We call a metric space $X$ *totally bounded* if for every $\varepsilon > 0$ there exist finitely many points $x_i \in X$, $i = 1, \dots, m$, such that $(B(x_i, \varepsilon))_{i=1,\dots,m}$ is an open cover of $X$.

It turns out that all the above definitions are equivalent, at least in metric spaces (but not in general topological spaces).

**4.4 Theorem** *For a metric space $(X, d)$ the following statements are equivalent:*

(i) *$X$ is compact;*

(ii) *$X$ is sequentially compact;*

(iii) *$X$ is complete and totally bounded.*

*Proof.* To prove that (i) implies (ii) assume that $X$ is compact and that $(x_n)$ is a sequence in $X$. We set $C_n := \overline{\{x_j : j \geq n\}}$ and $U_n := X \setminus C_n$. Then $U_n$ is open for all $n \in \mathbb{N}$ as $C_n$ is closed. By Proposition 3.8 the sequence $(x_n)$ has a point of accumulation if

$$\bigcap_{n\in\mathbb{N}} C_n \neq \emptyset,$$

which is equivalent to

$$\bigcup_{n\in\mathbb{N}} U_n = \bigcup_{n\in\mathbb{N}} X \setminus C_n = X \setminus \bigcap_{n\in\mathbb{N}} C_n \neq X$$

Clearly $C_0 \supset C_1 \supset \cdots \supset C_n \neq \emptyset$ for all $n \in \mathbb{N}$. Hence every finite intersection of sets $C_n$ is nonempty. Equivalently, every finite union of sets $U_n$ is strictly smaller than $X$, so that $X$ cannot be covered by finitely many of the sets $U_n$. As $X$ is compact it is impossible that $\bigcup_{n\in\mathbb{N}} U_n = X$ as otherwise a finite number would cover $X$ already, contradicting what we just proved. Hence $(x_n)$ must have a point of accumulation.

Now assume that (ii) holds. If $(x_n)$ is a Cauchy sequence it follows from (ii) that it has a point of accumulation. By Theorem 3.9 we conclude that it has a limit, showing that $X$ is complete. Suppose now that $X$ is not totally bounded. Then, there exists $\varepsilon > 0$ such that $X$ cannot be covered by finitely many balls of radius $\varepsilon$. If we let $x_0$ be arbitrary we can therefore choose $x_1 \in X$ such that $d(x_0, x_1) > \varepsilon$. By induction we may construct a sequence $(x_n)$ such that $d(x_j, x_n) \geq \varepsilon$ for all $j = 1, \dots, n-1$. Indeed, suppose we have $x_0, \dots, x_n \in X$ with $d(x_j, x_n) \geq \varepsilon$ for all $j = 1, \dots, n-1$. Assuming that $X$ is not totally bounded $\bigcup_{j=1}^n B(x_j, \varepsilon) \neq X$, so we can choose $x_{n+1}$ not in that union. Hence $d(x_j, x_{n+1}) \geq \varepsilon$ for $j = 1, \dots, n$. By construction it follows that $d(x_n, x_m) \geq \varepsilon/2$ for all $n, m \in \mathbb{N}$, showing that $(x_n)$ does not contain a Cauchy subsequence, and thus has no point of accumulation. As this contradicts (ii), the space $X$ must be totally bounded.

Suppose now that (iii) holds, but $X$ is not compact. Then there exists an open cover $(U_\alpha)_{\alpha \in A}$ not having a finite sub-cover. As $X$ is totally bounded, for every $n \in \mathbb{N}$ there exist finite sets $F_n \subseteq X$ such that

$$X = \bigcup_{x \in F_n} B(x, 2^{-n}). \tag{4.1}$$

Assuming that $(U_\alpha)_{\alpha \in A}$ does not have a finite sub-cover, there exists $x_1 \in F_1$ such that $B(x_1, 2^{-1})$ and thus $K_1 := \overline{B(x_1, 3 \cdot 2^{-1})}$ cannot be covered by finitely many $U_\alpha$. By (4.1) it follows that there exists $x_2 \in F_2$ such that $B(x_1, 2^{-1}) \cap B(x_2, 2^{-2})$ and therefore $K_2 := \overline{B(x_2, 3 \cdot 2^{-2})}$ is not finitely covered by $(U_\alpha)_{\alpha \in A}$. We can continue this way and choose $x_{n+1} \in F_{n+1}$ such that $B(x_n, 2^{-n}) \cap B(x_{n+1}, 2^{-(n+1)})$ and therefore $K_{n+1} := \overline{B(x_2, 3 \cdot 2^{-(n+1)})}$ is not finitely covered by $(U_\alpha)_{\alpha \in A}$. Note that $B(x_n, 2^{-n}) \cap B(x_{n+1}, 2^{-(n+1)}) \neq \emptyset$ since otherwise the intersection is finitely covered by $(U_\alpha)_{\alpha \in A}$. Hence if $x \in K_{n+1}$, then

$$d(x_n, x) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x) \leq \frac{1}{2^n} + \frac{1}{2^{n+1}} + \frac{3}{2^{n+1}} = \frac{6}{2^{n+1}} = \frac{3}{2^n},$$

implying that $x \in K_n$. Also $\operatorname{diam} K_n \leq 3 \cdot 2^{n-1} \to 0$. Since $X$ is complete, by Cantor's intersection Theorem 3.11 there exists $x \in \bigcap_{n \in \mathbb{N}} K_n$. As $(U_\alpha)$ is a cover of $X$ we have $x \in U_{\alpha_0}$ for some $\alpha_0 \in A$. Since $U_{\alpha_0}$ is open there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U_{\alpha_0}$. Choose now $n$ such that $6/2^n < \varepsilon$ and fix $y \in K_n$. Since $x \in K_n$ we have $d(x, y) \leq d(x, x_n) + d(x_n, y) \leq 6/2^n < \varepsilon$. Hence $K_n \subseteq B(x, \varepsilon) \subseteq U_{\alpha_0}$, showing that $K_n$ is covered by $U_{\alpha_0}$. However, by construction $K_n$ cannot be covered by finitely many $U_\alpha$, so we have a contradiction. Hence $X$ is compact, completing the proof of the theorem. ■

The last part of the proof is modelled on the usual proof of the Heine-Borel theorem asserting that bounded and closed sets are the compact sets in $\mathbb{R}^N$. Hence it is not a surprise that the Heine-Borel theorem easily follows from the above characterisations of compactness.

**4.5 Theorem (Heine-Borel)** *A subset of $\mathbb{R}^N$ is compact if and only if it is closed and bounded.*

*Proof.* Suppose $A \subseteq \mathbb{R}^N$ is compact. By Theorem 4.4 the set $A$ is totally bounded, and thus may be covered by finitely many balls of radius one. A finite union of such balls is clearly bounded, so $A$ is bounded. Again by Theorem 4.4, the set $A$ is complete, so in particular it is closed. Now assume $A$ is closed and bounded. As $\mathbb{R}^N$ is complete it follows that $A$ is complete. Next we show that $A$ is totally bounded. We let $M$ be such that $A$ is contained in the cube $[-M, M]^N$. Given $\varepsilon > 0$ the interval $[-M, M]$ can be covered by $m := [2M/\varepsilon] + 1$ closed intervals of length $\varepsilon/2$ (here $[2M/\varepsilon]$ is the integer part of $2M/\varepsilon$). Hence $[-M, M]^N$ can be covered by $m^N$ cubes with edges $\varepsilon/2$ long. Such cubes are contained in open balls of radius $\varepsilon$, so we can cover $[-M, M]^N$ and thus $A$ by a finite number of balls of radius $\varepsilon$. Hence $A$ is complete and totally bounded. By Theorem 4.4 the set $A$ is compact. ■

We can also look at subsets of metric spaces. As they are metric spaces with the metric induced on them we can talk about compact subsets of a metric space. It follows from the above theorem that compact subsets of a metric space are always closed (as they are complete). Often in applications one has sets that are not compact, but their closure is compact.

**4.6 Definition (Relatively Compact Sets)** We call a subset of a metric space *relatively compact* if its closure is compact.

**4.7 Proposition** *Closed subsets of compact metric spaces are compact.*

*Proof.* Suppose $C \subseteq X$ is closed and $X$ is compact. If $(U_\alpha)_{\alpha\in A}$ is an open cover of $C$ then we get an open cover of $X$ if we add the open set $X \setminus C$ to the $U_\alpha$. As $X$ is compact there exists a finite sub-cover of $X$, and as $X \setminus C \cap C = \emptyset$ also a finite sub-cover of $C$. Hence $C$ is compact. ■

Next we show that finite products of compact metric spaces are compact.

**4.8 Proposition** *Let $(X_i, d_i)$, $i = 1, \dots, n$, be compact metric spaces. Then the product $X := X_1 \times \cdots \times X_n$ is compact with respect to the product metric introduced in Proposition 2.12.*

*Proof.* By Proposition 3.12 it follows that the product space $X$ is complete. By Theorem 4.4 is is therefore sufficient to show that $X$ is totally bounded. Fix $\varepsilon > 0$. Since $X_i$ is totally bounded there exist $x_{ik} \in X_i$, $k = 1, \dots m_i$ such that $X_i$ is covered by the balls $B_{ik}$ of radius $\varepsilon/n$ and centre $x_{ik}$. Then $X$ is covered by the balls of radius $\varepsilon$ with centres $(x_{1k_1}, \dots, x_{ik_i}, \dots x_{nk_n})$, where $k_i = 1, \dots m_i$. Indeed, suppose that $x = (x_1, x_2, \dots, x_n) \in X$ is arbitrary. By assumption, for every $i = 1, \dots n$ there exist $1 \leq k_i \leq m_i$ such that $d(x_i, x_{ik_i}) < \varepsilon/n$. By definition of the product metric the distance between $(x_{1k_1}, \dots, x_{nk_n})$ and $x$ is no larger than $d(x_1, x_{1k_1}) + \cdots + d(x_n, x_{nk_n}) \leq n\varepsilon/n = \varepsilon$. Hence $X$ is totally bounded and thus $X$ is compact. ■

# 5 Continuous Functions

We give a brief overview on continuous functions between metric spaces. Throughout, let $X = (X, d)$ denote a metric space. We start with some basic definitions.

**5.1 Definition (Continuous Function)** A function $f : X \to Y$ between two metric spaces is called continuous at a point $x \in X$ if for every neighbourhood $V \subseteq Y$ of $f(x)$

there exists a neighbourhood $U \subseteq X$ of $x$ such that $f(U) \subseteq V$. The map $f: X \to Y$ is called continuous if it is continuous at all $x \in X$. Finally we set

$$C(X, Y) := \{f : X \to Y \mid f \text{ is continuous}\}.$$

The above is equivalent to the usual $\varepsilon$-$\delta$ definition.

**5.2 Theorem** *Let $X, Y$ be metric spaces and $f : X \to Y$ a function. Then the following assertions are equivalent:*

(i) *$f$ is continuous at $x \in X$;*

(ii) *For every $\varepsilon > 0$ there exists $\delta > 0$ such that $d_Y\big(f(x), f(y)\big) \leq \varepsilon$ for all $y \in X$ with $d_X(x, y) < \delta$;*

(iii) *For every sequence $(x_n)$ in $X$ with $x_n \to x$ we have $f(x_n) \to f(x)$ as $n \to \infty$.*

*Proof.* Taking special neighbourhoods $V = B(f(x), \varepsilon)$ and $U := B(x, \delta)$ then (ii) is clearly necessary for $f$ to be continuous. To show the (ii) is sufficient let $V$ be an arbitrary neighbourhood of $f(x)$. Then there exists $\varepsilon > 0$ such that $B(f(x), \varepsilon) \subseteq V$. By assumption there exists $\delta > 0$ such that $d_Y\big(f(x), f(y)\big) \leq \varepsilon$ for all $y \in X$ with $d_X(x, y) < \delta$, that is, $f(U) \subseteq V$ if we let $U := B(x, \delta)$. As $U$ is a neighbourhood of $x$ it follows that $f$ is continuous. Let now $f$ be continuous and $(x_n)$ a sequence in $X$ converging to $x$. If $\varepsilon > 0$ is given then there exists $\delta > 0$ such that $d_Y(f(x), f(y)) < \varepsilon$ for all $y \in X$ with $d_X(x, y) < \delta$. As $x_n \to x$ there exists $n_0 \in \mathbb{N}$ such that $d_X(x, x_n) < \delta$ for all $n \geq n_0$. Hence $d_Y(f(x), f(x_n)) < \varepsilon$ for all $n \geq n_0$. As $\varepsilon > 0$ was arbitrary $f(x_n) \to f(x)$ as $n \to \infty$. Assume now that (ii) does not hold. Then there exists $\varepsilon > 0$ such that for each $n \in \mathbb{N}$ there exists $x_n \in X$ with $d_X(x, x_n) < 1/n$ but $d_Y(f(x), f(x_n)) \geq \varepsilon$ for all $n \in \mathbb{N}$. Hence $x_n \to x$ in $X$ but $f(x_n) \not\to f(x)$ in $Y$, so (iii) does not hold. By contrapositive (iii) implies (ii), completing the proof of the theorem. ■

Next we want to give various equivalent characterisations of continuous maps (without proof).

**5.3 Theorem (Characterisation of Continuity)** *Let $X, Y$ be metric spaces. Then the following statements are equivalent:*

(i) *$f \in C(X, Y)$;*

(ii) *$f^{-1}[O] := \{x \in X : f(x) \in O\}$ is open for every open set $O \subseteq Y$;*

(iii) *$f^{-1}[C]$ is closed for every closed set $C \subseteq Y$;*

(iv) *For every $x \in X$ and every neighbourhood $V \subseteq Y$ of $f(x)$ there exists a neighbourhood $U \subseteq X$ of $x$ such that $f(U) \subseteq V$;*

(v) *For every $x \in X$ and every $\varepsilon > 0$ there exists $\delta > 0$ such that $d_Y\big(f(x), f(y)\big) < \varepsilon$ for all $y \in X$ with $d_X(x, y) < \delta$.*

**5.4 Definition (Distance to a Set)** Let $A$ be a nonempty subset of $X$. We define the distance between $x \in X$ and $A$ by

$$\operatorname{dist}(x, A) := \inf_{a \in A} d(x, a)$$

**5.5 Proposition** *For every nonempty set $A \subseteq X$ the map $X \to \mathbb{R}$, $x \mapsto \operatorname{dist}(x, A)$, is continuous.*

*Proof.* By the properties of a metric $d(x, a) \leq d(x, y) + d(y, a)$. By first taking an infimum on the left hand side and then on the right hand side we get $\operatorname{dist}(x, A) \leq d(x, y) + \operatorname{dist}(y, A)$ and thus

$$\operatorname{dist}(x, A) - \operatorname{dist}(y, A) \leq d(x, y)$$

for all $x, y \in X$. Interchanging the roles of $x$ and $y$ we get $\operatorname{dist}(y, A) - \operatorname{dist}(x, A) \leq d(x, y)$, and thus

$$\left|\operatorname{dist}(x, A) - \operatorname{dist}(y, A)\right| \leq d(x, y),$$

implying the continuity of $\operatorname{dist}(\cdot\,, A)$. ■

We continue to discuss properties of continuous functions on compact sets.

**5.6 Theorem** *If $f \in C(X, Y)$ and $X$ is compact then the image $f(X)$ is compact in $Y$.*

*Proof.* Suppose that $(U_\alpha)$ is an open cover of $f(X)$ then by continuity $f^{-1}[U_\alpha]$ are open sets, and so $(f^{-1}[U_\alpha])$ is an open cover of $X$. By the compactness of $X$ it has a finite sub-cover. Clearly the image of that finite sub-cover is a finite sub-cover of $f(X)$ by $(U_\alpha)$. Hence $f(X)$ is compact. ■

From the above theorem we can deduce an important property of real valued continuous functions.

**5.7 Theorem (Extreme value theorem)** *Suppose that $X$ is a compact metric space and $f \in C(X, \mathbb{R})$. Then $f$ attains its maximum and minimum, that is, there exist $x_1, x_2 \in X$ such that $f(x_1) = \inf\{f(x) : x \in X\}$ and $f(x_2) = \sup\{f(x) : x \in X\}$.*

*Proof.* By Theorem 5.6 the image of $f$ is compact, and so by the Heine-Borel theorem (Theorem 4.5) closed and bounded. Hence the image $f(X) = \{f(x) : x \in X\}$ contain its infimum and supremum, that is, $x_1$ and $x_2$ as required exist. ■

Continuous functions on compact sets have other nice properties. To discuss these properties we introduce a stronger notion of continuity.

**5.8 Definition (Uniform continuity)** We say a function $f : X \to Y$ is uniformly continuous if for all $\varepsilon > 0$ there exists $\delta > 0$ such that $d_Y\big(f(x), f(y)\big) < \varepsilon$ for all $x, y \in X$ satisfying $d_X(x, y) < \delta$.

The difference to continuity is that $\delta$ does not depend on the point $x$, but can be chosen to be the same for *all* $x \in X$, that is *uniformly* with respect to $x \in X$. In fact, not all functions are uniformly continuous. For instance the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is not uniformly continuous. To see this note that

$$|f(x) - f(y)| = |x^2 - y^2| = |x + y||x - y|$$

for all $x, y \in \mathbb{R}$. Hence, no matter how small $|x - y|$ is, $|f(x) - f(y)|$ can be as big as we like by choosing $|x + y|$ large enough. However, the above also shows that $f$ is uniformly continuous on any *bounded* set of $\mathbb{R}$.

We next show that continuous functions on compact metric spaces are automatically uniformly continuous. The direct proof based on standard properties of continuous functions is taken from [6].

**5.9 Theorem (Uniform continuity)** *Let $X, Y$ be complete metric spaces. If $X$ is compact, then every function $f \in C(X, Y)$ is uniformly continuous.*

*Proof.* First note that the function $F : X \times X \to \mathbb{R}$ given by $F(x, y) := d_Y\big(f(x), f(y)\big)$ is continuous with respect to the product metric on $X \times X$. Fix $\varepsilon > 0$ and consider the inverse image

$$A_\varepsilon := F^{-1}\big[[\varepsilon, \infty)\big] := \big\{(x, y) \in X \times X : F(x, y) \geq \varepsilon\big\}.$$

As $F$ is continuous and $[\varepsilon, \infty)$ is closed, $A_\varepsilon$ is closed by Theorem 5.3. By Proposition 4.8 $X \times X$ is compact and so by Proposition 4.7 $A_\varepsilon$ is compact. Assume now that $A_\varepsilon \neq \emptyset$. Because the real valued function $(x, y) \mapsto d_X(x, y)$ is continuous on $X \times X$ it attains a minimum at some point $(x_0, y_0) \in A_\varepsilon$; see Theorem 5.7. In particular,

$$\delta := d_X(x_0, y_0) \leq d_X(x, y)$$

for all $(x, y) \in A_\varepsilon$. We have $\delta > 0$ as otherwise $x_0 = y_0$ and hence $(x_0, y_0) \notin A_\varepsilon$ by definition of $A_\varepsilon$. Furthermore, if $d_X(x, y) < \delta$, then $(x, y) \in A_\varepsilon^c$ and therefore

$$d_X(x, y) < \delta \quad \Rightarrow \quad d_Y\big(f(x), f(y)\big) = F(x, y) < \varepsilon. \tag{5.1}$$

This is exactly what is required for uniform continuity. If $A_\varepsilon = \emptyset$, then (5.1) holds for every $\delta > 0$. As the arguments work for every choice of $\varepsilon > 0$ this proves the uniform continuity of $f$. ∎

One could give an alternative proof of the above theorem using the covering property of compact sets, or a contradiction proof based on the sequential compactness.

Chapter II

# Banach Spaces

The purpose of this chapter is to introduce a class of vector spaces modelled on $\mathbb{R}^N$. Besides the algebraic properties we have a "norm" on $\mathbb{R}^N$ allowing us to measure distances between points. We generalise the concept of a norm to general vector spaces and prove some properties of these "normed spaces." We will see that all finite dimensional normed spaces are essentially $\mathbb{R}^N$ or $\mathbb{C}^N$. The situation becomes more complicated if the spaces are infinite dimensional. Functional analysis mainly deals with infinite dimensional vector spaces.

## 6 Normed Spaces

We consider a class of vector spaces with an additional topological structure. The underlying field is always $\mathbb{R}$ or $\mathbb{C}$. Most of the theory is developed simultaneously for vector spaces over the two fields. Throughout, $\mathbb{K}$ will be one of the two fields.

**6.1 Definition (Normed space)** Let $E$ be a vector space. A map $E \to \mathbb{R}$, $x \mapsto \|x\|$ is called a *norm on* $E$ if

(i) $\|x\| \geq 0$ for all $x \in E$ and $\|x\| = 0$ if and only if $x = 0$;

(ii) $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in E$ and $\alpha \in \mathbb{K}$;

(iii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in E$ (triangle inequality).

We call $(E, \|\cdot\|)$ or simply $E$ a *normed space*.

There is a useful consequence of the above definition.

**6.2 Proposition (Reversed triangle inequality)** *Let $(E, \|\cdot\|)$ be a normed space. Then, for all $x, y \in E$,*

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

*Proof.* By the triangle inequality $\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|$, so $\|x - y\| \geq \|x\| - \|y\|$. Interchanging the roles of $x$ and $y$ and applying (ii) we get $\|x - y\| = |-1|\|(-1)(x - y)\| = \|y - x\| \geq \|y\| - \|x\|$. Combining the two inequalities, the assertion of the proposition follows. ■

**6.3 Lemma** *Let* $(E, \|\cdot\|)$ *be a normed space and define*

$$d(x, y) := \|x - y\|$$

*for all* $x, y \in E$. *Then* $(E, d)$ *is a metric space.*

*Proof.* By (i) $d(x, y) = \|x - y\| \geq 0$ for all $x \in E$ and $d(x, y) = \|x - y| = 0$ if and only if $x - y = 0$, that is, $x = y$. By (ii) we have $d(x, y) = \|x - y\| = \|(-1)(y - x)\| = |-1|\|y - x\| = \|y - x\| = d(y, x)$ for all $x, y \in E$. Finally, for $x, y, z$ it follows from (iii) that $d(x, y) = \|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y)$, proving that $d(\cdot\,,\cdot)$ satisfies the axioms of a metric (see Definition 2.1). ■

We will always equip a normed space with the topology of $E$ generated by the metric induced by the norm. Hence it makes sense to talk about continuity of functions. It turns out the that topology is compatible with the vector space structure as the following theorem shows.

**6.4 Theorem** *Given a normed space* $(E, \|\cdot\|)$*, the following maps are continuous (with respect to the product topologies).*

(1) $E \to \mathbb{R}$, $x \mapsto \|x\|$ *(continuity of the norm);*

(2) $E \times E \to E$, $(x, y) \mapsto x + y$ *(continuity of addition);*

(3) $\mathbb{K} \times E \to E$, $(\alpha, x) \mapsto \alpha x$ *(continuity of multiplication by scalars).*

*Proof.* (1) By the reversed triangle inequality $\big|\|x\| - \|y\|\big| \leq \|x - y\|$, implying that $\|y\| \to \|y\|$ as $x \to y$ (that is, $d(x, y) = \|x - y\| \to 0$). Hence the norm is continuous as a map from $E$ to $\mathbb{R}$.

(2) If $x, y, a, b \in E$ then $\|(x+y)-(a+b)\| = \|(x-a)+(y-b)\| \leq \|x-a\|+\|y-b\| \to 0$ as $x \to a$ and $y \to b$, showing that $x + y \to a + b$ as $x \to a$ and $y \to b$. This proves the continuity of the addition.

(3) For $\alpha, \xi \in \mathbb{K}$ and $a, x \in E$ we have, by using the properties of a norm,

$$\begin{aligned}\|\xi x - \alpha a\| &= \|\xi(x - a) + (\xi - \alpha)a\| \\ &\leq \|\xi(x - a)\| + \|(\xi - \alpha)x\| = |\xi|\|x - a\| + |\xi - \alpha\|x\|.\end{aligned}$$

As the last expression goes to zero as $x \to a$ in $E$ and $\xi \to \alpha$ in $\mathbb{K}$ we have also proved the continuity of the multiplication by scalars. ■

When looking at metric spaces we discussed a rather important class of metric spaces, namely complete spaces. Similarly, complete spaces play a special role in functional analysis.

**6.5 Definition (Banach space)** A normed space which is complete with respect to the metric induced by the norm is called a *Banach space*.

**6.6 Example** The simplest example of a Banach space is $\mathbb{R}^N$ or $\mathbb{C}^N$ with the Euclidean norm.

We give a characterisation of Banach spaces in terms of properties of series. We recall the following definition.

**6.7 Definition (absolute convergence)** A series $\sum_{k=0}^{\infty} a_k$ in $E$ is called *absolutely convergent* if $\sum_{k=0}^{\infty} \|a_k\|$ converges.

**6.8 Theorem** *A normed space $E$ is complete if and only every absolutely convergent series in $E$ converges.*

*Proof.* Suppose that $E$ is complete. Let $\sum_{k=0}^{\infty} a_k$ an absolutely convergent series in $E$, that is, $\sum_{k=0}^{\infty} \|a_k\|$ converges. By the Cauchy criterion for the convergence of a series in $\mathbb{R}$, for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that

$$\sum_{k=m+1}^{n} \|a_k\| < \varepsilon$$

for all $n > m > n_0$. (This simply means that the sequence of partial sums $\sum_{k=0}^{n} \|a_k\|$, $n \in \mathbb{N}$ is a Cauchy sequence.) Therefore, by the triangle inequality

$$\Big\|\sum_{k=0}^{n} a_k - \sum_{k=0}^{m} a_k\Big\| = \Big\|\sum_{k=m+1}^{n} a_k\Big\| \leq \sum_{k=m+1}^{n} \|a_k\| < \varepsilon$$

for all $n > m > n_0$. Hence the sequence of partial sums $\sum_{k=0}^{n} a_k$, $n \in \mathbb{N}$ is a Cauchy sequence in $E$. Since $E$ is complete,

$$\sum_{k=0}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=0}^{n} a_k$$

exists. Hence if $E$ is complete, every absolutely convergent series converges in $E$. Next assume that every absolutely convergent series converges. We have to show that every Cauchy sequence $(x_n)$ in $E$ converges. By Theorem 3.9 it is sufficient to show that $(x_n)$ has a convergent subsequence. Since $(x_n)$ is a Cauchy sequence, for every $k \in \mathbb{N}$ there exists $m_k \in \mathbb{N}$ such that $\|x_n - x_m\| < 2^{-k}$ for all $n, m > m_k$. Now set $n_1 := m_1 + 1$. Then inductively choose $n_k$ such that $n_{k+1} > n_k > m_k$ for all $k \in \mathbb{N}$. Then by the above

$$\|x_{n_{k+1}} - x_{n_k}\| < \frac{1}{2^k}$$

for all $k \in \mathbb{N}$. Now observe that

$$x_{n_k} - x_{n_1} = \sum_{j=1}^{k-1} (x_{n_{j+1}} - x_{n_j})$$

for all $k \in \mathbb{N}$. Hence $(x_{n_k})$ converges if and only the series $\sum_{j=1}^{\infty} (x_{n_{j+1}} - x_{n_j})$ converges. By choice of $n_k$ we have

$$\sum_{j=1}^{\infty} \|x_{n_{j+1}} - x_{n_j}\| \leq \sum_{j=1}^{\infty} \frac{1}{2^j} = 1 < \infty.$$

Hence $\sum_{j=1}^{\infty} (x_{n_{j+1}} - x_{n_j})$ is absolutely convergent. By assumption every absolutely convergent series converges and therefore $(x_{n_k})$ converges, completing the proof of the theorem. ∎

# 7 Examples of Banach Spaces

In this section we give examples of Banach spaces. They will be used throughout the course. We start by elementary inequalities and a family of norms in $\mathbb{K}^N$. They serve as a model for more general spaces of sequences or functions.

## 7.1 Elementary Inequalities

In this section we discuss inequalities arising when looking at a family of norms on $\mathbb{K}^N$. For $1 \leq p \leq \infty$ we define the $p$-norms of $x := (x_1, \dots, x_N) \in \mathbb{R}^N$ by

$$|x|_p := \begin{cases} \Big(\sum_{i=1}^{N} |x_i|^p\Big)^{1/p} & \text{if } 1 \leq p < \infty, \\ \max_{i=1,\dots,N} |x_i| & \text{if } p = \infty. \end{cases} \tag{7.1}$$

At this stage we do not know whether $|\cdot|_p$ is a norm. We now prove that the $p$-norms are norms, and derive some relationships between them. First we need *Young's inequality*.

**7.1 Lemma (Young's inequality)** *Let $p, p' \in (1, \infty)$ such that*

$$\frac{1}{p} + \frac{1}{p'} = 1. \tag{7.2}$$

*Then*

$$ab \leq \frac{1}{p}a^p + \frac{1}{p'}b^{p'}$$

*for all $a, b \geq 0$.*

*Proof.* The inequality is obvious if $a = 0$ or $b = 0$, so we assume that $a, b > 0$. As ln is concave we get from (7.2) that

$$\ln(ab) = \ln a + \ln b = \frac{1}{p}\ln a^p + \frac{1}{p'}\ln b^{p'} \leq \ln\Big(\frac{1}{p}a^p + \frac{1}{p'}b^{p'}\Big).$$

Hence, as exp is increasing we have

$$ab = \exp(\ln(ab)) \leq \exp\Big(\ln\Big(\frac{1}{p}a^p + \frac{1}{p'}b^{p'}\Big)\Big) = \frac{1}{p}a^p + \frac{1}{p'}b^{p'},$$

proving our claim. ∎

The relationship (7.2) is rather important and appears very often. We say that $p'$ is the *exponent dual to p*. If $p = 1$ we set $p' := \infty$ and if $p = \infty$ we set $p' := 1$.

From Young's inequality we get *Hölder's inequality*.

**7.2 Proposition** *Let $1 \leq p \leq \infty$ and $p'$ the exponent dual to $p$. Then for all $x, y \in \mathbb{K}^N$*

$$\sum_{i=1}^{N} |x_i||y_i| \leq |x|_p |y|_{p'}.$$

*Proof.* If $x = 0$ or $y = 0$ then the inequality is obvious. Also if $p = 1$ or $p = \infty$ then the inequality is also rather obvious. Hence assume that $x, y \neq 0$ and $1 < p < \infty$. By Young's inequality (Lemma 7.1) we have

$$\begin{aligned}
\frac{1}{|x|_p |y|_{p'}} \sum_{i=1}^{N} |x_i||y_i| &= \sum_{i=1}^{N} \frac{|x_i|}{|x|_p} \frac{|y_i|}{|y|_{p'}} \leq \sum_{i=1}^{N} \Big( \frac{1}{p} \Big( \frac{|x_i|}{|x|_p} \Big)^p + \frac{1}{p'} \Big( \frac{|y_i|}{|y|_{p'}} \Big)^{p'} \Big) \\
&= \frac{1}{p} \sum_{i=1}^{N} \Big( \frac{|x_i|}{|x|_p} \Big)^p + \frac{1}{p'} \sum_{i=1}^{N} \Big( \frac{|y_i|}{|y|_{p'}} \Big)^{p'} \\
&= \frac{1}{p} \frac{1}{|x|_p^p} \sum_{i=1}^{N} |x_i|^p + \frac{1}{p'} \frac{1}{|y|_{p'}^{p'}} \sum_{i=1}^{N} |y_i|^{p'} \\
&= \frac{1}{p} \frac{|x|_p^p}{|x|_p^p} + \frac{1}{p'} \frac{|y|_{p'}^{p'}}{|y|_{p'}^{p'}} = \frac{1}{p} + \frac{1}{p'} = 1,
\end{aligned}$$

from which the required inequality readily follows. ■

Now we prove the main properties of the $p$-norms.

**7.3 Theorem** *Let $1 \leq p \leq \infty$. Then $|\cdot|_p$ is a norm on $\mathbb{K}^N$. Moreover, if $1 \leq p \leq q \leq \infty$ then*

$$|x|_q \leq |x|_p \leq N^{\frac{q-p}{pq}} |x|_q \tag{7.3}$$

*for all $x \in \mathbb{K}^N$. (We set $(q-p)/pq := 1/p$ if $q = \infty$.) Finally, the above inequalities are optimal.*

*Proof.* We first prove that $|\cdot|_p$ is a norm. The cases $p = 1, \infty$ are easy, and left to the reader. Hence assume that $1 < p < \infty$. By definition $|x|_p \geq 0$ and $|x|_p = 0$ if and only if $|x_i| = 0$ for all $i = 1, \dots, N$, that is, if $x = 0$. Also, if $\alpha \in \mathbb{K}$, then

$$|\alpha x|_p = \Big( \sum_{i=1}^{N} |\alpha x_i|^p \Big)^{1/p} = \Big( \sum_{i=1}^{N} |\alpha|^p |x_i|^p \Big)^{1/p} = |\alpha||\alpha x|_p.$$

Thus it remains to prove the triangle inequality. For $x, y \in \mathbb{K}^N$ we have, using Hölder's inequality (Proposition 7.2), that

$$\begin{aligned}
|x+y|_p^p &= \sum_{i=1}^{N} |x_i + y_i|^p = \sum_{i=1}^{N} |x_i + y_i||x_i + y_i|^{p-1} \\
&\leq \sum_{i=1}^{N} |x_i||x_i + y_i|^{p-1} + \sum_{i=1}^{N} |y_i||x_i + y_i|^{p-1} \\
&\leq |x|_p \Big( \sum_{i=1}^{N} |x_i + y_i|^{(p-1)p'} \Big)^{1/p'} + |y|_p \Big( \sum_{i=1}^{N} |x_i + y_i|^{(p-1)p'} \Big)^{1/p'} \\
&= \big( |x|_p + |y|_p \big) \Big( \sum_{i=1}^{N} |x_i + y_i|^{(p-1)p'} \Big)^{1/p'}.
\end{aligned}$$

Now, observe that
$$p' = \left(1 - \frac{1}{p}\right)^{-1} = \frac{p}{p-1},$$
so we get from the above that
$$|x+y|_p^p \leq \left(|x|_p + |y|_p\right)\left(\sum_{i=1}^N |x_i + y_i|^p\right)^{(p-1)/p} = \left(|x|_p + |y|_p\right)|x+y|_p^{p-1}.$$
Hence if $x + y \neq 0$ we get the triangle inequality $|x+y|_p \leq |x|_p + |y|_p$. The inequality is obvious if $x + y = 0$, so $|\cdot|_p$ is a norm on $\mathbb{K}^N$.

Next we show the first inequality in (7.3). First let $p < q = \infty$. If $x \in \mathbb{K}^N$ we pick the component $x_j$ of $x$ such that $|x_j| = |x|_\infty$. Hence $|x|_\infty = |x_j| = \left(|x_j|^p\right)^{1/p} \leq |x|_p$, proving the first inequality in case $q = \infty$. Assume now that $1 \leq p \leq q < \infty$. If $x \neq 0$ then $|x_i|/|x|_p \leq 1$ Hence, as $1 \leq p \leq q < \infty$ we have
$$\left(\frac{|x_i|}{|x|_p}\right)^q \leq \left(\frac{|x_i|}{|x|_p}\right)^p$$
for all $x \in \mathbb{K}^N \setminus \{0\}$. Therefore,
$$\frac{|x|_q^q}{|x|_p^q} = \sum_{i=1}^N \left(\frac{|x_i|}{|x|_p}\right)^q \leq \sum_{i=1}^N \left(\frac{|x_i|}{|x|_p}\right)^p = \frac{|x|_p^p}{|x|_p^p} = 1$$
for all $x \in \mathbb{K}^N \setminus \{0\}$. Hence $|x|_q \leq |x|_p$ for all $x \neq 0$. For $x = 0$ the inequality is trivial. To prove the second inequality in (7.3) assume that $1 \leq p < q < \infty$. We define $s := q/p$. The corresponding dual exponent $s'$ is given by
$$s' = \frac{s}{s-1} = \frac{q}{q-p}.$$
Applying Hölder's inequality we get
$$|x|_p^p = \sum_{i=1}^N |x_i|^p \cdot 1 \leq \left(\sum_{i=1}^N |x_i|^{ps}\right)^{1/s}\left(\sum_{i=1}^N 1^{s'}\right)^{1/s'} = N^{\frac{q-p}{q}}\left(\sum_{i=1}^N |x_i|^q\right)^{p/q} = N^{\frac{q-p}{q}}|x|_q^p$$
for all $x \in \mathbb{K}^N$, from which the second inequality in (7.3) follows. If $1 \leq p < q = \infty$ and $x \in \mathbb{K}^N$ is given we pick $x_j$ such that $|x_j| = |x|_\infty$. Then
$$|x|_p = \left(\sum_{i=1}^N |x_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^N |x_j|^p\right)^{1/p} = N^{1/p}|x_j| = N^{1/p}|x|_\infty,$$
covering the last case.

We finally show that (7.3) is optimal. For the first inequality look at the standard basis of $\mathbb{K}^N$, for which we have equality. For the second choose $x = (1, 1, \dots, 1)$ and observe that $|x|_p = N^{1/p}$. Hence
$$N^{\frac{q-p}{pq}}|x|_q = N^{\frac{q-p}{pq}} N^{\frac{1}{q}} = N^{\frac{1}{p}} = |x|_p.$$
Hence we cannot decrease the constant $N^{\frac{q-p}{pq}}$ in the inequality. ∎

## 7.2 Spaces of Sequences

Here we discuss spaces of sequences. As most of you have seen this in *Metric Spaces* the exposition will be rather brief.

Denote by $\mathcal{S}$ the space of all sequences in $\mathbb{K}$, that is, the space of all functions from $\mathbb{N}$ into $\mathbb{K}$. We denote its elements by $x = (x_0, x_1, x_2, \dots) = (x_i)$. We define vector space operations "component" wise:

- $(x_i) + (y_i) := (x_i + y_i)$ for all $(x_i), (y_i) \in \mathcal{S}$;
- $\alpha(x_i) := (\alpha x_i)$ for all $\alpha \in \mathbb{K}$ and $(x_i) \in \mathcal{S}$.

With these operations $\mathcal{S}$ becomes a vector space. Given $(x_i) \in \mathcal{S}$ and $1 \leq p \leq \infty$ we define the "$p$-norms"

$$|(x_i)|_p := \begin{cases} \Big(\displaystyle\sum_{i=1}^{\infty} |x_i|^p\Big)^{1/p} & \text{if } 1 \leq p < \infty, \\ \sup_{i\in\mathbb{N}} |x_i| & \text{if } p = \infty. \end{cases} \tag{7.4}$$

These $p$-norms are not finite for all sequences. We define some subspaces of $\mathcal{S}$ in the following way:

- $\ell_p := \ell_p(\mathbb{K}) := \big\{x \in \mathcal{S} : |x|_p < \infty\big\}$ $(1 \leq p \leq \infty)$;
- $c_0 := c_0(\mathbb{K}) := \big\{(x_i) \in \mathcal{S} : \lim_{i\to\infty} |x_i| = 0\big\}$.

The $p$-norms for sequences have similar properties as the $p$-norms in $\mathbb{K}^N$. In fact most properties follow from the finite version given in Section 7.1.

**7.4 Proposition (Hölder's inequality)** *For* $1 \leq p \leq \infty$, $x \in \ell_p$ *and* $y \in \ell_{p'}$ *we have*

$$\sum_{i=1}^{\infty} |x_i||y_i| \leq |x|_p |y|_{p'},$$

*where* $p'$ *is the exponent dual to* $p$ *defined by* (7.2).

*Proof.* Apply Proposition 7.2 to partial sums and then pass to the limit. ■

**7.5 Theorem** *Let* $1 \leq p \leq \infty$. *Then* $(\ell_p, |\cdot|_p)$ *and* $(c_0, , |\cdot|_\infty)$ *are Banach spaces. Moreover, if* $1 < p < q < \infty$ *then*

$$\ell_1 \subsetneq \ell_p \subsetneq \ell_q \subsetneq c_0 \subsetneq \ell_\infty.$$

*Finally,* $|x|_q \leq |x|_p$ *for all* $x \in \ell_p$ *if* $1 \leq p < q \leq \infty$.

*Proof.* It readily follows that $|\cdot|_p$ is a norm by passing to the limit from the finite dimensional case in Theorem 7.3. In particular it follows that $\ell_p, c_0$ are subspaces of

$\mathcal{S}$. If $|(x_i)|_p < \infty$ for some $(x_i) \in \mathcal{S}$ and $p < \infty$ then we must have $|x_i| \to 0$. Hence $\ell_p \subset c_0$ for all $1 \leq p < \infty$. Clearly $c_0 \subsetneq \ell_\infty$. If $1 \leq p < q < \infty$ then by Theorem 7.3

$$\sup_{i=1,\dots,n} |x_i| \leq \Big(\sum_{i=1}^n |x_i|^q\Big)^{1/q} \leq \Big(\sum_{i=1}^n |x_i|^p\Big)^{1/p}.$$

Passing to the limit on the right hand side and then taking the supremum on the left hand side we get

$$|(x_i)|_\infty \leq |(x_i)|_q \leq |(x_i)|_p,$$

proving the inclusions and the inequalities. To show that the inclusions are proper we use the harmonic series

$$\sum_{i=1}^\infty \frac{1}{i} = \infty.$$

Clearly $(1/i) \in c_0$ but not in $\ell_1$. Similarly, $(1/i^{1/p}) \in \ell_q$ for $q > p$ but not in $\ell_p$.

We finally prove completeness. Suppose that $(x_n)$ is a Cauchy sequence in $\ell_p$. Then by definition of the $p$-norm

$$|x_{in} - x_{im}| \leq |x_n - x_m|_p$$

for all $i, m, n \in \mathbb{N}$. It follows that $(x_{in})$ is a Cauchy sequence in $\mathbb{K}$ for every $i \in \mathbb{N}$. Since $\mathbb{K}$ is complete

$$x_i := \lim_{n\to\infty} x_{in} \tag{7.5}$$

exists for all $i \in \mathbb{N}$. We set $x := (x_i)$. We need to show that $x_n \to x$ in $\ell_p$, that is, with respect to the $\ell_p$-norm. Let $\varepsilon > 0$ be given. By assumption there exists $n_0 \in \mathbb{N}$ such that for all $n, m > n_0$

$$\Big(\sum_{i=1}^N |x_{in} - x_{im}|^p\Big)^{1/p} \leq |x_n - x_m|_p < \frac{\varepsilon}{2}$$

if $1 \leq p < \infty$ and

$$\max_{i=1\dots,N} |x_{in} - x_{im}| \leq |x_n - x_m|_\infty < \frac{\varepsilon}{2}$$

if $p = \infty$. For fixed $N \in \mathbb{N}$ we can let $m \to \infty$, so by (7.5) and the continuity of the absolute value, for all $N \in \mathbb{N}$ and $n > n_0$

$$\Big(\sum_{i=1}^N |x_{in} - x_i|^p\Big)^{1/p} \leq \frac{\varepsilon}{2}$$

if $1 \leq p < \infty$ and

$$\max_{i=1\dots,N} |x_{in} - x_{im}| \leq |x_n - x|_\infty \leq \frac{\varepsilon}{2}$$

if $p = \infty$. Letting $N \to \infty$ we finally get

$$|x_n - x|_p \leq \frac{\varepsilon}{2} < \varepsilon$$

for all $n > n_0$. Since the above works for every $\varepsilon > 0$ it follows that $x_n - x \to 0$ in $\ell_p$ as $n \to \infty$. Finally, as $\ell_p$ is a vector space and $x_{n_0}, x_{n_0} - x \in \ell_p$ we have

$x = x_{n_0} - (x_{n_0} - x) \in \ell_p$. Hence $x_n \to x$ in $\ell_p$, showing that $\ell_p$ is complete for $1 \leq p \leq \infty$. To show that $c_0$ is complete we need to show that $x \in c_0$ if $x_n \in c_0$ for all $n \in \mathbb{N}$. We know that $x_n \to x$ in $\ell_\infty$. Hence, given $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $|x_n - x|_\infty < \varepsilon/2$ for all $n \geq n_0$. Therefore

$$|x_i| \leq |x_i - x_{in_0}| + |x_{in_0}| \leq |x - x_{n_0}|_\infty + |x_{in_0}| < \frac{\varepsilon}{2} + |x_{in_0}|$$

for all $i \in \mathbb{N}$. Since $x_{n_0} \in c_0$ there exists $i_0 \in \mathbb{N}$ such that $|x_{in_0}| < \varepsilon/2$ for all $i > i_0$. Hence $|x_i| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$ for all $i > i_0$, so $x \in c_0$ as claimed. This completes the proof of completeness of $\ell_p$ and $c_0$. ■

**7.6 Remark** The proof of completeness in many cases follows similar steps as the one above.

(1) Take a an arbitrary Cauchy sequence $(x_n)$ in the normed space $E$ and show that $(x_n)$ converges not in the norm of $E$, but in some weaker sense to some $x$. (In the above proof it is, "component-wise" convergence, that is $x_{in}$ converges for each $i \in \mathbb{N}$ to some $x_i$.);

(2) Show that $\|x_n - x\|_E \to 0$;

(3) Show that $x \in E$ by using that $E$ is a vector space.

## 7.3 Lebesgue Spaces

The Lebesgue or simply $L_p$-spaces are familiar to all those who have taken "Lebesgue Integration and Fourier Analysis." We give an outline on the main properties of these spaces. They are some sort of continuous version of the $\ell_p$-spaces.

Suppose that $X \subset \mathbb{R}^N$ is an open (or simply measurable) set. If $u\colon X \to \mathbb{K}$ is a measurable function we set

$$\|u\|_p := \begin{cases} \Bigl(\displaystyle\int_X |u(x)|^p \,dx\Bigr)^{1/p} & \text{if } 1 \leq p < \infty, \\ \operatorname*{ess\,sup}_{x\in X} |u(x)| & \text{if } p = \infty. \end{cases} \tag{7.6}$$

We then let $L_p(X, \mathbb{K}) = L_p(X)$ be the space of all measurable functions on $X$ for which $\|u\|_p$ is finite. Two such functions are equal if they are equal almost everywhere.

**7.7 Proposition (Hölder's inequality)** *Let $1 \leq p \leq \infty$. If $u \in L_p(X)$ and $v \in L_{p'}(X)$ then*

$$\int_X |u||v| \,dx \leq \|u\|_p \|v\|_{p'}.$$

Moreover there are the following facts.

**7.8 Theorem** *Let $1 \leq p \leq \infty$. Then $(L_p(X), \|\cdot\|_p)$ is a Banach space. If $X$ has finite measure then*

$$L_\infty(X) \subset L_q(X) \subset L_p(X) \subset L_1(X)$$

*if $1 < p < q < \infty$. If $X$ has infinite measure there are no inclusions between $L_p(X)$ and $L_q(X)$ for $p \neq q$.*

## 7.4 Spaces of Bounded and Continuous Functions

Suppose that $X$ is a set and $E = (E, \|\cdot\|)$ a normed space. For a function $u\colon X \to E$ we let

$$\|u\|_\infty := \sup_{x\in X} \|u(x)\|$$

We define the space of bounded functions by

$$B(X, E) := \{u\colon X \to E \mid \|u\|_\infty < \infty\}.$$

This space turns out to be a Banach space if $E$ is a Banach space. Completeness of $B(X, E)$ is equivalent to the fact that the uniform limit of a bounded sequence of functions is bounded. The proof follows the steps outlined in Remark 7.6.

**7.9 Theorem** *If $X$ is a set and $E$ a Banach space, then $B(X, E)$ is a Banach space with the supremum norm.*

*Proof.* Let $(u_n)$ be a Cauchy sequence in $B(X, E)$. As

$$\|u_n(x) - u_m(x)\|_E \leq \|u_n - u_m\|_\infty \tag{7.7}$$

for all $x \in X$ and $m, n \in \mathbb{N}$ it follows that $\big(u_n(x)\big)$ is a Cauchy sequence in $E$ for every $x \in X$. Since $E$ is complete

$$u(x) := \lim_{n\to\infty} u_n(x)$$

exists for all $x \in X$. We need to show that $u_n \to u$ in $B(X, E)$, that is, with respect to the supremum norm. Let $\varepsilon > 0$ be given. Then by assumption there exists $n_0 \in \mathbb{N}$ such that $\|u_n - u_m\|_\infty < \varepsilon/2$ for all $m, n > n_0$. Using (7.7) we get

$$\|u_n(x) - u_m(x)\| < \frac{\varepsilon}{2}$$

for all $m, n > n_0$ and $x \in X$. For fixed $x \in X$ we can let $m \to \infty$, so by the continuity of the norm

$$\|u_n(x) - u(x)\| < \frac{\varepsilon}{2}$$

for all $x \in X$ and all $n > n_0$. Hence $\|u_n - u\|_\infty \leq \varepsilon/2 < \varepsilon$ for all $n > n_0$. Since the above works for every $\varepsilon > 0$ it follows that $u_n - u \to 0$ in $E$ as $n \to \infty$. Finally, as $B(X, E)$ is a vector space and $u_{n_0}, u_{n_0} - u \in B(X, E)$ we have $u = u_{n_0} - (u_{n_0} - u) \in B(X, E)$. Hence $u_n \to u$ in $B(X, E)$, showing that $B(X, E)$ is complete. ■

If $X$ is a metric space we denote the vector space of all continuous functions by $C(X, E)$. This space does not carry a topology or norm in general. However,

$$BC(X, E) := B(X, E) \cap C(X, E)$$

becomes a normed space with norm $\|\cdot\|_\infty$. Note that if $X$ is compact then $C(X, E) = BC(X, E)$. The space $BC(X, E)$ turns out to be a Banach space if $E$ is a Banach space. Note that the completeness of $BC(X, E)$ is equivalent to the fact that the uniform limit of continuous functions is continuous. Hence the language of functional analysis provides a way to rephrase standard facts from analysis in a concise and unified way.

**7.10 Theorem** *If $X$ is a metric space and $E$ a Banach space, then $BC(X, E)$ is a Banach space with the supremum norm.*

*Proof.* Let $u_n$ be a Cauchy sequence in $BC(X, E)$. Then it is a Cauchy sequence in $B(X, E)$. By completeness of that space $u_n \to u$ in $B(X, E)$, so we only need to show that $u$ is continuous. Fix $x_0 \in X$ arbitrary. We show that $f$ is continuous at $x_0$. Clearly

$$\|u(x) - u(x_0)\|_E \le \|u(x) - u_n(x)\|_E + \|u_n(x) - u_n(x_0)\|_E + \|u_n(x_0) - u(x_0)\|_E \quad (7.8)$$

for all $x \in X$ and $n \in \mathbb{N}$. Fix now $\varepsilon > 0$ arbitrary. Since $u_n \to u$ in $B(X, E)$ there exists $n_0 \in \mathbb{N}$ such that $\|u_n(x) - u(x)\|_E < \varepsilon/4$ for all $n > n_0$ and $x \in X$. Hence (7.8) implies that

$$\|u(x) - u(x_0)\|_E < \frac{\varepsilon}{2} + \|u_n(x) - u_n(x_0)\|_E. \quad (7.9)$$

Since $u_{n_0+1}$ is continuous at $x_0$ there exists $\delta > 0$ such that $\|u_{n_0+1}(x) - u_{n_0+1}(x_0)\|_E < \varepsilon/2$ for all $x \in X$ with $d(x, x_0) < \delta$ and $n > n_0$. Using (7.9) we get $\|u(x) - u(x_0)\|_E < \varepsilon/2 + \varepsilon/2$ if $d(x, x_0) < \delta$ and so $u$ is continuous at $x_0$. As $x_0$ was arbitrary, $u \in C(X, E)$ as claimed. ■

Note that the above is a functional analytic reformulation of the fact that a uniformly convergent sequence of bounded functions is bounded, and similarly that a uniformly convergent sequence of continuous functions is continuous.

# 8 Basic Properties of Bounded Linear Operators

One important aim of functional analysis is to gain a deep understanding of properties of linear operators. We start with some definitions.

**8.1 Definition (bounded sets)** A subset $U$ of a normed space $E$ is called bounded if there exists $M > 0$ such that $U \subset B(0, M)$.

We next define some classes of linear operators.

**8.2 Definition** Let $E, F$ be two normed spaces.

(a) We denote by $\operatorname{Hom}(E, F)$ the set of all linear operators from $E$ to $F$. ("Hom" because linear operators are *homomorphisms* between vector spaces.) We also set $\operatorname{Hom}(E) := \operatorname{Hom}(E, E)$.

(b) We set

$$\mathcal{L}(E, F) := \{T \in \operatorname{Hom}(E, F) : T \text{ continuous}\}$$

and $\mathcal{L}(E) := \mathcal{L}(E, E)$.

(c) We call $T \in \operatorname{Hom}(E, F)$ *bounded* if $T$ maps every bounded subset of $E$ onto a bounded subset of $F$.

In the following theorem we collect the main properties of continuous and bounded linear operators. In particular we show that a linear operator is bounded if and only if it is continuous.

**8.3 Theorem** *For $T \in \operatorname{Hom}(E, F)$ the following statements are equivalent:*

(i) *$T$ is uniformly continuous;*

(ii) *$T \in \mathcal{L}(E, F)$;*

(iii) *$T$ is continuous at $x = 0$;*

(iv) *$T$ is bounded;*

(v) *There exists $\alpha > 0$ such that $\|Tx\|_F \leq \alpha \|x\|_E$ for all $x \in E$.*

*Proof.* The implications (i)⇒(ii) and (ii)⇒(iii) are obvious. Suppose now that $T$ is continuous at $x = 0$ and that $U$ is an arbitrary bounded subset of $E$. As $T$ is continuous at $x = 0$ there exists $\delta > 0$ such that $\|Tx\|_F \leq 1$ whenever $\|x\|_E \leq \delta$. Since $U$ is bounded $M := \sup_{x \in U} \|x\|_E < \infty$, so for every $x \in U$

$$\Big\| \frac{\delta}{M} x \Big\|_E = \frac{\delta}{M} \|x\|_E \leq \delta.$$

Hence by the linearity of $T$ and the choice of $\delta$

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

showing that

$$\|Tx\|_F \leq \frac{M}{\delta}$$

for all $x \in U$. Therefore, the image of $U$ under $T$ is bounded, showing that (iii) implies (iv). Suppose now that $T$ is bounded. Then there exists $\alpha > 0$ such that $\|Tx\|_F \leq \alpha$ whenever $\|x\|_E \leq 1$. Hence, using the linearity of $T$

$$\frac{1}{\|x\|_E} \|Tx\|_F = \Big\| T\Big(\frac{x}{\|x\|_E}\Big) \Big\|_F \leq \alpha$$

for all $x \in E$ with $x \neq 0$. Since $T0 = 0$ it follows that $\|Tx\|_F \leq \alpha \|x\|_E$ for all $x \in E$. Hence (iv) implies (v). Suppose now that there exists $\alpha > 0$ such that $\|Tx\|_F \leq \alpha \|x\|_E$ for all $x \in E$. Then by the linearity of $T$

$$\|Tx - Ty\|_F = \|T(x - y)\|_F \leq \alpha \|x - y\|_E,$$

showing the uniform continuity of $T$. Hence (v) implies (i), completing the proof of the theorem. ∎

Obviously, $\mathcal{L}(E, F)$ is a vectors space. We will show that it is a *normed* space if we define an appropriate norm.

**8.4 Definition (operator norm)** For $T \in \mathcal{L}(E, F)$ we define

$$\|T\|_{\mathcal{L}(E,F)} := \inf\{\alpha > 0 \colon \|Tx\|_F \leq \alpha \|x\|_E \text{ for all } x \in E\}$$

We call $\|T\|_{\mathcal{L}(E,F)}$ the *operator norm* of $E$.

**8.5 Remark** We could define $\|T\|_{\mathcal{L}(E,F)}$ for all $T \in \mathrm{Hom}(E,F)$, but Theorem 8.3 shows that $T \in \mathcal{L}(E,F)$ if and only if $\|T\|_{\mathcal{L}(E,F)} < \infty$.

Before proving that the operator norm is in fact a norm, we first give other characterisations.

**8.6 Proposition** *Suppose that* $T \in \mathcal{L}(E,F)$. *Then* $\|Tx\|_F \leq \|T\|_{\mathcal{L}(E,F)}\|x\|_E$ *for all* $x \in E$. *Moreover,*

$$\|T\|_{\mathcal{L}(E,F)} = \sup_{x\in E\setminus\{0\}} \frac{\|Tx\|_F}{\|x\|_E} = \sup_{\|x\|_E=1} \|Tx\|_F = \sup_{\|x\|_E<1} \|Tx\|_F = \sup_{\|x\|_E\leq 1} \|Tx\|_F. \tag{8.1}$$

*Proof.* Fix $T \in \mathcal{L}(E,F)$ and set $A := \{\alpha > 0 \colon \|Tx\|_F \leq \alpha\|x\|_E \text{ for all } x \in E\}$. By definition $\|T\|_{\mathcal{L}(E,F)} = \inf A$. If $\alpha \in A$, then $\|Tx\|_F \leq \alpha\|x\|_E$ for all $x \in E$. Hence, for every $x \in E$ we have $\|Tx\|_F \leq (\inf A)\|x\|_E$, proving the first claim. Set now

$$\lambda := \sup_{x\in E\setminus\{0\}} \frac{\|Tx\|_F}{\|x\|_E}.$$

Then, $\|Tx\|_F \leq \lambda\|x\|_E$ for all $x \in E$, and so $\lambda \geq \inf A = \|T\|_{\mathcal{L}(E,F)}$. By the above we have $\|Tx\|_F \leq \|T\|_{\mathcal{L}(E,F)}\|x\|_E$ and thus

$$\frac{\|Tx\|_F}{\|x\|_E} \leq \|T\|_{\mathcal{L}(E,F)}$$

for all $x \in E \setminus \{0\}$, implying that $\lambda \leq \|T\|_{\mathcal{L}(E,F)}$. Combining the inequalities $\lambda = \|T\|_{\mathcal{L}(E,F)}$, proving the first equality in (8.1). Now by the linearity of $T$

$$\sup_{x\in E\setminus\{0\}} \frac{\|Tx\|_F}{\|x\|_E} = \sup_{x\in E\setminus\{0\}} \Big\|T\frac{x}{\|x\|_E}\Big\|_F = \sup_{\|x\|_E=1} \|Tx\|_F,$$

proving the second equality in (8.1). To prove the third equality note that

$$\beta := \sup_{\|x\|_E<1} \|Tx\|_F \leq \sup_{\|x\|_E<1} \frac{\|Tx\|_F}{\|x\|_E} \leq \sup_{x\in E\setminus\{0\}} \frac{\|Tx\|_F}{\|x\|_E} = \lambda.$$

On the other hand, we have for every $x \in E$ and $\varepsilon > 0$

$$\Big\|T\frac{x}{\|x\|_E + \varepsilon}\Big\|_F \leq \beta$$

and thus $\|Tx\|_F \leq \beta(\|x\|_E + \varepsilon)$ for all $\varepsilon > 0$ and $x \in E$. Hence $\|Tx\|_F \leq \beta\|x\|_E$ for all $x \in E$, implying that $\beta \geq \lambda$. Combining the inequalities $\beta = \lambda$, which is the third inequality in (8.1). For the last equality note that $\|Tx\|_F \leq \|T\|_{\mathcal{L}(E,F)}\|x\|_E \leq \|T\|_{\mathcal{L}(E,F)}$ whenever $\|x\|_E \leq 1$. Hence

$$\sup_{\|x\|_E\leq 1} \|Tx\|_F \leq \|T\|_{\mathcal{L}(E,F)} = \sup_{\|x\|_E=1} \|Tx\|_F \leq \sup_{\|x\|_E\leq 1} \|Tx\|_F,$$

implying the last inequality. ∎

We next show that $\|\cdot\|_{\mathcal{L}(E,F)}$ is a norm.

**8.7 Proposition** *The space* $\big(\mathcal{L}(E,F), \|\cdot\|_{\mathcal{L}(E,F)}\big)$ *is a normed space.*

*Proof.* By definition $\|T\|_{\mathcal{L}(E,F)} \geq 0$ for all $T \in (E,F)$. Let now $\|T\|_{\mathcal{L}(E,F)} = 0$. Then by (8.1) we have

$$\sup_{x\in E\setminus\{0\}} \frac{\|Tx\|_F}{\|x\|_E} = 0,$$

so in particular $\|Tx\|_F = 0$ for all $x \in E$. Hence $T = 0$ is the zero operator. If $\lambda \in \mathbb{K}$, then by (8.1)

$$\begin{aligned}\|\lambda T\|_{\mathcal{L}(E,F)} = \sup_{\|x\|_E=1} \|\lambda Tx\|_F &= \sup_{\|x\|_E=1} |\lambda|\|Tx\|_F \\ &= |\lambda| \sup_{\|x\|_E=1} \|Tx\|_F = |\lambda|\|T\|_{\mathcal{L}(E,F)}.\end{aligned}$$

If $S, T \in \mathcal{L}(E,F)$, then again by (8.1)

$$\begin{aligned}\|S+T\|_{\mathcal{L}(E,F)} &= \sup_{\|x\|_E=1} \|(S+T)x\|_F \leq \sup_{\|x\|_E=1} \big(\|Sx\|_F + \|Tx\|_F\big) \\ &\leq \sup_{\|x\|_E=1} \big(\|S\|_{\mathcal{L}(E,F)} + \|T\|_{\mathcal{L}(E,F)}\big)\|x\|_E = \|S\|_{\mathcal{L}(E,F)} + \|T\|_{\mathcal{L}(E,F)},\end{aligned}$$

completing the proof of the proposition. ■

From now on we will *always* assume that $\mathcal{L}(E,F)$ is equipped with the operator norm. Using the steps outlined in Remark 7.6 we prove that $\mathcal{L}(E,F)$ is complete if $F$ is complete.

**8.8 Theorem** *If $F$ is a Banach space, then $\mathcal{L}(E,F)$ is a Banach space with respect to the operator norm.*

*Proof.* To simplify notation we let $\|T\| := \|T\|_{\mathcal{L}(E,F)}$ for all $T \in \mathcal{L}(E,F)$. Suppose that $F$ is a Banach space, and that $(T_n)$ is a Cauchy sequence in $\mathcal{L}(E,F)$. By Proposition 8.6 we have

$$\|T_n x - T_m x\|_F = \|(T_n - T_m)x\|_F \leq \|T_n - T_m\|\|x\|_E$$

for all $x \in E$ and $n, m \in \mathbb{N}$. As $(T_n)$ is a Cauchy sequence in $\mathcal{L}(E,F)$ it follows that $(T_n x)$ is a Cauchy sequence in $F$ for all $x \in E$. As $F$ is complete

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

exists for all $x \in E$. If $x, y \in E$ and $\lambda, \mu \in \mathbb{K}$, then

$$\begin{array}{ccc} T_n(\lambda x + \mu y) & = & \lambda T_n x + \mu T_n y \\ \big\downarrow{\scriptstyle n\to\infty} & & \big\downarrow{\scriptstyle n\to\infty} \\ T(\lambda x + \mu y) & = & \lambda Tx + \mu Ty. \end{array}$$

Hence, $T\colon E \to F$ is a linear operator. It remains to show that $T \in \mathcal{L}(E,F)$ and that $T_n \to T$ in $\mathcal{L}(E,F)$. As $(T_n)$ is a Cauchy sequence, for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that

$$\|T_n x - T_m x\|_F \leq \|T_n - T_m\|\|x\|_E \leq \varepsilon\|x\|_E$$

for all $n, m \geq n_0$ and all $x \in E$. Letting $m \to \infty$ and using the continuity of the norm we see that

$$\|T_n x - Tx\|_F \leq \varepsilon \|x\|_E$$

for all $x \in E$ and $n \geq n_0$. By definition of the operator norm $\|T_n - T\| \leq \varepsilon$ for all $n \geq n_0$. In particular, $T_n - T \in \mathcal{L}(E, F)$ for all $n \geq n_0$. Since $\varepsilon > 0$ was arbitrary, $\|T_n - T\| \to 0$ in $\mathcal{L}(E, F)$. Finally, since $\mathcal{L}(E, F)$ is a vector space and $T_{n_0}, T_{n_0} - T \in \mathcal{L}(E, F)$, we have $T = T_{n_0} - (T_{n_0} - T) \in \mathcal{L}(E, F)$, completing the proof of the theorem. ■

# 9 Equivalent Norms

Depending on the particular problem we look at, it may be convenient to work with different norms. Some norms generate the same topology as the original norm, others may generate a different topology. Here are some definitions.

**9.1 Definition (Equivalent norms)** Suppose that $E$ is a vector space, and that $\|\cdot\|_1$ and $\|\cdot\|_2$ are norms on $E$.

- We say that $\|\cdot\|_1$ is *stronger* than $\|\cdot\|_2$ if there exists a constant $C > 0$ such that

  $$\|x\|_2 \leq C\|x\|_1$$

  for all $x \in E$. In that case we also say that $\|\cdot\|_2$ is weaker than $\|\cdot\|_1$.

- We say that the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are *equivalent* if there exist two constants $c, C > 0$ such that

  $$c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1$$

  for all $x \in E$.

**9.2 Examples** (a) Let $|\cdot|_p$ denote the $p$-norms on $\mathbb{K}^N$, where $1 \leq p \leq \infty$ as defined in Section 7.2. We proved in Theorem 7.3 that

$$|x|_q \leq |x|_p \leq N^{\frac{q-p}{pq}} |x|_q$$

for all $x \in \mathbb{K}^N$ if $1 \leq p \leq q \leq \infty$. Hence all $p$-norms on $\mathbb{K}^N$ are equivalent.

(b) Consider $\ell_p$ for some $p \in [1, \infty)$. Then by Theorem 7.5 we have $|x|_q \leq |x|_p$ for all $x \in \ell_p$ if $1 \leq p < q \leq \infty$. Hence the $p$-norm is stronger than the $q$-norm considered as a norm on $\ell_p$. Note that in contrast to the finite dimensional case considered in (a) there is no equivalence of norms!

The following worthwhile observations are easily checked.

**9.3 Remarks** (a) Equivalence of norms is an equivalence relation.

(b) Equivalent norms generate the same topology on a space.

(c) If $\|\cdot\|_1$ is stronger than $\|\cdot\|_2$, then the topology $\mathcal{T}_1$ on $E$ induced by $\|\cdot\|_1$ is stronger than the topology $\mathcal{T}_2$ induced by $\|\cdot\|_2$. This means that $\mathcal{T}_1 \supseteq \mathcal{T}_2$, that is, sets open with respect to $\|\cdot\|_1$ are open with respect to $\|\cdot\|_2$ but not necessarily vice versa.

(d) Consider the two normed spaces $E_1 := (E, \|\cdot\|_1)$ and $E_2 := (E, \|\cdot\|_2)$. Clearly $E_1 = E_2 = E$ as sets, but not as normed (or metric) spaces. By Theorem 8.3 it is obvious that $\|\cdot\|_1$ is stronger than $\|\cdot\|_2$ if and only if the linear map $i(x) := x$ is a bounded linear operator $i \in \mathcal{L}(E_1, E_2)$. If the two norms are equivalent then also $i^{-1} = i \in \mathcal{L}(E_2, E_1)$.

**9.4 Lemma** *If $\|\cdot\|_1$ and $\|\cdot\|_2$ are two equivalent norms, then $E_1 := (E, \|\cdot\|_1)$ is complete if and only if $E_2 := (E, \|\cdot\|_2)$ is complete.*

*Proof.* Let $(x_n)$ be a sequence in $E$. Since $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent there exists $c, C > 0$ such that

$$c\|x_n - x_m\|_1 \leq \|x_n - x_m\|_2 \leq C\|x_n - x_m\|_1$$

for all $n, m \in \mathbb{N}$. Hence $(x_n)$ is a Cauchy sequence in $E_1$ if and only if it is a Cauchy sequence in $E_2$. Denote by $i(x) := x$ the identity map. If $x_n \to x$ in $E_1$, then $x_n \to x$ in $E_2$ since $i \in \mathcal{L}(E_1, E_2)$ by Remark 9.3(d). Similarly $x_n \to x$ in $E_1$ if $x_n \to x$ in $E_2$ since $i \in \mathcal{L}(E_2, E_1)$. ∎

Every linear operator between normed spaces induces a norm on its domain. We show under what circumstances it is equivalent to the original norm.

**9.5 Definition (Graph norm)** Suppose that $E, F$ are normed spaces and $T \in \mathrm{Hom}(E, F)$. We call

$$\|u\|_T := \|u\|_E + \|Tu\|_F$$

the *graph norm* on $E$ associated with $T$.

Let us now explain the term “graph norm.”

**9.6 Remark** From the linearity of $T$ it is rather evident that $\|\cdot\|_T$ is a norm on $E$. It is called the “graph norm” because it really is a norm on the *graph* of $T$. The graph of $T$ is the set

$$\mathrm{graph}(T) := \{(u, Tu) \colon u \in E\} \subset E \times F.$$

By the linearity of $T$ that graph is a linear subspace of $E \times F$, and $\|\cdot\|_T$ is a norm, making $\mathrm{graph}(T)$ into a normed space. Hence the name graph norm.

Since $\|u\|_E \leq \|u\|_E + \|Tu\|_F = \|u\|_T$ for all $u \in E$ the graph norm of any linear operator is stronger than the norm on $E$. We have equivalence if and only if $T$ is bounded!

**9.7 Proposition** *Let $T \in \mathrm{Hom}(E, F)$ and denote by $\|\cdot\|_T$ the corresponding graph norm on $E$. Then $T \in \mathcal{L}(E, F)$ if and only if $\|\cdot\|_T$ is equivalent to $\|\cdot\|_E$.*

*Proof.* If $T \in \mathcal{L}(E, F)$, then

$$\|u\|_E \leq \|u\|_T = \|u\|_E + \|Tu\|_F \leq \|u\|_E + \|T\|_{\mathcal{L}(E,F)}\|u\|_E = \big(1 + \|T\|_{\mathcal{L}(E,F)}\big)\|u\|_E$$

for all $u \in E$, showing that $\|\cdot\|_T$ and $\|\cdot\|_E$ are equivalent. Now assume that the two norms are equivalent. Hence there exists $C > 0$ such that $\|u\|_T \leq C\|u\|_E$ for all $u \in E$. Therefore,

$$\|Tu\|_F \leq \|u\|_E + \|Tu\|_F = \|u\|_T \leq C\|u\|_E$$

for all $u \in E$. Hence $T \in \mathcal{L}(E, F)$ by Theorem 8.3. ∎

# 10 Finite Dimensional Normed Spaces

In the previous section we did not make any assumption on the dimension of a vector space. We prove that all norms on such spaces are equivalent.

**10.1 Theorem** *Suppose that $E$ is a finite dimensional vector space. Then all norms on $E$ are equivalent. Moreover, $E$ is complete with respect to every norm.*

*Proof.* Suppose that $\dim E = N$. Given a basis $(e_1, \dots, e_N)$, for every $x \in E$ there exist unique scalars $\xi_1, \dots, \xi_N \in \mathbb{K}$ such that

$$x = \sum_{i=1}^{N} \xi_i e_i.$$

In other words, the map $T : \mathbb{K}^N \to E$, given by

$$T\xi := \sum_{i=1}^{N} \xi_i e_i \tag{10.1}$$

for all $\xi = (\xi_1, \dots, \xi_N)$, is an isomorphism between $\mathbb{K}^N$ and $E$. Let now $\|\cdot\|_E$ be a norm on $E$. Because equivalence of norms is an equivalence relation, it is sufficient to show that the graph norm of $T^{-1}$ on $E$ is equivalent to $\|\cdot\|_E$. By Proposition 9.7 this is the case if $T^{-1} \in \mathcal{L}(E, \mathbb{K}^N)$. First we show that $T \in \mathcal{L}(\mathbb{K}^N, E)$. Using the properties of a norm and the Cauchy-Schwarz inequality for the dot product in $\mathbb{K}^N$ (see also Proposition 7.2 for $p = q = 2$) we get

$$\|T\xi\|_E = \Big\|\sum_{n=1}^{N} \xi_i e_i\Big\|_E \leq \sum_{n=1}^{N} |\xi_i| \|e_i\|_E \leq \Big(\sum_{n=1}^{N} \|e_i\|_E^2\Big)^{1/2} |\xi|_2 = C|\xi|_2$$

for all $\xi \in \mathbb{K}^N$ if we set $C := \left(\sum_{n=1}^{N} \|e_i\|_E^2\right)^{1/2}$. Hence, $T \in \mathcal{L}(\mathbb{K}^N, E)$. By the continuity of a norm the map $\xi \mapsto \|T\xi\|_E$ is a continuous map from $\mathbb{K}^N$ to $\mathbb{R}$. In particular it is continuous on the unit sphere $S = \{\xi \in \mathbb{K}^N : |\xi|_2 = 1\}$. Clearly $S$ is a compact subset of $\mathbb{K}^N$. We know from Theorem 5.7 that continuous functions attain a minimum on such a set. Hence there exists $\beta \in S$ such that $\|T\beta\|_E \leq \|T\xi\|_E$ for all $\xi \in S$. Since $\|\beta\| = 1 \neq 0$ and $T$ is an isomorphism, property (i) of a norm (see Definition 6.1) implies that $c := \|T\beta\|_E > 0$. If $\xi \neq 0$, then since $\xi/|\xi|_2 \in S$,

$$c \leq \Big\|T\frac{\xi}{|\xi|_2}\Big\|_E.$$

Now by the linearity of $T$ and property (ii) of a norm $c|\xi|_2 \leq \|T\xi\|_E$ and therefore $|T^{-1}x|_2 \leq c^{-1}\|x\|_E$ for all $x \in E$. Hence $T^{-1} \in \mathcal{L}(E, \mathbb{K}^N)$ as claimed. We finally need to prove completeness. Given a Cauchy sequence in $E$ with respect to $\|\cdot\|_E$ we have from what we just proved that

$$|T^{-1}x_n - T^{-1}x_m|_2 \leq c\|x_n - x_m\|_E,$$

showing that $(T^{-1}x_n)$ is a Cauchy sequence in $\mathbb{K}^N$. By the completeness of $\mathbb{K}^N$ we have $T^{-1}x_n \to \eta$ in $\mathbb{K}^N$. By continuity of $T$ proved above we get $x_n \to T\eta$, so $(x_n)$ converges. Since the above arguments work for every norm on $E$, this shows that $E$ is complete with respect to every norm. ∎

There are some useful consequences to the above theorem. The first is concerned with finite dimensional subspaces of an arbitrary normed space.

**10.2 Corollary** *Every finite dimensional subspace of a normed space $E$ is closed and complete in $E$.*

*Proof.* If $F$ is a finite dimensional subspace of $E$, then $F$ is a normed space with the norm induced by the norm of $E$. By the above theorem $F$ is complete with respect to that norm, so in particular it is closed in $E$. ∎

The second shows that any linear operator on a finite dimensional normed space is continuous.

**10.3 Corollary** *Let $E, F$ be normed spaces and* $\dim E < \infty$. *If $T : E \to F$ is linear, then $T$ is bounded.*

*Proof.* Consider the graph norm $\|x\|_T := \|x\|_E + \|Tx\|_F$, which is a norm on $E$. By Theorem 10.1 that norm is equivalent to $\|\cdot\|_E$. Hence by Proposition 9.7 we conclude that $T \in \mathcal{L}(E, F)$ as claimed. ∎

We finally prove a counterpart to the Heine-Borel Theorem (Theorem 4.5) for general finite dimensional normed spaces. In the next section we will show that the converse is true as well, providing a topological characterisation of finite dimensional normed spaces.

**10.4 Corollary (General Heine-Borel Theorem)** *Let $E$ be a finite dimensional normed space. Then $A \subset E$ is compact if and only if $A$ is closed and bounded.*

*Proof.* Suppose that $\dim E = N$. Given a basis $(e_1, \dots, e_N)$ define $T \in \mathcal{L}(\mathbb{K}^N, E)$ as in (10.1). We know from the above that $T$ and $T^{-1}$ are continuous and therefore map closed sets onto closed sets (see Theorem 5.3). Also $T$ and $T^{-1}$ map bounded sets onto bounded sets (see Theorem 8.3) and compact sets onto compact sets (see Theorem 5.6). Hence the assertion of the corollary follows. ∎

# 11 Infinite Dimensional Normed Spaces

The purpose of this section is to characterise finite and infinite dimensional vector spaces by means of *topological properties*.

**11.1 Theorem (Almost orthogonal elements)** *Suppose $E$ is a normed space and $M$ a proper closed subspace of $E$. Then for every $\varepsilon \in (0, 1)$ there exists $x_\varepsilon \in E$ with $\|x_\varepsilon\| = 1$ and*

$$\operatorname{dist}(x_\varepsilon, M) := \inf_{x \in M} \|x - x_\varepsilon\| \geq 1 - \varepsilon.$$

*Proof.* Fix an arbitrary $x \in E \setminus M$ which exists since $M$ is a proper subspace of $E$. As $M$ is closed $\operatorname{dist}(x, M) := \alpha > 0$ as otherwise $x \in \bar{M} = M$. Let $\varepsilon \in (0, 1)$ be arbitrary and note that $(1-\varepsilon)^{-1} > 1$. Hence by definition of an infimum there exists $m_\varepsilon \in M$ such that

$$\|x - m_\varepsilon\| \leq \frac{\alpha}{1-\varepsilon}. \tag{11.1}$$

We define

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

Then clearly $\|x_\varepsilon\| = 1$ and by (11.1) we have

$$\begin{aligned}\|x_\varepsilon - m\| = \left\|\frac{x - m_\varepsilon}{\|x - m_\varepsilon\|} - m\right\| &= \frac{1}{\|x - m_\varepsilon\|}\Big\|x - \big(m_\varepsilon + \|x - m_\varepsilon\|m\big)\Big\| \\ &\geq \frac{1-\varepsilon}{\alpha}\Big\|x - \big(m_\varepsilon + \|x - m_\varepsilon\|m\big)\Big\|\end{aligned}$$

for all $m \in M$. As $m_\varepsilon \in M$ and $M$ is a subspace of $E$ we clearly have

$$m_\varepsilon + \|x - m_\varepsilon\|m \in M$$

for all $m \in M$. Thus by our choice of $x$

$$\|x_\varepsilon - m\| \geq \frac{1-\varepsilon}{\alpha}\alpha = 1 - \varepsilon$$

for all $m \in M$. Hence $x_\varepsilon$ is as required in the theorem. ∎

**11.2 Corollary** *Suppose that $E$ has closed subspaces $M_i$, $i \in \mathbb{N}$. If*

$$M_1 \subsetneq M_2 \subsetneq M_3 \subsetneq \cdots \subsetneq M_n \subsetneq M_{n+1}$$

*for all $n \in \mathbb{N}$ then there exist $m_n \in M_n$ such that $\|m_n\| = 1$ and $\operatorname{dist}(m_n, M_{n-1}) \geq 1/2$ for all $n \in \mathbb{N}$. Likewise, if*

$$M_1 \supsetneq M_2 \supsetneq M_3 \supsetneq \cdots \supsetneq M_n \supsetneq M_{n+1}$$

*for all $n \in \mathbb{N}$, then there exist $m_n \in M_n$ such that $\|m_n\| = 1$ and $\operatorname{dist}(m_n, M_{n+1}) \geq 1/2$ for all $n \in \mathbb{N}$.*

*Proof.* Consider the first case. As $M_{n-1}$ is a proper closed subspace of $M_n$ we can apply Theorem 11.1 and select $m_n \in M_n$ such that $\operatorname{dist}(m_n, M_{n-1}) \geq 1/2$. Doing so inductively for all $n \in \mathbb{N}$ we get the required sequence $(m_n)$. In the second case we proceed similarly: There exists $m_1 \in M_1$ such that $\operatorname{dist}(m_1, M_2) \geq 1/2$. Next choose $m_2 \in M_2$ such that $\operatorname{dist}(m_2, M_3) \geq 1/2$, and so on. ∎

With the above we are able to give a topological characterisation of finite and infinite dimensional spaces.

**11.3 Theorem** *A normed space $E$ is finite dimensional if and only if the unit sphere $S = \{x \in E : \|x\| = 1\}$ is compact.*

*Proof.* First assume that $\dim E = N < \infty$. Since the unit sphere is closed and bounded, by the general Heine-Borel Theorem (Corollary 10.4) it is compact. Now suppose that $E$ is infinite dimensional. Then there exists a countable linearly independent set $\{e_n : n \in \mathbb{N}\}$. We set $M_n := \operatorname{span}\{e_k : k = 1, \dots, n\}$. Clearly $\dim M_n = n$ and thus by Corollary 10.2 $M_n$ is closed for all $n \in \mathbb{N}$ $M_n$. As $\dim M_{n+1} > \dim M_n$ the sequence $(M_n)$ satisfies the assumptions of Corollary 11.2. Hence there exist $m_n \in M_n$ such that $\|m_n\| = 1$ and $\operatorname{dist}(m_n, M_{n-1}) \geq 1/2$ for all $n \in \mathbb{N}$. However, this implies that $\|m_n - m_k\| \geq 1/2$ whenever $n \neq k$, showing that there is a sequence in $S$ which does not have a convergent subsequence. Hence $S$ cannot be compact, completing the proof of the theorem. ■

**11.4 Corollary** *Let $E$ be a normed vector space. Then the following assertions are equivalent:*

(i) $\dim E < \infty$*;*

(ii) *The unit sphere in $E$ is compact;*

(iii) *The unit ball in $E$ is relatively compact;*

(iv) *Every closed and bounded set in $E$ is compact.*

*Proof.* Assertions (i) and (ii) are equivalent by Theorem 11.3. Suppose that (ii) is true. Denote the unit sphere in $E$ by $S$. Define a map $f : [0, 1] \times S \to E$ by setting $f(t, x) := tx$. Then by Theorem 6.4 the map $f$ is continuous and its image is the closed unit ball in $E$. By Proposition 4.8 the set $[0, 1] \times S$ is compact. Since the image of a compact set under a continuous function is compact it follows that the closed unit ball in $E$ is compact, proving (iii). Now assume that (iii) holds. Let $M$ be an arbitrary closed and bounded set in $E$. Then there exists $R > 0$ such that $M \subset \overline{B(0, R)}$. Since the map $x \to Rx$ is continuous on $E$ and the closed unit ball is compact it follows that $\overline{B(0, R)}$ is compact. Now $M$ is compact because it is a closed subset of a compact set (see Proposition 4.7), so (iv) follows. If (iv) holds, then in particular the unit sphere is compact, so (ii) follows. ■

# 12 Quotient Spaces

Consider a vector space $E$ and a subspace $F$. We define an equivalence relation $\sim$ between elements $x, y$ in $E$ by $x \sim y$ if and only if $x - y \in F$. Denote by $[x]$ the equivalence class of $x \in E$ and set

$$E/F := \{[x] : x \in E\}.$$

As you probably know from Algebra, this is called the *quotient space of $E$ modulo $F$*. That quotient space is a vector space over $\mathbb{K}$ if we define the operations

$$[x] + [y] := [x + y]$$
$$\alpha[x] := [\alpha x]$$

for all $x, y \in E$ and $\alpha \in \mathbb{K}$. It is easily verified that these operations are well defined. If $E$ is a normed space we would like to show that $E/F$ is a normed space with norm

$$\|[x]\|_{E/F} := \inf_{z \in F} \|x - z\|_E. \tag{12.1}$$

This is a good definition since then $\|[x]\|_{E/F} \leq \|x\|_E$ for all $x \in E$, that is, the natural projection $E \to E/F$, $x \mapsto [x]$ is continuous. Geometrically, $\|[x]\|_{E/F}$ is the distance of the affine subspace $[x] = x + F$ from the origin, or equivalently the distance between the affine spaces $F$ and $x + F$ as Figure 12.1 shows in the situation of two dimensions. Unfortunately, $\|\cdot\|_{E/F}$ is not always a norm, but only if $F$ is a *closed* subspace of $E$.

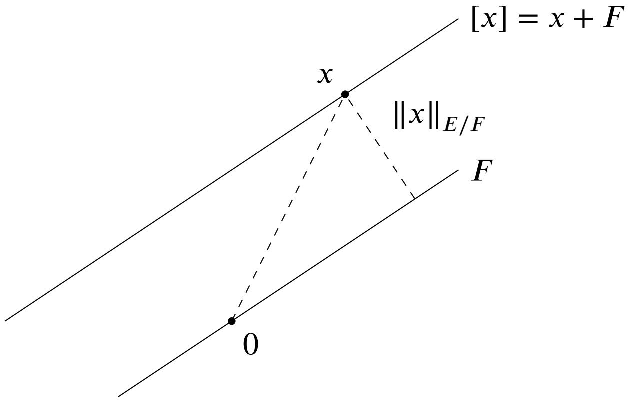



Figure 12.1: Distance between affine subspaces

**12.1 Proposition** *Let $E$ be a normed space. Then $E/F$ is a normed space with norm* (12.1) *if and only if $F$ is a closed subspace of $E$.*

*Proof.* Clearly $\|[x]\|_{E/F} \geq 0$ for all $x \in E$ and $\|[x]\|_{E/F} = 0$ if $[x] = [0]$, that is, $x \in F$. Now suppose that $\|[x]\|_{E/F} = 0$. We want to show that then $[x] = [0]$ if and only if $F$ is closed. First suppose that $F$ is closed. If $\|[x]\|_{E/F} = 0$, then by definition there exist $z_n \in F$ with $\|x - z_n\| \to 0$. Hence $z_n \to x$, and since $F$ is closed $x \in F$. But then $[x] = [0]$ proving what we want. Suppose now $F$ is not closed. Then there exists a sequence $z_n \in F$ with $z_n \to x$ and $x \notin F$. Hence $[x] \neq [0]$, but

$$0 \leq \|[x]\|_{E/F} \leq \lim_{n \to \infty} \|x - z_n\| = 0,$$

that is $\|[x]\|_{E/F} = 0$ even though $[x] \neq [0]$. Hence (12.1) does not define a norm. The other properties of a norm are valid no matter whether $F$ is closed or not. First note that for $\alpha \in \mathbb{K}$ and $x \in E$

$$\|\alpha[x]\|_{E/F} = \|[\alpha x]\|_{E/F} \leq \|\alpha(x - z)\|_E = |\alpha|\|x - z\|_E$$

for all $z \in F$. Hence $\|\alpha[x]\|_{E/F} \leq |\alpha|\|[x]\|_{E/F}$ with equality if $\alpha = 0$. If $\alpha \neq 0$, then by the above

$$\|[x]\|_{E/F} = \|\alpha^{-1}\alpha[x]\|_{E/F} \leq |\alpha|^{-1}\|\alpha[x]\|_{E/F},$$

so $\|\alpha[x]\|_{E/F} \geq |\alpha|\|x\|_{E/F}$, showing that $\|\alpha[x]\|_{E/F} = |\alpha|\|x\|_{E/F}$. Finally let $x, y \in E$ and fix $\varepsilon > 0$ arbitrary. By definition of the quotient norm there exist $z, w \in F$ such that $\|x - z\|_E \leq \|[x]\|_{E/F} + \varepsilon$ and $\|y - w\|_E \leq \|[y]\|_{E/F} + \varepsilon$. Hence

$$\|[x]+[y]\|_{E/F} \leq \|x+y-z-w\|_E \leq \|x-z\|_E+\|y-w\|_E \leq \|[x]\|_{E/F}+\|[y]\|_{E/F}+2\varepsilon.$$

As $\varepsilon > 0$ was arbitrary $\|[x]+[y]\|_{E/F} \leq \|[x]\|_{E/F}+\|[y]\|_{E/F}$, so the triangle inequality holds. ■

The above proposition justifies the following definition.

**12.2 Definition (quotient norm)** If $F$ is a closed subspace of the normed space $E$, then the norm $\|\cdot\|_{E/F}$ is called the *quotient norm* on $E/F$.

We next look at completeness properties of quotient spaces.

**12.3 Theorem** *Suppose $E$ is a Banach space and $F$ a closed subspace. Then $E/F$ is a Banach space with respect to the quotient norm.*

*Proof.* The only thing left to prove is that $E/F$ is complete with respect to the quotient norm. We use the characterisation of completeness of a normed space given in Theorem 6.8. Hence let $\sum_{n=1}^{\infty}[x_n]$ be an absolutely convergent series in $E/F$, that is,

$$\sum_{n=1}^{\infty} \|[x_n]\|_{E/F} \leq M < \infty.$$

By definition of the quotient norm, for each $n \in \mathbb{N}$ there exists $z_n \in F$ such that

$$\|x_n - z_n\|_E \leq \|[x_n]\|_{E/F} + \frac{1}{2^n}.$$

Hence,

$$\sum_{n=1}^{m} \|x_n - z_n\|_E \leq \sum_{n=1}^{m} \|[x_n]\|_{E/F} + \sum_{n=1}^{m} \frac{1}{2^n} \leq M + 2 < \infty$$

for all $m \in \mathbb{N}$. This means that $\sum_{n=1}^{\infty}(x_n - z_n)$ is absolutely convergent and therefore convergent by Theorem 6.8 and the assumption that $E$ be complete. We set

$$s := \sum_{n=1}^{\infty}(x_n - z_n).$$

Now by choice of $z_n$ and the definition of the quotient norm

$$\Big\|\Big(\sum_{n=1}^{m}[x_n]\Big) - [s]\Big\|_{E/F} = \Big\|\Big(\sum_{n=1}^{m}[x_n - z_n]\Big) - [s]\Big\|_{E/F} \leq \Big\|\Big(\sum_{n=1}^{m}(x_n - z_n)\Big) - s\Big\|_E \to 0$$

as $m \to \infty$ by choice of $s$. Hence $\sum_{n=1}^{\infty}[x_n] = [s]$ converges with respect to the quotient norm, and so by Theorem 6.8 $E/F$ is complete. ■

Next we look at factorisations of bounded linear operators. Given normed spaces $E, F$ and an operator $T \in \mathcal{L}(E, F)$ it follows from Theorem 5.3 that $\ker T := \{x \in$

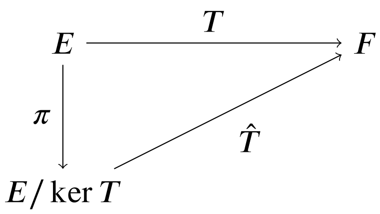
$E : Tx = 0\}$ is a closed subspace of $E$. Hence $E/\ker T$ is a normed space with the quotient norm. We then define a linear operator $\hat{T} : E/\ker T \to F$ by setting

$$\hat{T}[x] := Tx$$

for all $x \in E$. It is easily verified that this operator is well defined and linear. Moreover, if we set $\pi(x) := [x]$, then by definition of the quotient norm $\|\pi(x)\|_{E/\ker T} \leq \|x\|_E$, so $\pi \in \mathcal{L}(E, E/\ker T)$ with $\|\pi\|_{\mathcal{L}(E,E/\ker T)} \leq 1$. Moreover, we have the factorisation

$$T = \hat{T} \circ \pi,$$

meaning that the following diagram is commutative.




We summarise the above in the following Theorem.

**12.4 Theorem** *Suppose that $E, F$ are normed spaces and that $T \in \mathcal{L}(E, F)$. If $\hat{T}, \pi$ are defined as above, then $\hat{T} \in \mathcal{L}(E/\ker T, F)$, $\|T\|_{\mathcal{L}(E,F)} = \|\hat{T}\|_{\mathcal{L}(E/\ker T,F)}$ and we have the factorisation $T = \hat{T} \circ \pi$.*

*Proof.* The only thing left to prove is that $\hat{T} \in \mathcal{L}(E/\ker T, F)$, and that $\|T\| := \|T\|_{\mathcal{L}(E,F)} = \|\hat{T}\|_{\mathcal{L}(E/\ker T,F)} =: \|\hat{T}\|$. First note that

$$\|\hat{T}[x]\|_F = \|Tx\|_F = \|T(x - z)\|_F \leq \|T\|\|x - z\|_E$$

for all $z \in \ker T$. Hence by definition of the quotient norm

$$\|\hat{T}[x]\|_F \leq \|T\|\|[x]\|_{E/\ker T}.$$

Now by definition of the operator norm $\|\hat{T}\| \leq \|T\| < \infty$. In particular, $\hat{T} \in \mathcal{L}(E/\ker T, F)$. To show equality of the operator norms observe that

$$\|Tx\|_F = \|\hat{T}[x]\|_F \leq \|\hat{T}\|\|[x]\|_{E/\ker T} \leq \|\hat{T}\|\|x\|_E$$

by definition of the operator and quotient norms. Hence $\|T\| \leq \|\hat{T}\|$, showing that $\|T\| = \|\hat{T}\|$. ■

Chapter III

# Banach algebras and the Stone-Weierstrass Theorem

## 13 Banach algebras

Some Banach spaces have an additional structure. For instance if we consider the space of continuous functions $C(K)$ with $K$ compact, then we can *multiply* the functions as well and define

$$(fg)(x) := f(x)g(x)$$

for all $x \in K$. The natural norm on $C(K)$ is the supremum norm, and we easily see that

$$\|fg\|_\infty \leq \|f\|_\infty \|g\|_\infty,$$

which in particular implies that multiplication is continuous. Also, there is a neutral element for multiplication, namely the constant function with value one. Vector spaces with this additional multiplicative structure are called algebras. Here we have an additional topological structure. We introduce the following definition.

**13.1 Definition** A normed space $E$ is called a (commutative) *normed algebra* if there is an operation $E \times E \to E$ called multiplication with the following properties:

(i) $xy = yx$ for all $x, y \in E$ (commutative law)

(ii) $x(yz) = (xy)z$ for all $x, y, z \in E$ (associative law);

(iii) $x(y + z) = xy + xz$ for all $x, y \in E$ (distributive law);

(iv) $\alpha(xy) = (\alpha x)y$ for all $x, y \in E$ and $\alpha, \beta \in \mathbb{K}$;

(v) $\|xy\| \leq \|x\|\|y\|$ for all $x, y \in E$.

We call $e \in E$ a unit if $ex = x$ for all $x \in E$. If $E$ is complete we call $E$ a *Banach algebra*.

Note that not every Banach algebra has a unit. For instance $L^1(\mathbb{R}^N)$ is a Banach algebra with multiplication defined as *convolution*, that is,

$$(f * g)(x) := \int_{\mathbb{R}^N} f(x - y)g(y)\, dy.$$

Young's inequality implies that $\|f * g\|_1 \leq \|f\|_1\|g\|_1$, but there is no unity (see MATH3969). Unity would correspond to the "delta function" which is really a measure, not a function.

# 14 The Stone-Weierstrass Theorem

In this section we consider properties of subspaces of the Banach algebra $C(K)$, where $K$ is a compact set. A classical theorem in real analysis, the *Weierstrass approximation theorem* asserts that every continuous function on a compact interval in $\mathbb{R}$ can be uniformly approximated by a sequence of polynomials. In other words, the space of polynomials on a compact interval $I$ is dence in $C(I)$.

The aim of this section is to prove a generalisation of this theorem to Banach algebras with unity, called the Stone-Weierstrass theorem. It is named after Stone who published the generalisation in [12] in 1948 with a huge success.

The standard proof of the Stone-Weierstrass theorem requires a very special case of the Weierstrass approximation theorem, namely the fact that the absolute value function $|x|$ can be uniformly approximated by polynomials in the interval $[-1, 1]$. To achieve that we write

$$|x| = \sqrt{x^2} = \sqrt{1 - (1 - x^2)}.$$

Using the binomial theorem we can write

$$|x| = \sum_{k=0}^{\infty}(-1)^k\binom{1/2}{k}(1 - x^2)^k \tag{14.1}$$

which is a power series in $t = (1 - x^2)$, converging for $|t| < 1$, that is, $x \in (-1, 1)$. If we can prove that it converges absolutely and uniformly with respect to $t \in [-1, 1]$, then the Weierstrass $M$-test shows that (14.1) converges absolutely and uniformly with respect to $x \in [-1, 1]$. Hence the sequence of partial sums of (14.1) provides a uniform polynomial approximation of $|x|$ on $[-1, 1]$.

**14.1 Lemma** *Let $\alpha \in (0, 1)$. Then the binomial series*

$$(1 - t)^\alpha = \sum_{k=0}^{\infty}(-1)^k\binom{\alpha}{k}t^k \tag{14.2}$$

*converges absolutely and uniformly with respect to t in the closed ball $\overline{B(0, 1)} \subseteq \mathbb{C}$.*

*Proof.* We know that (14.2) holds for $t$ in the open ball $B(0, 1)$. As $\alpha \in (0, 1)$ we have

$$\binom{\alpha}{k} = \frac{\alpha(\alpha - 1)\dots(\alpha - k + 1)}{k!} = (-1)^{k-1}\left|\binom{\alpha}{k}\right|$$

for all $k \geq 1$. Hence, for $t \in (0, 1)$ the series (14.2) can be rewritten in the form

$$(1 + t)^\alpha = 1 - \sum_{k=0}^{\infty}\left|\binom{\alpha}{k}\right|t^k. \tag{14.3}$$

Rearranging we obtain

$$\sum_{k=0}^{n} \left|\binom{\alpha}{k}\right| t^k \leq \sum_{k=0}^{\infty} \left|\binom{\alpha}{k}\right| t^k = 1 - (1-t)^\alpha \leq 1$$

for all $n \in \mathbb{N}$ and all $t \in (0,1)$. Letting $t \uparrow 1$ we see that

$$\sum_{k=0}^{n} \left|\binom{\alpha}{k}\right| \leq 1$$

for all $n \in \mathbb{N}$. These are the partial sums of a series with non-negative terms and hence,

$$\sum_{k=0}^{\infty} \left|\binom{\alpha}{k}\right| \leq 1$$

converges. The Weierstrass $M$-Test now implies that (14.2) converges absolutely and uniformly on $\overline{B(0,1)}$. ■

We now state and prove the main result of this section about the density of sub-algebras of $C(K)$. A sub-algebra is simply a subspace of $C(K)$ that is also a normed algebra.

**14.2 Theorem (Stone-Weierstrass)** *Let $K$ be a compact set and suppose that $B$ is a sub-algebra of $C(K)$ over $\mathbb{K}$ satisfying the following conditions:*

(i) *$B$ contains the constant function 1;*

(ii) *$B$ separates points, that is, for every pair of distinct points $x, y \in K$ there exist $f \in B$ such that $f(x) \neq f(y)$.*

(iii) *If $f \in B$, then also $\bar{f} \in B$, where $\bar{f}$ is the complex conjugate of $f$.*

*Then, $B$ is dense in $C(K)$.*

*Proof.* (a) We first assume that $\mathbb{K} = \mathbb{R}$. By Lemma 14.1 there exists a sequence of polynomials $p_n$ such that $\left|q_n(s) - |s|\right| < 1/n^2$ for all $n \in \mathbb{N}$ and all $s \in [-1,1]$. We rescale the inequality and obtain

$$\left|nq_n(s) - |ns|\right| < 1/n$$

for all $n \in \mathbb{N}$ and $s \in [-1,1]$. Note that we can rewrite $nq_n(s)$ in the form $p_n(ns)$ for some polynomial $p_n$. Given that $q_n(s) = a_0 + a_1 s + \cdots + a_m s^m$, the polynomial

$$p_n(ns) := na_0 + n\frac{a_1}{n}(ns) + \cdots + n\frac{a_m}{n^m}(ns)^m$$

satisfies $p_n(ns) = nq_n(s)$ for all $s \in \mathbb{R}$. In particular, setting $t = ns$, we have a sequence of polynomials so that

$$\left|p_n(t) - |t|\right| < 1/n \tag{14.4}$$

for all $n \in \mathbb{N}$ and all $t \in [-n, n]$. Let $f \in B$. Since $B$ is an algebra $p_n \circ f \in B$ for all $n \in \mathbb{N}$, and that

$$\left|(p_n \circ f)(x) - |f(x)|\right| < 1/n \tag{14.5}$$

for all $n \geq \|f\|_\infty < \infty$ and all $x \in K$. Hence $p_n \circ f \to |f|$ in $C(K)$, which means that $|f| \in \overline{B}$ for all $f \in B$. As a consequence also

$$\max\{f, g\} = \frac{1}{2}\big(f + g + |f - g|\big) \quad \text{and} \quad \min\{f, g\} = \frac{1}{2}\big(f + g - |f - g|\big) \tag{14.6}$$

are in the closure of $B$ for all $f, g \in B$.

Fix now $f \in C(K)$ and $x, y \in K$ with $x \neq y$. By assumption (ii) there exists $g \in B$ such that $g(x) \neq g(y)$. We set

$$h_{xy}(z) := f(x)1 + \big(f(y) - f(x)\big)\frac{g(z) - g(x)}{g(y) - g(x)}$$

for all $z \in K$. As $1 \in B$ by assumption we have $h_{xy} \in B$ and $h_{xy}(x) = f(x)$ and $h_{xy}(y) = f(y)$. Let now $\varepsilon > 0$ be fixed. Using that $h_{xy}$ and $f$ are continuous the sets

$$\begin{aligned} U_{xy} &:= \big\{z \in K : h_{xy}(z) < f(z) + \varepsilon\big\} \\ V_{xy} &:= \big\{z \in K : h_{xy}(z) > f(z) - \varepsilon\big\} \end{aligned}$$

are open in $K$; see Theorem 5.3. Both sets are non-empty since $h_{xy}(x) = f(x)$ and $h_{xy}(y) = f(y)$ and thus $x, y \in U_{xy} \cap V_{xy}$. The family $(U_{xy})_{x,y \in K}$ is clearly an open cover of $K$. Fix now $y \in K$. By the compactness of $K$ there exist finitely many points $x_j \in K$, $k = 1, \dots, m$, so that $K = \bigcup_{j=1}^m U_{x_j y}$. Applying (14.6) repeatedly we see that

$$f_y := \min_{j=1,\dots,m} h_{x_j y} \in \overline{B}.$$

By choice of $x_j$ it follows that

$$\begin{aligned} &f_y(z) < f(z) + \varepsilon \quad \text{for all } z \in K; \\ &f_y(z) > f(z) - \varepsilon \quad \text{for all } z \in V_y := \bigcup_{j=1}^m V_{x_j y}. \end{aligned}$$

As $V_y$ is the intersection of finitely many open sets all containing $y$ it is open and non-empty. Again by compactness of $K$ we can choose $y_1, \dots, y_\ell$ such that $K = \bigcup_{j=1}^\ell U_{y_j}$. Again applying (14.6) we see that

$$h := \max_{j=1,\dots,m} f_y \in \overline{B},$$

and that

$$f(z) - \varepsilon < h(z) < f(z) + \varepsilon$$

for all $z \in K$. In particular $\|f - h\|_\infty \leq \varepsilon$. As $h \in \overline{B}$ and the argument above works for any $\varepsilon >$, we conclude that $f \in \overline{B}$ as well.

(b) Let us now consider the case of $\mathbb{K} = \mathbb{C}$. Given $f \in C(K, \mathbb{C}) = C(K, \mathbb{R}) + iC(K, \mathbb{R})$, assumption (iii) implies that real and imaginary parts given by

$$\operatorname{Re} f = \frac{1}{2}(f + \bar{f}) \qquad \text{and} \qquad \operatorname{Re} f = \frac{1}{2i}(f - \bar{f})$$

are both in $B$. Set $B_{\mathbb{R}} := B \cap C(K, \mathbb{R})$. We can apply the real version proved already to conclude that $\overline{B_{\mathbb{R}}} = C(K, \mathbb{R})$. Hence, we also have $\overline{B} = C(K, \mathbb{C})$. ■

We can deduce the Weierstrass approximation theorem as a corollary.

**14.3 Theorem (Weierstrass approximation theorem)** *Let $I$ be a compact subset of $\mathbb{R}$. Then the space of polynomials $\mathbb{R}[x]$ is dense in $C(I)$.*

*Proof.* The space of polynomials is clearly a sub-algebra containing 1 (the constant polynomial). It also contains the identity $p(x) = x$, and therefore separates points. Now the Stone-Weierstrass theorem 14.2 applies. ■

**14.4 Remark** The Weierstrass approximation theorem does not apply to $C(K)$ if $K$ is a compact subset of $\mathbb{C}$ containing some interior. In fact, polynomials are naturally analytic funtions and a theorem from complex analysis asserts that the uniform limit of analytic functions are analytic. Hence it is impossible to approximate an arbitrary continuous function uniformly by polynomials unless it is analytic. All assumptions of Theorem 14.2 are fulfilled except for the last, the one on the complex conjugates: if $p(z)$ is a polynomial, then its complex conjugate $\overline{p(z)}$ is no longer a polynomial as $\bar{z}$ cannot be writen as a polynomial in $z$.

Chapter IV

# Hilbert Spaces

Hilbert spaces are in some sense a direct generalisation of finite dimensional Euclidean spaces, where the norm has some geometric meaning and angles can be defined by means of the dot product. The dot product can be used to define the norm and prove many of its properties. Hilbert space theory is doing this in a similar fashion, where an *inner product* is a map with properties similar to the dot product in Euclidean space. We will emphasise the analogies and see how useful they are to find proofs in the general context of inner product spaces.

## 15 Inner Product Spaces

Throughout we let $E$ denote a vector space over $\mathbb{K}$.

**15.1 Definition (Inner product, inner product space)** A function $(\cdot \mid \cdot)\colon E \times E \to \mathbb{K}$ is called an *inner product* or *scalar product* if

(i) $(u \mid v) = \overline{(v \mid u)}$ for $u, v \in E$,

(ii) $(u \mid u) \geq 0$ for all $u \in E$ and $(u \mid u) = 0$ if and only if $u = 0$.

(iii) $(\alpha u + \beta v \mid w) = \alpha(u \mid w) + \beta(v \mid w)$ for all $u, v, w \in E$ and $\alpha, \beta \in \mathbb{K}$,

We say that $E$ equipped with $(\cdot \mid \cdot)$ is an *inner product space*.

**15.2 Remark** As an immediate consequence of the above definition, inner products have the following properties:

(a) By property (i) we have $(u \mid u) = \overline{(u \mid u)}$ and therefore $(u \mid u) \in \mathbb{R}$ for all $u \in E$. Hence property (ii) makes sense.

(b) Using (i) and (iii) we have

$$(u \mid \alpha v + \beta w) = \overline{\alpha}(u \mid v) + \overline{\beta}(u \mid w)$$

for all $u, v, w \in E$ and $\alpha, \beta \in \mathbb{K}$. In particular we have

$$(u \mid \lambda v) = \overline{\lambda}(u \mid v)$$

for all $u, v \in E$ and $\lambda \in \mathbb{K}$.

Next we give some examples of Banach and Hilbert spaces.

**15.3 Examples** (a) The space $\mathbb{C}^N$ equipped with the Euclidean scalar product given by

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

for all $x := (x_1, \dots, x_N), y := (y_1, \dots, y_N) \in \mathbb{C}^N$ is an inner product space. More generally, if we take a positive definite Hermitian matrix $A \in \mathbb{C}^{N\times N}$, then

$$(x \mid y)_A := x^T A \bar{y}$$

defines an inner product on $\mathbb{C}^N$.

(b) An infinite dimensional version is $\ell_2$ defined in Section 7.2. An inner product is defined by

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

for all $(x_i), (y_i) \in \ell_2$. The series converges by Hölder's inequality (Proposition 7.4).

(c) For $u, v \in L_2(X)$ we let

$$(u \mid v) := \int_X u(x)\overline{v(x)}\, dx.$$

By Hölder's inequality Proposition 7.7 the integral is finite and easily shown to be an inner product.

The Euclidean norm on $\mathbb{C}^N$ is defined by means of the dot product, namely by $\|x\| = \sqrt{x \cdot x}$ for $x \in \mathbb{C}^N$. We make a similar definition in the context of general inner product spaces.

**15.4 Definition (induced norm)** If $E$ is an inner product space with inner product $(\cdot \mid \cdot)$ we define

$$\|u\| := \sqrt{(u \mid u)} \tag{15.1}$$

for all $u \in E$.

Note that from Remark 15.2 we always have $(x \mid x) \geq 0$, so $\|x\|$ is well defined. We call $\|\cdot\|$ a "norm," but at the moment we do not know whether it really is a norm in the proper sense of Definition 6.1. We now want to work towards a proof that $\|\cdot\|$ is a norm on $E$. On the way we look at some geometric properties of inner products and establish the Cauchy-Schwarz inequality.

By the algebraic properties of the inner products in a space over $\mathbb{R}$ and the definition of the norm we get

$$\|v - u\|^2 = \|u\|^2 + \|v\|^2 - 2(u \mid v).$$

On the other hand, by the law of cosines we know that for vectors $u, v \in \mathbb{R}^2$

$$\|v - u\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\| \cos\theta.$$

if we form a triangle from $u, v$ and $v - u$ as shown in Figure 15.1. Therefore

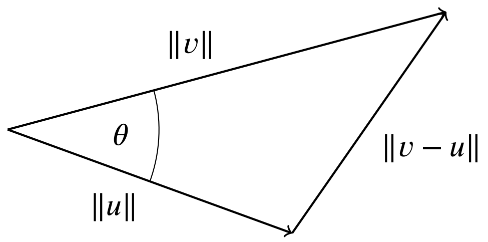



Figure 15.1: Triangle formed by $u$, $v$ and $v-u$.

$$u \cdot v = \|u\|\|v\|\cos\theta$$

and thus

$$|u \cdot v| \leq \|u\|\|v\|.$$

The latter inequality has a counterpart in general inner product spaces. We give a proof *inspired* by (but not relying on) the geometry in the plane. All arguments used purely depend on the algebraic properties of an inner product and the definition of the induced norm.

**15.5 Theorem (Cauchy-Schwarz inequality)** *Let $E$ be an inner product space with inner product $(\cdot \mid \cdot)$. Then*

$$|(u \mid v)| \leq \|u\|\|v\| \tag{15.2}$$

*for all $u, v \in E$ with equality if and only if $u$ and $v$ are linearly dependent.*

*Proof.* If $u = 0$ or $v = 0$ the inequality is obvious and $u$ and $v$ are linearly dependent. Hence assume that $u \neq 0$ and $v \neq 0$. We can then define

$$n = v - \frac{(u \mid v)}{\|u\|^2}\, u.$$

Note that the vector

$$p := \frac{(u \mid v)}{\|u\|^2}\, u$$

is the projection of $v$ in the direction of $u$, and $n$ is the projection of $v$ orthogonal to $u$ as shown in Figure 15.2. Using the algebraic rules for the inner product and the definition

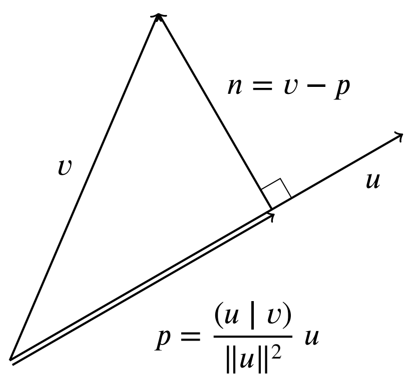



Figure 15.2: Geometric interpretation of $n$.

of the norm we get

$$0 \le \|n\|^2 = v \cdot v - 2\frac{(u \mid v)(v \mid u)}{\|u\|^2} + \frac{(u \mid v)\overline{(u \mid v)}}{\|u\|^4}(u \mid u)$$
$$= \|v\|^2 - 2\frac{|(u \mid v)|^2}{\|u\|^2} + \frac{|(u \mid v)|^2}{\|u\|^4}\|u\|^2 = \|v\|^2 - \frac{|(u \mid v)|^2}{\|u\|^2}.$$

Therefore $|(u \mid v)|^2 \le \|u\|^2\|v\|^2$, and by taking square roots we find (15.2). Clearly equality holds if and only if $\|n\| = 0$, that is, if

$$v = \frac{(u \mid v)}{\|u\|^2}\, u.$$

Hence we have equality in (15.2) if and only if $u$ and $v$ are linearly dependent. This completes the proof of the theorem. ■

As a consequence we get a different characterisation of the induced norm.

**15.6 Corollary** *If $E$ is an inner product space and $\|\cdot\|$ the induced norm, then*

$$\|u\| = \sup_{\|v\|\le 1} |(u \mid v)| = \sup_{\|v\|=1} |(u \mid v)|$$

*for all $u \in E$.*

*Proof.* If $u = 0$ the assertion is obvious, so assume that $u \neq 0$. If $\|v\| \le 1$, then $|(u \mid v)| \le \|u\|\|v\| = \|u\|$ by the Cauchy-Schwarz inequality. Hence

$$\|u\| \le \sup_{\|v\|\le 1} |(u \mid v)|.$$

Choosing $v := u/\|u\|$ we have $|(u \mid v)| = \|u\|^2/\|u\| \le \|u\|$, so equality holds in the above inequality. Since the supremum over $\|v\| = 1$ is larger or equal to that over $\|v\| \le 1$, the assertion of the corollary follows. ■

Using the Cauchy-Schwarz inequality we can now prove that $\|\cdot\|$ is in fact a norm.

**15.7 Theorem** *If $E$ is an inner product space, then* (15.1) *defines a norm on $E$.*

*Proof.* By property (ii) of an inner product (see Definition 15.1 we have $\|u\| = \sqrt{(u \mid u)} \ge 0$ with equality if and only if $u = 0$. If $u \in E$ and $\lambda \in \mathbb{K}$, then

$$\|\lambda u\| = \sqrt{(\lambda u \mid \lambda u)} = \sqrt{\lambda\bar{\lambda}(u \mid u)} = \sqrt{|\lambda|^2\|u\|^2} = |\lambda|\|u\|$$

as required. To prove the triangle inequality let $u, v \in E$. By the algebraic properties of an inner product and the Cauchy-Schwarz inequality we have

$$\|u + v\|^2 = (u + v \mid u + v) = \|u\|^2 + (u \mid v) + (v \mid u) + \|v\|^2$$
$$\le \|u\|^2 + 2|(u \mid v)|^2 + \|v\|^2 \le \|u\|^2 + 2\|u\|^2\|v\|^2 + \|v\|^2 = \big(\|u\| + \|v\|\big)^2.$$

Taking square roots the triangle inequality follows. Hence $\|\cdot\|$ defines a norm. ■

As a matter of convention we always consider inner product spaces as normed spaces.

**15.8 Convention** Since every inner product induces a norm we will always assume that an inner product space is a normed space with the *norm induced by the inner product*.

Once we have a norm we can talk about convergence and completeness. Note that not every inner product space is complete, but those which are play a special role.

**15.9 Definition (Hilbert space)** An inner product space which is complete with respect to the induced norm is called a *Hilbert space*.

The inner product is a map on $E \times E$. We show that this map is continuous with respect to the induced norm.

**15.10 Proposition (Continuity of inner product)** *Let $E$ be an inner product space. Then the inner product $(\cdot \mid \cdot)\colon E \times E \to \mathbb{K}$ is continuous with respect to the induced norm.*

*Proof.* If $x_n \to x$ and $y_n \to y$ in $E$ (with respect to the induced norm), then using the Cauchy-Schwarz inequality

$$\begin{aligned}|(x_n \mid y_n) - (x \mid y)| &= |(x_n - x \mid y_n) + (x \mid y_n - y)| \\ &\leq |(x_n - x \mid y_n)| + |(x \mid y_n - y)| \leq \|x_n - x\| \|y_n\| + \|x\| \|y_n - y\| \to 0\end{aligned}$$

as $n \to \infty$. Note that we also use the continuity of the norm in the above argument to conclude that $\|y_n\| \to \|y\|$ (see Theorem 6.4). Hence the inner product is continuous. ∎

The lengths of the diagonals and edges of a parallelogram in the plane satisfy a relationship. The norm in an inner product space satisfies a similar relationship, called the parallelogram identity. The identity will play an essential role in the next section.

**15.11 Proposition (Parallelogram identity)** *Let $E$ be an inner product space and $\|\cdot\|$ the induced norm. Then*

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2 \tag{15.3}$$

*for all $u, v \in E$.*

*Proof.* By definition of the induced norm and the properties of an inner product

$$\begin{aligned}\|u + v\|^2 + \|u - v\|^2 = \|u\|^2 &+ (u \mid v) + (v \mid u) + \|v\|^2 \\ &+ \|u\|^2 - (u \mid v) - (v \mid u) + \|v\|^2 = 2\|u\|^2 + 2\|v\|^2\end{aligned}$$

for all $u, v \in E$ as required. ∎

It turns out that the converse is true as well. More precisely, if a norm satisfies (15.3) for all $u, v \in E$, then there is an inner product inducing that norm (see [13, Section I.5] for a proof).

# 16 Projections and Orthogonal Complements

In this section we discuss the existence and properties of "nearest point projections" from a point onto a set, that is, the points that minimise the distance from a closed set to a given point.

**16.1 Definition (Projection)** Let $E$ be a normed space and $M$ a non-empty closed subset. We define the *set of projections of $x$ onto $M$* by

$$P_M(x) := \{m \in M : \|x - m\| = \operatorname{dist}(x, M)\}.$$

The meaning of $P_M(x)$ is illustrated in Figure 16.1 for the Euclidean norm in the plane. If the set is not convex, $P_M(x)$ can consist of several points, if it is convex, it is precisely one.

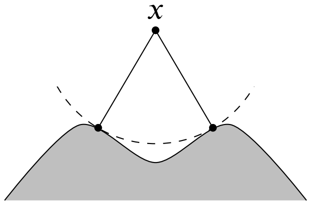



Figure 16.1: The set of nearest point projections $P_M(x)$.

We now look at some example. First we look at subsets of $\mathbb{R}^N$, and show that then $P_M(x)$ is never empty.

**16.2 Example** Suppose that $M \subset \mathbb{R}^N$ is non-empty and closed and $x \in M$. If we fix $\alpha > \operatorname{dist}(x, M)$ and $x \in \mathbb{R}^N$, then the set $K := M \cap \overline{B(x, \alpha)}$ is a closed and bounded, and $\operatorname{dist}(x, M) = \operatorname{dist}(x, K)$. We know from Proposition 5.5 that the distance function $x \mapsto \operatorname{dist}(x, K)$ is continuous. Since $K$ is compact by the Heine-Borel theorem, the continuous map $y \mapsto d(x, y)$ attains a minimum on $K$. Hence there exists $y \in K$ such that $d(x, y) = \inf_{z \in K} d(y, z) = \operatorname{dist}(x, K) = \operatorname{dist}(x, M)$, which means that $y \in P_M(x)$. Hence $P_M(x)$ is non-empty if $M \subset \mathbb{R}^N$. The same applies to any finite dimensional space.

The argument to prove that $P_M(x)$ is non-empty used above very much depends on the set $K$ to be compact. In the following example we show that $P_M(x)$ can be empty. It is evident that $P_M(x)$ can be empty if $E$ is not complete. It may be more surprising and counter intuitive that even if $E$ is complete, $P_M(x)$ may be empty! There is no mystery about this, it just shows how much our intuition relies on bounded and closed sets to be compact.

**16.3 Example** Let $E := C([0, 1])$ with norm $\|u\| := \|u\|_\infty + \|u\|_1$. We claim that $E$ is complete. Clearly $\|u\|_\infty \le \|u\|$ for all $u \in E$. Also

$$\|u\|_1 = \int_0^1 |u(x)|\, dx \le \|u\|_\infty \int_0^1 1\, dx = \|u\|_\infty$$

for all $x \in E$. Hence $\|\cdot\|$ is equivalent to the supremum norm $\|\cdot\|_\infty$, so convergence with respect to $\|\cdot\|$ is uniform convergence. By Section 7.4 and Lemma 9.4 the space $E$ is complete with respect to the norm $\|\cdot\|$. Now look at the subspace

$$M := \{u \in C([0,1]) : u(0) = 0\}.$$

Since uniform convergence implies pointwise convergence $M$ is closed. (Note that by closed we mean closed as a subset of the metric space $E$, not algebraically closed.) Denote by $\mathbf{1}$ the constant function with value 1. If $u \in M$, then

$$\|\mathbf{1} - u\| \geq \|\mathbf{1} - u\|_\infty \geq |1 - u(0)| = 1, \tag{16.1}$$

so $\operatorname{dist}(\mathbf{1}, M) \geq 1$. If we set $u_n(x) := \sqrt[n]{x}$, then $0 \leq 1 - u_n(x) \to 0$ for all $x \in (0,1]$, so by the dominated convergence theorem

$$\|\mathbf{1} - u_n\| = \|\mathbf{1} - u_n\|_\infty + \int_0^1 1 - u_n(x)\,dx \geq 1 + \int_0^1 1 - u_n(x)\,dx \to 1$$

as $n \to \infty$. Hence $\operatorname{dist}(\mathbf{1}, M) = 1$. We show now that $P_M(\mathbf{1}) = \emptyset$, that is, there is no nearest element in $M$ from $\mathbf{1}$. If $u \in M$, then $u(0) = 0$ and thus by continuity of $u$ there exists an interval $[0, \delta]$ such that $|\mathbf{1} - u(x)| > 1/2$ for all $x \in [0, \delta]$. Hence for all $u \in M$ we have $\|\mathbf{1} - u\|_1 > 0$. Using (16.1) we have

$$\|\mathbf{1} - u\| = \|\mathbf{1} - u\|_\infty + \|\mathbf{1} - u\|_1 \geq 1 + \|\mathbf{1} - u\|_1 > 1$$

for all $u \in M$, so $P_M(\mathbf{1})$ is empty.

In the light of the above example it is a non-trivial fact that $P_M(x)$ is not empty in a Hilbert space, at least if $M$ is closed and convex. By a convex set, as usual, we mean a subset $M$ such that $tx + (1-t)y \in M$ for all $x, y \in M$ and $t \in [0,1]$. In other words, if $x, y$ are in $M$, so is the line segment connecting them. The essential ingredient in the proof is the parallelogram identity from Proposition 15.11.

**16.4 Theorem (Existence and uniqueness of projections)** *Let $H$ be a Hilbert space and $M \subset H$ non-empty, closed and convex. Then $P_M(x)$ contains precisely one element which we also denote by $P_M(x)$.*

*Proof.* Let $M \subset H$ be non-empty, closed and convex. If $x \in M$, then $P_M(x) = x$, so there is existence and also uniqueness of an element of $P_M(x)$. Hence we assume that $x \notin M$ and set

$$\alpha := \operatorname{dist}(x, M) = \inf_{m \in M} \|x - m\|.$$

Since $M$ is closed and $x \notin M$ we have $\alpha > 0$. From the parallelogram identity Proposition 15.11 we get

$$\begin{aligned}\|m_1 - m_2\|^2 &= \|(m_1 - x) - (m_2 - x)\|^2 \\ &= 2\|m_1 - x\|^2 + 2\|m_1 - x\|^2 - \|(m_1 - x) + (m_2 - x)\|^2.\end{aligned}$$

If $m_1, m_2 \in M$, then $\|m_i - x\| \geq \alpha$ for $i = 1, 2$ and by the convexity of $M$ we have $(m_1 + m_2)/2 \in M$. Hence

$$\|(m_1 - x) + (m_2 - x)\| = \|m_1 + m_2 - 2x\| = 2\left\|\frac{m_1 + m_2}{2} - x\right\| \geq 2\alpha.$$

and by using the above

$$\|m_1 - m_2\|^2 \leq 2\|m_1 - x\|^2 + 2\|m_1 - x\|^2 - 4\alpha^2. \tag{16.2}$$

for all $m_1, m_2 \in M$. We can now prove uniqueness. Given $m_1, m_2 \in P_M(x)$ we have by definition $\|m_i - x\| = \alpha$ ($i = 1, 2$), and so by (16.2)

$$\|m_1 - m_2\|^2 \leq 4\alpha^2 - 4\alpha^2 = 0.$$

Hence $\|m_1 - m_2\| = 0$, that is, $m_1 = m_2$ proving uniqueness. As a second step we prove the existence of an element in $P_M(x)$. By definition of an infimum there exists a sequence $(x_n)$ in $M$ such that

$$\|x_n - x\| \to \alpha := \operatorname{dist}(x, M).$$

This obviously implies that $(x_n)$ a bounded sequence in $H$, but since $H$ is not necessarily finite dimensional, we cannot conclude it is converging without further investigation. We show that $(x_n)$ is a Cauchy sequence and therefore converges by the completeness of $H$. Fix now $\varepsilon > 0$. Since $\alpha \leq \|x_n - x\| \to \alpha$ there exists $n_0 \in \mathbb{N}$ such that

$$\alpha \leq \|x_n - x\| \leq \alpha + \varepsilon$$

for all $n > n_0$. Hence using (16.2)

$$\|x_k + x_n\|^2 \leq 2\|x_k - x\|^2 + 2\|x_n - x\|^2 - 4\alpha^2 \leq 4(\alpha + \varepsilon)^2 - 4\alpha^2 = 4(2\alpha + \varepsilon)\varepsilon$$

for all $n, k > n_0$. Hence $(x_n)$ is a Cauchy sequence as claimed. ■

We next derive a geometric characterisation of the projection onto a convex set. If we look at a convex set $M$ in the plane and the nearest point projection $m_x$ from a point $x$ onto $M$, then we expect the angle between $x - m_x$ and $m_x - m$ to be larger or equal than $\pi/2$. This means that the inner product $(x - m_x \mid m_x - m) \leq 0$. We also expect the converse, that is, if the angle is larger or equal to $\pi/2$ for all $m \in M$, then $m_x$ is the projection. Look at Figure 16.2 for an illustration. A similar fact remains true in an arbitrary Hilbert space, except that we have to be careful in a complex Hilbert space because $(x - m_x \mid m_x - m)$ does not need to be real.

**16.5 Theorem** *Suppose $H$ is a Hilbert space and $M \subset H$ a non-empty closed and convex subset. Then for a point $m_x \in M$ the following assertions are equivalent:*

(i) $m_x = P_M(x)$;

(ii) $\operatorname{Re}(m - m_x \mid x - m_x) \leq 0$ *for all* $m \in M$.

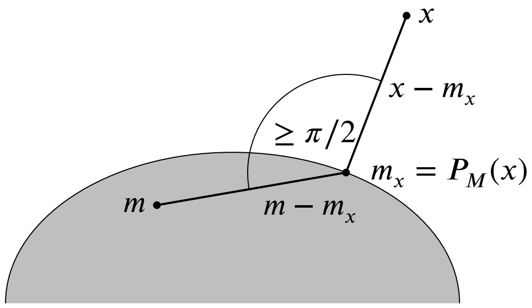



Figure 16.2: Projection onto a convex set

*Proof.* By a translation we can assume that $m_x = 0$. Assuming that $m_x = 0 = P_M(x)$ we prove that $\operatorname{Re}(m \mid x) \leq 0$ for all $m \in M$. By definition of $P_M(x)$ we have $\|x\| = \|x - 0\| = \inf_{m \in M} \|x - m\|$, so $\|x\| \leq \|x - m\|$ for all $m \in M$. As $0, m \in M$ and $M$ is convex we have

$$\|x\|^2 \leq \|x - tm\|^2 = \|x\|^2 + t^2\|m\|^2 - 2t \operatorname{Re}(m \mid x)$$

for all $m \in M$ and $t \in (0, 1]$. Hence

$$\operatorname{Re}(m \mid x) \leq \frac{t}{2}\|m\|^2$$

for all $m \in M$ and $t \in (0, 1]$. If we fix $m \in M$ and let $t$ go to zero, then $\operatorname{Re}(m \mid x) \leq 0$ as claimed. Now assume that $\operatorname{Re}(m \mid x) \leq 0$ for all $m \in M$ and that $0 \in M$. We want to show that $0 = P_M(x)$. If $m \in M$ we then have

$$\|x - m\|^2 = \|x\|^2 + \|m\|^2 - 2 \operatorname{Re}(x \mid m) \geq \|x\|^2$$

since $\operatorname{Re}(m \mid x) \leq 0$ by assumption. As $0 \in M$ we conclude that

$$\|x\| = \inf_{m \in M} \|x - m\|,$$

so $0 = P_M(x)$ as claimed. ■

Every vector subspace $M$ of a Hilbert space is obviously convex. If it is closed, then the above characterisation of the projection can be applied. Due to the linear structure of $M$ it simplifies and the projection turns out to be linear. From Figure 16.3 we expect that $(x - m_x \mid m) = 0$ for all $m \in M$ if $m_x$ is the projection of $x$ onto $M$ and vice versa. The corollary also explains why $P_M$ is called the *orthogonal projection* onto $M$.

**16.6 Corollary** *Let $M$ be a closed subspace of the Hilbert space $H$. Then $m_x = P_M(x)$ if and only if $m_x \in M$ and $(x - m_x \mid m) = 0$ for all $m \in M$. Moreover, $P_M \colon H \to M$ is linear.*

*Proof.* By the above theorem $m_x = P_M(x)$ if and only if $\operatorname{Re}(m_x - x \mid m - m_x) \leq 0$ for all $m \in M$. Since $M$ is a subspace $m + m_x \in M$ for all $m \in M$, so using $m + m_x$ instead of $m$ we get that

$$\operatorname{Re}(m_x - x \mid (m + m_x) - m_x) = \operatorname{Re}(m_x - x \mid m) \leq 0$$

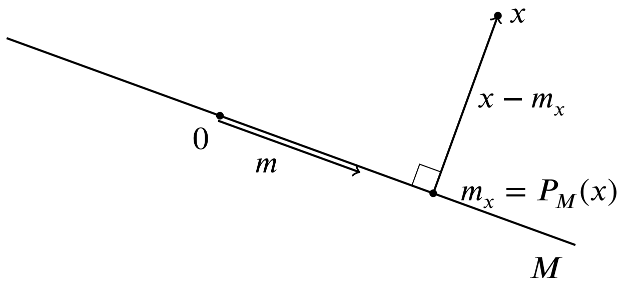



Figure 16.3: Projection onto a closed subspace

for all $m \in M$. Replacing $m$ by $-m$ we get $-\operatorname{Re}(m_x - x \mid m) = \operatorname{Re}(m_x - x \mid -m) \leq 0$, so we must have $\operatorname{Re}(m_x - x \mid m) = 0$ for all $m \in M$. Similarly, replacing $m = \pm im$ if $H$ is a complex Hilbert space we have

$$\pm \operatorname{Im}(m_x - x \mid im) = \operatorname{Re}(m_x - x \mid \pm m) \leq 0,$$

so also $\operatorname{Im}(m_x - x \mid m) = 0$ for all $m \in M$. Hence $(m_x - x \mid m) = 0$ for all $m \in M$ as claimed. It remains to show that $P_M$ is linear. If $x, y \in H$ and $\lambda, \mu \in \mathbb{R}$, then by what we just proved

$$0 = \lambda(x - P_M(x) \mid m) + \mu(x - P_M(y) \mid m) = (\lambda x + \mu y - (\lambda P_M(x) + \mu P_M(y)) \mid m)$$

for all $m \in M$. Hence again by what we proved $P_M(\lambda x + \mu y) = \lambda P_M(x) + \mu P_M(y)$, showing that $P_M$ is linear. ∎

We next connect the projections discussed above with the notion of orthogonal complements.

**16.7 Definition (Orthogonal complement)** For an arbitrary non-empty subset $M$ of an inner product space $H$ we set

$$M^\perp := \{x \in H : (x \mid m) = 0 \text{ for all } m \in M\}.$$

We call $M^\perp$ the *orthogonal complement of $M$ in $H$*.

We now establish some elementary but very useful properties of orthogonal complements.

**16.8 Lemma** *Suppose $M$ is a non-empty subset of the inner product space $H$. Then $M^\perp$ is a closed subspace of $H$ and $M^\perp = \overline{M}^\perp = (\operatorname{span} M)^\perp = (\operatorname{span} \overline{M})^\perp$.*

*Proof.* If $x, y \in M^\perp$ and $\lambda, \mu \in \mathbb{K}$, then

$$(\lambda x + \mu y \mid m) = \lambda(x \mid m) + \mu(y \mid m) = 0,$$

for all $m \in M$, so $M^\perp$ is a subspace of $H$. If $x$ is from the closure of $M^\perp$, then there exist $x_n \in M^\perp$ with $x_n \to x$. By the continuity of the inner product

$$(x \mid m) = \lim_{n\to\infty}(x_n \mid m) = \lim_{n\to\infty} 0 = 0$$

for all $m \in M$. Hence $x \in M^\perp$, showing that $M^\perp$ is closed. We next show that $M^\perp = \overline{M}^\perp$. Since $M \subset \overline{M}$ we have $\overline{M}^\perp \subset M^\perp$ by definition the orthogonal complement. Fix $x \in M^\perp$ and $m \in \overline{M}$. Then there exist $m_n \in M$ with $m_n \to m$. By the continuity of the inner product

$$(x \mid m) = \lim_{n\to\infty}(x \mid m_n) = \lim_{n\to\infty} 0 = 0.$$

Hence $x \in \overline{M}^\perp$ and thus $\overline{M}^\perp \supset M^\perp$, showing that $\overline{M}^\perp = M^\perp$. Next we show that $M^\perp = (\operatorname{span} M)^\perp$. Clearly $(\operatorname{span} M)^\perp \subset M^\perp$ since $M \subset \operatorname{span} M$. Suppose now that $x \in M^\perp$ and $m \in \operatorname{span} M$. Then there exist $m_i \in M$ and $\lambda_i \in \mathbb{K}$, $i = 1, \dots, n$, such that $m = \sum_{i=1}^n \lambda_i m_i$. Hence

$$(x \mid m) = \overline{\lambda}_i \sum_{i=1}^{n} (x \mid m_i) = 0,$$

and thus $x \in (\operatorname{span} M)^\perp$. Therefore $(\operatorname{span} M)^\perp \supset M^\perp$ and so $(\operatorname{span} M)^\perp = M^\perp$ as claimed. The last assertion of the lemma follows by what we have proved above. Indeed we know that $M^\perp = \overline{M}^\perp$ and that $\overline{M}^\perp = (\operatorname{span} \overline{M})^\perp$. ■

We are now ready to prove the main result on orthogonal projections. It is one of the most important and useful facts on Hilbert spaces.

**16.9 Theorem (orthogonal complements)** *Suppose that $M$ is a closed subspace of the Hilbert space $H$. Then*

(i) *$H = M \oplus M^\perp$;*

(ii) *$P_M$ is the projection of $H$ onto $M$ parallel to $M^\perp$ (that is, $P_M(M^\perp) = \{0\}$)*

(iii) *$P_M \in \mathcal{L}(H, M)$ with $\|P_M\|_{\mathcal{L}(H,M)} \leq 1$.*

*Proof.* (i) By Corollary 16.6 we have $(x - P_M(x) \mid m) = 0$ for all $x \in H$ and $m \in M$. Hence $x - P_M(x) \in M^\perp$ for all $x \in H$ and therefore

$$x = P_M(x) + (I - P_M)(x) \in M + M^\perp,$$

and thus $H = M + M^\perp$. If $x \in M \cap M^\perp$, then $(x \mid x) = 0$, so $x = 0$, showing that $H = M \oplus M^\perp$ is a direct sum.

(ii) By Corollary 16.6 the map $P_M$ is linear. Since $P_M(x) = x$ for $x \in M$ we have $P_M^2 = P_M$ and $P_M(M^\perp) = \{0\}$. Hence $P_M$ is a projection.

(iii) By (i) we have $(P_M(x) \mid x - P_M(x)) = 0$ and so

$$\begin{aligned}\|x\|^2 &= \|P_M(x) + (I - P_M)(x)\|^2 \\ &= \|P_M(x)\|^2 + \|x - P_M(x)\|^2 + 2\operatorname{Re}(P_M(x) \mid x - P_M(x)) \geq \|P_M(x)\|^2\end{aligned}$$

for all $x \in H$. Hence $P_M \in \mathcal{L}(H, M)$ with $\|P_M\|_{\mathcal{L}(H,M)} \leq 1$ as claimed. ■

**16.10 Remark** The above theorem in particular implies that for every closed subspace $M$ of a Hilbert space $H$ there exists a *closed* subspace $N$ such that $H = M \oplus N$. We call $M$ a *complemented subspace*. The proof used the existence of a projection. We know from Example 16.3 that projections onto closed subspaces do not necessarily exist in a Banach space, so one may not expect every subspace of a Banach space to be complemented. A rather recent result [9] shows that if every closed subspace of a Banach space is complemented, then its norm is equivalent to a norm induced by an inner product! Hence the above theorem provides a unique property of Hilbert spaces.

The above theorem can be used to prove some properties of orthogonal complements. The first is a very convenient criterion for a subspace of a Hilbert space to be dense.

**16.11 Corollary** *A subspace $M$ of a Hilbert space $H$ is dense in $H$ if and only if $M^\perp = \{0\}$.*

*Proof.* Since $M^\perp = \overline{M}^\perp$ by Lemma 16.8 it follows from Theorem 16.9 that

$$H = \overline{M} \oplus M^\perp$$

for every subspace $M$ of $H$. Hence if $M$ is dense in $H$, then $\overline{M} = H$ and so $M^\perp = \{0\}$. Conversely, if $M^\perp = \{0\}$, then $\overline{M} = H$, that is, $M$ is dense in $H$. ∎

We finally use Theorem 16.9 to get a characterisation of the second orthogonal complement of a set.

**16.12 Corollary** *Suppose $M$ is a non-empty subset of the Hilbert space $H$. Then*

$$M^{\perp\perp} := (M^\perp)^\perp = \overline{\operatorname{span} M}.$$

*Proof.* By Lemma 16.8 we have $M^\perp = (\operatorname{span} M)^\perp = (\overline{\operatorname{span} M})^\perp$. Hence by replacing $M$ by $\overline{\operatorname{span} M}$ we can assume without loss of generality that $M$ is a closed subspace of $H$. We have to show that $M = M^{\perp\perp}$. Since $(x \mid m) = 0$ for all $x \in M$ and $m \in M^\perp$ we have $M \subset M^{\perp\perp}$. Set now $N := M^\perp \cap M^{\perp\perp}$. Since $M$ is a closed subspace it follows from Theorem 16.9 that $M^{\perp\perp} = M \oplus N$. By definition $N \subset M^\perp \cap M^{\perp\perp} = \{0\}$, so $N = \{0\}$, showing that $M = M^{\perp\perp}$. ∎

# 17 Orthogonal Systems

In $\mathbb{R}^N$, the standard basis or any other basis of mutually orthogonal vectors of length one play a special role. We look at generalisations of such bases. Recall that two $u, v$ of an inner product space are called *orthogonal* if $(u \mid v) = 0$.

**17.1 Definition (orthogonal systems)** Let $H$ be an inner product space with inner product $(\cdot \mid \cdot)$ and induced norm $\|\cdot\|$. Let $M \subset H$ be a non-empty subset.

(i) $M$ is called an *orthogonal system* if $(u \mid v) = 0$ for all $u, v \in M$ with $u \neq v$.

(ii) $M$ is called an *orthonormal system* if it is an orthogonal system and $\|u\| = 1$ for all $u \in M$.

(iii) $M$ is called a *complete orthonormal system* or *orthonormal basis* of $H$ if it is an orthogonal system and $\overline{\operatorname{span} M} = H$.

Note that the notion of orthogonal system depends on the particular inner product, so we always have to say with respect to which inner product it is orthogonal.

**17.2 Example** (a) The standard basis in $\mathbb{K}^N$ is a complete orthonormal system in $\mathbb{K}^N$ with respect to the usual dot product.

(b) The set

$$M := \{(2\pi)^{-1/2} e^{inx} : n \in \mathbb{Z}\}$$

forms an orthonormal system in $L_2((-\pi, \pi), \mathbb{C})$. Indeed,

$$\|(2\pi)^{-1/2} e^{inx}\|_2^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx} e^{-inx}\, dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} 1\, dx = 1$$

for all $n \in \mathbb{N}$. Moreover, if $n \neq m$, then

$$\begin{aligned}\Big(\frac{1}{\sqrt{2\pi}} e^{inx} \Big| \frac{1}{\sqrt{2\pi}} e^{imx}\Big) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx} e^{-imx}\, dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)x}\, dx = \frac{1}{2\pi} \frac{1}{i(n-m)} e^{i(n-m)x}\Big|_{-\pi}^{\pi} = 0\end{aligned}$$

since the exponential function is $2\pi i$-periodic. The system is in fact complete. To see this we use the Stone-Weierstrass Theorem 14.2. Clearly $e^{inx}$ are continuous functions. The span of these functions is a sub-algebra of $C([-\pi, \pi])$ because the products $e^{inx} e^{imx} = e^{i(n+m)x}$ are in the sysem for all $n, m \in \mathbb{Z}$. For $n = 0$ we have the constant function. Also the family separates points. Hence the span these functions is dense in $C([-\pi, \pi])$, which in turn is dense in $L^2((-\pi, \pi))$.

(c) The set of real valued functions

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos nx, \frac{1}{\sqrt{\pi}} \sin nx, \qquad n \in \mathbb{N} \setminus \{0\}$$

forms an orthonormal system on $L_2((-\pi, \pi), \mathbb{R})$. Again it turns out that this system is complete. The proof of the orthogonality is a consequence of the trigonometric identities

$$\begin{aligned}\sin mx \sin nx &= \frac{1}{2}\big(\cos(m-n)x - \cos(m+n)x\big) \\ \cos mx \cos nx &= \frac{1}{2}\big(\cos(m-n)x + \cos(m+n)x\big) \\ \sin mx \cos nx &= \frac{1}{2}\big(\sin(m-n)x + \sin(m+n)x\big)\end{aligned}$$

which easily follow from using the standard addition theorems for $\sin(m \pm n)x$ and $\cos(m \pm n)x$

We next show that orthogonal systems are linearly independent if we remove the zero element. Recall that by definition an infinite set is linearly independent if every finite subset is linearly independent. We also prove a generalisation of Pythagoras' theorem.

**17.3 Lemma (Pythagoras theorem)** *Suppose that $H$ is an inner product space and $M$ an orthogonal system in $H$. Then the following assertions are true:*

(i) *$M \setminus \{0\}$ is linearly independent.*

(ii) *If $(x_n)$ is a sequence in $M$ with $x_n \neq x_m$ for $n \neq m$ and $H$ is complete, then $\sum_{k=0}^{\infty} x_k$ converges if and only if $\sum_{k=0}^{\infty} \|x_k\|^2$ converges. In that case*

$$\Big\|\sum_{k=0}^{\infty} x_k\Big\|^2 = \sum_{k=0}^{\infty} \|x_k\|^2. \tag{17.1}$$

*Proof.* (i) We have to show that every finite subset of $M \setminus \{0\}$ is linearly independent. Hence let $x_k \in M \setminus \{0\}$, $k = 1, \dots, n$ be a finite number of distinct elements. Assume that $\lambda_k \in \mathbb{K}$ are such that

$$\sum_{k=0}^{n} \lambda_k x_k = 0.$$

If we fix $x_m$, $m \in \{0, \dots, n\}$, then by the orthogonality

$$0 = \Big(\sum_{k=0}^{n} \lambda_k x_k \Big| x_m\Big) = \sum_{k=0}^{n} \lambda_k (x_k \mid x_m) = \lambda_m \|x_m\|^2.$$

Since $x_m \neq 0$ it follows that $\lambda_m = 0$ for all $m \in \{0, \dots, n\}$, showing that $M \setminus \{0\}$ is linearly independent.

(ii) Let $(x_n)$ be a sequence in $M$ with $x_n \neq x_m$. (We only look at the case of an infinite set because otherwise there are no issues on convergence). We set $s_n := \sum_{k=1}^{n} x_k$ and $t_n := \sum_{k=1}^{n} \|x_k\|^2$ the partial sums of the series under consideration. If $1 \leq m < n$, then by the orthogonality

$$\begin{aligned}\|s_n - s_m\|^2 = \Big\|\sum_{k=m+1}^{n} x_k\Big\|^2 = \Big(\sum_{k=m+1}^{n} x_k \Big| \sum_{j=m+1}^{n} x_j\Big)\\ = \sum_{k=m+1}^{n} \sum_{j=m+1}^{n} (x_k \mid x_j) = \sum_{k=m+1}^{n} \|x_k\|^2 = |t_n - t_m|.\end{aligned}$$

Hence $(s_n)$ is a Cauchy sequence in $H$ if and only if $t_n$ is a Cauchy sequence in $\mathbb{R}$, and by the completeness they either both converge or diverge. The identity (17.1) now follows by setting $m = 0$ in the above calculation and then letting $n \to \infty$. ■

In the case of $H = \mathbb{K}^N$ and the standard basis $e_i$, $i = 1, \dots, N$, we call $x_i = (x \mid e_i)$ the components of $x \in \mathbb{K}^N$. The Euclidean norm is given by

$$\|x\|^2 = \sum_{k=1}^{n} |x_i|^2 = \sum_{k=1}^{n} |(x \mid e_i)|^2.$$

If we do not sum over the full standard basis we may only get an inequality, namely

$$\sum_{k=1}^{m} |(x \mid e_i)|^2 \leq \sum_{k=1}^{n} |(x \mid e_i)|^2 \leq \|x\|^2.$$

if $m \leq n$. We now prove a similar inequality replacing the standard basis by an arbitrary orthonormal system $M$ in an inner product space $H$. From the above reasoning we expect that

$$\sum_{m \in M} |(x \mid m)|^2 \leq \|x\|$$

for all $x \in M$. The definition of an orthonormal system $M$ does not make any assumption on the cardinality of $M$, so it may be uncountable. However, if $M$ is uncountable, it is not clear what the series above means. To make sense of the above series we define

$$\sum_{m \in M} |(x \mid m)|^2 := \sup_{N \subset M \text{ finite}} \sum_{m \in N} |(x \mid m)|^2 \tag{17.2}$$

We now prove the expected inequality.

**17.4 Theorem (Bessel's inequality)** *Let $H$ be an inner product space and $M$ an orthonormal system in $H$. Then*

$$\sum_{m \in M} |(x \mid m)|^2 \leq \|x\|^2 \tag{17.3}$$

*for all $x \in H$. Moreover, the set $\{m \in M : (x \mid m) \neq 0\}$ is at most countable for every $x \in H$.*

*Proof.* Let $N = \{m_k : k = 1, \dots n\}$ be a finite subset of the orthonormal set $M$ in $H$. Then, geometrically,

$$\sum_{k=1}^{n} (x \mid m_k) m_k$$

is the projection of $x$ onto the span of $N$. By Pythagoras theorem (Lemma 17.3) and since $\|m_k\| = 1$ we have

$$\Big\| \sum_{k=1}^{n} (x \mid m_k) m_k \Big\|^2 = \sum_{k=1}^{n} |(x \mid m_k)|^2 \|m_k\|^2 = \sum_{k=1}^{n} |(x \mid m_k)|^2.$$

We expect the norm of the projection to be smaller than the norm of $\|x\|$. To see that we use the properties of the inner product and the above identity to get

$$\begin{aligned} 0 \leq \Big\| x - \sum_{k=1}^{n} (x \mid m_k) m_k \Big\|^2 &= \|x\|^2 + \Big\| \sum_{k=1}^{n} (x \mid m_k) m_k \Big\|^2 \\ &\quad - \sum_{k=1}^{n} \overline{(x \mid m_k)} (x \mid m_k) - \sum_{k=1}^{n} (x \mid m_k)(m_k \mid x) \\ &= \|x\|^2 + \sum_{k=1}^{n} |(x \mid m_k)|^2 - 2 \sum_{k=1}^{n} |(x \mid m_k)|^2 \\ &= \|x\|^2 - \sum_{k=1}^{n} |(x \mid m_k)|^2. \end{aligned}$$

Hence we have shown that

$$\sum_{m\in N} |(x \mid m)|^2 \le \|x\|^2$$

for every finite set $N \subset M$. Taking the supremum over all such finite sets (17.3) follows. To prove the second assertion note that for every given $x \in H$ the sets $M_n := \{m \in M : |(x \mid m)| \ge 1/n\}$ is finite for every $n \in \mathbb{N}$ as otherwise (17.3) could not be true. Since countable unions of finite sets are countable, the set

$$\{m \in M : (x \mid m) \ne 0\} = \bigcup_{n\in\mathbb{N}} M_n$$

is countable as claimed. ■

**17.5 Remark** Since for every $x$ the set $\{m \in M : (x \mid m) \ne 0\}$ is countable we can choose an arbitrary enumeration and write $M_x := \{m \in M : (x \mid m) \ne 0\} = \{m_k : k \in \mathbb{N}\}$. Since the series $\sum_{k=1}^{\infty} |(x \mid m_k)|^2$ has non-negative terms and every such sequence is unconditionally convergent we have

$$\sum_{m\in M}^{\infty} |(x \mid m)|^2 = \sum_{k=1}^{\infty} |(x \mid m_k)|^2$$

no matter which enumeration we take. Recall that unconditionally convergent means that a series converges, and every rearrangement also converges to the same limit. We make this more precise in the next section.

# 18 Abstract Fourier Series

If $x$ is a vector in $\mathbb{K}^N$ and $e_i$ the standard basis, then we know that

$$\sum_{k=1}^{n} (x \mid e_k) e_k$$

is the orthogonal projection of $x$ onto the subspace spanned by $e_1, \dots, e_n$ if $n \le N$, and that

$$x = \sum_{k=1}^{N} (x \mid e_k) e_k.$$

We might therefore expect that the analogous expression

$$\sum_{m\in M} (x \mid m) m \tag{18.1}$$

is the orthogonal projection onto span $M$ if $M$ is an orthonormal system in a Hilbert space $H$. However, there are some difficulties. First of all, $M$ does not need to be countable, so the sum does not necessarily make sense. Since we are not working in $\mathbb{R}$, we cannot use a definition like (17.2). On the other hand, we know from Theorem 17.4 that the set

$$M_x := \{m \in M : (x \mid M) \ne 0\} \tag{18.2}$$

is at most countable. Hence $M_x$ is finite or its elements can be enumerated. If $M_x$ is finite (18.1) makes perfectly good sense. Hence let us assume that $m_k$, $k \in \mathbb{N}$ is an enumeration of $M_x$. Hence, rather than (18.1), we could write

$$\sum_{k=0}^{\infty}(x \mid m_k)m_k.$$

This does still not solve all our problems, because the limit of the series may depend on the particular enumeration chosen. The good news is that this is not the case, and that the series is unconditionally convergent, that is, the series converges and for every bijection $\sigma \colon \mathbb{N} \to \mathbb{N}$ we have

$$\sum_{k=0}^{\infty}(x \mid m_k)m_k = \sum_{k=0}^{\infty}(x \mid m_{\sigma(k)})m_{\sigma(k)}.$$

Recall that the series on the right hand side is called a rearrangement of the series on the left. We now show that (18.1) is actually a projection, not onto span $M$, but onto its closure.

**18.1 Theorem** *Suppose that $M$ is an orthonormal system in a Hilbert space $H$ and set $N := \overline{\operatorname{span} M}$. Let $x \in H$ and $m_k$, $k \in \mathbb{N}$ an enumeration of $M_x$. Then $\sum_{k=0}^{\infty}(x \mid m_k)m_k$ is unconditionally convergent, and*

$$P_N(x) = \sum_{k=0}^{\infty}(x \mid m_k)m_k, \tag{18.3}$$

*where $P_N(x)$ is the orthogonal projection onto $N$ as defined in Section 16.*

*Proof.* Fix $x \in H$. By Theorem 17.4 the set $M_x$ is either finite or countable. We let $m_k$, $k \in \mathbb{N}$ an enumeration of $M_x$, setting for convenience $m_k := 0$ for $k$ larger than the cardinality of $M_x$ if $M_x$ is finite. Again by Theorem 17.4

$$\sum_{k=0}^{\infty}|(x \mid m_k)|^2 \leq \|x\|^2,$$

so by Lemma 17.3 the series

$$y := \sum_{k=0}^{\infty}(x \mid m_k)m_k$$

converges in $H$ since $H$ is complete. We now use the characterisation of projections from Corollary 16.6 to show that $y = P_N(x)$. For $m \in M$ we consider

$$s_n(m) := \Big(\sum_{k=0}^{n}(x \mid m_k)m_k - x \Big| m\Big) = \sum_{k=0}^{n}(x \mid m_k)(m_k \mid m) - (x \mid m).$$

Since the series is convergent, the continuity of the inner product shows that

$$(y - x \mid m) = \lim_{n\to\infty} s_n(m) = \sum_{k=0}^{\infty}(x \mid m_k)(m_k \mid m) - (x \mid m)$$

exists for all $m \in M$. If $m \in M_x$, that is, $m = m_j$ for some $j \in \mathbb{N}$, then by the orthogonality

$$(y - x \mid m) = (x \mid m_j) - (x \mid m_j) = 0.$$

If $m \in M \setminus M_x$, then $(x \mid m) = (m_k \mid m) = 0$ for all $k \in \mathbb{N}$ by definition of $M_x$ and the orthogonality. Hence again $(y - x \mid m) = 0$, showing that $y - x \in M^\perp$. By Lemma 16.8 it follows that $y - x \in \overline{\operatorname{span} M}^\perp$. Now Corollary 16.6 implies that $y = P_N(x)$ as claimed. Since we have worked with an arbitrary enumeration of $M_x$ and $P_N(x)$ is independent of that enumeration, it follows that the series is unconditionally convergent. ■

We have just shown that (18.3) is unconditionally convergent. For this reason we can make the following definition, giving sense to (18.1).

**18.2 Definition (Fourier series)** Let $M$ be an orthonormal system in the Hilbert space $H$. If $x \in H$ we call $(x \mid m)$, $m \in M$, the *Fourier coefficients* of $x$ with respect to $M$. Given an enumeration $m_k$, $k \in \mathbb{N}$ of $M_x$ as defined in (18.2) we set

$$\sum_{m \in M} (x \mid m)m := \sum_{k=0}^{\infty} (x \mid m_k)m_k$$

and call it the *Fourier series* of $x$ with respect to $M$. (For convenience here we let $m_k = 0$ for $k$ larger than the cardinality of $M_x$ if it is finite.)

With the above definition, Theorem 18.1 shows that

$$\sum_{m \in M} (x \mid m)m = P_N(x)$$

for all $x \in H$ if $N := \overline{\operatorname{span} M}$. As a consequence of the above theorem we get the following characterisation of complete orthonormal systems.

**18.3 Theorem (orthonormal bases)** *Suppose that $M$ is an orthonormal system in the Hilbert space $H$. Then the following assertions are equivalent:*

(i) *$M$ is complete;*

(ii) $x = \sum_{m \in M} (x \mid m)m$ *for all $x \in H$ (Fourier series expansion);*

(iii) $\|x\|^2 = \sum_{m \in M} |(x \mid m)|^2$ *for all $x \in H$ (Parseval's identity).*

*Proof.* (i)⇒(ii): If $M$ is complete, then by definition $N := \overline{\operatorname{span} M} = H$ and so by Theorem 18.1

$$x = P_N(x) = \sum_{m \in M} (x \mid m)m$$

for all $x \in H$, proving (ii).

(ii)⇒(iii): By Lemma 17.3 and since $M_x$ is countable we have

$$\|x\|^2 = \Big\| \sum_{m \in M} (x \mid m)m \Big\|^2 = \sum_{m \in M} |(x \mid m)|^2$$

if (ii) holds, so (iii) follows.

(iii)⇒(i): Let $N := \overline{\operatorname{span} M}$ and fix $x \in N^\perp$. By assumption, Theorem 16.9 and 18.1 as well as Lemma 17.3 we have

$$0 = \|P_N(x)\|^2 = \Big\|\sum_{m\in M}(x \mid m)m\Big\|^2 = \sum_{m\in M}|(x \mid m)|^2 = \|x\|^2.$$

Hence $x = 0$, showing that $\overline{\operatorname{span} M}^\perp = \{0\}$. By Corollary 16.11 $\overline{\operatorname{span} M} = H$, that is, $M$ is complete, proving (i). ■

We next provide the connection of the above "abstract Fourier series" to the "classical" Fourier series you may have seen elsewhere. To do so we look at the expansions with respect to the orthonormal systems considered in Example 17.2.

**18.4 Example** (a) Let $e_i$ be the standard basis in $\mathbb{K}^N$. The Fourier "series" of $x \in \mathbb{K}^N$ with respect to $e_i$ is

$$x = \sum_{i=1}^{N}(x \mid e_i)e_i.$$

Of course we do not usually call this a "Fourier series" but say $x_i := (x \mid e_i)$ are the components of the vector $x$ and the above sum the representation of $x$ with respect to the basis $e_i$. The example should just illustrate once more the parallels of Hilbert space theory to various properties of Euclidean spaces.

(b) The Fourier coefficients of $u \in L_2((-\pi,\pi),\mathbb{C})$ with respect to the orthonormal system

$$\frac{1}{\sqrt{2\pi}}e^{inx}, \qquad n \in \mathbb{Z},$$

are given by

$$c_n := \frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi} e^{-inx}u(x)\,dx.$$

Hence the Fourier series of $u$ with respect to the above system is

$$u = \sum_{n\in\mathbb{Z}} c_n e^{inx} = \sum_{n\in\mathbb{Z}} \frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi} e^{-inx}u(x)\,dx e^{inx}.$$

This is precisely the complex form of the classical Fourier series of $u$. As the orthonormal system under considerationis complete, our theory tells us that the series converges in $L_2((-\pi,\pi),\mathbb{C})$, but we do not get any information on pointwise or uniform convergence.

(c) We now look at $u \in L_2((-\pi,\pi),\mathbb{R})$ and its expansion with respect to the orthonormal system given by

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos nx, \frac{1}{\sqrt{\pi}}\sin nx, \qquad n \in \mathbb{N}\setminus\{0\}.$$

The Fourier coefficients are

$$a_0 = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} u(x)\,dx$$
$$a_n = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} u(x)\cos nx\,dx$$
$$b_n = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} u(x)\sin nx\,dx$$

Hence the Fourier series with respect to the above system is

$$u = a_0 + \sum_{n=0}^{\infty}(a_n \cos nx + b_n \sin nx),$$

which is the classical cosine-sine Fourier series. Again convergence is guaranteed in $L_2((-\pi,\pi),\mathbb{R})$, but not pointwise or uniform. The completeness proof is the same as in (b) as we just consider a “change of basis” here.

Orthonormal bases in linear algebra come from diagonalising symmetric matrices associated with a particular problem from applications or otherwise. Similarly, orthogonal systems of functions come by solving partial differential equations by separation of variables. There are many such systems like Legendre and Laguerre polynomials, spherical Harmonics, Hermite functions, Bessel functions and so on. They all fit into the framework discussed in this section if we choose the right Hilbert space of functions with the appropriate inner product. The most difficult issue is to prove the completeness of a particular system. This is often handled very casually as pointed out in a letter by Bôcher to Lord Rayleigh who extensively worked with orthogonal expansions; see [1].

**18.5 Remark** One can also get orthonormal systems from any finite or countable set of linearly independent elements of an inner product space by means of the *Gram-Schmidt orghogonalisation process* as seen in second year algebra.

We have mentioned the possibility of uncountable orthonormal systems or bases. They can occur, but in practice all orthogonal bases arising from applications (like partial differential equations) are countable. Recall that a metric space is separable if it has a countable dense subset.

**18.6 Theorem** *A Hilbert space is separable if and only if it has a countable orthonormal basis.*

*Proof.* If the space $H$ is finite dimensional and $e_i$, $i = 1, \dots N$, is an orthonormal basis of $H$, then the set

$$\operatorname{span}_{\mathbb{Q}}\{e_1, \dots, e_N\} := \Big\{ \sum_{k=1}^{N} \lambda_k e_k : \lambda_k \in \mathbb{Q}(+i\mathbb{Q}) \Big\}$$

is dense in $H$ since $\mathbb{Q}$ is dense in $\mathbb{R}$, so every finite dimensional Hilbert space is separable. Now assume that $H$ is infinite dimensional and that $H$ has a complete countable orthonormal system $M = \{e_k : k \in \mathbb{N}\}$. For every $N \in \mathbb{N}$ we let $H_N :=$

span$\{e_1, \dots, e_N\}$. Then $\dim H_N = N$ and by what we just proved, $H_N$ is separable. Since countable unions of countable sets are countable it follows that countable unions of separable sets are separable. Hence

$$\operatorname{span} M = \bigcup_{N \in \mathbb{N}} H_N$$

is separable. Since $M$ is complete span $M$ is dense. Hence any dense subset of span $M$ is dense in $H$ as well, proving that $H$ is separable. Assume now that $H$ is a separable Hilbert space and let $D := \{x_k : k \in \mathbb{N}\}$ be a dense subset of $H$. We set $H_n := \operatorname{span}\{x_k : k = 1, \dots, n\}$. Then $H_n$ is a nested sequence of finite dimensional subspaces of $H$ whose union contains $D$ and therefore is dense in $H$. We have $\dim H_n \leq \dim H_{n+1}$, possibly with equality. We inductively construct a basis for span $D$ by first choosing a basis of $H_1$. Given a basis for $H_n$ we extend it to a basis of $H_{n+1}$ if $\dim H_{n+1} > \dim H_n$, otherwise we keep the basis we had. Doing that inductively from $n = 1$ will give a basis for $H_n$ for each $n \in \mathbb{N}$. The union of all these bases is a countable linearly independent set spanning span $D$. Applying the Gram-Schmidt orthonormalisation process we can get a countable orthonormal system spanning span $D$. Since span $D$ is dense, it follows that $H$ has a complete countable orthonormal system. ■

Using the above theorem we show that there is, up to an isometric isomorphism, there is only one separable Hilbert space, namely $\ell_2$. Hence $\ell_2$ plays the same role as $\mathbb{K}^N$ is isomorphic to an arbitrary $N$-dimensional space.

**18.7 Corollary** *Every separable infinite dimensional Hilbert space is isometrically isomorphic to $\ell_2$.*

*Proof.* Let $H$ be a separable Hilbert space. Then by Theorem 18.6 $H$ has a countable orthonormal basis $\{e_k : k \in \mathbb{N}\}$. We define a linear map $T : H \to \ell_2$ by setting

$$(Tx)_i := (x \mid e_i)$$

for $x \in H$ and $i \in \mathbb{N}$. (This corresponds to the components of $x$ in case $H = \mathbb{K}^N$.) By Parseval's identity from Theorem 18.3 we have

$$\|x\|^2 = \sum_{i=1}^{\infty} |(x \mid e_i)|^2 = |Tx|_2^2$$

Hence $T$ is an isometry. Hence it remains to show that $T$ is surjective. Let $(\xi_i) \in \ell_2$ and set

$$x := \sum_{i=1}^{\infty} \xi_k e_i$$

Since $(\xi_i) \in \ell_2$ we have

$$\sum_{i=1}^{\infty} |\xi_i|^2 \|e_i\|^2 = \sum_{i=1}^{\infty} |\xi_i|^2 < \infty$$

By Lemma 17.3 the series defining $x$ converges in $H$. Also, by orthogonality, $(x \mid e_i) = \xi_i$, so $Tx = (\xi_i)$. Hence $T$ is surjective and thus an isometric isomorphism between $H$ and $\ell_2$. ■

Chapter V

# Linear Operators

We have discussed basic properties of linear operators already in Section 8. Here we want to prove some quite unexpected results on bounded linear operator on Banach spaces. The essential ingredient to prove these results is Baire's Theorem from topology.

## 19 Baire's Theorem

To prove some fundamental properties of linear operators on Banach spaces we need Baire's theorem on the intersection of dense open sets.

**19.1 Theorem (Baire's Theorem)** *Suppose that $X$ is a complete metric space and that $O_n$, $n \in \mathbb{N}$, are open dense subsets of $X$. Then $\bigcap_{n\in\mathbb{N}} O_n$ is dense in $X$.*

*Proof.* We want to show that for every $x \in X$ and $\varepsilon > 0$ the ball $B(x, \varepsilon)$ and $\bigcap_{n\in\mathbb{N}} O_n$ have non-empty intersection. Fix $x \in X$ and $\varepsilon > 0$. We show that there exist sequences $(x_n)$ in $X$ and $(\varepsilon_n)$ in $\mathbb{R}$ such that $x_1 = x$, $\varepsilon_1 = \varepsilon$, $\varepsilon_n \to 0$ and

$$\overline{B(x_{n+1}, \varepsilon_{n+1})} \subseteq B(x_n, \varepsilon_n) \cap O_n; \tag{19.1}$$

for all $n \in \mathbb{N}$. We prove the existence of such sequences inductively. Fix $x \in X$ and $\varepsilon > 0$. We set $x_1 := x$ and $\varepsilon_1 := \min\{1, \varepsilon\}$. Suppose that we have chosen $x_n$ and $\varepsilon_n$ already. Since $O_n$ is open and dense in $X$ the set $B(x_n, \varepsilon_n) \cap O_n$ is non-empty and open. Hence we can choose $x_{n+1} \in B(x_n, \varepsilon_n)$ and $0 < \varepsilon_{n+1} < \min\{\varepsilon_n, 1/n\}$ such that (19.1) is true.

From the construction we have $\overline{B(x_{n+1}, \varepsilon_{n+1})} \subseteq \overline{B(x_n, \varepsilon_n)}$ for all $n \in \mathbb{N}$ and $\varepsilon_n \to 0$. Since $X$ is complete Cantor's intersection theorem (Theorem 3.11) implies that $\bigcap_{n\in\mathbb{N}} \overline{B(x_n, \varepsilon_n)}$ is non-empty and so we can choose $y \in \bigcap_{n\in\mathbb{N}} \overline{B(x_n, \varepsilon_n)}$. By (19.1) we have

$$y \in \overline{B(x_{n+1}, \varepsilon_{n+1})} \subseteq B(x_n, \varepsilon_n) \cap O_n$$

for every $n \in \mathbb{N}$ and therefore $y \in O_n$ for all $n \in \mathbb{N}$. Hence $y \in \bigcap_{n\in\mathbb{N}} O_n$. By construction $B(x_n, \varepsilon_n) \subseteq B(x, \varepsilon)$ for all $n \in \mathbb{N}$ and so $y \in B(x, \varepsilon)$ as well. Since $x$ and $\varepsilon > 0$ were chosen arbitrarily we conclude that $\bigcap_{n\in\mathbb{N}} O_n$ is dense in $X$. ∎

**19.2 Remark** In Baire's Theorem it is essential that the sets $O_n$ be open in $X$ (or satisfy a suitable other condition). If we set $O_1 = \mathbb{Q}$ and $O_2 = \mathbb{R} \setminus \mathbb{Q}$, then $O_1$ and $O_2$ are dense subsets of $\mathbb{R}$, but $O_1 \cap O_2 = \emptyset$.

We can also reformulate Baire's theorem in terms of properties of closed sets.

**19.3 Corollary** *Suppose that $X$ is a non-empty complete metric space and that $C_n$, $n \in \mathbb{N}$, is a family of closed sets in $X$ with $X = \bigcup_{n\in\mathbb{N}} C_n$. Then there exists $n \in \mathbb{N}$ such that $C_n$ has non-empty interior.*

*Proof.* If $C_n$ has empty interior for all $n \in \mathbb{N}$, then $O_n := X \setminus C_n$ is open and dense in $X$ for all $n \in \mathbb{N}$. By Baire's Theorem the intersection $\bigcap_{n\in\mathbb{N}} O_n$ is dense in $X$. Hence

$$\bigcup_{n\in\mathbb{N}} C_n = \bigcup_{n\in\mathbb{N}} (X \setminus O_n) = X \setminus \Big(\bigcap_{n\in\mathbb{N}} O_n\Big) \neq X.$$

Therefore, $C_n$ must have non-empty interior for some $n \in \mathbb{N}$. ∎

# 20 The Open Mapping Theorem

We know that a map between metric spaces is continuous if and only pre-images of open sets are open. We do not want to look at pre-images, but the images of open sets and look at maps for which such images are open.

**20.1 Definition (open map)** Suppose that $X, Y$ are metric spaces. We call a map $f : X \to Y$ *open* if $f$ maps every open subset of $X$ onto an open subset of $Y$.

**20.2 Remark** To be open is a very strong property of a map. Homeomorphisms are open because the inverse function is continuous. On the other hand, if a continuous function is not surjective such as the constant function, we do not expect it to be open. However, even surjective continuous functions are not necessarily open. As an example consider $f(x) := x(x^2 - 1)$ which is surjective from $\mathbb{R}$ to $\mathbb{R}$. Indeed, $f((-1, 1)) = [-2/3\sqrt{3}, 2/3\sqrt{3}]$ is closed since the function has a maximum and a minimum in $(-1, 1)$.

Our aim is to show that linear surjective maps are open. We start by proving a characterisation of open linear maps.

**20.3 Proposition** *Suppose $E$ is a Banach space and $F$ a normed space. For $T \in \mathcal{L}(E, F)$ the following assertions are equivalent:*

(i) *$T$ is open;*

(ii) *There exists $r > 0$ such that $B(0, r) \subseteq T\big(\overline{B(0, 1)}\big)$;*

(iii) *There exists $r > 0$ such that $B(0, r) \subseteq \overline{T\big(\overline{B(0, 1)}\big)}$.*

*Proof.* We prove (i) implies (ii) and hence also (iii). As $B(0,1)$ is open, the set $T\big(B(0,1)\big)$ is open in $F$ by (i). Since $0 \in T(B(0,1))$ there exists $r > 0$ such that

$$B(0,r) \subseteq T(B(0,1)) \subseteq T\big(\overline{B(0,1)}\big) \subseteq \overline{T\big(\overline{B(0,1)}\big)},$$

so (iii) follows.

We next prove that (iii) implies (ii). This is the most difficult part of the proof. Assume that there exists $r > 0$ such that

$$B(0,r) \subseteq \overline{T\big(\overline{B(0,1)}\big)}.$$

We show that $B(0,r/2) \subseteq T\big(\overline{B(0,1)}\big)$ which proves (ii). Hence let $y \in B(0,r/2)$. Then $2y \in B(0,r)$ and since $B(0,r) \subseteq \overline{T\big(\overline{B(0,1)}\big)}$ there exists $x_1 \in \overline{B(0,1)}$ such that

$$\|2y - Tx_1\| \leq \frac{r}{2}.$$

Hence $4y - 2Tx_1 \in B(0,r)$ and by the same argument as before there exists $x_2 \in \overline{B(0,1)}$ such that

$$\|4y - 2Tx_1 - Tx_2\| \leq \frac{r}{2}.$$

Continuing this way we can construct a sequence $(x_n)$ in $\overline{B(0,1)}$ such that

$$\|2^n y - 2^{n-1}Tx_1 - \cdots - 2Tx_{n-1} - Tx_n\| \leq \frac{r}{2}$$

for all $n \in \mathbb{N}$. Dividing by $2^n$ we get

$$\Big\|y - \sum_{k=1}^{n} 2^{-k}Tx_k\Big\| \leq \frac{r}{2^{n+1}}$$

for all $n \in \mathbb{N}$. Hence

$$y = \sum_{k=1}^{\infty} 2^{-k}Tx_k.$$

Since $\|x_k\| \leq 1$ for all $k \in \mathbb{N}$ we have that

$$\sum_{k=1}^{\infty} 2^{-k}\|x_k\| \leq \sum_{k=1}^{\infty} 2^{-k} = 1$$

and so the series

$$x := \sum_{k=1}^{\infty} 2^{-k}x_k$$

converges absolutely in $E$ because $E$ is complete. Moreover, $\|x\| \leq 1$ and so $x \in \overline{B(0,1)}$. Because $T$ is continuous we have

$$Tx = \lim_{n\to\infty} T\Big(\sum_{k=1}^{n} 2^{-k}x_k\Big) = \lim_{n\to\infty} \sum_{k=1}^{n} 2^{-k}Tx_k = y$$

by construction of $x$. Hence $y \in T\big(\overline{B(0,1)}\big)$ and (ii) follows.

We finally prove that (ii) implies (i). By (ii) and the linearity of $T$ we have

$$T\big(\overline{B(0,\varepsilon)}\big) = \varepsilon T\big(\overline{B(0,1)}\big)$$

for every $\varepsilon > 0$. Because the map $x \to \varepsilon x$ is a homeomorphism on $F$ the set $T\big(\overline{B(0,\varepsilon)}\big)$ is a neighbourhood of zero for every $\varepsilon > 0$. Let now $U \subseteq E$ be open and $y \in T(U)$. As $U$ is open there exists $\varepsilon > 0$ such that

$$\overline{B(x,\varepsilon)} = x + \overline{B(0,\varepsilon)} \subseteq U,$$

where $y = Tx$. Since $z \to x + z$ is a homeomorphism and $T$ is linear we get

$$T\big(\overline{B(x,\varepsilon)}\big) = Tx + T\big(\overline{B(0,\varepsilon)}\big) = y + T\big(\overline{B(0,\varepsilon)}\big) \subseteq T(U).$$

Hence $T\big(\overline{B(x,\varepsilon)}\big)$ is a neighbourhood of $y$ in $T(U)$. As $y$ was an arbitrary point in $T(U)$ it follows that $T(U)$ is open. ■

We next prove a lemma on convex "balanced" sets.

**20.4 Lemma** *Let $E$ be a normed space and $S \subseteq E$ convex with $S = -S := \{-x \colon x \in S\}$. If $\bar{S}$ has non-empty interior, then $\bar{S}$ is a neighbourhood of zero.*

*Proof.* We first show that $\bar{S}$ is convex. If $x, y \in \bar{S}$ and $x_n, y_n \in S$ with $x_n \to x$ and $y_n \to x$, then $tx_n + (1-t)y_n \in S$ for all $n \in \mathbb{N}$ and all $t \in [0,1]$. Letting $n \to \infty$ we get $tx + (1-t)y \in \bar{S}$ for all $t \in [0,1]$, so $\bar{S}$ is convex. Also, if $-x_n \in -S$, then passing to the limit $-x \in -\bar{S}$. Hence also $\bar{S} = -\bar{S}$. If $\bar{S}$ has non-empty interior, then there exist $z \in \bar{S}$ and $\varepsilon > 0$ such that $B(z,\varepsilon) \subseteq \bar{S}$. Therefore $z \pm h \in \bar{S}$ whenever $\|h\| < \varepsilon$ and since $\bar{S} = -\bar{S}$ we also have $-(z \pm h) \in \bar{S}$. By the convexity of $\bar{S}$ we have

$$y = \frac{1}{2}\big((x+h) + (-x+h)\big) \in \bar{S}$$

whenever $\|h\| < \varepsilon$. Hence $B(0,\varepsilon) \subseteq \bar{S}$, so $\bar{S}$ is a neighbourhood of zero. ■

We can now prove our main theorem of this section.

**20.5 Theorem (open mapping theorem)** *Suppose that $E$ and $F$ are Banach spaces. If $T \in \mathcal{L}(E,F)$ is surjective, then $T$ is open.*

*Proof.* As $T$ is surjective we have

$$F = \bigcup_{n \in \mathbb{N}} \overline{T\big(B(0,n)\big)}$$

with $\overline{T\big(B(0,n)\big)}$ closed for all $n \in \mathbb{N}$. Since $F$ is complete, by Corollary 19.3 to Baire's theorem there exists $n \in \mathbb{N}$ such that $\overline{T\big(B(0,n)\big)}$ has non-empty interior. Since the map

$x \to nx$ is a homeomorphism and $T$ is linear, the set $\overline{T\left(\overline{B(0,1)}\right)}$ has non-empty interior as well. Now $\overline{B(0,1)}$ is convex and $\overline{B(0,1)} = -\overline{B(0,1)}$. The linearity of $T$ implies that

$$T\left(\overline{B(0,1)}\right) = -T\left(\overline{B(0,1)}\right)$$

is convex as well. Since we already know that $\overline{T\left(\overline{B(0,1)}\right)}$ has non-empty interior, Lemma 20.4 implies that $\overline{T\left(\overline{B(0,1)}\right)}$ is a neighbourhood of zero, that is, there exists $r > 0$ such that

$$B(0,r) \subseteq \overline{T\left(\overline{B(0,1)}\right)}.$$

Since $E$ is complete Proposition 20.3 implies that $T$ is open. ■

**20.6 Corollary (bounded inverse theorem)** *Suppose that $E$ and $F$ are Banach spaces and that $T \in \mathcal{L}(E,F)$ is bijective. Then $T^{-1} \in \mathcal{L}(F,E)$.*

*Proof.* We know that $T^{-1}\colon F \to E$ is linear. If $U \subseteq E$ is open, then by the open mapping theorem $(T^{-1})^{-1}(U) = T(U)$ is open in $F$ since $T$ is surjective. Hence $T^{-1}$ is continuous by Theorem 5.3, that is, $T^{-1} \in \mathcal{L}(F,E)$. ■

**20.7 Corollary** *Suppose $E$ is a Banach space with respect to the norms $\|\cdot\|_1$ and $\|\cdot\|_2$. If $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$, then the two norms are equivalent.*

*Proof.* Let $E_1 := (E, \|\cdot\|_1)$ and $E_2 := (E, \|\cdot\|_2)$. Since $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$ the map $I\colon x \mapsto x$ is bounded and bijective from $E_2$ into $E_1$. Since $E_1$ and $E_2$ are Banach spaces the inverse map $I^{-1}\colon x \mapsto x$ is bounded as well by the bounded inverse theorem (Corollary 20.6). Hence there exists $c > 0$ such that $\|x\|_2 \leq c\|x\|_1$ for all $x \in E$, proving that the two norms are equivalent. ■

# 21 The Closed Graph Theorem

Suppose that $T\colon E \to F$ is a linear map between the normed spaces $E$ and $F$. Consider the graph

$$G(T) = \{(x, Tx)\colon x \in E\} \subseteq E \times F$$

which is a vector space with norm

$$\|(x, Tx)\|_T = \|x\| + \|Tx\|.$$

If $T$ is continuous, then $G(T)$ is closed as a subset of $E \times F$. The question is whether the converse is true as well. The answer is no in general. An example was given in Assignment 1. In that example we had $E := C^1([-1,1])$ to $F := C([-1,1])$ both with the supremum norm and $Tu = u'$ was a differential operator. We also saw in that assignment that $E$ was not complete with respect to the supremum norm. We now show that this non-completeness makes it possible that an operator with closed graph is unbounded (discontinuous). We prove that this cannot happen if $E$ is complete! The proof is based on the open mapping theorem.

**21.1 Theorem (closed graph theorem)** *Suppose that $E$ and $F$ are Banach spaces and that $T \in \operatorname{Hom}(E, F)$. Then $T \in \mathcal{L}(E, F)$ if and only if $T$ has closed graph.*

*Proof.* If $T \in \mathcal{L}(E, F)$ and $(x_n, Tx_n) \to (x, y)$ then by the continuity $Tx_n \to Tx = y$, so $(x, y) = (x, Tx) \in G(T)$. This shows that $G(T)$ is closed. Now suppose that $G(T)$ is closed. We define linear maps

$$E \xrightarrow{S} G(T) \xrightarrow{P} F, \quad x \mapsto (x, Tx) \mapsto Tx.$$

Then obviously

$$\|P(x, Tx)\|_F = \|Tx\|_F \leq \|(x, Tx)\|_T.$$

Hence $P \in \mathcal{L}(G(T), F)$. Next we note that $S : E \to G(T)$ is an isomorphism because clearly $S$ is surjective and also injective since $Sx = (x, Tx) = (0, 0)$ implies that $x = 0$. Moreover, $S^{-1}(x, Tx) = x$, and therefore

$$\|S^{-1}(x, Tx)\|_E = \|x\|_E \leq \|(x, Tx)\|_T$$

for all $x \in E$. Hence $S^{-1} \in \mathcal{L}(G(T), E)$. Since every closed subspace of a Banach space is a Banach space, $G(T)$ is a Banach space with respect to the graph norm. By the bounded inverse theorem (Corollary 20.6) $S = (S^{-1})^{-1} \in \mathcal{L}(E, G(T))$ and so $T = P \circ S \in \mathcal{L}(E, F)$ as claimed. ■

# 22 The Uniform Boundedness Principle

We show that pointwise boundedness of a family of bounded linear operators between Banach spaces implies the boundedness of the operator norms. Virtually all textbooks give a proof based on Baire's theorem. We give a short elementary proof which is a simplified version of the original proof by Hahn and Banach. This proof was published very recently in [11] and does not rely on Baire's theorem.

**22.1 Theorem (uniform boundedness principle)** *Suppose that $E$ and $F$ are Banach spaces. Let $(T_\alpha)_{\alpha \in A}$ be a family of linear operators in $\mathcal{L}(E, F)$ such that $\sup_{\alpha \in A} \|T_\alpha x\|_F < \infty$ for all $x \in E$. Then $\sup_{\alpha \in A} \|T_\alpha\|_{\mathcal{L}(E,F)} < \infty$.*

*Proof.* Fix $x_0 \in E$ and $r > 0$. Then by definition of the operator norm

$$\begin{aligned}\|Th\| &= \frac{1}{2}\|T(x_0 + h) - T(x_0 - h)\| \leq \frac{1}{2}\big(\|T(x_0 + h)\| + \|T(x_0 - h)\|\big) \\ &\leq \max\big\{\|T(x_0 + h)\|, \|T(x_0 - h)\|\big\} \leq \sup_{\|h\| \leq r} \|T(x_0 + h)\|\end{aligned}$$

for all $h \in E$ with $\|h\| < r$. By definition of the operator norm we therefore get

$$r\|T\|_{\mathcal{L}(E,F)} = \sup_{\|h\| \leq r} \|Th\| \leq \sup_{y \in B(x_0, r)} \|Ty\|. \tag{22.1}$$

Assume now that $\sup_{\alpha \in A} \|T_\alpha\|_{\mathcal{L}(E,F)} = \infty$. Choose a sequence $\alpha_n \in A$ such that

$$\|T_{\alpha_n}\|_{\mathcal{L}(E,F)} > \frac{1}{4^n}. \tag{22.2}$$

Let $x_0 = 0$ and use (22.1) to choose $x_1 \in E$ such that

$$\|x_1 - x_0\| < \frac{1}{3} \qquad \text{and} \qquad \|T_{\alpha_1} x_1\| \geq \frac{2}{3}\frac{1}{3}\|T_{\alpha_1}\|.$$

Then by (22.1) it is possible to choose $x_2 \in E$ such that

$$\|x_2 - x_1\| < \frac{1}{3^2} \qquad \text{and} \qquad \|T_{\alpha_2} x_2\| \geq \frac{2}{3}\frac{1}{3^2}\|T_{\alpha_2}\|.$$

Continuing that way we inductively choose $x_n \in E$ such that

$$\|x_n - x_{n-1}\| < \frac{1}{3^n} \qquad \text{and} \qquad \|T_{\alpha_n} x_n\| \geq \frac{2}{3}\frac{1}{3^n}\|T_{\alpha_n}\|.$$

The sequence $(x_n)$ is a Cauchy sequence since for $n > m \geq 0$

$$\|x_n - x_m\| \leq \sum_{k=m+1}^{n} \|x_k - x_{k-1}\| \leq \sum_{k=m+1}^{n} \frac{1}{3^k} = \frac{1}{2}\Big(\frac{1}{3^m} - \frac{1}{3^n}\Big).$$

By completeness of $E$ that sequence converges, so $x_n \to x$ in $E$. Letting $n \to \infty$ in the above estimate we get

$$\|x - x_m\| \leq \frac{1}{2}\frac{1}{3^m}$$

for all $m \in \mathbb{N}$. Hence by construction of $x_m$ and the reversed triangle inequality

$$\begin{aligned}
\|T_{\alpha_m} x\| &= \|T_{\alpha_m}(x - x_m) + T_{\alpha_m} x_m\| \\
&\geq \|T_{\alpha_m} x_m\| - \|T_{\alpha_m}(x - x_m)\| \geq \frac{2}{3}\frac{1}{3^m}\|T_{\alpha_m}\| - \|T_{\alpha_m}\|\|x - x_m\| \\
&\geq \frac{2}{3}\frac{1}{3^m}\|T_{\alpha_m}\| - \frac{1}{2}\frac{1}{3^m}\|T_{\alpha_m}\| = \frac{1}{3^m}\|T_{\alpha_m}\|\Big(\frac{2}{3} - \frac{1}{2}\Big) = \frac{1}{6}\frac{1}{3^m}\|T_{\alpha_m}\|.
\end{aligned}$$

Using (22.2) we finally get

$$\|T_{\alpha_m} x\| \geq \frac{1}{6}\Big(\frac{4}{3}\Big)^m \to \infty$$

as $m \to \infty$. Hence there exists $x \in E$ such that $\sup_{\alpha \in A} \|T_\alpha x\| = \infty$. Hence the assertion of the theorem follows by contrapositive. ■

**22.2 Remark** There are stronger versions of the uniform boundedness theorem: Either the family is bounded in the operator norm, or it is pointwise unbounded for a rather large set of points (see [10, Theorem 5.8] for a more precise statement).

We can apply the above to countable families of linear operators, and in particular to sequences which converge pointwise. We prove that the limit operator is bounded if it is the pointwise limit of bounded linear operators between Banach spaces. Note that in general, the pointwise limit of continuous functions is not continuous.

**22.3 Corollary** *Suppose that $E$ and $F$ are Banach spaces and that $(T_n)$ is a sequence in $\mathcal{L}(E, F)$ such that $Tx := \lim_{n\to\infty} T_n x$ exists for all $x \in E$. Then $T \in \mathcal{L}(E, F)$ and*

$$\|T\|_{\mathcal{L}(E,F)} \leq \liminf_{n\to\infty} \|T_n\|_{\mathcal{L}(E,F)} < \infty. \tag{22.3}$$

*Proof.* $T$ is linear because for $x, y \in E$ and $\lambda, \mu \in \mathbb{K}$ we have

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

By assumption the sequence $(T_n x)$ is bounded in $F$ for every $x \in E$. Hence the uniform boundedness principle (Theorem 22.1) implies that $\big(\|T_n\|_{\mathcal{L}(E,F)}\big)$ is a bounded sequence as well. By definition of the limit inferior we get

$$\begin{aligned}\|Tx\| = \lim_{n\to\infty} \|T_n x\|_F = \liminf_{n\to\infty} \|T_n x\|_F &= \lim_{n\to\infty}\big(\inf_{k\geq n} \|T_k x\|_F\big) \\ &\leq \lim_{n\to\infty}\big(\inf_{k\geq n} \|T_k\|_{\mathcal{L}(E,F)}\big)\|x\|_E = \big(\liminf_{n\to\infty} \|T_n\|_{\mathcal{L}(E,F)}\big)\|x\|_E < \infty\end{aligned}$$

for all $x \in E$. Hence by definition of the operator norm (22.3) follows. ∎

Note that in the above proof we cannot in general replace the limit inferior by a limit because we only know that the sequence $\|T_n\|$ is bounded.

As an application of the uniform boundedness principles we can prove that there are Fourier series of continuous functions not converging at a given point, say at zero.

**22.4 Example** Let $C_{2\pi}([-\pi,\pi],\mathbb{C})$ be the subspace of $u \in C([-\pi,\pi],\mathbb{C})$ with $u(-\pi) = u(\pi)$ equipped with the supremum norm. It represents the space of continuous functions on $\mathbb{R}$ that are $2\pi$-periodic. As $C_{2\pi}([-\pi,\pi],\mathbb{C}) \subseteq L^2((-\pi,\pi),\mathbb{C})$ it follows from Example 18.4 that the Fourier series of every $f \in C([-\pi,\pi],\mathbb{C})$ converges in $L^2((-\pi,\pi),\mathbb{C})$. The question about pointwise convergence is much more delicate. We use the uniform boundedness principle to show that there are continuous functions, where the Fourier series diverges. To do so let $u \in C([-\pi,\pi],\mathbb{C})$ with Fourier series

$$u(x) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\int_{-\pi}^{\pi} u(t)e^{-ikt}\,dt e^{ikx}$$

We will show that the operators $T_n\colon C_{2\pi}([-\pi,\pi],\mathbb{C}) \to \mathbb{C}$ given by the partial sum

$$T_n u := u_n(0) := \frac{1}{2\pi}\sum_{k=-n}^{n}\int_{-\pi}^{\pi} u(t)e^{-ikt}\,dt \tag{22.4}$$

of the Fourier series evaluated at $x = 0$ is a bounded linear operator. It turns out that $\|T_n\| \to \infty$ as $n \to \infty$, so the family of operators $(T_n)_{n\in\mathbb{N}}$ is not bounded. Hence, by the uniform boundedness principle there must exist $u \in C_{2\pi}([-\pi,\pi],\mathbb{C})$ so that $T_n u = u_n(0)$ is not bounded, that is, the Fourier series of $u$ does not converge at $x = 0$. One can replace $x = 0$ by an arbitrary $x \in [-\pi,\pi]$. One can also use Baire's Theorem to show that for most $u \in C_{2\pi}([-\pi,\pi],\mathbb{C})$, the Fourier series diverges for instance on $[-\pi,\pi] \cap \mathbb{Q}$. We do not do that here, but refer to [10, Section 5.11].

We now show determine $\|T_n\|$ and show that $\|T_n\| \to \infty$ as $n \to \infty$. To do so we set

$$D_n(t) := \sum_{k=-n}^{n} e^{ikt}.$$

We can then write the partial sum of the Fourier series of $u$ in the form

$$u_n(x) = \sum_{k=-n}^{n} c_k e^{ikx} = \frac{1}{2\pi} \sum_{k=-n}^{n} e^{ikx} \int_{-\pi}^{\pi} u(t)e^{-ikt}\,dt$$
$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} u(t) \sum_{k=-n}^{n} e^{ik((x-t)}\,dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(t)D_n(x-t)\,dt.$$

Using the formula for the partial sum of a geometric series we also see that

$$D_n(t) = \sum_{k=-n}^{n} e^{ikt} = e^{-int} \sum_{k=0}^{2n} (e^{it})^k$$
$$= e^{-int} \frac{e^{i(2n+1)t} - 1}{e^{it} - 1} = \frac{e^{i(n+1/2)t} - e^{-i(n+1/2)t}}{e^{it/2} - e^{-it/2}} = \frac{\sin(n+1/2)t}{\sin(t/2)}.$$

The $2\pi$-periodic function $D_n$ is called the *Dirichlet kernel*. To summarise, we obtain

$$D_n(t) = \frac{\sin(n+1/2)t}{\sin(t/2)} \tag{22.5}$$

with a removable singluarity $\lim_{t\to 0} D_n(t) = 2n+1$. Using the Dirichlet kernel we obtain a closed expression for the partial sum of a Fourier series, namely,

$$u_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(t)D_n(x-t)\,dt. \tag{22.6}$$

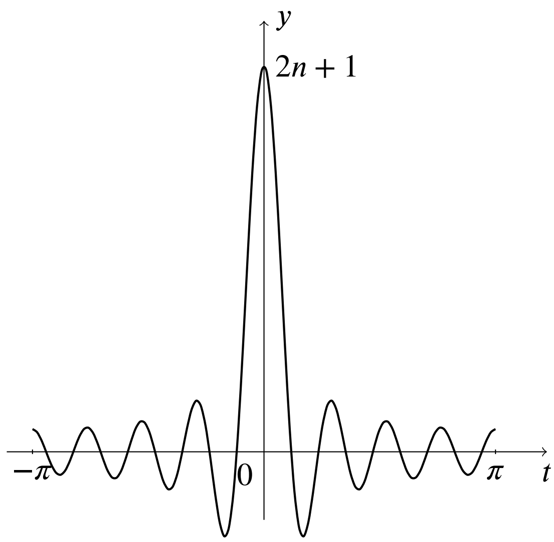



Figure 22.1: The Dirichlet kernel $D_n$.

The operators $T_n : C_{2\pi}([-\pi,\pi],\mathbb{C}) \to \mathbb{C}$ introduced in (22.4) is hence given by

$$T_n u := u_n(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(t)D_n(t)\,dt \tag{22.7}$$

are bounded since, using (22.6).

$$|T_n u| \leq \frac{1}{2\pi} \|D_n\|_1 \|u\|_\infty.$$

Hence,

$$\|T_n\| \leq \frac{1}{2\pi} \|D_n\|_1.$$

We would like to show equality. If we set $u := \operatorname{sign} D_n$, then $u(t)D_n(t) = |D_n(t)|$ and thus

$$|T_n u| = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)|\, dt = \frac{1}{2\pi} \|D_n\|_1.$$

Unfortunately, $u = \operatorname{sign} D_n$ is not continous, otherwise the above would imply that

$$\|T_n\| \geq \frac{1}{2\pi} \|D_n\|_1 \tag{22.8}$$

However, we can approximate $u$ pointwise by functions $u_k \in C_{2\pi}([-\pi, \pi], \mathbb{R})$ with values in $[-1, 1]$ by putting in a line at every jump of with slope $\pm k$ as indicated in Figure 22.2. The dominated convergence theorem then implies that

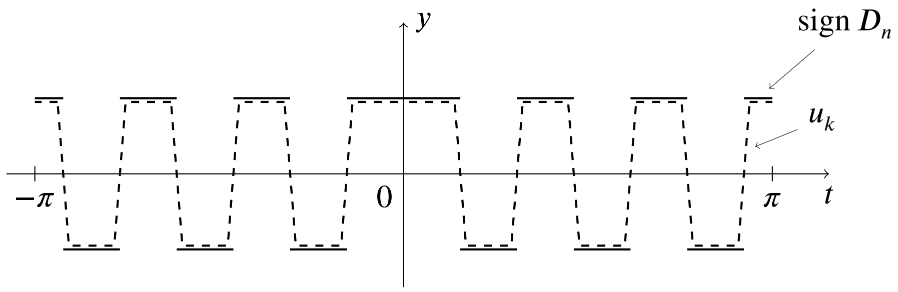



Figure 22.2: Approximating sign $D_n$ by continuous functions.

$$\lim_{k\to\infty} \int_{-\pi}^{\pi} u_k(t) D_n(t)\, dt = \int_{-\pi}^{\pi} \operatorname{sign}(D_n(t)) D_n(t)\, dt = \int_{-\pi}^{\pi} |D_n(t)|\, dt = \|D_n\|_1$$

and hence (22.8) is valid. In particular $2\pi\|T_n\| = \|D_n\|_1$ for all $n \in \mathbb{N}$. We finally need to show that $\|T_n\| \to \infty$ or equivalently that $\|D_n\|_1 \to \infty$ as $n \to \infty$. As $\sin(t/2) \leq |t|/2$ for all $t \in \mathbb{R}$ we first note that

$$\begin{aligned} \int_{-\pi}^{\pi} |D_n(t)|\, dt = 2 \int_0^{\pi} |D_n(t)|\, dt = 2 \int_0^{\pi} \frac{|\sin(n+1/2)t|}{|\sin(t/2)|}\, dt \\ \geq 4 \int_0^{\pi} \frac{|\sin(n+1/2)t|}{t}\, dt =: 4I_n \end{aligned}$$

for all $n \in \mathbb{N}$. Using the substitution $s = (n + 1/2)t$ we see that

$$\begin{aligned} I_n = \int_0^{\pi} \frac{|\sin(n+1/2)t|}{t}\, dt = \int_0^{(n+1/2)\pi} \frac{|\sin s|}{s}\, ds \geq \sum_{k=1}^{n} \int_{(k-1)\pi}^{k\pi} \frac{|\sin s|}{s}\, ds \\ \geq \sum_{k=1}^{n} \frac{1}{k\pi} \int_{(k-1)\pi}^{k\pi} |\sin s|\, ds = \frac{1}{\pi} \sum_{1}^{n} \frac{1}{k} \xrightarrow{n\to\infty} \infty \end{aligned}$$

because of the divergence of the harmonic series. Hence, $\|D_n\|_1 \to \infty$ as claimed.

# 23 Closed Operators

In this section we look at a class of operators which are not necessarily bounded, but still sharing many properties with bounded operators.

**23.1 Definition (closed operator)** Let $E, F$ be Banach spaces. We call $A$ a *closed operator* from $E$ to $F$ with domain $D(A)$ if $D(A)$ is a subspace of $E$ and the graph

$$G(A) := \{(x, Ax) : x \in D(A)\} \subseteq E \times F$$

is closed in $E \times F$. We write $A : D(A) \subseteq E \to F$.

**23.2 Remark** (a) The graph of a linear operator $A : D(A) \subseteq E \to F$ is closed if and only if $x_n \to x$ in $E$ and $Ax_n \to y$ in $F$ imply that $x \in D(A)$ and $Ax = y$. The condition looks similar to continuity. The difference is that continuity means that $x_n \to x$ in $E$ implies that the sequence $(Ax_n)$ automatically converges, whereas this is assumed in case of a closed operator.

(b) Note that $A : D(A) \to F$ is always continuous if we consider $D(A)$ with the graph norm $\|x\|_A := \|x\|_E + \|Ax\|_F$.

Classical examples of closed but unbounded operators are differential operators. We demonstrate this with the simplest possible example.

**23.3 Example** Consider the vector space of continuous functions $E = C([0,1])$. Define the operator $Au := u'$ with domain

$$D(A) := C^1([0,1]) := \{u : [0,1] \to \mathbb{R} \mid u' \in C([0,1])\}.$$

We show that $A : D(A) \subseteq E \to E$ is a closed operator which is not bounded.

A standard result from analysis asserts that if $u_n \to u$ pointwise and $u_n' \to v$ uniformly, then $u$ is differentiable with $u' = v \in C([0,1])$. Suppose that $u_n \to u$ and $Au_n \to v$ in $E$. Since convergence in the supremum norm is the same as uniform convergence we have in particular that $u_n \to u$ pointwise and $Au_n = u_n' \to v$ uniformly. Hence $u \in C^1([0,1]) = D(A)$ and $Au = u' = v$. Hence $A$ is a closed operator. We next show that $A$ is not bounded. To do so let $u_n(x) := x^n$. Then $\|u_n\|_\infty = 1$ for all $n \in \mathbb{N}$, but $\|u_n'\|_\infty = n \to \infty$ as $n \to \infty$. Hence $A$ maps a bounded set of $D(A)$ onto an unbounded set in $E$ and therefore $A$ is not continuous by Theorem 8.3.

We next prove some basic properties of closed operators:

**23.4 Theorem** *Suppose that $E$ and $F$ are Banach spaces and that $A : D(A) \subseteq E \to F$ is an injective closed operator.*

(i) *Then $A^{-1} : \operatorname{im}(A) \subseteq F \to E$ is a closed operator with domain $D(A^{-1}) = \operatorname{im}(A)$.*

(ii) *If $\operatorname{im}(A) = F$, then $A^{-1} \in \mathcal{L}(F, E)$ is bounded.*

(iii) *If $\overline{\operatorname{im}(A)} = F$ and $A^{-1} : \operatorname{im}(A) \subseteq F \to E$ is bounded, then $\operatorname{im}(A) = F$ and $A^{-1} \in \mathcal{L}(F, E)$.*

*Proof.* (i) Note that

$$G(A) = \{(x, Ax) : x \in D(A)\} = \{(A^{-1}y, y) : y \in \operatorname{im}(A)\}$$

is closed in $E \times F$. Since the map $(x, y) \to (y, x)$ is an isometric isomorphism from $E \times F$ to $F \times E$, the graph of $A^{-1}$ is closed in $F \times E$.

(ii) By part (i) the operator $A^{-1} : \operatorname{im}(A) = F \to E$ is closed. Since $E, F$ are Banach spaces, the closed graph theorem (Theorem 21.1) implies that $A^{-1} \in \mathcal{L}(F, E)$.

(iii) Suppose $y \in F$. Since $\operatorname{im}(A)$ is dense there exist $y_n \in \operatorname{im}(A)$ such that $y_n \to y$ in $F$. If we set $x_n := A^{-1}y_n$, then by the boundedness of $A^{-1}$ we have $\|x_n - x_m\|_F \leq \|A^{-1}\|_{\mathcal{L}(\operatorname{im}(A),E)}\|y_n - y_m\|_E$ for all $n, m \in \mathbb{N}$. Since $y_n \to y$ it follows that $(x_n)$ is a Cauchy sequence in $E$, so by completeness of $E$ we have $x_n = A^{-1}y_n \to x$ in $E$. Since $A^{-1}$ is closed by (i) we get that $y \in D(A^{-1}) = \operatorname{im}(A)$ and $A^{-1}y = x$. Since $y \in F$ was arbitrary we get $\operatorname{im}(A) = F$. ■

We can also characterise closed operators in terms of properties of the graph norm on $D(A)$ as introduced in Definition 9.5.

**23.5 Proposition** *Suppose that $E$ and $F$ are Banach spaces. Then, the operator $A : D(A) \subseteq E \to F$ is closed if and only if $D(A)$ is a Banach space with respect to the graph norm $\|x\|_A := \|x\|_E + \|Ax\|_F$.*

*Proof.* We first assume that $A$ is closed and show that $D(A)$ is complete with respect to the graph norm. Let $(x_n)$ be a Cauchy sequence in $(D(A), \|\cdot\|_A)$. Fix $\varepsilon > 0$. Then there exists $n_0 \in \mathbb{N}$ such that

$$\|x_n - x_m\|_A = \|x_n - x_m\|_E + \|Ax_n - Ax_m\|_F < \varepsilon$$

for all $m, n > n_0$. Hence $(x_n)$ and $(Ax_n)$ are Cauchy sequences in $E$ and $F$, respectively. By the completeness of $E$ and $F$ we conclude that $x_n \to x$ and $Ax_n \to y$. Since $A$ is closed we have $x \in D(A)$ and $Ax = y$, so

$$\|x_n - x\|_A = \|x_n - x\|_E + \|Ax_n - Ax\|_F \to 0.$$

Hence, $x_n \to x$ in $(D(A), \|\cdot\|_A)$, proving that $(D(A), \|\cdot\|_A)$ is a Banach space.

Suppose now that $(D(A), \|\cdot\|_A)$ is a Banach space. Suppose that $x_n \to x$ in $E$ and $Ax_n \to y$ in $F$. Since $(x_n)$ is a Cauchy sequence in $E$ and $(Ax_n)$ a Cauchy sequence in $F$, the definition of the graph norm

$$\|x_n - x_m\|_A = \|x_n - x_m\|_E + \|Ax_n - Ax_m\|_F$$

implies that $(x_n)$ is a Cauchy sequence with respect to the graph norm in $D(A)$. By the completeness of $D(A)$ with respect to the graph norm we have $x_n \to x$ in the graph norm. By the definition of the graph norm this in particular implies $x \in D(A)$ and $Ax_n \to Ax$, so $Ax = y$. Hence $A$ is closed. ■

In the following corollary we look at the relationship between continuous and closed operators. Continuous means with respect to the norm on $D(A)$ induced by the norm in $E$.

**23.6 Corollary** *Suppose that $E, F$ are Banach spaces and that $A\colon D(A) \subseteq E \to F$ is a linear operator. Then $A \in \mathcal{L}(D(A), F)$ is closed if and only if $D(A)$ is closed in $E$.*

*Proof.* Suppose that $A \in \mathcal{L}(D(A), F)$. By Proposition 9.7, the graph norm on $D(A)$ is equivalent to the norm of $E$ restricted to $D(A)$. Since $A$ is closed we know from Proposition 23.5 that $D(A)$ is complete with respect to the graph norm, and therefore with respect to the norm in $E$. Hence $D(A)$ is closed in $E$. Suppose now that $D(A)$ is closed in $E$. Since $E$ is a Banach space, also $D(A)$ is a Banach space. Hence, the closed graph theorem (Theorem 21.1) implies that $A \in \mathcal{L}(D(A), F)$. ■

As a consequence of the above result the domain of a closed and unbounded operator $A\colon D(A) \subseteq E \to F$ is never a closed subset of $E$.

# 24 Closable Operators and Examples

Often operators are initially defined on a relatively small domain with the disadvantage that they are not closed. We show now that under a suitable assumption there exists a closed operator extending such a linear operator.

**24.1 Definition (extension/closure of a linear operator)** (i) Suppose $A\colon D(A) \subseteq E \to F$ is a linear operator. We call the linear operator $B\colon D(B) \subseteq E \to F$ and *extension of* $A$ if $D(A) \subseteq D(B)$ and $Ax = Bx$ for all $x \in D(A)$. We write $B \supseteq A$ if that is the case.

(ii) The linear operator $A\colon D(A) \subseteq E \to F$ is called closable if there exists a closed operator $B\colon D(B) \subseteq E \to F$ with $B \supseteq A$.

(iii) We call $\bar{A}$ the closure of $A$ if $\bar{A}$ is closed, $\bar{A} \supseteq \bar{A}$, and $B \supseteq \bar{A}$ for every closed operator with $B \supseteq A$.

**24.2 Theorem (closure of linear operator)** *Suppose that $A\colon D(A) \subseteq E \to F$ is a linear operator. Then the following assertions are equivalent:*

(i) *$A$ is closable.*

(ii) *If $x_n \to 0$ in $E$ and $Ax_n \to y$ in $F$, then $y = 0$.*

*In that case $\overline{G(A)} = G(\bar{A}) \subseteq E \times F$ is the graph of the closure $\bar{A}$, and $D(\bar{A}) = \{x \in E\colon (x, y) \in \overline{G(A)}\}$ its domain.*

*Proof.* Suppose that (i) is true and that $B$ is a closed operator with $B \supseteq A$. If $x_n \in D(A)$, $x_n \to 0$ and $Ax_n \to y$, then $x \in D(B)$ and $Bx_n = Ax_n \to y$ as well. Since $B$ is closed $B0 = 0 = y$, so (ii) follows.

Suppose now that (ii) holds. We show that $\overline{G(A)}$ is the graph of a linear operator. To do so we let $(x, y), (x, \tilde{y}) \in \overline{G(A)}$ and show that $y = \tilde{y}$. By choice of $(x, y)$ there exist $x_n \in D(A)$ such that $x_n \to x$ in $E$ and $Ax_n \to y$ in $F$. Similarly there exist $z_n \in D(A)$ with $z_n \to x$ and $Az_n \to \tilde{y}$. Then $x_n - z_n \to 0$ and $A(x_n - z_n) \to (y - \tilde{y})$. By assumption (ii) we conclude that $y - \tilde{y} = 0$, that is, $y = \tilde{y}$. Hence we can define an operator $\bar{A}$ by setting $\bar{A}x := y$ for all $(x, y) \in \overline{G(A)}$. Its domain is given by

$$D(\bar{A}) := \{x \in E\colon (x, y) \in \overline{G(A)}\}$$

and its graph by $G(\bar{A}) = \overline{G(A)}$. If $x, z \in D(\bar{A})$, then by the above construction there exist sequences $(x_n)$ and $(z_n)$ in $D(A)$ with $Ax_n \to \bar{A}x$ and $Az_n \to \bar{A}z$. Hence

$$A(\lambda x_n + \mu z_n) = \lambda A x_n + \mu A z_n \to \lambda \bar{A} x + \mu \bar{A} z$$

Since $\lambda x_n + \mu z_n \to \lambda x + \mu z$, by definition of $\bar{A}$ we get $(\lambda x + \mu z, \lambda \bar{A} x + \mu \bar{A} z) \in G(\bar{A})$ and so $\bar{A}(\lambda x + \mu z) = \lambda \bar{A} x + \mu \bar{A} z$. Hence $D(\bar{A})$ is a subspace of $E$ and $\bar{A}$ is linear. Because $G(\bar{A}) = \overline{G(A)}$ is closed, it follows that $\bar{A}$ is a closed operator. Moreover, $\bar{A}$ is the closure of $A$ because $\overline{G(A)}$ must be contained in the graph of any closed operator extending $A$. ■

We can use the above in particular to construct extensions of bounded linear operators defined on dense subspaces.

**24.3 Corollary (Extension of linear operators)** *Suppose that $E, F$ are Banach spaces, and that $E_0$ is a dense subspace of $E$. If $T_0 : \mathcal{L}(E_0, F)$, then there exists a unique operator $T \in \mathcal{L}(E, F)$ such that $Tx = T_0 x$ for all $x \in E_0$. Moreover, $\|T_0\|_{\mathcal{L}(E_0,F)} = \|T\|_{\mathcal{L}(E,F)}$.*

*Proof.* We consider $T_0$ as a linear operator on $E$ with domain $D(T_0) = E_0$. Suppose that $x_n \in E_0$ with $x_n \to 0$ and $T_0 x_n \to y$. Since $T_0$ is bounded we have $\|T_0 x_n\|_F \leq \|T_0\|_{\mathcal{L}(E_0,F)} \|x_n\|_E \to 0$, so $y = 0$. By Theorem 24.2 $T_0$ has a closure which we denote by $T$. We show that $D(T) = E$ and that $T$ is bounded. If $x_n \in E_0$ and $x_n \to x$ in $E$, then by definition of $Tx$ we have

$$\|Tx\| = \lim_{n\to\infty} \|T_0 x_n\|_F \leq \|T_0\|_{\mathcal{L}(E_0,F)} \lim_{n\to\infty} \|x_n\|_E \leq \|T_0\|_{\mathcal{L}(E_0,F)} \|x\|_E.$$

Hence $\|T\|_{\mathcal{L}(E,F)} \leq \|T_0\|_{\mathcal{L}(E_0,F)}$. Since clearly $\|T\|_{\mathcal{L}(E,F)} \geq \|T_0\|_{\mathcal{L}(E_0,F)}$, the equality of the operator norms follow. Now Corollary 23.6 implies that $D(T)$ is closed in $E$, and since $E_0 \subseteq D(T)$ is dense in $E$ we get $D(T) = E$. ■

We complete this section by an example of a closable operator arising in reformulating a Sturm-Liouville boundary value problem in a Hilbert space setting. The approach can be used for much more general boundary value problems for partial differential equations. We need one class of functions to deal with such problems.

**24.4 Definition (support/test function)** Let $\Omega \subseteq \mathbb{R}^N$ be an open set and $u : \Omega \to \mathbb{K}$ a function. The set

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

is called the *support* of $u$. We let

$$C^\infty(\Omega) := \{u : \Omega \to \mathbb{K} : u \text{ has partial derivatives of all orders}\}$$

and

$$C_c^\infty(\Omega) := \{u \in C^\infty(\Omega) : \operatorname{supp}(u) \subseteq \Omega \text{ is compact}\}.$$

The elements of $C_c^\infty(\Omega)$ are called *test functions*.

**24.5 Example** Let $I = [a, b]$ be a compact interval, $p \in C^1(I)$, $q \in C(I)$ with $p(x) > 0$ for all $x \in I$ and $\alpha_i, \beta_i \in \mathbb{R}$ with $\alpha_i^2 + \beta_i^2 \neq 0$ $(i = 1, 2)$. We want to reformulate the Sturm-Liouville boundary value problem

$$\begin{aligned} -(pu')' + qu &= f \quad \text{in } (a, b), \\ \alpha_1 u(a) - \beta_1 u'(a) &= 0 \\ \alpha_2 u(a) + \beta_2 u'(a) &= 0 \end{aligned} \tag{24.1}$$

in a Hilbert space setting as a problem in $L_2((a, b), \mathbb{R})$. Functions in $L_2((a, b), \mathbb{R})$ are not generally differentiable, so we first define a linear operator on a smaller domain by setting

$$A_0 := -(pu')' + qu$$

for all $u \in D(A_0)$, where

$$D(A_0) := \{u \in C^\infty(I) : \alpha_1 u(a) - \beta_1 u'(a) = \alpha_2 u(a) + \beta_2 u'(a) = 0\}.$$

The idea is to deal with the boundary conditions by incorporating them into the definition of the domain of the differential operator. We note that if $A_0 u = f$ with $u \in D(A_0)$, then $u$ is a solution of (24.1).

We would like to deal with a closed operator, but unfortunately, $A_0$ is not closed. We will show that $A_0$ is closable as an operator in $L_2((a, b), \mathbb{R})$. Clearly $D(A_0) \subseteq L_2((a, b), \mathbb{R})$. Moreover, $C_c^\infty((a, b)) \subseteq D(A_0)$. Using integration by parts twice we see that

$$\int_a^b \varphi A_0 u \, dx = \int_a^b u A_0 \varphi \, dx + p(u\varphi' - \varphi u')\Big|_a^b = \int_a^b u A_0 \varphi \, dx \tag{24.2}$$

for all $u \in D(A_0)$ and $\varphi \in C_c^\infty((a, b))$ because $\varphi(a) = \varphi'(a) = \varphi(b) = \varphi'(b) = 0$. Suppose now that $u_n \in D(A_0)$ such that $u_n \to 0$ in $L_2((a, b), \mathbb{R})$ and $A_0 u_n \to v$ in $L_2((a, b), \mathbb{R})$. Using (24.2) we and the Cauchy-Schwarz inequality we get

$$\Big|\int_a^b \varphi A_0 u_n \, dx\Big| = \Big|\int_a^b u_n A_0 \varphi \, dx\Big| \leq \|u_n\|_2 \|A_0 \varphi\|_2 \to 0$$

as $n \to \infty$. Therefore,

$$\int_a^b \varphi A_0 u_n \, dx \to \int_a^b \varphi v \, dx = 0$$

for all $\varphi \in C_c^\infty((a, b))$. One can show that $C_c^\infty((a, b))$ is dense in $L_2((a, b), \mathbb{R})$, so that $v = 0$ in $L_2((a, b), \mathbb{R})$. By Theorem 24.2 the operator $A_0$ is closable. Denote its closure by $A$ with domain $D(A)$. Since $C_c^\infty((a, b))$ is dense in $L_2((a, b), \mathbb{R})$ and $C_c^\infty((a, b)) \subseteq D(A_0)$ it follows that $D(A)$ is a dense subset of $L_2((a, b), \mathbb{R})$.

Instead of (24.1) it is common to study the abstract equation

$$Au = f$$

in $L_2((a, b), \mathbb{R})$. The solutions may not be differentiable, but if they are, then they are solutions to (24.1). We call solutions of the abstract equation *generalised solutions* of (24.1). The advantage of that setting is that we can use all the Hilbert space theory including Fourier series expansions of solutions.

Chapter VI

# Duality

## 25 Dual Spaces

In this section we discuss the space of linear operators from a normed space over $\mathbb{K}$ into $\mathbb{K}$. This is called the "dual space" and plays an important role in many applications. We give some precise definition and then examples.

**25.1 Definition (Dual space)** If $E$ is a normed vector space over $\mathbb{K}$, then we call

$$E' := \mathcal{L}(E, \mathbb{K})$$

the *dual space* of $E$. The elements of $E'$ are often referred to as *bounded linear functionals on* $E$. The operator norm on $\mathcal{L}(E, \mathbb{K})$ is called the dual norm on $E'$ and is denoted by $\|\cdot\|_{E'}$. If $f \in E'$, then we define

$$\langle f, x\rangle := f(x).$$

**25.2 Remark** (a) As $\mathbb{K}$ (= $\mathbb{R}$ or $\mathbb{C}$) is complete it follows from Theorem 8.8 that $E'$ is a Banach space, even if $E$ is not complete.

(b) For fixed $x \in E$, the map

$$E' \to \mathbb{K},\ f \to \langle f, x\rangle$$

is obviously linear. Hence

$$\langle \cdot\,, \cdot\rangle\colon\ E' \times E \to \mathbb{K}$$

is a bilinear map. It is called the *duality map* on $E$. Despite the similarity in notation, $\langle \cdot\,, \cdot\rangle$ is not an inner product, and we want to use different notation for the two concepts.

(c) By definition of the operator norm, the dual norm is given by

$$\|f\|_{E'} = \sup_{x\in E\setminus\{0\}} \frac{|\langle f, x\rangle|}{\|x\|_E} = \sup_{\|x\|_E\leq 1} |\langle f, x\rangle|$$

and as a consequence

$$|\langle f, x\rangle| \leq \|f\|_{E'}\|x\|_E$$

for all $f \in E'$ and $x \in E$. Hence the duality map $\langle \cdot\,, \cdot\rangle\colon\ E' \times E \to \mathbb{K}$ is continuous.

We next give some examples of dual spaces. We also look at specific representations of dual spaces and the associated duality map, which is useful in many applications. These representations are often not easy to get.

**25.3 Example** Let $E = \mathbb{K}^N$. Then the dual space is given by all linear maps from $\mathbb{K}^N$ to $\mathbb{K}$. Such maps can be represented by a $1 \times N$ matrix. Hence the elements of $\mathbb{K}^N$ are column vectors, and the elements of $(\mathbb{K}^N)'$ are row vectors, and the duality map is simply matrix multiplication. More precisely, if

$$f = [f_1, \dots, f_N] \in (\mathbb{K}^N)' \qquad \text{and} \qquad x = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \in \mathbb{K}^N,$$

then

$$\langle f, x\rangle = [f_1, \dots, f_N] \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = \sum_{k=1}^{N} f_k x_k,$$

which is almost like the dot product except for the complex conjugate in the second argument. By identifying the row vector $f$ with the column vector $f^T$ we can identify the dual of $\mathbb{K}^N$ with itself.

Note however, that the dual norm depends by definition on the norm of the original space. We want to compute the dual norm for the above identification by looking at $E_p = \mathbb{K}^N$ with the $p$-norm as defined in (7.1). We are going to show that, with the above identification,

$$E_p' = E_{p'} \qquad \Bigl(\frac{1}{p} + \frac{1}{p'} = 1\Bigr)$$

with equal norms. By Hölder's inequality (Proposition 7.2)

$$|\langle f, x\rangle| \leq |f|_{p'} |x|_p, \tag{25.1}$$

for all $x \in \mathbb{K}^N$ and $1 \leq p \leq \infty$. Hence by definition of the dual norm $\|f\|_{E_p'} \leq |f|_{p'}$. We show that there is equality for $1 \leq p \leq \infty$, that is,

$$\|f\|_{E_p'} = |f|_{p'}. \tag{25.2}$$

We do this by choosing a suitable $x$ in (25.1). Let $\lambda_k \in \mathbb{K}$ such that $|\lambda_k| = 1$ and $|f_k| = \lambda_k f_k$. First assume that $p \in (1, \infty]$. For $k = 1, \dots, N$ we set

$$x_k := \begin{cases} \lambda_k |f_k|^{1/(p-1)} & \text{if } p \in (1, \infty) \\ \lambda_k & \text{if } p = \infty. \end{cases}$$

Then, noting that $p' = p/(p-1)$ if $p \in (1, \infty)$ and $p' = 1$ if $p = \infty$, we have

$$f_k x_k = \begin{cases} f_k \lambda_k |f_k|^{1/(p-1)} = |f_k|^{p/(p-1)} = |f_k|^{p'} & \text{if } p \in (1, \infty), \\ f_k \lambda_k = |f_k| = |f_k|^{p'} & \text{if } p = \infty, \end{cases}$$

and hence for that choice of $x$

$$\langle f, x\rangle = \sum_{k=1}^{n} |f_k|^{p'} = |f|_{p'} |f|_{p'}^{p'-1}.$$

Now if $p = \infty$ then $|f|_{p'}^{p'-1} = |f|_1^0 = 1$, so $\|f\|_{E_p'} = |f|_1$. If $p \in (1, \infty)$ then, by definition of $x$ and since $p' = p/(p-1)$, we have

$$|f_k|^{p'-1} = |f_k|^{1/(p-1)} = |x_k|$$

and so

$$|f|_{p'}^{p'-1} = \Big(\sum_{k=1}^{N} |f_k|^{p/(p-1)}\Big)^{1-1/p'} = \Big(\sum_{k=1}^{N} |x_k|^p\Big)^{1/p} = |x|_p.$$

Together with (25.1) we conclude that (25.2) is true if $1 < p \leq \infty$. In case $p = 1$ we let $j \in \{1, \dots, N\}$ such that $|f_j| = \max_{k=1,\dots,N} |f_k|$. Then set $x_j := \lambda_j$ ($\lambda_j$ as defined above) and $x_k = 0$ otherwise. Then

$$\langle f, x\rangle = \lambda_j f_j = |f_j| = \max_{k=1,\dots,N} |f_k| = |f|_\infty,$$

so $\|f\|_{E_p'} = |f|_\infty$ if $p = 1$ as well.

We next look at an infinite dimensional analogue to the above example, namely the sequence spaces $\ell_p$ from Section 7.2.

**25.4 Example** For two sequences $x = (x_k)$ and $y = (y_k)$ in $\mathbb{K}$ we set

$$\langle x, y\rangle := \sum_{k=0}^{\infty} x_k y_k$$

whenever the series converges. By Hölder's inequality (see Proposition 7.4)

$$|\langle x, y\rangle| \leq |x|_p |y|_{p'}$$

for all $x \in \ell_p$ and $y \in \ell_{p'}$, where and $1 \leq p \leq \infty$. This means that for every fixed $y \in \ell_{p'}$, the linear functional $x \mapsto \langle y, x\rangle$ is bounded, and therefore an element of the dual space $(\ell_p)'$. Hence, if we identify $y \in \ell_{p'}$ with the map $x \mapsto \langle y, x\rangle$, then we naturally have

$$\ell_{p'} \subseteq \ell_p'$$

for $1 \leq p \leq \infty$. The above identification is essentially the same as the one in the previous example on finite dimensional spaces. We prove now that

(i) $\ell_p' = \ell_{p'}$ if $1 \leq p < \infty$ with equal norm;

(ii) $c_0' = \ell_1$ with equal norm.

Most of the proof works for (i) and (ii). Given $f \in \ell_p'$ we need to find $y \in \ell_{p'}$ such that $\langle f, x\rangle = \sum_{k=0}^{\infty} y_k x_k$ for all $x \in \ell_p$. We set $e_k := (\delta_{jk}) \in \ell_p$. We define

$$y_k := \langle f, e_k\rangle$$

and show that $y := (y_k) \in \ell_{p'}$. Further we define $f_n \in \ell_p'$ by

$$\langle f_n, x\rangle := \sum_{k=0}^{n} y_k x_k$$

If we set $x^{(n)} := \sum_{k=0}^{n} x_k e_k$, then

$$\langle f, x^{(n)}\rangle = \Big\langle f, \sum_{k=0}^{n} x_k e_k \Big\rangle = \sum_{k=0}^{n} x_k \langle f, e_k\rangle = \sum_{k=0}^{n} y_k x_k = \langle f_n, x\rangle \tag{25.3}$$

by definition of $y_k$ for all $n \in \mathbb{N}$. Next note that $x^{(n)} \to x$ in $\ell_p$. Indeed, if $1 \leq p < \infty$ and $x \in \ell_p$, then

$$|x^{(n)} - x|_p^p = \sum_{k=n+1}^{\infty} |x_k|^p \to 0$$

as $n \to \infty$. If $p = \infty$ and $x \in c_0$, then also

$$|x^{(n)} - x|_\infty = \sup_{k \geq n} |x_k| \to 0$$

as $n \to \infty$, so $x^{(n)} \to x$ in $c_0$. Hence, by definition of $y$ and (25.3) we obtain

$$\sum_{k=0}^{\infty} x_k y_k = \lim_{n\to\infty} \langle f_n, x\rangle = \lim_{n\to\infty} \langle f, x^{(n)}\rangle = \langle f, x\rangle,$$

and so $f_n \to f$ pointwise and $\langle y, x\rangle = \langle f, x\rangle$. By the previous example we know that $|(y_0, \dots, y_n)|_{p'} = \|f_n\|_{\ell_p'}$ for all $n \in \mathbb{N}$. Hence the uniform boundedness principle (Theorem 22.1) implies that

$$|y|_{p'} = \sup_{n\in\mathbb{N}} |(y_0, \dots, y_n)|_{p'} = \sup_{n\in\mathbb{N}} \|f_n\|_{\ell_p'} < \infty.$$

The first equality holds since $|(y_1, \dots, y_n)|_{p'}$ defines a monotone increasing sequence with limit $|y|_{p'}$. In particular we conclude that $y \in \ell_{p'}$.

We next show that $|y|_{p'} = \|f\|_{\ell_p'}$. By Hölder's inequality $|\langle f, x\rangle| = |\langle y, x\rangle| \leq |y|_{p'} |x|_p$, and so $\|f\|_{\ell_p'} \leq |y|_{p'}$. To prove the reverse inequality choose $x^{(n)} \in \ell_p$ such that $\langle f, x^{(n)}\rangle = |(y_0, \dots, y_n)|_{p'} |x^{(n)}|_p$, which is possible by the construction in the previous example. Thus by definition of the dual norm, $|(y_0, \dots, y_n)|_{p'} \leq \|f\|_{\ell_p'}$ for all $n \in \mathbb{N}$ and hence,

$$|y|_{p'} = \sup_{n\in\mathbb{N}} |(y_0, \dots, y_n)|_{p'} \leq \|f\|_{\ell_p'}$$

as required. Hence (i) and (ii) follow.

We finally mention an even more general version of the above, namely the dual spaces of the $L_p$ spaces. Proofs can be found in [13, Section IV.9].

**25.5 Example** If $\Omega \subseteq \mathbb{R}^N$ is an open set we can consider $L_p(\Omega)$. By Hölder's inequality we have

$$\Big| \int_\Omega f(x) g(x)\, dx \Big| \leq \|f\|_{p'} \|g\|_p$$

for $1 \leq p \leq \infty$. Hence, if we identify $f \in L_{p'}(\Omega)$ with the linear functional

$$g \to \int_\Omega f(x) g(x)\, dx,$$

then we have $f \in \left(L_p(\Omega)\right)'$. With that identification $L_{p'}(\Omega) \subseteq \left(L_p(\Omega)\right)'$ for $1 \leq p \leq \infty$. Using the Radon-Nikodym Theorem from measure theory one can show that

$$\left(L_p(\Omega)\right)' = L_{p'}(\Omega)$$

if $1 \leq p < \infty$ and

$$L_1(\Omega) \subset \left(L_\infty(\Omega)\right)'$$

with proper inclusion.

# 26 The Hahn-Banach Theorem

We have introduced the dual space to a normed vector space, but do not know how big it is in general. The Hahn-Banach theorem in particular shows the existence of many bounded linear functionals on every normed vector space. For the most basic version for vector spaces over $\mathbb{R}$ we do not even need to have a norm on the vector space, but just a semi-norm.

**26.1 Theorem (Hahn-Banach, $\mathbb{R}$-version)** *Suppose that $E$ is a vector space over $\mathbb{R}$ and that $p\colon E \to \mathbb{R}$ is a map satisfying*

(i) *$p(x + y) \leq p(x) + p(y)$ for all $x \in E$ (sub-additivity);*

(ii) *$p(\alpha x) = \alpha p(x)$ for all $x \in E$ and $\alpha \geq 0$ (positive homogeneity).*

*Let $M$ be a subspace and $f_0 \in \operatorname{Hom}(M, \mathbb{R})$ with $f_0(x) \leq p(x)$ for all $x \in M$. Then there exists an extension $f \in \operatorname{Hom}(E, \mathbb{R})$ such that $f|_M = f_0$ and $f(x) \leq p(x)$ for all $x \in E$.*

*Proof.* (a) We first show that $f_0$ can be extended as required if $M$ has co-dimension one. More precisely, let $x_0 \in E \setminus M$ and assume that $\operatorname{span}(M \cup \{x_0\}) = E$. As $x_0 \notin M$ we can write every $x \in E$ in the form

$$x = m + \alpha x_0$$

with $\alpha \in \mathbb{R}$ and $m \in M$ uniquely determined by $x$. Hence for every $c \in \mathbb{R}$, the map $f_c \in \operatorname{Hom}(E, \mathbb{R})$ given by

$$f_c(m + \alpha x) := f_0(m) + c\alpha$$

is well defined, and $f_c(m) = f_0(m)$ for all $m \in M$. We want to show that it is possible to choose $c \in \mathbb{R}$ such that $f_c(x) \leq p(x)$ for all $x \in E$. Equivalently

$$f_0(m) + c\alpha \leq p(m + \alpha x_0)$$

for all $m \in M$ and $\alpha \in \mathbb{R}$. By the positive homogeneity of $p$ and the linearity of $f$ the above is equivalent to

$$\begin{aligned} f_0(m/\alpha) + c &\leq p(x_0 + m/\alpha) &\qquad &\text{if } \alpha > 0, \\ f_0(-m/\alpha) - c &\leq p(-x_0 - m/\alpha) &\qquad &\text{if } \alpha < 0. \end{aligned}$$

Hence we need to show that we can choose $c$ such that

$$c \leq p(x_0 + m/\alpha) - f_0(m/\alpha)$$
$$c \geq -p(-x_0 - m/\alpha) + f_0(-m/\alpha)$$

for all $m \in M$ and $\alpha \in \mathbb{R}$. Note that $\pm m/\alpha$ is just an arbitrary element of $M$, so the above conditions reduce to

$$c \leq p(x_0 + m) - f_0(m)$$
$$c \geq -p(-x_0 + m) + f_0(m)$$

for all $m \in M$. Such a choice of $c$ is therefore possible if

$$-p(-x_0 + m_1) + f_0(m_1) \leq p(x_0 + m_2) - f_0(m_2)$$

for all $m_1, m_2 \in M$. By the linearity of $f_0$ this is equivalent to

$$f_0(m_1 + m_2) = f_0(m_1) + f_0(m_2) \leq p(-x_0 + m_1) + p(x_0 + m_2)$$

for all $m_1, m_2 \in M$. Using the sub-additivity of $p$ we can verify this condition since

$$f_0(m_1 + m_2) \leq p(m_1 + m_2) = p(m_1 - x_0 + m_2 + x_0) \leq p(m_1 - x_0) + p(m_2 + x_0)$$

for all $m_1, m_2 \in M$. Hence $c$ can be chosen as required.

(b) To prove the assertion of the theorem in the general case we use Zorn's Lemma discussed in Section 1. We consider the set $X$ of all extensions $g$ of $f_0$ such that $g(x) \leq p(x)$ for all $x$ in the domain $D(g)$ of $g$. We write $g_2 \supseteq g_1$ if $g_2$ an extension $g_1$. Then $\subseteq$ defines a partial ordering of $X$. Moreover, $f_0 \in X$, so that $X \neq \emptyset$. Suppose now that $C = \{g_\alpha : \alpha \in A\}$ is an arbitrary chain in $X$ and set $D(g) := \bigcup_{\alpha \in A} D(g_\alpha)$ and $g(x) := g_\alpha(x)$ if $x \in D(g_\alpha)$. Then $g$ is an upper bound of $C$ in $X$. By Zorn's Lemma (Theorem 1.7) $X$ has a maximal element $f$ with domain $D(f)$. We claim that $D(f) = E$. If not, then by part (a) of this proof $f$ has a proper extension in $X$, so $f$ was not maximal. Hence $D(f) = E$, and $f$ is a functional as required. ■

We next want to get a version of the Hahn-Banach theorem for vector spaces over $\mathbb{C}$. To do so we need to strengthen the assumptions on $p$.

**26.2 Definition (semi-norm)** Suppose that $E$ is a vector space over $\mathbb{K}$. A map $p : E \to \mathbb{R}$ is called a *semi-norm* on $E$ if

(i) $p(x + y) \leq p(x) + p(y)$ for all $x, y \in E$;

(ii) $p(\alpha x) = |\alpha| p(x)$ for all $x \in E$ and all $\alpha \in \mathbb{K}$.

**26.3 Lemma** *Let $p$ be a semi-norm on $E$. Then $p(x) \geq 0$ for all $x \in E$ and $p(0) = 0$. Moreover, $|p(x) - p(y)| \leq p(x - y)$ for all $x, y \in E$ (reversed triangle inequality).*

*Proof.* Firstly by (ii) we have $p(0) = p(0x) = 0p(x) = 0$ for all $x \in E$, so $p(0) = 0$. If $x, y \in E$, then by (i)

$$p(x) = p(x - y + y) \leq p(x - y) + p(y),$$

so that

$$p(x) - p(y) \le p(x - y).$$

Interchanging the roles of $x$ and $y$ and using (ii) we also have

$$p(y) - p(x) \le p(y - x) = p(-(x - y)) = p(x - y).$$

Combining the two inequalities we get $|p(x) - p(y)| \le p(x - y)$ for all $x, y \in E$. Choosing $y = 0$ we get $0 \le |p(x)| \le p(x)$, so $p(x) \ge 0$ for all $x \in E$. ■

We can now prove a version of the Hahn-Banach Theorem for vector spaces over $\mathbb{C}$.

**26.4 Theorem (Hahn-Banach, $\mathbb{C}$-version)** *Suppose that $p$ is a semi-norm on the vector space $E$ over $\mathbb{K}$ and let $M$ be a subspace of $E$. If $f_0 \in \mathrm{Hom}(M, \mathbb{K})$ is such that $|f_0(x)| \le p(x)$ for all $x \in M$, then there exists an extension $f \in \mathrm{Hom}(E, \mathbb{K})$ such that $f|_M = f_0$ and $|f(x)| \le p(x)$ for all $x \in E$*

*Proof.* If $\mathbb{K} = \mathbb{R}$, then the assertion of the theorem follows from Lemma 26.3 and Theorem 26.1. Let now $E$ be a vector space over $\mathbb{C}$. We split $f_0 \in \mathrm{Hom}(M, \mathbb{C})$ into real and imaginary parts

$$f_0(x) := g_0(x) + ih_0(x)$$

with $g_0, h_0$ real valued and $g_0, h_0 \in \mathrm{Hom}(M_\mathbb{R}, \mathbb{R})$, where $M_\mathbb{R}$ is $M$ considered as a vector space over $\mathbb{R}$. Due to the linearity of $f_0$ we have

$$\begin{aligned} 0 = if_0(x) - f_0(ix) &= ig_0(x) - h_0(x) - g_0(ix) - ih_0(ix) \\ &= -\big(g_0(ix) + h_0(x)\big) + i\big(g_0(x) - h_0(ix)\big) \end{aligned}$$

for all $x \in M$. In particular, $h_0(x) = -g_0(ix)$ and so $g_0(ix) + h_0(x) = 0$ for all $x \in M$. Therefore,

$$f_0(x) = g_0(x) - ig_0(ix) \tag{26.1}$$

for all $x \in M$. We now consider $E$ as a vector space over $\mathbb{R}$. We denote that vector space by $E_\mathbb{R}$. We can consider $M_\mathbb{R}$ as a vector subspace of $E_\mathbb{R}$. Now clearly $g_0 \in \mathrm{Hom}(M_\mathbb{R}, \mathbb{R})$ and by assumption

$$g_0(x) \le |f_0(x)| \le p(x)$$

for all $x \in M_\mathbb{R}$. By Theorem 26.1 there exists $g \in \mathrm{Hom}(E_\mathbb{R}, \mathbb{R})$ such that $g|_{M_\mathbb{R}} = g_0$ and

$$g(x) \le p(x)$$

for all $x \in E_\mathbb{R}$. We set

$$f(x) := g(x) - ig(ix)$$

for all $x \in E_\mathbb{R}$, so that $f|_M = f_0$ by (26.1). To show that $f$ is linear from $E$ to $\mathbb{C}$ we only need to look at multiplication by $i$ because $f$ is linear over $\mathbb{R}$. We have

$$\begin{aligned} f(ix) &= g(ix) - ig(i^2x) = g(ix) - ig(-x) \\ &= g(ix) + ig(x) = i\big(g(x) - ig(ix)\big) = if(x) \end{aligned}$$

for all $x \in E$. It remains to show that $|f(x)| \leq p(x)$. For fixed $x \in E$ we choose $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ such that $\lambda f(x) = |f(x)|$. Then since $|f(x)| \in \mathbb{R}$ and using the definition of $f$

$$|f(x)| = \lambda f(x) = f(\lambda x) = g(\lambda x) \leq p(\lambda x) = |\lambda| p(x) = p(x).$$

This completes the proof of the theorem. ■

From the above version of the Hahn-Banach theorem we can prove the existence of bounded linear functionals with various special properties.

**26.5 Corollary** *Suppose that $E$ is a normed space. Then for every $x \in E$, $x \neq 0$, there exists $f \in E'$ such that $\|f\|_{E'} = 1$ and $\langle f, x\rangle = \|x\|_E$.*

*Proof.* Fix $x \in E$ with $x \neq 0$ and set $M := \operatorname{span}\{x\}$. Define $f_0 \in \operatorname{Hom}(M, \mathbb{K})$ by setting $\langle f_0, \alpha x\rangle := \alpha\|x\|_E$ for all $\alpha \in \mathbb{K}$. Setting $p(y) := \|y\|$ we have

$$|\langle f_0, \alpha x\rangle| = |\alpha|\|x\|_E = \|\alpha x\|_E = p(\alpha x)$$

for all $\alpha \in \mathbb{K}$. By Theorem 26.4 there exists $f \in \operatorname{Hom}(E, \mathbb{K})$ such that $f(\alpha x) = f_0(\alpha x)$ for all $\alpha \in \mathbb{K}$ and $|\langle f, y\rangle| \leq p(y) = \|y\|_E$ for all $y \in E$. Hence $f \in E'$ and $\|f\|_{E'} \leq 1$. Since $|\langle f, x\rangle| = \|x\|_E$ by definition of $f$ we have $\|f\|_{E'} = 1$, completing the proof of the corollary. ■

**26.6 Remark** By the above corollary $E' \neq \{0\}$ for all normed spaces $E \neq \{0\}$.

Theorem 26.4 also allows us to get continuous extensions of linear functionals.

**26.7 Theorem (extension of linear functionals)** *Let $E$ be normed space and $M \subset E$ a proper subspace. Then for every $f_0 \in M'$ there exists $f \in E'$ such that $f|_M = f_0$ and $\|f\|_{E'} = \|f_0\|_{M'}$.*

*Proof.* The function $p\colon E \to \mathbb{R}$ given by $p(x) := \|f_0\|_{M'}\|x\|_E$ defines a semi-norm on $E$. Clearly

$$|\langle f_0, x\rangle| \leq \|f_0\|_{M'}\|x\|_E$$

for all $x \in M$, so by Theorem 26.4 there exists $f \in \operatorname{Hom}(E, \mathbb{K})$ with $f|_M = f_0$ and

$$|\langle f, x\rangle| \leq p(x) = \|f_0\|_{M'}\|x\|_E$$

for all $x \in E$. In particular $f \in E'$ and $\|f\|_{E'} \leq \|f_0\|_{M'}$. Moreover,

$$\|f\|_{M'} = \sup_{\substack{x\in M\\ \|x\|_E\leq 1}} |\langle f, x\rangle| \leq \sup_{\substack{x\in E\\ \|x\|_E\leq 1}} |\langle f, x\rangle| = \|f\|_{E'},$$

so that $\|f\|_{E'} = \|f_0\|_{M'}$. ■

We can use the above theorem to obtain linear functionals with special properties.

**26.8 Corollary** *Let $E$ be a normed space and $M \subset E$ a proper closed subspace. Then for every $x_0 \in E \setminus M$ there exists $f \in E'$ such that $f|_M = 0$ and $\langle f, x_0\rangle = 1$.*

*Proof.* Fix $x_0 \in E \setminus M$ be arbitrary and let $M_1 := M \oplus \operatorname{span}\{x_0\}$. For $m + \alpha x_0 \in M_1$ define

$$\langle f_0, m + \alpha x_0 \rangle := \alpha.$$

for all $\alpha \in \mathbb{K}$ and $m \in M$. Then clearly $\langle f_0, x_0 \rangle = 1$ and $\langle f_0, m \rangle = 0$ for all $m \in M$. Moreover, since $M$ is closed $d := \operatorname{dist}(x_0, M) > 0$ and

$$\inf_{m \in M} \|m + \alpha x_0\| = |\alpha| \inf_{m \in M} \| - \alpha^{-1} m - x_0\| = |\alpha| \inf_{m \in M} \|m - x_0\| = d|\alpha|.$$

Hence

$$|\langle f_0, m + \alpha x_0 \rangle| = |\alpha| \leq d^{-1} \|m + \alpha x_0\|$$

for all $m \in M$ and $\alpha \in \mathbb{K}$. Hence $f_0 \in M_1'$. Now we can apply Theorem 26.7 to conclude the proof. ■

# 27 Reflexive Spaces

The dual of a normed space is again a normed space. Hence we can look at the dual of that space as well. The question is whether or not the second dual coincides with the original space.

**27.1 Definition (Bi-dual space)** If $E$ is a normed space we call

$$E'' := (E')'$$

the *Bi-dual* or *second dual* space of $E$.

We have seen that for every fixed $x \in E$ the map $f \to \langle f, x \rangle$ is linear and

$$|\langle f, x \rangle| \leq \|f\|_{E'} \|x\|_E$$

for all $f \in E'$. Hence we can naturally identify $x \in E$ with $\langle \cdot, x \rangle \in E''$. With that canonical identification we have $E \subseteq E''$ and

$$\|\langle \cdot, x \rangle\|_{E''} \leq \|x\|_E.$$

By the Hahn-Banach Theorem (Corollary 26.5) there exists $f \in E'$ such that $\langle f, x \rangle = \|x\|$, so we have

$$\|\langle \cdot, x \rangle\|_{E''} = \|x\|_E.$$

This proves the following proposition. It also turns out that the norms are the same.

**27.2 Proposition** *The map $x \to \langle \cdot, x \rangle$ is an isometric embedding of $E$ into $E''$, that is, $E \subseteq E''$ with equal norms.*

For some spaces there is equality.

**27.3 Definition (reflexive space)** We say that a normed space $E$ is reflexive if $E = E''$ with the canonical identification $x \to \langle \cdot, x \rangle$.

**27.4 Remark** (a) As dual spaces are always complete (see Remark 25.2(a)), every reflexive space is a Banach space.

(b) Not every Banach space is reflexive. We have shown in Example 25.4 that $c_0' = \ell_1$ and $\ell_1' = \ell_\infty$, so that

$$c_0'' = \ell_1' = \ell_\infty \neq c_0$$

Other examples of non-reflexive Banach spaces are $L_1(\Omega)$, $L_\infty(\Omega)$ and $BC(\Omega)$.

**27.5 Examples** (a) Every finite dimensional normed space is reflexive.

(b) $\ell_p$ is reflexive for $p \in (1, \infty)$ as shown in Example 25.4.

(c) $L_p(\Omega)$ is reflexive for $p \in (1, \infty)$ as shown in Example 25.5.

(d) Every Hilbert space is reflexive as we will show in Section 30.

# 28 Weak convergence

So far, in a Banach space we have only looked at sequences converging with respect to the norm in the space. Since we work in infinite dimensions there are more notions of convergence of a sequence. We introduce one involving dual spaces.

**28.1 Definition (weak convergence)** Let $E$ be a Banach space and $(x_n)$ a sequence in $E$. We say $(x_n)$ *converges weakly* to $x$ in $E$ if

$$\lim_{n\to\infty} \langle f, x_n \rangle = \langle f, x \rangle$$

for all $f \in E'$. We write $x_n \rightharpoonup x$ in that case.

**28.2 Remark** (a) Clearly, if $x_n \to x$ with respect to the norm, then $x_n \rightharpoonup x$ weakly.

(b) The converse is not true, that is, weak convergence does not imply convergence with respect to the norm. As an example consider the sequence $e_n := (\delta_{kn})_{k\in\mathbb{N}} \in \ell_2$. Then $|e_n|_2 = 1$ and $\langle f, e_n \rangle = f_n$ for all $f = (f_k)_{k\in\mathbb{N}} \in \ell_2' = \ell_2$. Hence for all $f \in \ell_2'$

$$\langle f, e_n \rangle = f_n \to 0$$

as $n \to \infty$, so $e_n \rightharpoonup 0$ weakly. However, $(e_n)$ does not converge with respect to the norm.

(c) The weak limit of a sequence is unique. To see this note that if $\langle f, x \rangle = 0$ for all $f \in E'$, then $x = 0$ since otherwise the Hahn-Banach Theorem (Corollary 26.5) implies the existence of $f \in E'$ with $\langle f, x \rangle \neq 0$.

From the example in Remark 28.2 we see that in general $\|x_n\| \not\to \|x\|$ if $x_n \rightharpoonup x$ weakly. However, we can still get an upper bound for $\|x\|_E$ using the uniform boundedness principle.

**28.3 Proposition** *Suppose that $E$ is a Banach space and that $x_n \rightharpoonup x$ weakly in $E$. Then $(x_n)$ is bounded and*

$$\|x\|_E \leq \liminf_{n\to\infty} \|x_n\|_E.$$

*Proof.* We consider $x_n$ as an element of $E''$. By assumption $\langle f, x_n\rangle \to \langle f, x\rangle$ for all $f \in E'$. Hence by the uniform boundedness principle (Corollary 22.3) and Proposition 27.2

$$\|x\|_E = \|x\|_{E''} \leq \liminf_{n\to\infty} \|x_n\|_{E''} = \liminf_{n\to\infty} \|x_n\|_E < \infty.$$

as claimed. ■

# 29 Dual Operators

For every linear operator $T$ from a normed space $E$ into a normed space $F$ we can define an operator from the dual $F'$ into $E'$ by composing a linear functional on $f$ with $T$ to get a linear functional on $E$.

**29.1 Definition** Suppose that $E$ and $F$ are Banach spaces and that $T \in \mathcal{L}(E, F)$. We define the *dual operator* $T' : F' \to E'$ by

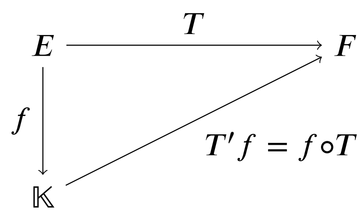
$$T'f := f \circ T$$

for all $f \in F'$.

Note that for $f \in F'$ and $x \in E$ by definition

$$\langle T'f, x\rangle = \langle f, Tx\rangle.$$

The situation is also given in the following diagram:




**29.2 Remark** We can also look at the second dual of an operator $T \in \mathcal{L}(E, F)$. Then $T'' = (T')' : E'' \to F''$. By definition of $T''$ and the canonical embedding of $E \subseteq E''$ we always haven $T''|_E = T$. If $E$ and $F$ are reflexive, then $T'' = T$.

We next show that an operator and its dual have the same norm.

**29.3 Theorem** *Let $E$ and $F$ be normed spaces and $T \in \mathcal{L}(E, F)$. Then $T' \in \mathcal{L}(F', E')$ and $\|T\|_{\mathcal{L}(E,F)} = \|T'\|_{\mathcal{L}(F',E')}$.*

*Proof.* By definition of the dual norm we have

$$\|T'f\|_{E'} = \sup_{\|x\|_E \leq 1} |\langle T'f, x\rangle| = \sup_{\|x\|_E \leq 1} |\langle f, Tx\rangle|$$
$$\leq \sup_{\|x\|_E \leq 1} \|f\|_{F'} \|Tx\|_F = \|T\|_{\mathcal{L}(E,F)} \|f\|_{F'},$$

so $\|T'\|_{\mathcal{L}(F',E')} \leq \|T\|_{\mathcal{L}(E,F)}$. For the opposite inequality we use the Hahn-Banach Theorem (Corollary 26.5) which guarantees that for every $x \in E$ there exists $f \in F'$ such that $\|f\|_{F'} = 1$ and $\langle f, Tx\rangle = \|Tx\|_F$. Hence, for that choice of $f$ we have

$$\begin{aligned}\|Tx\|_F = \langle f, Tx\rangle = \langle T'f, x\rangle \leq \|T'f\|_{E'}\|x\|_E \\ \leq \|T'\|_{\mathcal{L}(F',E')}\|f\|_{F'}\|x\|_E = \|T'\|_{\mathcal{L}(F',E')}\|x\|_E.\end{aligned}$$

Hence $\|T'\|_{\mathcal{L}(F',E')} \geq \|T\|_{\mathcal{L}(E,F)}$ and the assertion of the theorem follows. ■

We can use the dual operator to obtain properties of the original linear operator.

**29.4 Proposition** *Let $E$ and $F$ be normed spaces and $T \in \mathcal{L}(E, F)$. Then the image of $T$ is dense in $F$ if and only if* $\ker T' = \{0\}$.

*Proof.* Suppose that $\overline{\operatorname{im}(T)} = F$ and that $f \in \ker T'$. Then

$$0 = \langle T'f, x\rangle = \langle f, Tx\rangle$$

for all $x \in E$. Since $\operatorname{im}(T)$ is dense in $F$ we conclude that $f = 0$, so $\ker T' = \{0\}$. We prove the converse by contrapositive, assuming that $M := \overline{\operatorname{im}(T)}$ is a proper subspace of $F$. By the Hahn-Banach Theorem (Corollary 26.8) there exists $f \in F'$ with $\|f\|_{F'} = 1$ and $\langle f, y\rangle = 0$ for all $y \in M$. In particular

$$\langle T'f, x\rangle = \langle f, Tx\rangle = 0$$

for all $x \in E$. Hence $T'f = 0$ but $f \neq 0$ and so $\ker T' \neq \{0\}$ as claimed. ■

We apply the above to show that the dual of an isomorphism is also an isomorphism.

**29.5 Proposition** *Let $E$ and $F$ be normed spaces and $T \in \mathcal{L}(E, F)$ is an isomorphism with $T^{-1} \in \mathcal{L}(F, E)$. Then $T' \in \mathcal{L}(F', E')$ is an isomorphism and $(T^{-1})' = (T')^{-1}$.*

*Proof.* By Proposition 29.4 the dual $T'$ is injective since $\operatorname{im}(T) = F$. For given $g \in E'$ we set $f := (T^{-1})'g$. Then

$$\langle T'f, x\rangle = \langle f, Tx\rangle = \langle (T^{-1})'g, Tx\rangle = \langle g, T^{-1}Tx\rangle = \langle g, x\rangle$$

for all $x \in E$. Hence $T'$ is surjective and $(T')^{-1} = (T^{-1})'$ as claimed. ■

We can apply the above to "embeddings" of spaces.

**29.6 Definition (continuous embedding)** Suppose that $E$ and $F$ are normed spaces and that $E \subseteq F$. If the natural injection $i\colon E \to F$, $x \mapsto i(x) := x$ is continuous, then we write $E \hookrightarrow F$. We say *$E$ is continuously embedded into $F$*. If in addition $E$ is dense in $F$, then we write $E \overset{d}{\hookrightarrow} F$ and say $E$ is densely embedded into $F$.

If $E \hookrightarrow F$ and $i \in \mathcal{L}(E, F)$ is the natural injection, then the dual is given by

$$\langle i'(f), x\rangle = \langle f, i(x)\rangle$$

for all $f \in F'$ and $x \in E$. Hence $i'(f) = f|_E$ is the restriction of $f$ to $E$. Hence $F'$ in some sense is a subset of $E'$. However, $i'$ is an injection if and only if $E \overset{d}{\hookrightarrow} F$ by Proposition 29.4. Hence we can prove the following proposition.

**29.7 Proposition** *Suppose that* $E \overset{d}{\hookrightarrow} F$*. Then* $F' \hookrightarrow E'$*. If* $E$ *and* $F$ *are reflexive, then* $F' \overset{d}{\hookrightarrow} E'$*.*

*Proof.* The first statement follows from the comments before. If $E$ and $F$ are reflexive, then $i'' = i$ by Remark 29.2. Hence $(i')' = i$ is injective, so by Proposition 29.4 $i'$ has dense range, that is, $F' \overset{d}{\hookrightarrow} E'$. ■

# 30 Duality in Hilbert Spaces

Suppose now that $H$ is a Hilbert space. By the Cauchy-Schwarz inequality

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

for all $x, y \in H$. As a consequence

$$J(y) := (\cdot \mid y) \in H'$$

is naturally in the dual space of $H$ and $\|J(y)\|_{H'} \leq \|y\|$ for every $y \in H$.

**30.1 Proposition** *The map* $J : H \to H'$ *defined above has the following properties:*

(i) $J(\lambda x + \mu y) = \bar{\lambda} J(x) + \bar{\mu} J(y)$ *for all* $x, y \in H$ *and* $\lambda, \mu \in \mathbb{K}$*. We say* $J$ *is* conjugate linear.

(ii) $\|J(y)\|_{H'} = \|y\|_H$ *for all* $y \in H$*.*

(iii) $J : H \to H'$ *is continuous.*

*Proof.* (i) is obvious from the definition of $J$ and the properties of an inner product. (ii) follows from Corollary 15.6 and the definition of the dual norm. Finally (iii) follows from (ii) since $\|J(x) - J(y)\|_{H'} = \|J(x - y)\|_{H'} = \|x - y\|_H$. ■

We now show that $J$ is bijective, that is, $H = H'$ if we identify $y \in H$ with $J(y) \in H'$.

**30.2 Theorem (Riesz representation theorem)** *The map* $J : H \to H'$ *is a conjugate linear isometric isomorphism. It is called the* Riesz isomorphism.

*Proof.* By the above proposition, $J$ is a conjugate linear isometry. To show that $J$ is surjective let $f \in H'$ with $f \neq 0$. By continuity of $f$ the kernel $\ker f$ is a proper closed subspace of $H$, and so by Theorem 16.9 we have

$$H = \ker f \oplus (\ker f)^{\perp}$$

and there exists $z \in (\ker f)^{\perp}$ with $\|z\| = 1$. Clearly

$$-\langle f, x \rangle z + \langle f, z \rangle x \in \ker f$$

for all $x \in H$. Since $z \in (\ker f)^\perp$ we get

$$0 = \big(-\langle f, x\rangle z + \langle f, z\rangle x \big| z\big) = -\langle f, x\rangle (z \mid z) + \langle f, z\rangle (x \mid z) = -\langle f, x\rangle + \langle f, z\rangle (x \mid z)$$

for all $x \in H$. Hence

$$\langle f, x\rangle = \langle f, z\rangle (x \mid z) = \big(x \big| \overline{\langle f, z\rangle} z\big)$$

for all $x \in H$, so $J(y) = f$ if we set $y = \overline{\langle f, z\rangle} z$. ■

We use the above to show that Hilbert spaces are reflexive.

**30.3 Corollary (reflexivity of Hilbert spaces)** *If $H$ is a Hilbert space, then $H'$ is a Hilbert space with inner product $(f \mid g)_{H'} := (J^{-1}(g) \mid J^{-1}(f))_H$. The latter inner product induces the the dual norm. Moreover, $H$ is reflexive.*

*Proof.* By the properties of $J$, $(f \mid g)_{H'} := (J^{-1}(g) \mid J^{-1}(f))_H$ clearly defines an inner product on $H'$. Moreover, again by the properties of $J$

$$(f \mid f)_{H'} = (J^{-1}(f) \mid J^{-1}(f))_H = \|f\|_{H'}^2,$$

so $H'$ is a Hilbert space. To prove that $H$ is reflexive let $x \in H''$ and denote the Riesz isomorphisms on $H$ and $H'$ by $J_H$ and $J_{H'}$, respectively. Then by the definition of the inner product on $H'$ and $J_H$

$$\langle f, x\rangle_H = \langle x, f\rangle_{H'} = (f \mid J_{H'}^{-1}(x))_{H'} = (J_H^{-1} J_{H'}^{-1}(x) \mid J_H^{-1}(f))_H = \langle f, J_H^{-1} J_{H'}^{-1}(x)\rangle$$

for all $f \in H'$. Hence by definition of reflexivity $H$ is reflexive. ■

Let finally $H_1$ and $H_2$ be Hilbert spaces and $T \in \mathcal{L}(H_1, H_2)$. Moreover, let $J_1$ and $J_2$ be the Riesz isomorphisms on $H_1$ and $H_2$, respectively. If $T'$ is the dual operator of $T$, then we set

$$T^* = J_1^{-1} \circ T' \circ J_2 \in \mathcal{L}(H_2, H_1).$$

The situation is depicted in the commutative diagram below:

$$\begin{array}{ccc} H_2' & \xrightarrow{\quad T' \quad} & H_1' \\ {\scriptstyle J_2}\big\uparrow & & \big\uparrow{\scriptstyle J_1} \\ H_2 & \xrightarrow{\quad T^* \quad} & H_1 \end{array}$$

**30.4 Definition (adjoint operators)** The operator $T^* \in \mathcal{L}(H_2, H_1)$ as defined above is called the *adjoint operator* of $T$. If $H_1 = H_2$ and $T = T^*$, then $T$ is called *self-adjoint*, and if $TT^* = T^*T$, then $T$ is called *normal*.

**30.5 Remark** If $T \in \mathcal{L}(H_1, H_2)$ and $T^*$ its adjoint, then by definition

$$(Tx \mid y)_{H_2} = (x \mid T^* y)_{H_1}$$

for all $x \in H_1$ and $y \in H_2$. To see this we use the definition of $T^*$:

$$(Tx \mid y) = (J_2 y \mid J_2 Tx) = (J_2 y \mid T' J_1 x) = (y \mid J_1^{-1} T' J_1 x) = (y \mid T^* x)$$

for all $x \in H_1$ and $y \in H_2$.

# 31 The Lax-Milgram Theorem

In the previous section we looked at the representation of linear functionals on a Hilbert space by means of the scalar product. Now we look at sesqui-linear forms.

**31.1 Definition (sesqui-linear form)** Let $H$ be a Hilbert space over $\mathbb{K}$. A map $b\colon H\times H\to\mathbb{K}$ is called a *sesqui-linear form* on $H$ if

(i) $u\mapsto b(u,v)$ is linear for all $v\in H$;

(ii) $v\mapsto b(u,v)$ is conjugate linear for all $u\in H$.

The sesqui-linear form $b(\cdot\,,\cdot)$ is called a *Hermitian form* if

(iii) $b(u,v)=\overline{b(v,u)}$ for all $u,v\in H$.

Moreover, the sesqui-linear form $b(\cdot\,,\cdot)$ is called *bounded* if there exists $M>0$ such that

$$|b(u,v)|\le M\|u\|_H\|v\|_H \tag{31.1}$$

for all $u,v\in H$ and *coercive* if there exists $\alpha>0$ such that

$$\alpha\|u\|_H^2\le \operatorname{Re} b(u,u) \tag{31.2}$$

for all $u\in H$.

If $A\in\mathcal{L}(H)$, then

$$b(u,v):=(u\mid Av) \tag{31.3}$$

defines a bounded sesqui-linear form on $H$. We next show that the converse is also true, that is, for every bounded sesqui-linear form there exists a linear operator such that (31.3) holds for all $u,v\in H$.

**31.2 Proposition** *For every bounded sesqui-linear form $b(\cdot\,,\cdot)$ there exists a unique linear operator $A\in\mathcal{L}(H)$ such that (31.3) holds for all $u,v\in H$. Moreover, $\|A\|_{\mathcal{L}(H)}\le M$, where $M$ is the constant in (31.1).*

*Proof.* By (31.1) it follows that for every fixed $v\in H$ the map $u\to b(u,v)$ is a bounded linear functional on $H$. The Riesz representation theorem implies that there exists a unique element $Av$ in $H$ such that (31.3) holds. The map $v\to Av$ is linear because

$$\begin{aligned}(w\mid A(\lambda u+\mu v))=b(w,\lambda u+\mu v)&=\bar\lambda b(w,u)+\bar\mu b(w,v)\\&=\bar\lambda(w\mid Au)+\bar\mu(w\mid Av)=(w\mid \lambda Au+\mu Av)\end{aligned}$$

for all $u,v,w\in H$ and $\lambda,\mu\in\mathbb{K}$, and so by definition of $A$ we get $A(\lambda u+\mu v)=\lambda Au+\mu Av$. To show that $A$ is bounded we note that by Corollary 15.6 and (31.1) we have

$$\|Av\|_H=\sup_{\|u\|_H=1}|(u\mid Av)|=\sup_{\|u\|_H=1}|b(u,v)|\le\sup_{\|u\|_H=1}M\|u\|_H\|v\|_H=M\|v\|_H$$

for all $v\in H$. Hence $A\in\mathcal{L}(H)$ and $\|A\|_{\mathcal{L}(H)}\le M$ as claimed. ■

We next show that if the form is coercive, then the operator constructed above is invertible.

**31.3 Theorem (Lax-Milgram theorem)** *Suppose $b(\cdot\,,\cdot)$ is a bounded coercive sesquilinear form and $A \in \mathcal{L}(H)$ the corresponding operator constructed above. Then $A$ is invertible and $A^{-1} \in \mathcal{L}(H)$ with $\|A^{-1}\|_{\mathcal{L}(H)} \leq \alpha^{-1}$, where $\alpha > 0$ is the constant from* (31.2).

*Proof.* By the coercivity of $b(\cdot\,,\cdot)$ we have

$$\alpha\|u\|_H^2 \leq \operatorname{Re} b(u,u) \leq |b(u,u)| = |(u \mid Au)| \leq \|u\|_H \|Au\|_H$$

for all $u \in H$. Therefore $\alpha\|u\|_H \leq \|Au\|_H$ for all $u \in H$ and so $A$ is injective, $A^{-1} \in \mathcal{L}(\operatorname{im}(A), H)$ with $\|A^{-1}\|_{\mathcal{L}(H)} \leq \alpha^{-1}$. We next show that $\operatorname{im}(A)$ is dense in $H$ by proving that its orthogonal complement is trivial. If $u \in (\operatorname{im}(A))^{\perp}$, then in particular

$$0 = (u \mid Au) = b(u,u) \geq \alpha\|u\|_H^2$$

by definition of $A$ and the coercivity of $b(\cdot\,,\cdot)$. Hence $u = 0$ and so Corollary 16.11 implies that $\operatorname{im}(A)$ is dense in $H$. Since $A$ is in particular a closed operator with bounded inverse, Theorem 23.4(iii) implies that $\operatorname{im}(A) = H$. ■

Chapter VII

# Spectral Theory

Spectral theory is a generalisation of the theory of eigenvalues and eigenvectors of matrices to operators on infinite dimensional spaces.

## 32 Resolvent and Spectrum

Consider an $N \times N$ matrix $A$. The set of $\lambda \in \mathbb{C}$ for which $(\lambda I - A)$ is not invertible is called the set of eigenvalues of $A$. Due to the fact that $\dim(\operatorname{im}(A)) + \dim(\ker(A)) = N$ the matrix $(\lambda I - A)$ is only invertible if it has a trivial kernel. It turns out that there are at most $N$ such $\lambda$ called the eigenvalues of $A$. We consider a similar set of $\lambda$ for closed operators on a Banach space $E$ over $\mathbb{C}$.

**32.1 Definition (resolvent and spectrum)** Let $E$ be a Banach space over $\mathbb{C}$ and $A\colon D(A) \subseteq E \to E$ a closed operator. We call

(i) $\varrho(A) := \{\lambda \in \mathbb{C} \mid (\lambda I - A)\colon D(A) \to E$ is bijective and $(\lambda I - A)^{-1} \in \mathcal{L}(E)\}$ the *resolvent set* of $A$;

(ii) $R(\lambda, A) := (\lambda I - A)^{-1}$ the *resolvent of* $A$ at $\lambda \in \varrho(A)$;

(iii) $\sigma(A) := \mathbb{C} \setminus \varrho(A)$ the *spectrum* of $A$.

**32.2 Remarks** (a) If $A\colon D(A) \subseteq E \to E$ is closed, then $(\lambda I - A)\colon D(A) \subseteq E \to E$ is closed as well for all $\lambda \in \mathbb{K}$. To see this assume that $x_n \in D(A)$ with $x_n \to x$ and $(\lambda I - A)x_n \to y$ in $E$. Then $Ax_n = \lambda x_n - (\lambda I - A)x_n \to \lambda x - y$ in $E$. Since $A$ is closed $x \in D(A)$ and $Ax = \lambda x - y$, that is, $(\lambda I - A)x = y$. Hence $\lambda I - A$ is closed.

(b) The graph norms of $(\lambda I - A)$ are equivalent for all $\lambda \in \mathbb{K}$. This follows since for every $x \in D(A)$ and $\lambda \in \mathbb{K}$

$$\|x\|_{\lambda I - A} = \|x\| + \|(\lambda I - A)x\| \leq (1 + |\lambda|)(\|x\| + \|Ax\|) = (1 + |\lambda|)\|x\|_A$$

and

$$\begin{aligned}\|x\|_A = \|x\| + \|Ax\| &= \|x\| + \|\lambda x - (\lambda I - A)x\| \\ &\leq (1 + |\lambda|)(\|x\| + \|(\lambda x - A)x\|) = (1 + |\lambda|)\|x\|_{\lambda I - A}.\end{aligned}$$

We next get some characterisations of the points in the resolvent set.

**32.3 Proposition** *Let $A\colon D(A)\subseteq E\to E$ be a closed operator on the Banach space $E$. Then the following assertions are equivalent.*

(i) $\lambda\in\varrho(A)$;

(ii) $\lambda I-A\colon D(A)\to E$ *is bijective;*

(iii) $\lambda I-A$ *is injective,* $\overline{\operatorname{im}(\lambda I-A)}=E$ *and* $(\lambda I-A)^{-1}\in\mathcal{L}(\operatorname{im}(\lambda I-A),E)$.

*Proof.* By definition of the resolvent set (i) implies (ii). Since $\lambda I-A$ is a closed operator by the above remark, the equivalence of (i) to (iii) follows from Theorem 23.4. ■

We now want to derive properties of $R(\lambda,A)=(\lambda I-A)^{-1}$ as a function of $\lambda$. If $A$ is not an operator but a complex number we can use the geometric series to find various series expansions for $\dfrac{1}{\lambda-a}=(\lambda-a)^{-1}$. To get a more precise formulation we need a lemma.

**32.4 Lemma** *Let $T\in\mathcal{L}(E)$. Then*

$$r:=\lim_{n\to\infty}\|T^n\|^{1/n}\tag{32.1}$$

*exists and $r\leq\|T\|$.*

*Proof.* First note that

$$\|T^{n+m}x\|=\|T^nT^mx\|\leq\|T^n\|\|T^mx\|\leq\|T^n\|\|T^m\|\|x\|$$

for all $x\in E$, so by definition of the operator norm

$$\|T^{n+m}\|\leq\|T^n\|\|T^m\|\tag{32.2}$$

for all $n,m\geq 0$. Fix now $m\in\mathbb{N}$, $m>1$. If $n>m$, then there exist $q_n,r_n\in\mathbb{N}$ with $n=mq_n+r_n$ and $0\leq r_n<m$. Hence

$$\frac{q_n}{n}=\frac{1}{m}\Bigl(1+\frac{r_n}{n}\Bigr)\to\frac{1}{m}$$

as $n\to\infty$. From (32.2) we conclude that $\|T^n\|^{1/n}\leq\|T^m\|^{q_n/n}\|T\|^{r_n/n}$ for all $n>m$. Hence

$$\liminf_{n\to\infty}\|T^n\|^{1/n}\leq\|T^m\|^{1/m}$$

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

$$\liminf_{n\to\infty}\|T^n\|^{1/n}\leq\limsup_{n\to\infty}\|T^n\|^{1/n}\leq\|T\|.$$

Hence the limit (32.1) exists, and our proof is complete. ■

We now look at an analogue of the "geometric series" for bounded linear operators.

**32.5 Proposition (Neumann series)** *Let $T \in \mathcal{L}(E)$ and set*

$$r := \lim_{k\to\infty} \|T^k\|^{1/k}.$$

*Then the following assertions are true.*

(i) $\sum_{k=0}^{\infty} T^k$ *converges absolutely in $\mathcal{L}(E)$ if $r < 1$ and diverges if $r > 1$.*

(ii) *If $\sum_{k=0}^{\infty} T^k$ converges, then $\sum_{k=0}^{\infty} T^k = (I - T)^{-1}$.*

*Proof.* Part (i) is a consequence of the root test for the absolute convergence of a series. For part (ii) we use a similar argument as in case of the geometric series. Clearly

$$(I - T)\sum_{k=0}^{n} T^k = \sum_{k=0}^{n} T^k(I - T) = \sum_{k=0}^{n} T^k - \sum_{k=0}^{n+1} T^k = I - T^{n+1}$$

for all $n \in \mathbb{N}$. Since the series $\sum_{k=0}^{n} T^k$ converges, in particular $T^n \to 0$ in $\mathcal{L}(E)$ and therefore

$$(I - T)\sum_{k=0}^{\infty} T^k x = \sum_{k=0}^{\infty} T^k(I - T)x = x$$

for all $x \in E$. Hence (ii) follows. ∎

We next use the above to show that the resolvent is analytic on $\varrho(A)$. The series expansions we find are very similar to corresponding expansions of $(\lambda - a)^{-1}$ based on the geometric series.

**32.6 Definition (analytic function)** Let $E$ be a Banach space over $\mathbb{C}$ and $U \subseteq \mathbb{C}$ an open set. A map $f : U \to E$ is called *analytic* (or *holomorphic*) if for every $\lambda_0 \in U$ there exists $r > 0$ and $a_k \in E$ such that

$$f(\lambda) = \sum_{k=0}^{\infty} a_k(\lambda - \lambda_0)^k$$

for all $\lambda \in U$ with $|\lambda - \lambda_0| < r$.

**32.7 Theorem** *Let $E$ be a Banach space over $\mathbb{C}$ and $A\colon D(A) \subseteq E \to E$ a closed operator.*

(i) *The resolvent set $\varrho(A)$ is open in $\mathbb{C}$ and the resolvent $R(\cdot\,, A)\colon \varrho(A) \to \mathcal{L}(E)$ is analytic;*

(ii) *For all $\lambda_0 \in \varrho(A)$ we have $\|R(\lambda_0, A)\|_{\mathcal{L}(E)} \geq \dfrac{1}{\operatorname{dist}(\lambda_0, \sigma(A))}$.*

(iii) *For all $\lambda, \mu \in \varrho(A)$*

$$R(\lambda, A) - R(\mu, A) = (\mu - \lambda)R(\lambda, A)R(\mu, A)$$

*(resolvent equation);*

(iv) *For all* $\lambda, \mu \in \varrho(A)$

$$R(\lambda, A)R(\mu, A) = R(\mu, A)R(\lambda, A)$$

*and*

$$AR(\lambda, A)x = R(\lambda, A)Ax = (\lambda R(\lambda, A) - I)x$$

*for all* $x \in D(A)$.

*Proof.* (i) If $\varrho(A) = \emptyset$, then $\varrho(A)$ is open. Hence suppose that $\varrho(A) \neq \emptyset$ and fix $\lambda_0 \in \varrho(A)$. Then

$$\lambda I - A = (\lambda - \lambda_0) + (\lambda_0 I - A)$$

for all $\lambda \in \mathbb{C}$. Hence

$$(\lambda I - A) = (\lambda_0 I - A)\big(I + (\lambda - \lambda_0)R(\lambda_0, A)\big) \tag{32.3}$$

for all $\lambda \in \mathbb{C}$. Now fix $r \in (0, 1)$ and suppose that

$$|\lambda - \lambda_0| \leq \frac{r}{\|R(\lambda_0, A)\|_{\mathcal{L}(E)}}.$$

Then $|\lambda - \lambda_0|\|R(\lambda_0, A)\|_{\mathcal{L}(E)} < 1$ and by Proposition 32.5 $I + (\lambda - \lambda_0)R(\lambda_0, A)$ is invertible and

$$\big(I + (\lambda - \lambda_0)R(\lambda_0, A)\big)^{-1} = \sum_{k=0}^{\infty}(-1)^k R(\lambda_0, A)^k(\lambda - \lambda_0)^k.$$

Hence by (32.3) the operator $(\lambda I - A)\colon D(A) \to E$ has a bounded inverse and

$$\begin{aligned} R(\lambda, A) = (\lambda I - A)^{-1} &= \big(I + (\lambda - \lambda_0)R(\lambda_0, A)\big)^{-1}R(\lambda_0, A) \\ &= \sum_{k=0}^{\infty}(-1)^k R(\lambda_0, A)^{k+1}(\lambda - \lambda_0)^k \end{aligned}$$

for all $\lambda \in \mathbb{C}$ with

$$|\lambda - \lambda_0| < \frac{1}{\|R(\lambda_0, A)\|_{\mathcal{L}(E)}}.$$

Therefore $\varrho(A)$ is open and the map $\lambda \mapsto R(\lambda, A)$ analytic on $\varrho(A)$.

(ii) From the proof of (i) we have

$$|\lambda - \lambda_0| \geq \frac{1}{\|R(\lambda_0, A)\|_{\mathcal{L}(E)}}.$$

for all $\lambda \in \sigma(A)$. Hence

$$\operatorname{dist}(\lambda_0, \sigma(A)) = \inf_{\lambda \in \sigma(A)} |\lambda - \lambda_0| \geq \frac{1}{\|R(\lambda_0, A)\|_{\mathcal{L}(E)}}.$$

and so (ii) follows.

(iii) If $\lambda, \mu \in \varrho(A)$, then

$$(\mu I - A) = (\lambda I - A) + (\mu - \lambda)I.$$

If we multiply by $R(\lambda, A)$ from the left and by $R(\mu, A)$ from the right we get

$$R(\lambda, A) = R(\mu, A) + (\mu - \lambda)R(\lambda, A)R(\mu, A).$$

from which our claim follows.

(iv) The first identity follows from (iii) since we can interchange the roles of $\lambda$ and $\mu$ to get the same result. The last identity follows since

$$A = \lambda I - (\lambda I - A)$$

by applying $R(\lambda, A)$ from the left and from the right. ■

**32.8 Remarks** (a) Because $\varrho(A)$ is open, the spectrum $\sigma(A)$ is always a closed subset of $\mathbb{C}$.

(b) Part (ii) if the above theorem shows that $\varrho(A)$ is the natural domain of the analytic function given by the resolvent $R(\cdot\,, A)$. The function is Banach space valued, but nevertheless all theorems from complex analysis can be applied. In particular, one can analyse the structure of $A$ near an isolated point of the spectrum by using Laurent series expansions, see for instance [3].

Next we look at the spectrum of bounded operators.

**32.9 Definition** Let $E$ be a Banach space and $T \in \mathcal{L}(E)$. We call

$$\operatorname{spr}(T) := \sup\{|\lambda| : \lambda \in \sigma(T)\}$$

the *spectral radius* of $T$.

**32.10 Theorem** *Suppose that $E$ is a Banach space and $T \in \mathcal{L}(E)$. Then $\sigma(T)$ is non-empty and bounded. Moreover,* $\operatorname{spr}(T) = \lim_{n\to\infty} \|T^n\|^{1/n} \leq \|T\|$.

*Proof.* Using Proposition 32.5 we see that

$$R(\lambda, T) = (\lambda I - T)^{-1} = \frac{1}{\lambda}\Big(I - \frac{1}{\lambda}T\Big)^{-1} = \sum_{k=0}^{\infty} \frac{1}{\lambda^{k+1}} T^k \tag{32.4}$$

if $|\lambda| > r := \lim_{n\to\infty} \|T^n\|^{1/n}$. The above is the Laurent expansion of $R(\cdot\,, T)$ about zero for $|\lambda|$ large. We know that such an expansion converges in the largest annulus contained in the domain of the analytic function. Hence $r = \operatorname{spr}(T)$ as claimed. We next prove that $\sigma(T)$ is non-empty. We assume $\sigma(T)$ is empty and derive a contradiction. If $\sigma(T) = \emptyset$, then the function

$$g : \mathbb{C} \to \mathbb{C},\ \lambda \mapsto \langle f, R(\lambda, T)x\rangle$$

is analytic on $\mathbb{C}$ for every $x \in E$ and $f \in E'$. By (32.4) we get

$$|g(\lambda)| \leq \frac{1}{|\lambda|} \sum_{k=0}^{\infty} \frac{\|T\|^k}{2^k \|T\|^k} \|f\|_{E'} \|x\|_E = \frac{\|f\|_{E'} \|x\|_E}{|\lambda|}$$

for all $|\lambda| \geq 2\|T\|$. In particular, $g$ is bounded and therefore by Liouville's theorem $g$ is constant. Again from the above $g(\lambda) \to 0$ as $|\lambda| \to \infty$, and so that constant must be zero. Hence

$$\langle f, R(\lambda, T)x\rangle = 0$$

for all $x \in E$ and $f \in E'$. Hence by the Hahn-Banach theorem (Corollary 26.5) $R(\lambda, T)x = 0$ for all $x \in E$, which is impossible since $R(\lambda, T)$ is bijective. This proves that $\sigma(T) \neq \emptyset$. ■

Note that $\lambda I - A$ may fail to have a continuous inverse defined on $E$ for several reasons. We split the spectrum up accordingly.

**32.11 Definition** Let $A\colon D(A) \subseteq E \to E$ be a closed operator. We call

- $\sigma_p(A) := \{\lambda \in \mathbb{C}\colon \lambda I - A \text{ not injective}\}$ the *point spectrum* of $A$;
- $\sigma_c(A) := \{\lambda \in \mathbb{C}\colon \lambda I - A \text{ injective}, \overline{\operatorname{im}(\lambda I - A)} = E, (\lambda I - A)^{-1} \text{ unbounded}\}$ the *continuous spectrum* of $A$;
- $\sigma_r(A) := \{\lambda \in \mathbb{C}\colon \lambda I - A \text{ injective}, \overline{\operatorname{im}(\lambda I - A)} \neq E\}$ the *residual spectrum* of $A$;

**32.12 Remarks** (a) By definition of $\sigma(A)$ and Proposition 32.3

$$\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$$

is clearly a disjoint union.

(b) The elements of $\sigma_p(A)$ are called the *eigenvalues* of $A$, and the non-zero elements of $\ker(\lambda I - A)$ the corresponding *eigenvectors*.

(c) If $\dim E = N < \infty$ the spectrum consists of eigenvalues only. This follows since $\dim(\operatorname{im}(\lambda I - A)) + \dim(\ker(\lambda I - A)) = N$ and therefore $\lambda I - A$ is not invertible if and only if it has non-trivial kernel.

# 33 Projections, Complements and Reductions

In finite dimensions the eigenvalues are used to find basis under which the matrix associated with a given linear operator becomes as simple as possible. In particular, to every eigenvalue there is a maximal invariant subspace, the generalised eigenspace associated with an eigenvalue of a square matrix. We would like to have something similar in infinite dimensions. In this section we develop the concepts to formulate such results.

**33.1 Definition (projection)** Let $E$ be a vector space and $P \in \operatorname{Hom}(E)$. We call $P$ a projection if $P^2 = P$.

We next show that there is a one-to-one correspondence between projections and pairs of complemented subspaces.

**33.2 Lemma** (i) *If $P \in \mathrm{Hom}(E)$ is a projection, then $E = \mathrm{im}(P) \oplus \ker(P)$. Moreover, $I - P$ is a projection and* $\ker(P) = \mathrm{im}(I - P)$.

(ii) *If $E_1, E_2$ are subspaces of $E$ with $E = E_1 \oplus E_2$, then there exists a unique projection $P \in \mathrm{Hom}(E)$ with $E_1 = \mathrm{im}(P)$ and $E_2 = \mathrm{im}(I - P)$.*

*Proof.* (i) If $x \in E$, then since $P^2 = P$

$$P(I - P)x = Px - P^2x = Px - Px = 0,$$

so $\mathrm{im}(I-P) \subset \ker(P)$. On the other hand, if $x \in \ker(P)$, then $(I-P)x = x - Px = x$, so $x \in \mathrm{im}(I - P)$, so $\mathrm{im}(I - P) = \ker(P)$. Next note that $x = Px + (I - P)x$, so $E = \mathrm{im}(P) + \ker(P)$ by what we just proved. If $x \in \mathrm{im}(P) \cap \ker(P)$, then there exists $y \in E$ with $Py = x$. Since $P^2 = P$ and $x \in \ker(P)$ we have $x = Py = P^2y = Px = 0$. Hence we have $E = \mathrm{im}(P) \oplus \ker(P)$. Finally note that $(I - P)^2 = I - 2P + P^2 = I - 2P + P = I - P$, so $I - P$ is a projection as well.

(ii) Given $x \in E$ there are unique $x_i \in E_i$ $(i = 1, 2)$ such that $x = x_1 + x_2$. By the uniqueness of that decomposition the map $Px := x_1$ is linear, and also $P^2x = Px = x_1$ for all $x \in E$. Hence $P_1$ is a projection. Moreover, $x_2 = x - x_1 = x - Px$, so $E_2\, \mathrm{im}(I - P)$. Hence $P$ is the projection required. ■

We call $P$ the *projection of $E$ onto $E_1$ parallel to $E_2$*. Geometrically we have a situation as in Figure 33.1.

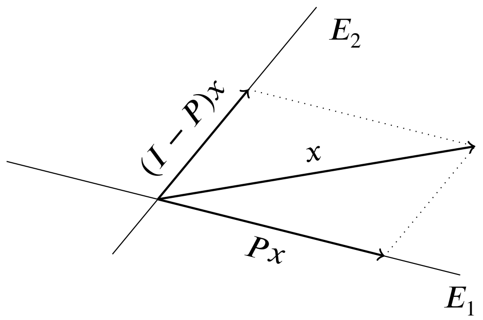



Figure 33.1: Projections onto complementary subspaces

**33.3 Proposition** *Let $E_1, E_2$ be subspaces of the Banach space $E$ with $E = E_1 \oplus E_2$ and $P \in \mathrm{Hom}(E)$ the associated projection onto $E_1$. Then $P \in \mathcal{L}(E)$ if and only if $E_1$ and $E_2$ are closed.*

*Proof.* By Lemma 33.2 we have $E_1 = \ker(I - P)$ and $E_2 = \ker(P)$. Hence if $P \in \mathcal{L}(E)$, then $E_1$ and $E_2$ are closed. Assume now that $E_1$ and $E_2$ are closed. To prove that $P$ is bounded we show $P$ has closed graph. Hence assume that $(x_n)$ is a sequence in $E$ with $x_n \to x$ and $Px_n \to y$ in $E$. Note that $Px_n \in E_1$ and $x_n - Px_n \in E_2$ for all $n \in \mathbb{N}$. Since $E_1$ and $E_2$ are closed, $Px_n \to y$ and $x_n - Px_n \to y - x \in E$ we conclude that $x \in E_1$ and $y - x \in E_2$. Now $x = y + (x - y)$, so $Px = y$ because $E = E_1 \oplus E_2$. Therefore $P$ has closed graph, and the closed graph theorem (Theorem 21.1) implies that $P \in \mathcal{L}(E)$. ■

**33.4 Definition (topological direct sum)** Suppose that $E_1, E_2$ are subspaces of the normed space $E$. We call $E = E_1 \oplus E_2$ a *topological direct sum* if the corresponding projection is bounded. In that case we say $E_2$ is a *topological complement* of $E_1$. Finally, a closed subspace of $E$ is called *complemented* if it has a topological complement.

**33.5 Remarks** (a) By Proposition 33.3 $E_1$ and $E_2$ must be closed if $E = E_1 \oplus E_2$ is a topological direct sum.

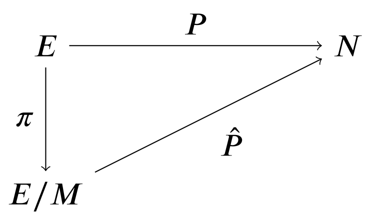
(b) By Theorem 16.9 every closed subspace of a Hilbert space is complemented. A rather recent result [9] shows that if every closed subspace of a Banach space is complemented, then its norm is equivalent to a norm induced by an inner product! Hence in general Banach spaces we do not expect every closed subspace to be complemented.

(c) Using Zorn's lemma 1.7 one can show that every subspace of a vector space has an algebraic complement.

We show that certain subspaces are always complemented.

**33.6 Lemma** *Suppose that $M$ is a closed subspace of the Banach space $E$ with finite co-dimension, that is,* $\dim(E/M) < \infty$. *Then $M$ is complemented.*

*Proof.* We know that $M$ has an algebraic complement $N$, that is, $E = M \oplus N$. Let $P$ denote the projection onto $N$. Then we have the following commutative diagram.




Since $M$ is a closed subspace, by Theorem 12.3 $E/M$ is a Banach space and $\hat{P}$ is an isomorphism. Since $\dim(E/M) < \infty$ the operator $\hat{P}$ is continuous, and so it is continuous. Therefore $P$ is continuous as well as a composition of two continuous operators. Hence we can apply Proposition 33.3 to conclude the proof. ∎

Suppose now that $E$ is a vector space and $E_1$ and $E_2$ are subspaces such that $E = E_1 \oplus E_2$. Denote by $P_1$ and $P_2$ the corresponding projections. If $A \in \mathrm{Hom}(E)$, then

$$\begin{aligned} Ax = P_1Ax + P_2Ax &= P_1A(P_1x + P_2x) + P_2A(P_1x + P_2x) \\ &= P_1AP_1x + P_1AP_2x + P_2AP_1x + P_2AP_2x \end{aligned}$$

for all $x \in E$. Setting

$$A_{ij} := P_iAP_j,$$

we have $A_{ij} \in \mathrm{Hom}(E_j, E_i)$. If we identify $x \in E$ with $(x_1, x_2) := (P_1x, P_2x) \in E_1 \times E_2$, then we can write

$$Ax = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

**33.7 Definition (complete reduction)** We say $E = E_1 \oplus E_2$ *completely reduces* $A$ if $A_{12} = A_{21} = 0$, that is,

$$A \sim \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$$

If that is the case we set $A_i := A_{ii}$ and write $A = A_1 \oplus A_2$.

We next characterise the reductions by means of the projections.

**33.8 Proposition** *Suppose that $E = E_1 \oplus E_2$ and that $P_1$ and $P_2$ are the corresponding projections. Let $A \in \mathrm{Hom}(E)$. Then $E = E_1 \oplus E_2$ completely reduces $A$ if and only if $P_1 A = AP_1$.*

*Proof.* Suppose that $E = E_1 \oplus E_2$ completely reduces $A$. Then

$$P_1 Ax = P_1 AP_1 x + P_1 AP_2 x = P_1 AP_1 x = P_1 AP_1 x + P_2 AP_1 x = AP_1 x.$$

To prove the converse Note that $P_1 P_2 = P_2 P_1 = 0$. Hence $A_{12} = P_1 AP_2 = AP_1 P_2 = 0$ and also $A_{21} = P_2 AP_1 = P_2 P_1 A = 0$, so $E = E_1 \oplus E_2$ completely reduces $A$. ■

Of course we could replace interchange the roles of $P_1$ and $P_2$ in the above proposition.

**33.9 Proposition** *Suppose that $E = E_1 \oplus E_2$ completely reduces $A \in \mathrm{Hom}(E)$. Let $A = A_1 \oplus A_2$ be this reduction. Then*

(i) $\ker A = \ker A_1 \oplus \ker A_2$*;*

(ii) $\operatorname{im} A = \operatorname{im} A_1 \oplus \operatorname{im} A_2$*;*

(iii) *$A$ is injective (surjective) if and only if $A_1$ and $A_2$ are injective (surjective);*

(iv) *if $A$ is bijective, then $E = E_1 \oplus E_2$ completely reduces $A^{-1}$ and $A^{-1} = A_1^{-1} \oplus A_2^{-1}$.*

*Proof.* (i) If $x = x_1 + x_2$ with $x_1 \in E_1$ and $x_2 \in E_2$, then $0 = Ax = A_1 x_1 + A_2 x_2$ if and only if $A_1 x_1 = 0$ and $A_2 x_2 = 0$. Hence (i) follows.

(ii) If $x = x_1 + x_2$ with $x_1 \in E_1$ and $x_2 \in E_2$, then $y = Ax = A_1 x_1 + A_2 x_2 = y_1 + y_2$ if and only if $A_1 x_1 = y_1$ and $A_2 x_2 = y_2$. Hence (ii) follows.

(iii) follows directly from (i) and (ii), and (iv) follows from the argument given in the proof of (ii). ■

We finally look at resolvent and spectrum of the reductions. It turns out that the spectral properties of $A$ can can be completely described by the operators $A_1$ and $A_2$.

**33.10 Proposition** *Suppose that $E$ is a Banach space and that $E = E_1 \oplus E_2$ a topological direct sum completely reducing $A \in \mathcal{L}(E)$. Let $A = A_1 \oplus A_2$ be this reduction. Then*

(i) $A_i \in \mathcal{L}(E_i)$ *($i = 1, 2$);*

(ii) $\sigma(A) = \sigma(A_1) \cup \sigma(A_2)$*;*

(iii) $\sigma_p(A) = \sigma_p(A_1) \cup \sigma_p(A_2)$.

*Proof.* By Proposition 33.3 the projections $P_1, P_2$ associated with $E = E_1 \oplus E_2$ are bounded. Hence $A_i = P_i A P_i$ is bounded for $i = 1, 2$. The equality in (ii) follows from Proposition 33.9(iv) which shows that $\varrho(A) = \varrho(A_1) \cap \varrho(A_2)$. Finally (iii) follows since by Proposition 33.9(i) shows that $\ker(\lambda I - A) = \{0\}$ if and only if $\ker(\lambda I_{E_1} - A_1) = \ker(\lambda I_{E_2} - A_2) = \{0\}$. ∎

## 34 The Ascent and Descent of an Operator

Let $E$ be a vector space over $\mathbb{K}$ and $T \in \operatorname{Hom}(E)$. We then consider two nested sequences of subspaces:

$$\{0\} = \ker T^0 \subseteq \ker T \subseteq \ker T^2 \subseteq \ker T^3 \subseteq \dots$$

and

$$E = \operatorname{im} T^0 \supseteq \operatorname{im} T \supseteq \operatorname{im} T^2 \supseteq \operatorname{im} T^3 \supseteq \dots.$$

We are interested in $n$ such that there is equality.

**34.1 Lemma** *If* $\ker T^n = \ker T^{n+1}$, *then* $\ker T^n = \ker T^{n+k}$ *for all* $k \in \mathbb{N}$. *Moreover, if* $\operatorname{im} T^n = \operatorname{im} T^{n+1}$, *then* $\operatorname{im} T^n = \operatorname{im} T^{n+k}$ *for all* $k \in \mathbb{N}$.

*Proof.* If $k \geq 1$ and $x \in \ker T^{n+k}$, then $0 = T^{n+k}x = T^{n+1}(T^{k-1}x)$ and so $T^{k-1}x \in \ker T^n$. By assumption $\ker T^n = \ker T^{n+1}$ and therefore $T^{n+k-1}x = T^n(T^{k-1}x) = 0$. Hence $\ker T^{n+k} = \ker T^{n+k-1}$ for all $k \geq 1$ from which the first assertion follows. The second assertion is proved similarly. If $k \geq 1$ and $x \in \operatorname{im} T^{n+k-1}$, then there exists $y \in E$ with $x = T^{n+k-1}y = T^{k-1}(T^n y)$ and so $T^n y \in \operatorname{im} T^n$. By assumption $\operatorname{im} T^n = \operatorname{im} T^{n+1}$ and therefore there exists $z \in \operatorname{im} T^{n+1}$ with $T^{n+1}z = T^n y$. Thus $T^{n+k}z = T^{k-1}(T^{n+1}z) = T^{k-1}(T^n y) = T^{n+k-1}y = x$ which implies that $\operatorname{im} T^{n+k-1} = \operatorname{im} T^{n+k}$ for all $k \geq 1$. This completes the proof of the lemma. ∎

The above lemma motivates the following definition.

**34.2 Definition (ascent and descent)** We call

$$\alpha(T) := \inf\{n \geq 1 : \ \ker T^n = \ker T^{n+1}\}$$

the *ascent* of $T$ and

$$\delta(T) := \inf\{n \geq 1 : \ \operatorname{im} T^n = \operatorname{im} T^{n+1}\}$$

the *descent* of $T$

If $\dim E = N < \infty$, then ascent and descent are clearly finite and since $\dim(\ker T^n) + \dim(\operatorname{im}(T^n)) = N$ for all $n \in \mathbb{N}$, ascent and descent are always equal. We now show that they are equal provided they are both finite. The arguments we use are similar to those to prove the Jordan decomposition theorem for matrices. The additional complication is that in infinite dimensions there is no dimension formula.

**34.3 Theorem** *Suppose that $\alpha(T)$ and $\delta(T)$ are both finite. Then $\alpha(T) = \delta(T)$. Moreover, if we set $n := \alpha(T) = \delta(T)$, then*

$$E = \ker T^n \oplus \operatorname{im} T^n$$

*and the above direct sum completely reduces $T$.*

*Proof.* Let $\alpha := \alpha(T)$ and $\delta := \delta(T)$. The plan is to show that

$$\operatorname{im} T^\alpha \cap \ker T^k = \{0\} \tag{34.1}$$

and that

$$\operatorname{im} T^k + \ker T^\delta = E \tag{34.2}$$

for all $k \geq 1$. Choosing $k = \delta$ in (34.1) and $k = \alpha$ in (34.2) we then get

$$\operatorname{im} T^\alpha \oplus \ker T^\delta = E. \tag{34.3}$$

To show that $\alpha = \delta$ assume that $x \in \ker T^{\delta+1}$. By (34.3) there is a unique decomposition $x = x_1 + x_2$ with $x_1 \in \operatorname{im} T^\alpha$ and $x_2 \in \ker T^\delta$. Because $\ker T^\delta \subseteq \ker T^{\delta+1}$ we get $x_1 = x - x_2 \in \ker T^{\delta+1}$. We also have $x_1 \in \operatorname{im} T^\alpha$, so by (34.1) we conclude that $x_1 = 0$. Hence $x = x_2 \in \ker T^\delta$, so that $\ker T^\delta = \ker T^{\delta+1}$. By definition of the ascent we have $\alpha \leq \delta$. Assuming that $\alpha < \delta$, then by definition of the descent $\operatorname{im} T^\delta \subsetneq \operatorname{im} T^\alpha$. By (34.2) we still have $\operatorname{im} T^\delta + \ker T^\delta = E$, contradicting (34.3). Hence $\alpha = \delta$.

To prove (34.1) fix $k \geq 1$ and let $x \in \operatorname{im} T^\alpha \cap \ker T^k$. Then $x = T^\alpha y$ for some $y \in E$ and $0 = T^k x = T^{\alpha+k} y$. As $\ker T^\alpha = \ker T^{\alpha+k}$ we also have $x = T^\alpha y = 0$. To prove (34.2) fix $k \geq 1$ and let $x \in E$. As $\operatorname{im} T^\delta = \operatorname{im} T^{\delta+k}$ there exists $y \in E$ with $T^\delta x = T^{\delta+k} y$. Hence $T^\delta(x - T^k y) = 0$, proving that $x - T^k y \in \ker T^\delta$. Hence we can write $x = T^k y + (x - T^k y)$ from which (34.2) follows.

We finally prove that (34.3) completely reduces $T$. By definition of $n = \alpha = \delta$ we have $T(\ker T^n) \subseteq \ker T^{n+1} = \ker T^n$ and $T(\operatorname{im} T^n) \subseteq \operatorname{im} T^{n+1} = \operatorname{im} T^n$. This completes the proof of the theorem. ■

We well apply the above to the operator $\lambda I - A$, where $A$ is a bounded operator on a Banach space and $\lambda \in \sigma_p(A)$.

**34.4 Definition (algebraic multiplicity)** Let $E$ be a Banach space and over $\mathbb{C}$. If $A \in \mathcal{L}(E)$ and $\lambda \in \sigma_p(A)$ is an eigenvalue we call

$$\dim \ker(\lambda I - A)$$

the *geometric multiplicity* of $\lambda$ and

$$\dim\Big(\bigcup_{n\in\mathbb{N}} \ker(\lambda I - A)^n\Big)$$

the *algebraic multiplicity* of $\lambda$. The ascent $\alpha(\lambda I - A)$ is called the *Riesz index* of the eigenvalue $\lambda$.

Note that the definitions here are consistent with the ones from linear algebra. As it turns out the algebraic multiplicity of the eigenvalue of a matrix as defined above coincides with the multiplicity of $\lambda$ as a root of the characteristic polynomial. In infinite dimensions there is no characteristic polynomial, so we use the above as the algebraic multiplicity. The only question now is whether there are classes of operators for which ascent and descent of $\lambda I - A$ are finite for every or some eigenvalues. It turns out that such properties are connected to compactness. This is the topic of the next section.

# 35 The Spectrum of Compact Operators

We want to study a class of operators appearing frequently in the theory of partial differential equations and elsewhere. They share many properties with operators on finite dimensional spaces.

**35.1 Definition (compact operator)** Let $E$ and $F$ be Banach spaces. Then $T \in \operatorname{Hom}(E, F)$ is called *compact if* $\overline{T(B)}$ is compact for all bounded sets $B \subset E$. We set

$$\mathcal{K}(E, F) := \{T \in \operatorname{Hom}(E, F) : T \text{ is compact}\}$$

and $\mathcal{K}(E) := \mathcal{L}(E, E)$.

**35.2 Remark** By the linearity $T$ is clearly compact if and only if $\overline{T(B(0,1))}$ is compact.

**35.3 Proposition** *Let $E, F$ be Banach spaces. Then $\mathcal{K}(E, F)$ is a closed subspace of $\mathcal{L}(E, F)$.*

*Proof.* If $B \subset E$ is bounded, then by assumption the closure of $T(B)$ is compact, and therefore bounded. To show that $\mathcal{K}(E, F)$ is a subspace of $\mathcal{L}(E, F)$ let $T, S \in \mathcal{K}(E, F)$ and $\lambda, \mu \in \mathbb{K}$. Let $B \subset E$ be bounded and let $(y_n)$ be a sequence in $(\lambda T + \mu S)(B)$. Then there exist $x_n \in B$ with $y_n = \lambda T x_n + \mu S x_n$. Since $B$ is bounded the sequence $(x_n)$ is bounded as well. By the compactness of $T$ there exists a subsequence $(x_{n_k})$ with $T x_{n_k} \to y$. By the compactness of $S$ we can select a further subsequence $(x_{n_{k_j}})$ with $S x_{n_{k_j}} \to z$. Hence $y_{n_{k_j}} = \lambda T x_{n_{k_j}} + \mu S x_{n_{k_j}} \to \lambda y + \mu z$. This shows that every sequence in $(\lambda T + \mu S)(B)$ has a convergent subsequence, so by Theorem 4.4 the closure of that set is compact. Hence $\lambda T + \mu S$ is a compact operator.

Suppose finally that $T_n \in \mathcal{K}(E, F)$ with $T_n \to T$ in $T \in \mathcal{L}(E, F)$. It is sufficient to show that $\overline{T(B)}$ is compact if $B = B(0, 1)$ is the unit ball. We show that $\overline{T(B)}$ is totally bounded. To do so fix $\epsilon > 0$. Since $T_n \to T$ there exists $n \in \mathbb{N}$ such that

$$\|T_n - T\|_{\mathcal{L}(E,F)} < \frac{\varepsilon}{4}. \tag{35.1}$$

By assumption $T_n$ is compact and so $\overline{T_n(B)}$ is compact. Therefore there exist $y_k \in \overline{T_n(B)}$, $k = 1, \dots, m$, such that

$$T_n(B) \subseteq \bigcup_{k=1}^{m} B(y_k, \varepsilon/4). \tag{35.2}$$

If we let $y \in T(B)$, then $y = Tx$ for some $x \in B$. By (35.2) there exists $k \in \{1, \dots, m\}$ such that $T_n x \in B(y_k, \varepsilon/4)$. By (35.1) and since $\|x\| \leq 1$ we get

$$\|y_k - y\| = \|y_k - T_n y\| + \|T_n y - Tx\| \leq \frac{\varepsilon}{4} + \frac{\varepsilon}{4}\|x\| \leq \frac{\varepsilon}{2},$$

so that $y \in B(y_k, \varepsilon/2)$. Since $y \in B$ was arbitrary this shows that

$$T(B) \subseteq \bigcup_{k=1}^{m} B(y_k, \varepsilon/2)$$

and therefore

$$\overline{T(B)} \subseteq \overline{\bigcup_{k=1}^{m} B(y_k, \varepsilon/2)} \subseteq \bigcup_{k=1}^{m} B(y_k, \varepsilon).$$

Since $\varepsilon > 0$ was arbitrary $\overline{T(B)}$ is totally bounded. Since $F$ is complete Theorem 4.4 implies that $\overline{T(B)}$ is compact. ∎

We now look at compositions of operators.

**35.4 Proposition** *Let $E, F, G$ be Banach spaces. If $K \in \mathcal{K}(F, G)$, then $KT \in \mathcal{K}(E, F)$ for all $T \in \mathcal{L}(E, F)$. Likewise, if $K \in \mathcal{K}(E, F)$, then $TK \in \mathcal{K}(E, G)$ for all $T \in \mathcal{L}(F, G)$.*

*Proof.* If $K \in \mathcal{K}(F, G)$, $T \in \mathcal{L}(E, F)$ and $B \subset E$ bounded, then $T(B)$ is bounded in $F$ and therefore $\overline{KT(B)}$ is compact in $G$. Hence $KT$ is compact. If $K \in \mathcal{K}(E, F)$, $T \in \mathcal{L}(F, G)$ and $B \subset E$ bounded, then $\overline{K(B)}$ is compact in $F$. Since $T$ is continuous $T(\overline{K(B)})$ is compact in $G$, so $\overline{TK(B)} \subset T(\overline{K(B)})$ is compact. Hence $TK$ is a compact operator. ∎

**35.5 Remark** If we define multiplication of operators in $\mathcal{L}(E)$ to be composition of operators, then $\mathcal{L}(E)$ becomes a non-commutative algebra. The above propositions show that $\mathcal{K}(E)$ is a closed ideal in that algebra.

The next theorem will be useful to characterise the spectrum of compact operators.

**35.6 Theorem** *Let $T \in \mathcal{K}(E)$ and $\lambda \in \mathbb{K} \setminus \{0\}$. Then for all $k \in \mathbb{N}$*

(i) $\dim\big(\ker(\lambda I - T)^k\big) < \infty$;

(ii) $\operatorname{im}(\lambda I - T)^k$ *is closed in $E$.*

*Proof.* Since $\ker(\lambda I - T)^k = \ker(I - \lambda^{-1}T)^k$ and $\operatorname{im}(\lambda I - T)^k = \operatorname{im}(I - \lambda^{-1}T)^k$ we can assume without loss of generality that $\lambda = 1$ by replacing $T$ by $\lambda T$. Also, we can reduce the proof to the case $k = 1$ because

$$(I - K)^k = \sum_{j=0}^{k} (-1)^j \binom{k}{j} T^j = I - \sum_{j=1}^{k} (-1)^{j-1} \binom{k}{j} T^j = I - \tilde{T}$$

if we set $\tilde{T} := \sum_{j=1}^{k} (-1)^{j-1} \binom{k}{j} T^j$. Note that by Proposition 35.4 the operator $\tilde{T}$ is compact.

(i) If $x \in \ker(I - T)$, then $x = Tx$ and so

$$S := \{x \in \ker(I - T) : \|x\| = 1\} \subseteq \overline{T(B(0, 1))}.$$

Since $T$ is compact $S$ is compact as well. Note that $S$ is the unit sphere in $\ker(I - T)$, so by Theorem 11.3 the kernel of $I - T$ is finite dimensional.

(ii) Set $S := I - T$. As $\ker S$ is closed $M := E/\ker S$ is a Banach space by Theorem 12.3 with norm $\|x\|_M := \inf_{y \in \ker S} \|y - x\|$. Now set $F := \overline{\operatorname{im}(S)}$. Let

$\hat{S} \in \mathcal{L}(M, F)$ be the map induced by $S$. We show that $\hat{S}^{-1}$ is bounded on $\operatorname{im}(S)$. If not, then there exist $y_n \in \operatorname{im}(S)$ such that $y_n \to 0$ but $\|\hat{S}^{-1} y_n\|_M = 1$ for all $n \in \mathbb{N}$. By definition of the quotient norm there exist $x_n \in E$ with $1 \leq \|x_n\|_E \leq 2$ such that

$$\hat{S}[x_n] = x_n - Tx_n = y_n \to 0.$$

Since $T$ is compact there exists a subsequence such that $Tx_{n_k} \to z$. Hence from the above

$$x_{n_k} = y_{n_k} + Tx_{n_k} \to z.$$

By the continuity of $T$ we have $Tx_{n_k} \to Tz$, so that $z = Tz$. Hence $z \in \ker(I - T)$ and $[x_{n_k}] \to [z] = [0]$ in $M$. Since $\|[x_n]\|_M = 1$ for all $n \in \mathbb{N}$ this is not possible. Hence $\hat{S}^{-1} \in \mathcal{L}(\operatorname{im}(S), M)$ and by Proposition 23.4 it follows that $F = \operatorname{im}(\hat{S}) = \operatorname{im}(I - T)$ is closed. ∎

We next prove that the ascent and descent of $\lambda I - T$ is finite if $T$ is compact and $\lambda \neq 0$.

**35.7 Theorem** *Let $T \in \mathcal{K}(E)$ and $\lambda \in \mathbb{K} \setminus \{0\}$. Then $n := \alpha(\lambda I - T) = \delta(\lambda I - T) < \infty$ and*

$$E = \ker(\lambda I - T)^n \oplus \operatorname{im}(\lambda I - T)^n$$

*is a topological direct sum completely reducing $\mu I - T$ for all $\mu \in \mathbb{C}$.*

*Proof.* Replacing $T$ by $\lambda T$ we can assume without loss of generality that $\lambda = 1$. Suppose that $\alpha(\lambda I - T) = \infty$. Then $N_k := \ker(I - T)^k$ is closed in $E$ for every $k \in \mathbb{N}$ and

$$N_0 \subset N_1 \subset N_2 \subset N_3 \subset \dots,$$

where all inclusions are proper inclusions. By Corollary 11.2 there exist $x_k \in N_k$ such that

$$\|x_k\| = 1 \qquad \text{and} \qquad \operatorname{dist}(x_k, N_{k-1}) \geq \frac{1}{2}$$

for all $k \geq 1$. Now

$$\begin{aligned} Tx_k - Tx_\ell &= (I - T)x_\ell - (I - T)x_k + x_k - x_\ell \\ &= x_k - \bigl(x_\ell - (I - T)x_\ell + (I - T)x_k\bigr) = x_k - z, \end{aligned}$$

where $z := x_\ell - (I - T)x_\ell + (I - T)x_k \in \ker(I - T)^{k-1} = N_{k-1}$ if $1 \leq \ell < k$. Hence by choice of $(x_k)$ we get

$$\|Tx_k - Tx_\ell\| = \|x_k - z\| \geq \frac{1}{2}$$

for all $1 \leq \ell < k$. Therefore the bounded sequence $Tx_k$ does not have a convergent subsequence, so $T$ is not compact. As a consequence, if $T \in \mathcal{K}(E, F)$ is compact, then $\alpha(I - T) < \infty$.

The proof for the descent works quite similarly. We assume that $\delta(I - T) = \infty$. Then by Theorem 35.6 $M_k := \operatorname{im}(I - T)^k$ is closed in $E$ for every $k \in \mathbb{N}$ and

$$M_0 \supset M_1 \supset M_2 \supset M_3 \supset \dots,$$

where all inclusions are proper inclusions. By Corollary 11.2 there exist $x_k \in M_k$ such that

$$\|x_k\| = 1 \qquad \text{and} \qquad \operatorname{dist}(x_k, M_{k+1}) \geq \frac{1}{2}$$

for all $k \geq 1$. Now

$$\begin{aligned} Tx_k - Tx_\ell = (I-T)x_\ell - (I-T)x_k - x_k + x_\ell \\ = x_\ell - \big(x_k - (I-T)x_\ell + (I-T)x_k\big) = x_k - z, \end{aligned}$$

where $z := x_k - (I-T)x_\ell + (I-T)x_k \in \operatorname{im}(I-T)^{k+1} = M_{k+1}$ if $1 \leq k < \ell$. Hence by choice of $(x_k)$ we get

$$\|Tx_k - Tx_\ell\| = \|x_\ell - z\| \geq \frac{1}{2}$$

for all $1 \leq k < \ell$. Therefore the bounded sequence $Tx_k$ does not have a convergent subsequence, so $T$ is not compact. As a consequence, if $T \in \mathcal{K}(E, F)$ is compact, then $\delta(I-T) < \infty$.

By Theorem 34.3 $n := \alpha(I-T) = \delta(I-T)$ and $E = \ker(I-T)^n \oplus \operatorname{im}(I-T)^n$ is a direct sum completely reducing $I - T$ and therefore $\mu I - T$ for all $\mu \in \mathbb{K}$. By Theorem 35.6 the subspaces $\ker(I-T)^n$ and $\operatorname{im}(I-T)^n$ are closed, so by 33.3 the above is a topological direct sum ■

As a consequence of the above theorem we prove that compact operators behave quite similarly to operators on finite dimensional spaces. The following corollary would be proved by the dimension formula in finite dimensions.

**35.8 Corollary** *Suppose that $T \in \mathcal{K}(E)$ and $\lambda \neq 0$. Then $\lambda I - T$ is injective if and only if it is surjective.*

*Proof.* If $\lambda I - T$ is injective, then $\{0\} = \ker(\lambda I - T) = \ker(\lambda I - T)^2$, and by Theorem 35.7 we get

$$E = \ker(\lambda I - T) \oplus \operatorname{im}(\lambda I - T) = \{0\} \oplus \operatorname{im}(\lambda I - T) = \operatorname{im}(\lambda I - T). \tag{35.3}$$

On the other hand, if $E = \operatorname{im}(\lambda I - T)$, then $\operatorname{im}(\lambda I - T) = \operatorname{im}(\lambda I - T)^2$, so that $\delta(\lambda I - T) = 1$. Hence $\ker(\lambda I - T) = \{0\}$ by Theorem 35.7 and (35.3). ■

We finally derive a complete description of the spectrum of closed operators.

**35.9 Theorem (Riesz-Schauder)** *Let $E$ be a Banach space over $\mathbb{C}$ and $T \in \mathcal{K}(E)$ compact. Then $\sigma(T) \setminus \{0\}$ consists of at most countably many isolated eigenvalues of finite algebraic multiplicity. The only possible accumulation point of these eigenvalues is zero.*

*Proof.* By Corollary 35.8 and Proposition 32.3 every $\lambda \in \sigma(T) \setminus \{0\}$ is an eigenvalue of $T$. By Theorems 35.6 and 35.7 these eigenvalues have finite algebraic multiplicity. Hence it remains to show that they are isolated. Fix $\lambda \in \sigma(T) \setminus \{0\}$ and let $n$ its Riesz index (see Definition 34.4). Then by Theorem 35.7

$$E = \ker(\lambda I - T)^n \oplus \operatorname{im}(\lambda I - T)^n$$

is a topological direct sum completely reducing $\mu I - T$ for all $\mu \in \mathbb{C}$. We set $N := \ker(\lambda I - T)^n$ and $M := \operatorname{im}(\lambda I - T)^n$. Let $T = T_N \oplus T_M$ be the reduction associated with the above direct sum decomposition. By construction $\lambda I - T_N$ is nilpotent, so by Theorem 32.10 $\operatorname{spr}(\lambda I - T_N) = 0$. In particular this means that $\sigma(T_N) = \{\lambda\}$. Moreover, by construction, $\lambda I - T_M$ is injective and so by Proposition 32.3 and Corollary 35.8 $\lambda \in \varrho(T_M)$. As $\varrho(T_M)$ is open by Theorem 32.7 there exists $r > 0$ such that $B(\lambda, r) \subset \varrho(T_M)$. By Theorem 33.10 $\varrho(T) = \varrho(T_N) \cap \varrho(T_M)$, so $B(\lambda, r) \setminus \{\lambda\} \subset \varrho(T)$, showing that $\lambda$ is an isolated eigenvalue. Since the above works for every $\lambda \in \sigma(T) \setminus \{0\}$ the assertion of the theorem follows. ∎

**35.10 Remarks** (a) If $T \in \mathcal{K}(E)$ and $\dim E = \infty$, then $0 \in \sigma(T)$. If not, then by Proposition 35.4 the identity operator $I = TT^{-1}$ is compact. Hence the unit sphere in $E$ is compact and so by Theorem 11.3 $E$ is finite dimensional.

(b) The above theorem allows us to decompose a compact operator $T \in \mathcal{L}(E)$ in the following way. Let $\lambda_k \in \sigma_p(T) \setminus \{0\}$, $k = 1, \dots, m$, and let $n_k$ be the corresponding Riesz indices. Then set

$$N_k := \ker(\lambda I - T)^{n_k}$$

and $T_k := T|_{N_k}$. Then there exists a closed subspace $M \subset E$ which is invariant under $T$ such that

$$E = N_1 \oplus N_2 \oplus \cdots \oplus N_m \oplus M$$

completely reduces $T = T_1 \oplus T_2 \oplus \cdots \oplus T_m \oplus T_M$. To get this reduction apply Theorem 35.9 inductively.

The above theory is also useful for certain classes of closed operators $A : D(A) \subseteq E \to E$, namely those for which $R(\lambda, A)$ is compact for some $\lambda \in \varrho(A)$. In that case the resolvent identity from Theorem 32.7 and Proposition 35.4 imply that $R(\lambda, A)$ is compact for all $\lambda \in \varrho(A)$. It turns out that $\sigma(A) = \sigma_p(A)$ consist of isolated eigenvalues, and the only possible point of accumulation is infinity. Examples of such a situation include boundary value problems for partial differential equations such as the one in Example 24.5.

# Bibliography


[1] G. Birkhoff and G.-C. Rota. "On the completeness of Sturm-Liouville expansions". In: *Amer. Math. Monthly* 67 (1960), pp. 835–841. DOI: 10.2307/2309440.

[2] E. Bishop and D. Bridges. *Constructive analysis*. Vol. 279. Grundlehren der Mathematischen Wissenschaften. Berlin: Springer-Verlag, 1985, pp. xii+477.

[3] A. P. Campbell and D. Daners. "Linear Algebra via Complex Analysis". In: *Amer. Math. Monthly* 120.10 (2013), pp. 877–892. DOI: 10.4169/amer.math.monthly.120.10.877.

[4] P. Cohen. "The independence of the continuum hypothesis". In: *Proc. Nat. Acad. Sci. U.S.A.* 50 (1963), pp. 1143–1148.

[5] P. J. Cohen. "The independence of the continuum hypothesis. II". In: *Proc. Nat. Acad. Sci. U.S.A.* 51 (1964), pp. 105–110.

[6] D. Daners. "Uniform continuity of continuous functions on compact metric spaces". In: *Amer. Math. Monthly* 122.6 (2015), p. 592. DOI: 10.4169/amer.math.monthly.122.6.592.

[7] J. Dugundji. *Topology*. Boston, Mass.: Allyn and Bacon Inc., 1966, pp. xvi+447.

[8] K. Gödel. "The consistency of the axiom of choice and of the generalized continuum-hypothesis". In: *Proc. Nat. Acad. Sci. U.S.A.* 24 (1938), pp. 556–557.

[9] J. Lindenstrauss and L. Tzafriri. "On the complemented subspaces problem". In: *Israel J. Math.* 9 (1971), pp. 263–269.

[10] W. Rudin. *Real and Complex Analysis*. 2nd. New York: McGraw-Hill Inc., 1974.

[11] A. D. Sokal. "A Really Simple Elementary Proof of the Uniform Boundedness Theorem". In: *Amer. Math. Monthly* 118.5 (2011), pp. 450–452. DOI: 10.4169/amer.math.monthly.118.05.450.

[12] M. H. Stone. "The generalized Weierstrass approximation theorem". In: *Math. Mag.* 21 (1948), pp. 167–184, 237–254. DOI: 10.2307/3029750.

[13] K. Yosida. *Functional analysis*. 6th. Vol. 123. Grundlehren der Mathematischen Wissenschaften. Berlin: Springer-Verlag, 1980, pp. xii+501.

[14] E. Zermelo. "Beweis, daß jede Menge wohlgeordnet werden kann". In: *Math. Ann.* 59 (1904), pp. 514–516.