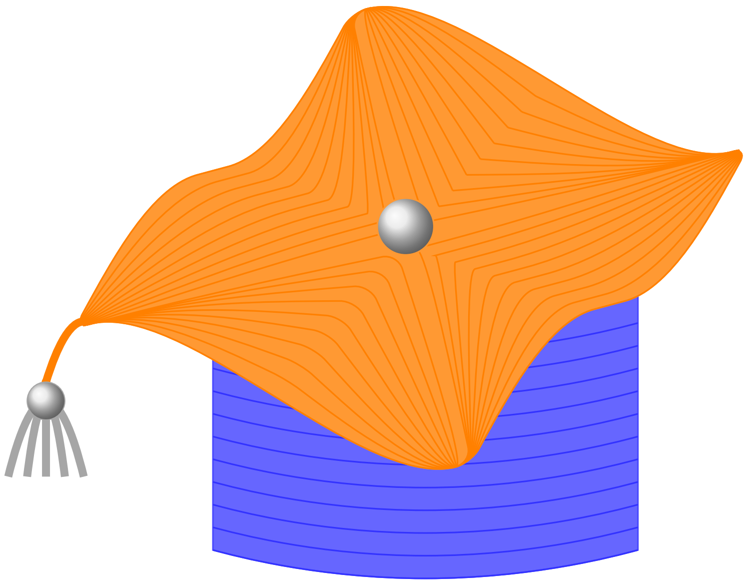
dr. R.R. van Hassel

Functional Analysis



March 11, 2014

Contents

1 Preface

You are reading the second edition of some lecture notes, which try to give an introduction to Functional Analysis. The second edition differs from the first edition. Further there are added more and more subjects and it becomes the question if there can be spoken about an introduction to the functional analysis. It becomes more and more a kind of overview of the functional analysis.

It is also possible that this is the ?-th edition of these lecture notes. My advice is, if you want to know something, look in the Index or the Contents and try to find everything that is needed to understand your particular problem. If you start with reading from the first sentence of these notes, it takes a long time before you come into the world of the functional analysis. As already said, these lectures notes become more and more a kind of overview of everything and nothing. I hope you can use these notes, more I can not do.

To me was asked is to treat the chapters 2 and 3 out of the book (Kreyszig, 1978). To understand these chapters, it is also needed to do parts out of chapter 1. These parts will be done if needed.
During the writing[1] of these lecture notes is made use[2] of the books of (Kreyszig, 1978), (Sutherland, 1975), (Griffel, 1981) and (Jain et al., 1996). Naturally there are used also other books and there is made use of lecture notes of various authors. Therefore here below a little invitation to look at internet. With "little" is meant, to be careful with your time and not to become enslaved in searching to books and lecture notes going about Functional Analysis. To search information is not so difficult, but to learn from the founded information is quite another discipline.
On the internet there are very much free available lectures notes, see for instance **Chen-1**. Before October 2009, there was also the site **geocities**, but this site is no longer available! Let's hope that something like geocities comes back! There are some initiatives to save the data of geocities, but at this moment: May 2011, I have no idea, where I can find the saved data.
It is also possible to download complete books, see for instance **esnips** or **kniga**. Searching with "functional analysis" and you will find the necessary documents, most of the time .djvu and/or .pdf files.
Be careful where you are looking, because there are two kinds of "functional analyses":

1. Mathematics:

[1] It still goes on, René.
[2] Also is made use of the wonderful TeX macro package ConTeXt, see **context**.

A branch of analysis which studies the properties of mappings of classes of functions from one topological vector space to another.

2. Systems Engineering:
   A part of the design process that addresses the activities that a system, software, or organization must perform to achieve its desired outputs, that is, the transformations necessary to turn available inputs into the desired outputs.

The first one will be studied.
Expressions or other things, which can be find in the Index, are given by a lightgray color in the text, such for instance functional analysis .
The internet gives a large amount of information about mathematics. It is worth to mention the wiki-encyclopedia **wiki-FA**. Within this encyclopedia there are made links to other mathematical sites, which are worth to read. Another site which has to be mentioned is **wolfram-index**, look what is written by Functional Analysis, **wolfram-FA**.
For cheap printed books about Functional Analysis look to **NewAge-publ**. The mentioned publisher has several books about Functional Analysis. The book of (Jain et al., 1996) is easy to read, the other books are going about a certain application of the Functional Analysis. The website of **Alibris** has also cheap books about Functional Analysis, used books as well as copies of books.
Problems with the mathematical analysis? Then it is may be good to look in **Math-Anal-Koerner**. From the last mentioned book, there is also a book with the answers of most of the exercises out of that book.
If there is need for a mathematical fitness program see then **Shankar-fitness**. Downloading the last two mentioned books needs some patience.

2 Preliminaries

A short overview will be given of all kind of terms, which are used in the chapters behind this one. It is not the intention to give a complete overview of the analysis on the $\mathbb{R}^n$, in this chapter.
Since the Functional Analysis is kind of generalisation of the analysis already known, so it is hard to present everything in one unbroken line. So there will be sometimes referred to paragraphs further on in the lectures notes. It is of importance to read these references.

2.1 Mappings

If X and Y are sets and $A \subseteq X$ any subset of X. A mapping $T : A \to Y$ is some relation, such that for each $x \in A$, there exists a single element $y \in Y$, such that $y = T(x)$. If $y = T(x)$ then y is called the image of x with respect to T.
Such a mapping T can also be called a function, a transformation or an operator. The name depends of the situation in which such a mapping is used. It also depends on the properties of the sets X and Y. If X and Y are vector spaces (Chapter 3.2), in particular normed spaces (Chapter 3.7), a map T is called an operator.
Such a mapping is may be not defined on the whole of X, but only a certain subset of X, such a subset is called the domain of T, denoted by $\mathcal{D}(T)$.
Some people make a distinction between a map and a function. If $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$, they speak about functions and not about maps. In these Lecture Notes there is not really made a strict distinction.
The set of all images of T is called the range of T, denoted by $\mathcal{R}(T)$,

$$\mathcal{R}(T) = \{y \in Y \mid y = T(x) \text{ for some } x \in \mathcal{D}(T)\}. \tag{2.1}$$

The set of all elements out of $x \in \mathcal{D}(T)$, such that $T(x) = 0$, is called the nullspace of T and denoted by $\mathcal{N}(T)$,

$$\mathcal{N}(T) = \{x \in \mathcal{D}(T) | T(x) = 0\}. \tag{2.2}$$

If $M \subset \mathcal{D}(T)$ then $T(M)$ is called the image of the subset M, note that $T(\mathcal{D}(T)) = \mathcal{R}(T)$.
Two properties called *one-to-one* and *onto* are of importance, if there is searched for a mapping from the range of T to the domain of T. Going back it is of importance

that every $y_0 \in \mathcal{R}(T)$ is the image of just one element $x_0 \in \mathcal{D}(T)$. This means that y_0 has a unique original.
A mapping T is called one-to-one if for every $x, y \in \mathcal{D}(T)$

$$x \neq y \Longrightarrow T(x) \neq T(y). \tag{2.3}$$

It is only a little bit difficult to use that definition. Another equivalent definition is

$$T(x) = T(y) \Longrightarrow x = y. \tag{2.4}$$

If T satisfies one of these properties, T is also called injective, T is an injection, or T is one-to-one.
A mapping $T : \mathcal{D}(T) \rightarrow Y$ is said to be onto if $\mathcal{R}(T) = Y$, or

$$\forall y \in Y \text{ there exists a } x \in \mathcal{D}(T), \text{ such that } y = T(x). \tag{2.5}$$

Note that $T : \mathcal{D}(T) \rightarrow \mathcal{R}(T)$ is always onto. If T is onto, it is also called surjective, T is an surjection or T is onto.
If $T : \mathcal{D}(T) \rightarrow Y$ is one-to-one and onto then T is called bijective, T is an bijection. This means that there exists an inverse mapping T^{-1} of T, with $T^{-1} : Y \rightarrow \mathcal{D}(T)$. Since for every $y \in Y$ there exists an unique $x \in \mathcal{D}(T)$, such that $T(x) = y$, the function T^{-1} is defined by $T^{-1}(y) = x$.
And so you have that $T^{-1}T = I$ with I the identity mapping on $\mathcal{D}(T)$ and $T\,T^{-1} = I$ with I the identity mapping on Y. Sometimes the identity mapping has some index, such that you know, about what identity is spoken. For instance I_X, the identity mapping on X and I_Y, the identity mapping on Y.

2.2 Bounded, open and closed subsets

The definitions will be given for subsets in $\mathbb{R}^n$ for some $n \in \mathbb{N}$. On $\mathbb{R}^n$, there is defined a mapping to measure distances between points in $\mathbb{R}^n$. A norm, notated by $\| \,.\, \|$, see definition 3.23, can be used to measure the distance between points. More general can be used a metric, notated by $d(.,.)$, see definition 3.18.
A subset $A \subset \mathbb{R}^n$ is bounded, if there exists a $K \in \mathbb{R}$ such that

$$\| \, x - y \, \| \leq K, \tag{2.6}$$

for all $x \in A$ and a fixed $y \in \mathbb{R}^n$.
An open ball, with radius $\epsilon > 0$ around some point $x_0 \in \mathbb{R}^n$ is written by $B_\epsilon(x_0)$ and defined by

$$B_\epsilon(x_0) = \{x \in \mathbb{R}^n | \parallel x - x_0 \parallel < \epsilon\}. \tag{2.7}$$

A subset $A \subset \mathbb{R}^n$ is open, if for every $x \in A$ there exists an $\epsilon > 0$, such that $B_\epsilon(x) \subset A$.
The complement of A is written by A^c and defined by

$$A^c = \{x \in \mathbb{R}^n | x \notin A\}. \tag{2.8}$$

A subset $A \subset \mathbb{R}^n$ is closed, if A^c is open.
If A and B are sets, then the relative complement of A in B is defined by

$$B \backslash A = \{x \in B \,|\, x \notin A\}, \tag{2.9}$$

in certain sense: set B minus set A.

2.3 Convergent and limits

Sequences $\{x_n\}_{n \in \mathbb{N}}$ are of importance to study the behaviour of all kind of different spaces and also mappings. Most of the time, there will be looked if a sequence is convergent or not? There will be looked if a sequence has a limit. The sequence $\{\lambda_n\}_{n \in \mathbb{N}}$ has limit λ if for every $\epsilon > 0$ there exists a $N(\epsilon)$ such that for every $n > N(\epsilon)$, $\parallel \lambda_n - \lambda \parallel < \epsilon$.
Sometimes it is difficult to calculate λ, and so also difficult to look if a sequence converges. But if a sequence converges, it is a Cauchy sequence. The sequence $\{\lambda_n\}_{n \in \mathbb{N}}$ is a Cauchy sequence, if for every $\epsilon > 0$ there exists a $N(\epsilon)$ such that for every $m, n > N(\epsilon)$, $\parallel \lambda_m - \lambda_n \parallel < \epsilon$. Only elements of the sequence are needed and not the limit of the sequence.
But be careful, a convergent sequence is a Cauchy sequence, but not every Cauchy sequence converges!
A space is called complete if every Cauchy sequence in that space converges.
If there is looked at a sequence, it is important to look to the tail of that sequence. In some cases the tail has to converge to a constant and in other situations it is of

importance that these terms become small. In some literature the authors define explicitly the tail of a sequence, see for instance in (Searcoid, 2007). In the lecture notes of (Melrose, 2004) is the term *tail* used, but nowhere is to find a definition of it.

Definition 2.1

Suppose that X is an non-empty set and let $x = \{x_n\}_{n \in \mathbb{N}}$ be a sequence in X. Let $m \in \mathbb{N}$, the set $\{x_n \mid n \in \mathbb{N} \text{ and } n \geq m\}$ is called the m-th tail of the sequence $\{x_n\}_{n \in \mathbb{N}}$, notated by $\text{tail}_m(x)$.

2.4 Rational and real numbers

There are several numbers, the natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$, the whole numbers $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, the rational numbers $\mathbb{Q} = \{\frac{p}{q} \mid p, q \in \mathbb{Z}\}$, the real numbers numbers $\mathbb{R}$ and the complex numbers $\mathbb{C} = \{a + \mathrm{i}\, b \mid a, b \in \mathbb{R} \text{ and } \mathrm{i}^2 = -1\}$.
Every real numbers is the limit of a sequence of rational numbers. The real numbers $\mathbb{R}$ is the completion of $\mathbb{Q}$. The real numbers $\mathbb{R}$ exist out of $\mathbb{Q}$ joined with all the limits of the Cauchy sequences in $\mathbb{Q}$.

2.5 Accumulation points and the closure of a subset

Let M be subset of some space X. Some point $x_0 \in X$ is called an accumulation point of M if every ball of x_0 contains at least a point $y \in M$, distinct from x_0.
The closure of M, denoted by $\overline{\mathrm{M}}$, is the union of M with all its accumulation points.

Theorem 2.1

$x \in \overline{\mathrm{M}}$ *if and only if* there is a sequence $\{x_n\}_{n\in\mathbb{N}}$ in M such that $\lim_{n\to\infty} x_n = x$.

Proof of Theorem 2.1

The proof exists out of two parts.

($\Rightarrow$) If $x \in \overline{\mathrm{M}}$ then $x \in M$ or $x \notin M$. If $x \in M$ take then $x_n = x$ for each n. If $x \notin M$, then x is an accumulation point of M, so for every $n \in \mathbb{N}$, the ball $B_{\frac{1}{n}}(x)$ contains a point $x_n \in M$. So there is constructed a sequence $\{x_n\}_{n\in\mathbb{N}}$ with
$\| x_n - x \| < \frac{1}{n} \to 0$, if $n \to \infty$.

($\Leftarrow$) If $\{x_n\}_{n\in\mathbb{N}} \subset M$ and $\| x_n - x \| \to 0$, if $n \to \infty$, then every neighbourhood of x contains points $x_n \neq x$, so x is an accumulation point of M.

$\square$

Theorem 2.2

M is closed *if and only if* the limit of every convergent sequence in M is an element of M.

Proof of Theorem 2.2

The proof exists out of two parts.

($\Rightarrow$) M is closed and there is a convergent sequence $\{x_n\}_{n\in\mathbb{N}}$ in M, $\lim_{n\to\infty} x_n = x$. If $x \notin M$ then $x \in M^c$. M^c is open, so there is a $\delta > 0$ such that $B_\delta(x) \subset M^c$, but then $\| x_n - x \| > \delta$. This means that the sequence is not convergent, but that is not the case, so $x \in M$.

($\Leftarrow$) If M is not closed, then is M^c not open. So there is an element $x \in M^c$, such that for every ball $B_{\frac{1}{n}}(x)$, with $n \in \mathbb{N}$, there exist an element $x_n \in M$.
Since $\lim_{n\to\infty} \frac{1}{n} = 0$, the sequence $\{x_n\}_{n\in\mathbb{N}}$ converges in M. The limit of every convergent sequence in M is an element of M, so $x \in M$, this gives a contradiction, so M is closed.

□

Theorem 2.3

M is closed *if and only if* $M = \overline{\mathrm{M}}$.

Proof of Theorem 2.3

The proof exists out of two parts.

(⇒) $M \subseteq \overline{\mathrm{M}}$, if there is some $x \in \overline{\mathrm{M}} \setminus M$, then x is an accumulation point of M, so there can be constructed a convergent sequence $\{x_n\}_{n \in \mathbb{N}}$ out of M with limit x. M is closed, so $x \in M$, so $\overline{\mathrm{M}} \setminus M = \varnothing$.

(⇐) Let $\{x_n\}_{n \in \mathbb{N}}$ be a convergent sequence in M, with $\lim_{n \to \infty} x_n = x$, since $\overline{\mathrm{M}} \setminus M = \varnothing$, the only possibility is that $x \in M$, so M is closed.

□

Theorem 2.4

Let K be the intersection of all closed sets containing M, then $K = \overline{\mathrm{M}}$. This means that $\overline{\mathrm{M}}$ is the smallest closed set containing M.

Proof of Theorem 2.4

The proof exists out of two parts.

(⇒) $\overline{\mathrm{M}}$ is closed and $M \subset \overline{\mathrm{M}}$, so $K \subset \overline{\mathrm{M}}$.

(⇐) If S is closed and $M \subset S$, then $\overline{\mathrm{M}} \subset \overline{\mathrm{S}} = S$ and so $\overline{\mathrm{M}} \subset K$.

□

2.6 Dense subset

Definition 2.2

The subset $Y \subset X$ is (everywhere) dense in X if $\overline{\mathrm{Y}} = X$.
This is the case if and only if $Y \cap B_r(x) \neq \varnothing$ for every $x \in X$ and every $r > 0$.

Most of the time, *dense* is used in the following sense:
Let Y and X be sets and $Y \subseteq X$. Y is a dense subset of X, if for every $x \in X$, there exists a sequence of elements $\{y_n\}_{n \in \mathbb{N}}$ in Y, such that $\lim_{n \to \infty} y_n = x$.
Or in other words, every point in X is a point of Y or a limit point of Y.
The rational numbers $\mathbb{Q}$ is a dense subset of real numbers $\mathbb{R}$, $\mathbb{Q}$ lies dense in $\mathbb{R}$.

2.7 Separable and countable space

With countable is meant that every element of a space X can be associated with an unique element of $\mathbb{N}$ and that every element out of $\mathbb{N}$ corresponds with an unique element out of X. The mathematical description of countable becomes, a set or a space X is called countable if there exists an injective function

$$f : X \to \mathbb{N}.$$

If f is also surjective, thus making f bijective, then X is called countably infinite or denumerable.
The space X is said to be separable if this space has a countable subset M of X, which is also dense in X. M is countable, means that $M = \{y_n | y_n \in X\}_{n \in \mathbb{N}}$. M is dense in X, means that $\overline{M} = X$. If $x \in X$ then there exists in every neighbourhood of x an element of M, so $\overline{\mathrm{span}}\{y_n \in M \,|\, n \in \mathbb{N}\} = X$.

The rational numbers $\mathbb{Q}$ are countable and are dense in $\mathbb{R}$, so the real numbers are separable.

2.8 Compact subset

There are several definitions of compactness of a subset M, out of another set X. These definitions are equivalent if (X, d) is a metric space (Metric Spaces, see section 3.5), but in non-metric spaces they have not to be equivalent, carefulness is needed in such cases.
Let $(S_\alpha)_{\alpha \in IS}$ be a family of subsets of X, with IS is meant an index set. This family of subsets is a cover of M, if

$$M \subset \cup_{\alpha \in IS} S_\alpha \tag{2.10}$$

and $(S_\alpha)_{\alpha \in IS}$ is a cover of X, if $\cup_{\alpha \in IS} S_\alpha = X$. Each element out of X belongs to a set S_α out of the cover of X.
If the sets S_α are open, there is spoken about a open cover.

- The subset M is said to be compact in X, if every open cover of M contains a finite subcover, a finite number of sets S_α which cover M.
- The subset M is said to be countable compact in X, if every countable open cover of M contains a finite subcover.
- The subset M is said to be sequentially compact in X, if every sequence in M has a convergent subsequence in M.

Example 2.1

The open interval $(0, 1)$ is not compact.

Explanation of Example 2.1

Consider the open sets $I_n = (\frac{1}{n+2}, \frac{1}{n})$ with $n \in \{1, 2, 3, \cdots\} = \mathbb{N}$. Look to the open cover $\{I_n \mid n \in \mathbb{N}\}$. Assume that this cover has a finite subcover

$F = \{(a_1, b_1), (a_2, b_2), \cdots, (a_{n_0}, b_{n_0})\}$, with $a_i < b_i$ and $1 \leq i \leq n_0$. Define $\alpha = \min(a_1, \cdots, a_{n_0})$ and $\alpha > 0$, because there are only a finite number of a_i. The points in the interval $(0, \alpha)$ are not covered by the subcover F, so the given cover has no finite subcover. □

Read the definition of compactness carefully: " Every open cover has to contain a finite subcover". Just finding a certain open cover, which has a finite subcover, is not enough!

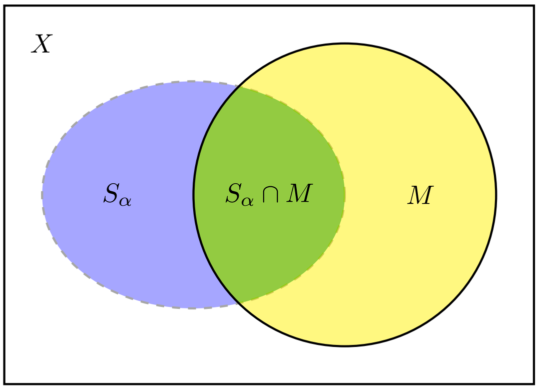


Figure 2.1 Compactness and open sets

Compactness is a topological property. In the situation of figure 2.1, there are two topologies, the topology on X and a topology on M. The topology on M is induced by the topology on X. Be aware of the fact that the set $S_\alpha \cap M$ is an open set of the topology on M.

Theorem 2.5

A compact subset M of a metric space (X, d) is closed and bounded.

Proof of Theorem 2.5

First will be proved that M is closed and then will be proved that M is bounded. Let $x \in \overline{M}$, then there exists a sequence $\{x_n\}$ in M, such that $x_n \rightarrow x$, see theorem 2.1. The subset M is compact in the metric space (X, d). In theorem 6.6 is proved, that in a metric space compactness is equivalent with sequentially compactness, so $x \in M$. Hence M is closed, because $x \in \overline{M}$ was arbitrary chosen.
The boundedness of M will be proved by a contradiction.

Suppose that M is unbounded, then there exists a sequence $\{y_n\} \subset M$ such that $d(y_n, a) > n$, with $a \in M$, a fixed element. This sequence has not a convergent subsequence, what should mean that M is not compact, what is not the case. Hence, M has to be bounded. □

The converse of **theorem 2.5** is in general not true, but for $\mathbb{R}^n$ the converse is true as well. The **Heine-Borel theorem 2.6** characterizes compact subsets of $\mathbb{R}^n$.

Theorem 2.6

The theorem of Heine-Borel:
In $\mathbb{R}^n$ with usual metric d, for any subset $A \subset \mathbb{R}^n$:

A is compact *if and only if* A is closed and bounded.

Proof of Theorem 2.6

The proof exists out of two parts.

($\Rightarrow$) The $(\mathbb{R}^n, d)$ is a metric space, A is a compact subset, so use **theorem 2.5**.

($\Leftarrow$) $A \subset \mathbb{R}^n$ is closed and bounded. Let $\{x_i\}_{i \in \mathbb{N}}$ be a sequence in A. Since A is bounded, any sequence in A must be bounded, so the sequence $\{x_i\}_{i \in \mathbb{N}}$ is a bounded sequence. The **theorem of Bolzano-Weierstrass 6.1** implies that there exists a convergent subsequence of $\{x_i\}_{i \in \mathbb{N}}$ in A. (Construct a convergent subsequence by taking coordinate wise subsequences of the original sequence $\{x_i\}_{i \in \mathbb{N}}$. The constructed convergent subsequence exists, since $\mathbb{R}^n$ is finite dimensional.) Let's call the limit point of this subsequence: x. Since A is closed, $x \in A$. So any sequence in A has a convergent subsequence with limit point in A, so A is sequentially compact. In a metric space sequentially compactness is equivalent with compactness, see **theorem 6.6**, so A is compact.

□

2.9 Supremum and infimum

Axiom 2.1 The Completeness Axiom for the real numbers
If a non-empty set $A \subset \mathbb{R}$ has an upper bound, it has a least upper bound.

A bounded subset $S \subset \mathbb{R}$ has a maximum or a supremum and has a mimimum or an infimum.
A supremum, denoted by sup, is the lowest upper bound of that subset S. If the lowest upper bound is an element of S then it is called a maximum, denoted by max.
An infimum, denoted by inf, is the greatest lower bound of that subset. If the greatest lower bound is an element of S then it is called a minimum, denoted by min.
There is always a sequence of elements $\{s_n\}_{n \in \mathbb{N}}$, with for $s_n \in S$ every $n \in \mathbb{N}$, which converges to a supremum or an infimum, if they exist.

Example 2.2

Look to the interval $S = (0, 1]$. Then $\inf\{S\} = 0$ and $\min\{S\}$ does not exist ($0 \notin S$) and $\sup\{S\} = \max\{S\} = 1 \in S$.

2.10 Continuous, uniformly continuous and Lipschitz continuous

Let $T : X \to Y$ be a mapping, from a space X with a norm $\| \, . \, \|_1$ to a space Y with a norm $\| \, . \, \|_2$. This mapping T is said to be continuous at a point $x_0 \in X$, if for every $\epsilon > 0$, there exists a $\delta(\epsilon) > 0$ such that for every $x \in B_\delta(x_0) = \{y \in X | \, \| y - x_0 \|_1 < \delta\}$, there is satisfied that $T(x) \in B_\epsilon(T(x_0))$, this means that $\| T(x) - T(x_0) \|_2 < \epsilon$, see figure 2.2.
The mapping T is said to be uniformly continuous, if for every $\epsilon > 0$, there exists a $\delta(\epsilon) > 0$ such that for every x and y in X, with $\| x - y \|_1 < \delta(\epsilon)$, there is satisfied

that $\| T(x) - T(y) \|_2 < \epsilon$.
If a mapping is continuous, the value of $\delta(\epsilon)$ depends on ϵ and on the point in the domain. If a mapping is uniformly continuous, the value of $\delta(\epsilon)$ depends only on ϵ and not on the point in the domain.

The mapping T is said to be Lipschitz continuous, if there exists a constant $L > 0$ such that $\| T(x) - T(y) \|_2 \leq L \| x - y \|_1$ for every x and y in X.

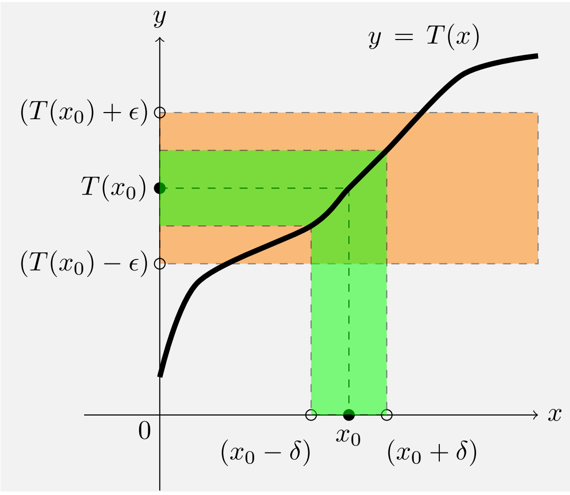


Figure 2.2 Continuous map

Theorem 2.7

A mapping $T : X \rightarrow Y$ of a normed space X with norm $\| \, . \, \|_1$ to a normed space Y with norm $\| \, . \, \|_2$ is continuous at $x_0 \in X$ *if and only if* for every sequence in $(x_n)_{n \in \mathbb{N}}$ in X with $\lim_{n \rightarrow \infty} x_n = x_0$ follows that $\lim_{n \rightarrow \infty} T(x_n) = T(x_0)$.

Proof of Theorem 2.7

The proof exists out of two parts.

($\Rightarrow$) Let $\epsilon > 0$ be given. Since T is continuous, then there exists a $\delta(\epsilon)$ such that $\| T(x_n) - T(x_0) \|_2 < \epsilon$ when $\| x_n - x_0 \|_1 < \delta(\epsilon)$. Known is that $x_n \rightarrow x_0$, so there exists an $N_\epsilon = N(\delta(\epsilon))$, such that $\| x_n - x_0 \|_1 < \delta(\epsilon)$ for every $n > N_\epsilon$. Hence $\| T(x_n) - T(x_0) \|_2 < \epsilon$ for $n > N_\epsilon$, so $T(x_n) \rightarrow T(x_0)$.

(⇐) Assume that T is not continuous. Then there exists a $\epsilon > 0$ such that for every $\delta > 0$, there exists an $x \in X$ with $\| x - x_0 \|_1 < \delta$ and $\| T(x_n) - T(x_0) \|_2 \geq \epsilon$. Take $\delta = \frac{1}{n}$ and there exists an $x_n \in X$ with $\| x_n - x_0 \|_1 < \delta = \frac{1}{n}$ with $\| T(x_n) - T(x_0) \|_2 \geq \epsilon$. So a sequence is constructed such that $x_n \rightarrow x_0$ but $T(x_n) \nrightarrow T(x_0)$ and this contradicts $T(x_n) \rightarrow T(x_0)$.

□

Remark 2.1

Theorem 2.7 can be generalised, so is Theorem 2.7 is also valid for a map $T : X \rightarrow Y$ between two Metric Spaces (X, d_1) and (Y, d_2). The proof is almost the same, replace $\| a - b \|_1$ by $d_1(a, b)$ and $\| c - d \|_2$ by $d_2(c, d)$ with respectively $a, b \in X$ and $c, d \in Y$.

2.11 Continuity and compactness

Important are theorems about the behaviour of continuous mappings with respect to compact sets.

Theorem 2.8

If $T : X \rightarrow Y$ is a continuous map
and $V \subset X$ is compact *then* is $T(V) \subset Y$ compact.

Proof of Theorem 2.8

(⇒) Let $\mathcal{U}$ be an open cover of $T(V)$. $T^{-1}(U)$ is open for every $U \in \mathcal{U}$, because T is continuous. The set $\{T^{-1}(U) \mid U \in \mathcal{U}\}$ is an open cover of V, since for every $x \in V$, $T(x)$ must be an element of some $U \in \mathcal{U}$. V is compact, so there exists a finite subcover $\{T^{-1}(U_1), \cdots, T^{-1}(U_{n_0})\}$, so $\{U_1, \cdots, U_{n_0}\}$ is a finite subcover of $\mathcal{U}$ for $T(V)$.

□

Theorem 2.9

Let (X, d_1) and (Y, d_2) be metric spaces and $T : X \to Y$ a continuous mapping then is the image $T(V)$, of a compact subset $V \subset X$, closed and bounded.

Proof of Theorem 2.9

The image $T(V)$ is compact, see **theorem 2.8** and a compact subset of a metric space is closed and bounded, see **theorem 2.5**. □

Definition 2.3

A Compact Metric Space X is a Metric Space in which every sequence has a subsequence that converges to a point in X.

In a Metric Space, sequentially compactness is equivalent to the compactness defined by open covers, see **section 2.8**.

Example 2.3

An example of a compact metric space is a bounded and closed interval $[a, b]$, with $a, b \in \mathbb{R}$ with the metric $d(x, y) = \mid x - y \mid$.

Theorem 2.10

Let (X, d_1) and (Y, d_2) be two Compact Metric Spaces, then every continuous function $f : X \to Y$ is uniformly continuous.

Proof of Theorem 2.10

The theorem will be proved by a contradiction.
Suppose that f is not uniformly continuous, but only continuous.
If f is not uniformly continous, then there exists an ϵ_0 such that for all $\delta > 0$, there are some $x, y \in X$ with $d_1(x, y) < \delta$ and $d_2(f(x), f(y)) \geq \epsilon_0$.
Choose two sequences $\{v_n\}$ and $\{w_n\}$ in X, such that

$$d_1(v_n, w_n) < \frac{1}{n} \text{ and } d_2(f(v_n), f(w_n)) \geq \epsilon_0.$$

The metric Space X is compact, so there exist two converging subsequences $\{v_{n_k}\}$ and $\{w_{n_k}\}$, ($v_{n_k} \to v_0$ and $w_{n_k} \to w_0$), so

$$d_1(v_{n_k}, w_{n_k}) < \frac{1}{n_k} \text{ and } d_2(f(v_{n_k}), f(w_{n_k})) \geq \epsilon 0. \tag{2.11}$$

The sequences $\{v_{n_k}\}$ and $\{w_{n_k}\}$ converge to the same point and since f is continuous, statement 2.11 is impossible.

The function f has to be uniformly continuous. □

2.12 Pointwise and uniform convergence

Pointwise convergence and uniform convergence are of importance when there is looked at sequences of functions.
Let $C[a, b]$, the space of continous functions on the closed interval $[a, b]$. A norm which is very much used on this space of functions is the so-called sup-norm, defined by $\sup_{t \in [a,b]} \mid f(t) \mid$

$$\| f \|_\infty = \sup_{t\in[a,b]} | f(t) | \tag{2.12}$$

with $f \in C[a,b]$. The fact that $[a,b]$ is a compact set of $\mathbb{R}$, means that the $\sup_{t\in[a,b]} | f(t) |= \max_{t\in[a,b]} | f(t) |$.
Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of functions, with $f_n \in C[a,b]$. If $x \in [a,b]$ then is $\{f_n(x)\}_{n\in\mathbb{N}}$ a sequence in $\mathbb{R}$.
If for each fixed $x \in [a,b]$ the sequence $\{f_n(x)\}_{n\in\mathbb{N}}$ converges, there can be defined the new function $f : [a,b] \to \mathbb{R}$, by $f(x) = \lim_{n\to\infty} f_n(x)$.
For each fixed $x \in [a,b]$ and every $\epsilon > 0$, there exist a $N(x,\epsilon)$ such that for every $n > N(x,\epsilon)$, the inequality $| f(x) - f_n(x) |< \epsilon$ holds.
The sequence $\{f_n\}_{n\in\mathbb{N}}$ converges pointwise to the function f. For each fixed $x \in [a,b]$, the sequence $\{f_n(x)\}_{n\in\mathbb{N}}$ converges to $f(x)$. Such limit function is not always continous.

Example 2.4

Let $f_n(x) = x^n$ and $x \in [0,1]$. The pointwise limit of this sequence of functions becomes

$$f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{if } x \in [0,1); \\ 1 & \text{if } x = 1. \end{cases}$$

Important to note is that the limit function f is not continuous, although the functions f_n are continous on the interval $[0,1]$.

If the sequence is uniform convergent, the limit function is continous. A sequence of functions $\{f_n\}_{n\in\mathbb{N}}$, with $f_n \in C[a,b]$, $n \in \mathbb{N}$, converges uniform to the function f, if for every $\epsilon > 0$, there exist a $N(\epsilon)$ such that for every $n > N(\epsilon)$ $\| f - f_n \|_\infty< \epsilon$. Note that $N(\epsilon)$ does not depend of x anymore. So $| f(x)-f_n(x) |< \epsilon$ for all $n > N(\epsilon)$ and for all $x \in [a,b]$.

Theorem 2.11

If the sequence of functions $\{f_n\}_{n\in\mathbb{N}}$, with $f_n \in C[a,b]$, $n \in \mathbb{N}$, converges uniform to the function f on the interval $[a,b]$, then the function f is continuous on $[a,b]$.

Proof of Theorem 2.11

Let $\epsilon > 0$ be given, and there is proved that the function f is continuous for some $x \in [a, b]$.
The sequence $\{f_n\}_{n \in \mathbb{N}}$ converges uniform on the interval $[a, b]$, so for every $s, x \in [a, b]$, there is a $N(\epsilon)$ such that for every $n > N(\epsilon)$,
$| f(s) - f_n(s) | < \frac{\epsilon}{3}$ and $| f_n(x) - f(x) | < \frac{\epsilon}{3}$ ($N(\epsilon)$ does not depend on the value of s or x).
Take some $n > N(\epsilon)$, the function f_n is continous in x, so there is a $\delta(\epsilon) > 0$, such that for every s, with $| s - x | < \delta(\epsilon)$, $| f_n(s) - f_n(x) | < \frac{\epsilon}{3}$. So the function f is continous in x, because

$$| f(s) - f(x) | < | f(s) - f_n(s) | + | f_n(s) - f_n(x) | + | f_n(x) - f(x) | < \epsilon,$$

for every s, with $| s - x | < \delta(\epsilon)$. □

2.13 Partially and totally ordered sets

On a non-empty set X, there can be defined a relation, denoted by $\preceq$, between the elements of that set. Important are partially ordered sets and totally ordered sets.

Definition 2.4

The relation $\preceq$ is called a partial order over the set X, if for all $a, b, c \in X$

PO 1: $a \preceq a$ (reflexivity),
PO 2: if $a \preceq b$ and $b \preceq a$ then $a = b$ (antisymmetry),
PO 3: if $a \preceq b$ and $b \preceq c$ then $a \preceq c$ (transitivity).

If $\preceq$ is a partial order over the set X then $(X, \preceq)$ is called a partial ordered set.

Definition 2.5

The relation $\preceq$ is called a total order over the set X, if for all $a, b, c \in X$

TO 1: if $a \preceq b$ and $b \preceq a$ then $a = b$ (antisymmetry),
TO 2: if $a \preceq b$ and $b \preceq c$ then $a \preceq c$ (transitivity),
TO 3: $a \preceq b$ or $b \preceq a$ (totality).

Totality implies reflexivity. Thus a total order is also a partial order.
If $\preceq$ is a total order over the set X then $(X, \preceq)$ is called a total ordered set.

Working with some order, most of the time there is searched for a maximal element or a minimal element.

Definition 2.6

Let $(X, \preceq)$ be partially ordered set and $Y \subset X$.

ME 1: $M \in Y$ is called a maximal element of Y if

$$M \preceq x \Rightarrow M = x, \text{ for all } x \in Y.$$

ME 2: $M \in Y$ is called a minimal element of Y if

$$x \preceq M \Rightarrow M = x, \text{ for all } x \in Y.$$

2.14 Equivalence relation

Definition 2.7

A given relation $\sim$ between two arbitrary elements of a set X is said to be an equivalence relation if and only if for every $a, b, c \in X$

EQ 1: $a \sim a$ (reflexivity),
EQ 2: if $a \sim b$ then $b \sim a$ (symmetry),
EQ 3: if $a \sim b$ and $b \sim c$ then $a \sim c$ (transitivity).

The equivalence class of a under $\sim$ is often denoted as

$$[a] = \{b \in X \mid b \sim a\},$$

but also quite often by $\widetilde{a}$.

2.15 Limit superior/inferior of sequences of numbers

If there is worked with the limit superior and the limit inferior, it is most of the time also necessary to work with the extended real numbers $\overline{\mathbb{R}} = \mathbb{R} \cup -\infty \cup \infty$.

Definition 2.8

Let $\{x_n\}$ be real sequence. The limit superior of $\{x_n\}$ is the extended real number

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \left(\sup_{k\geq n} x_k \right).$$

It can also be defined by the limit of the decreasing sequence $s_n = \sup \{x_k \mid k \geq n\}$.

Definition 2.9

Let $\{x_n\}$ be real sequence. The limit inferior of $\{x_n\}$ is the extended real number

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \left(\inf_{k\geq n} x_k \right).$$

It can also be defined by the limit of the increasing sequence $t_n = \inf \{x_k \mid k \geq n\}$.

To get an idea about the lim sup and lim inf, look to the sequence of maximum and minimum values of the wave of the function $f(x) = (1 + 4 \exp(-x/10)) \sin(5\,x)$ in figure 2.3.
The definitions of lim sup and lim inf, given the Definitions 2.8 and 2.9, are definitions for sequences of real numbers. But in the functional analysis, lim sup and lim inf, have also to be defined for sequences of sets.

2.16 Limit superior/inferior of sequences of sets

Let $(E_k \mid k \in \mathbb{N})$ be a sequence of subsets of an non-empty set S. The sequence of subsets $(E_k \mid k \in \mathbb{N})$ increases, written as $E_k \uparrow$, if $E_k \subset E_{k+1}$ for every $k \in \mathbb{N}$. The sequence of subsets $(E_k \mid k \in \mathbb{N})$ decreases, written as $E_k \downarrow$, if $E_k \supset E_{k+1}$

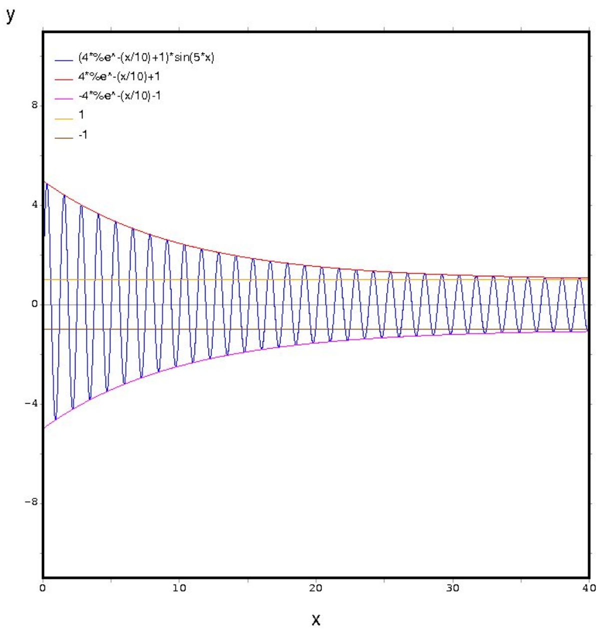


Figure 2.3 Illustration of $\limsup$ and $\liminf$.

for every $k \in \mathbb{N}$. The sequence $(E_k \mid k \in \mathbb{N})$ is a monotone sequence if it is either an increasing sequence or a decreasing sequence.

Definition 2.10

If the sequence $(E_k \mid k \in \mathbb{N})$ decreases then

$$\lim_{k \to \infty} E_k = \bigcap_{k \in \mathbb{N}} E_k = \{x \in S \mid x \in E_k \text{ for every } k \in \mathbb{N}\}.$$

If the sequence $(E_k \mid k \in \mathbb{N})$ increases then

$$\lim_{k \to \infty} E_k = \bigcup_{k \in \mathbb{N}} E_k = \{x \in S \mid x \in E_k \text{ for some } k \in \mathbb{N}\}.$$

For a monotone sequence $(E_k \mid k \in \mathbb{N})$, the $\lim_{k \to \infty} E_k$ always exists, but it may be $\varnothing$.

If $E_k \uparrow$ then $\lim_{k \to \infty} E_k = \varnothing \Leftrightarrow E_k = \varnothing$ for every $k \in \mathbb{N}$.
If $E_k \downarrow$ then $\lim_{k \to \infty} E_k = \varnothing$ can be the case, even if $E_k \neq \varnothing$ for every $k \in \mathbb{N}$. Take for instance $S = [0, 1]$ and $E_k = (0, \frac{1}{k})$ with $k \in \mathbb{N}$.

Definition 2.11

The limit superior and the limit inferior of a sequence $(E_k \mid k \in \mathbb{N})$ of subsets of a non-empty set S is defined by

$$\limsup_{n\to\infty} E_n = \bigcap_{n\in\mathbb{N}} (\bigcup_{k\geq n} E_k),$$

$$\liminf_{n\to\infty} E_n = \bigcup_{n\in\mathbb{N}} (\bigcap_{k\geq n} E_k),$$

both limits always exist, but they may be $\varnothing$.

It is easily seen that $D_n = \bigcup_{k\geq n} E_k$ is a decreasing sequence of subsets, so $\lim_{n\to\infty} D_n$ exists. Similarly is $I_n = \bigcap_{k\geq n} E_k$ is an increasing sequence of subsets, so $\lim_{n\to\infty} I_n$ exists.

Theorem 2.12

Let $(E_k \mid k \in \mathbb{N})$ be a sequence of subsets of an non-empty set S.

1. $\limsup_{k\to\infty} E_k = \{s \in S \mid s \in E_k \text{ for infinitely many } k \in \mathbb{N}\}$
2. $\liminf_{k\to\infty} E_k = \{s \in S \mid s \in E_k$ for every $k \in \mathbb{N}$, but with a finite number of exceptions$\}$,
3. $\liminf_{k\to\infty} E_k \subset \limsup_{k\to\infty} E_k$.

Proof of Theorem 2.12

Let $D_n = \bigcup_{k\geq n} E_k$ and $I_n = \bigcap_{k\geq n} E_k$.

1. $(\Rightarrow)$: Let $s \in \bigcap_{n\in\mathbb{N}} D_n$ and s is an element of only a finitely many E_k's. If there are only a finite number of E_k's then there is a maximum value of k. Let's call that maximum value k_0. Then $s \notin D_{k_0+1}$ and therefore $s \notin \bigcap_{n\in\mathbb{N}} D_n$, which is in contradiction with the assumption about s. So s belongs to infinitely many members of the sequence $(E_k \mid k \in \mathbb{N})$.
   $(\Leftarrow)$: $s \in S$ belongs to infinitely many E_k, so let $\phi(j)$ be the sequence, in increasing order, of these numbers k. For every arbitrary number $n \in \mathbb{N}$ there

exists a number α such that $\phi(\alpha) \geq n$ and that means that $s \in E_{\phi(\alpha)} \subseteq D_n$. So $s \in \bigcap_{n \in \mathbb{N}} D_n = \limsup_{k \to \infty} E_k$.

2. $(\Rightarrow)$: Let $s \in \bigcup_{n \in \mathbb{N}} I_n$ and suppose that there infinitely many k's such that $s \notin E_k$. Let $\psi(j)$ be the sequence, in increasing order, of these numbers k. For some arbitrary n there exists a β such that $\psi(\beta) > n$, so $s \notin E_{\psi(\beta)} \supseteq I_n$. Since n was arbitrary $s \notin \bigcup_{n \in \mathbb{N}} I_n$, which is in contradiction with the assumption about s. So s belongs to all the members of the sequence $(E_k \mid k \in \mathbb{N})$, but with a finite number of exceptions.
   $(\Leftarrow)$: Suppose that $s \in E_k$ for all $k \in \mathbb{N}$ but for a finite number values of k's not. Then there exists some maximum value K_0 such that $s \in E_k$, when $k \geq K_0$. So $s \in I_{K_0}$ and there follows that $s \in \bigcup_{n \in \mathbb{N}} I_n = \liminf_{k \to \infty} E_k$.
3. If $s \in \liminf_{k \to \infty} E_k$ then $s \notin E_k$ for a finite number of k's but then s is an element of infinitely many E_k's, so $s \in \limsup_{k \to \infty} E_k$, see the descriptions of $\liminf_{k \to \infty} E_k$ and $\limsup_{k \to \infty} E_k$ in **Theorem 2.12: 2** and **1**.

□

Example 2.5

A little example about the lim sup and lim inf of subsets is given by $S = \mathbb{R}$ and the sequence $(E_k \mid k \in \mathbb{N})$ of subsets of S, which is given by

$$\begin{cases} E_{2k} = [0, 2k] \\ E_{2k-1} = [0, \dfrac{1}{2k-1}] \end{cases}$$

with $k \in \mathbb{N}$. It is not difficult to see that $\limsup_{k \to \infty} E_k = [0, \infty)$ and $\liminf_{k \to \infty} E_k = \{0\}$.

With the lim sup and lim inf, it also possible to define a limit for an arbitrary sequence of subsets.

Definition 2.12

Let $(E_k \mid k \in \mathbb{N})$ be an arbitrary sequence of subsets of a set S. If $\limsup_{k \to \infty} E_k = \liminf_{k \to \infty} E_k$ then the sequence converges and

$$\lim_{k \to \infty} E_k = \limsup_{k \to \infty} E_k = \liminf_{k \to \infty} E_k$$

Example 2.6

It is clear that the sequence of subsets defined in **Example 2.5** has no limit, because $\limsup_{k\rightarrow\infty} E_k \neq \liminf_{k\rightarrow\infty} E_k$.
But the subsequence $(\, E_{2\,k} \mid k \in \mathbb{N})$ is an increasing sequence with

$$\lim_{k \rightarrow \infty} E_{2\,k} = [0, \infty),$$

and the subsequence $(\, E_{2\,k-1} \mid k \in \mathbb{N})$ is a decreasing sequence with

$$\lim_{k \rightarrow \infty} E_{2\,k-1} = \{0\}.$$

2.17 Essential supremum and essential infimum

Busy with limit superior and limit inferior, see the **Sections 2.15** and **2.16**, it is almost naturally also to write something about the essential supremum and the essential infimum. But the essential supremum and essential infimum have more to do with **Section 2.9**. It is a good idea to read first **Section 5.1.5**, to get a feeling where it goes about.There has to made use of some mathematical concepts, which are described later into detail, see at **page 270**.
Important is the triplet (Ω, Σ, μ), Ω is some set, Σ is some collection of subsets of Ω and with μ the sets out of Σ can be measured. (Σ has to satisfy certain conditions.) The triplet (Ω, Σ, μ) is called a measure space, see also **page 270**.
With the measure space, there can be said something about functions, which are not valid everywhere, but *almost everywhere*. And *almost everywhere* means that something is true, exept on a set of measure zero.

Example 2.7

A simple example is the interval $I = [-\sqrt{3}, \sqrt{3}] \subset \mathbb{R}$. If the subset $J = [-\sqrt{3}, \sqrt{3}] \cap \mathbb{Q}$ is measured with the Lebesque measure, see **Section 5.1.6**, the measure of J is zero. An important argument is that the numbers out of $\mathbb{Q}$ are countable and that is not the case for $\mathbb{R}$, the real numbers.

If there is measured with some measure, it gives also the possibility to define different bounds for a function $f : \Omega \to \mathbb{R}$.

A real number α is called an upper bound for f on Ω, if $f(x) \leq \alpha$ for all $x \in \Omega$. Another way to express that fact, is to say that

$$\{x \in \Omega \mid f(x) > \alpha\} = \varnothing.$$

But α is called an essential upper bound for f on Ω, if

$$\mu(\{x \in \Omega \mid f(x) > \alpha\}) = 0,$$

that means that $f(x) \leq \alpha$ *almost everywhere* on Ω. It is possible that there are some $x \in \Omega$ with $f(x) > \alpha$, but the measure of that set is zero.

And if there are essential upper bounds then there can also be searched to the smallest essential upper bound, which gives the essential supremum, so

$$ess\sup(f) = inf\{\alpha \in \mathbb{R} \mid \mu(\{x \in \Omega \mid f(x) > \alpha\}) = 0\},$$

if $\{\alpha \in \mathbb{R} \mid \mu(\{x \in \Omega \mid f(x) > \alpha\}) = 0\} \neq \varnothing$, otherwise $ess\sup(f) = \infty$.

At the same way, the essential infimum is defined as the largest essential lower bound, so the essential infimum is given by

$$ess\inf(f) = \sup\{\beta \in \mathbb{R} \mid \mu(\{x \in \Omega \mid f(x) < \beta\}) = 0\},$$

if $\{\alpha \in \mathbb{R} \mid \mu(\{x \in \Omega \mid f(x) < \beta\}) = 0\} \neq \varnothing$, otherwise $ess\sup(f) = -\infty$.

Example 2.8

This example is based on **Example 2.7**. Let's define the function f by

$$f(x) = \begin{cases} x & \text{if } x \in J \subset \mathbb{Q}, \\ \arctan(x) & \text{if } x \in (I \backslash J) \subset (\mathbb{R} \setminus \mathbb{Q}), \\ -4 & \text{if } x = 0. \end{cases}$$

Let's look to the values of the function f on the interval $[-\sqrt{3}, \sqrt{3}]$.
So are values less then -4 lower bounds of f and the infimum of f, the greatest lower bound, is equal to -4. A value β, with $-4 < \beta < -\frac{\pi}{3}$, is an essential lower bound of f. The greatest essential lower bound of f, the essential infimum, is equal to $\arctan(-\sqrt{3}) = -\frac{\pi}{3}$.
The value $\arctan(\sqrt{3}) = \frac{\pi}{3}$ is the essential supremum of f, the least essential upper bound. A value β with $\frac{\pi}{3} < \beta < \sqrt{3}$ is an essential upper bound of f. The least upper bound of f, the supremum, is equal to $\sqrt{3}$. Values greater then $\sqrt{3}$ are just upper bounds of f.

3 Spaces

Be careful in thinking about Vector Spaces and Topological Spaces. In a Vector Space there is looked at elements that can be added or subtracted of each other. In a Topological Space there is looked at the union or intersection of sets. So is a Metric Space not by definition a Vector Space, but a Normed Space has to be a Vector Spaces as well as a Topological Space. Be careful in reading the definitions of these different spaces!

3.1 Flowchart of spaces

In this chapter is given an overview of classes of spaces. A space is a particular set of objects, with which can be done specific actions and which satisfy specific conditions. Here are the different kind of spaces described in a very short way. It is the intention to make clear the differences between these specific classes of spaces. See the flowchart at page 37.
Let's start with a Vector Space and a Topological Space.
A Vector Space consists out of objects, which can be added together and which can be scaled (multiplied by a constant). The result of these actions is always an element in that specific Vector Space. Elements out of a Vector Space are called vectors.
A Topological Space consist out of sets, which can be intersected and of which the union can be taken. The union and the intersection of sets give always a set back in that specific Topological Space. This family of sets is most of the time called a topology. A topology is needed when there is be spoken about concepts as continuity, convergence and for instance compactness.
If there exist subsets of elements out of a Vector Space, such that these subsets satisfy the conditions of a Topological Space, then that space is called a Topological Vector Space. A Vector Space with a topology, the addition and the scaling become continuous mappings.
Topological Spaces can be very strange spaces. But if there exists a function, which can measure the distance between the elements out of the subsets of a Topological Space, then it is possible to define subsets, which satisfy the conditions of a Topological Space. That specific function is called a metric and the space in question is

then called a Metric Space. The topology of that space is described by a metric.
A metric measures the distance between elements, but not the length of a particular element. On the other hand, if the metric can also measure the length of an object, then that metric is called a norm.
A Topological Vector Space, together with a norm, that gives a Normed Space. With a norm it is possible to define a topology on a Vector Space.
If every Cauchy row in a certain space converges to an element of that same space then such a space is called complete.
A Metric Space, where all the Cauchy rows converges to an element of that space is called a Complete Metric Space. Be aware of the fact that for a Cauchy row, only the distance is measured between elements of that space. There is only needed a metric in first instance.
In a Normed Space it is possible to define a metric with the help of the norm. That is the reason that a Normed Space, which is complete, is called a Banach Space. With the norm still the length of objects can be calculated, which can not be done in a Complete Metric Space.
With a norm it is possible to measure the distance between elements, but it is not possible to look at the position of two different elements, with respect to each other. With an inner product, the length of an element can be measured and there can be said something about the position of two elements with respect to each other. With an inner products it is possible to define a norm and such Normed Spaces are called Inner Product Spaces. The norm of an Inner Product Space is described by an inner product.
An Inner Product Space which is complete, or a Banach Space of which the norm has the behaviour of an inner product, is called a Hilbert Space.
For the definition of the mentioned spaces, see the belonging chapters of this lecture note or click on the references given at the flowchart, see page **37**.

From some spaces can be made a completion, such that the enlarged space becomes complete. The enlarged space exist out of the space itself united with all the limits of the Cauchy rows. These completions exist from a metric space, normed space and an inner product space,

1. the completion of a metric space is called a complete metric space,
2. the completion of a normed space becomes a Banach space and
3. the completion of an inner product space becomes a Hilbert space.

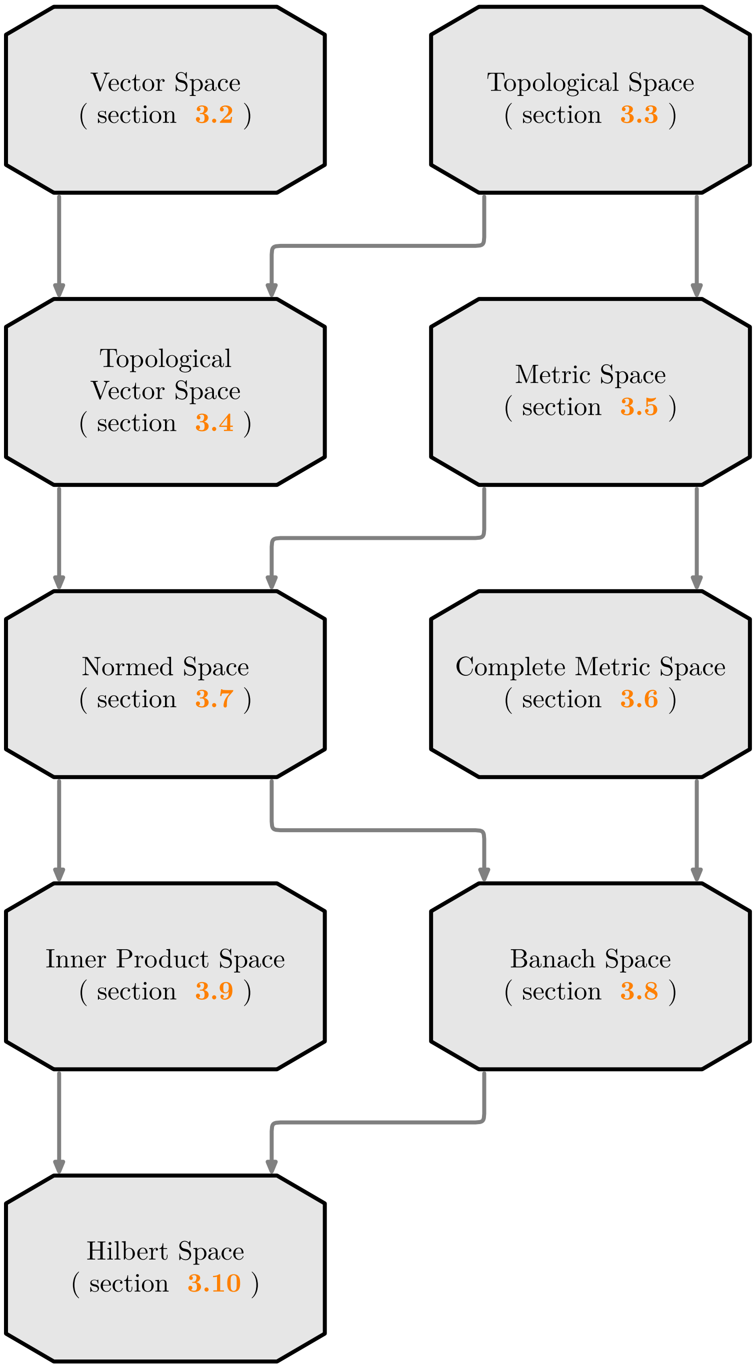


Figure 3.1 A flowchart of spaces.

3.2 Vector Spaces

A vector space is a set S of objects, which can be added together and multiplied by a scalar. The scalars are elements out of some field $\mathbb{K}$, most of the time, the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. The addition is written by $(+)$ and the scalar multiplication is written by $(\cdot)$.

Definition 3.1

A Vector Space VS is a set S, such that for every $x, y, z \in S$ and $\alpha, \beta \in \mathbb{K}$

VS 1: $x + y \in S$,

VS 2: $x + y = y + x$,

VS 3: $(x + y) + z = x + (y + z)$,

VS 4: there is an element $0 \in V$ with $x + 0 = x$,

VS 5: given x, there is an element $-x \in S$ with $x + (-x) = 0$,

VS 6: $\alpha \cdot x \in S$,

VS 7: $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$,

VS 8: $1 \cdot x = x$,

VS 9: $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$,

VS 10: $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$.

The quartet $(S, \mathbb{K}, (+), (\cdot))$ satisfying the above given conditions is called a Vector Space.
The different conditions have to do with: **VS 1** closed under addition, **VS 2** commutative, **VS 3** associative, **VS 4** identity element of addition, **VS 5** additive inverse, **VS 6** closed under scalar multiplication, **VS 7** compatible multiplications,

VS 8 identity element of multiplication, **VS 9** distributive: field addition, **VS 10** distributive: vector addition. For more information about a field , see **wiki-field**.

Remark 3.1

Let $x \in X$, $\gamma \in \mathbb{K}$ and let E, F be *subsets* of X, the following notations are adopted:

1. $x + F = \{x + y \mid y \in F\}$,
2. $E + F = \{x + y \mid x \in E, y \in F\}$,
3. $k\,E = \{k\,x \mid x \in E\}$.

3.2.1 Linear Subspaces

There will be worked very much with linear subspaces Y of a Vector Space X.

Definition 3.2

Let $\varnothing \neq Y \subseteq X$, with X a Vector Space. Y is a linear subspace of the Vector Space X if

LS 1: for every $y_1, y_2 \in Y$ holds that $y_1 + y_2 \in Y$,

LS 2: for every $y_1 \in Y$ and for every $\alpha \in \mathbb{K}$ holds that $\alpha\,y_1 \in Y$.

To look, if $\varnothing \neq Y \subseteq X$ could be a linear subspace of the Vector Space X, the following theorem is very useful.

Theorem 3.1

If $\varnothing \neq Y \subseteq X$ is a linear subspace of the Vector Space X then $0 \in Y$.

Proof of Theorem 3.1

Suppose that Y is a linear subspace of the Vector Space X. Take a $y_1 \in Y$ and take $\alpha = 0 \in \mathbb{K}$ then $\alpha\, y_1 = 0\, y_1 = 0 \in Y$. □

Furthermore it is good to realize that if Y is linear subspace of the Vector Space X that the quartet $(Y, \mathbb{K}, (+), (\cdot))$ is a Vector Space.
Sometimes there is worked with the sum of linear subspaces.

Definition 3.3

Let U and V be two linear subspaces of a Vector Space X. The sum $U + V$ is defined by

$$U + V = \{u + v \mid u \in U, v \in V\}.$$

It is easily verified that $U + V$ is a linear subspace of X.
If $X = U + V$ then X is said to be the sum of U and V. If $U \cap V = \varnothing$ then $x \in X$ can uniquely be written in the form $x = u + v$ with $u \in U$ and $v \in V$, then X is said to be the direct sum of U and V, denoted by $X = U \oplus V$.

Definition 3.4

A Vector Space X is said to be the direct sum of the linear subspaces U and V, denoted by

$$X = U \oplus V,$$

if $X = U + V$ and $U \cap V = \varnothing$. Every $x \in X$ has an unique representation

$$x = u + v,\, u \in U,\, v \in V.$$

If $X = U \oplus V$ then V is called the algebraic complement of U and vice versa.

3.2.2 Product Spaces

There will be very much worked with so-called products of Vector Spaces.

Definition 3.5

Let X_1 and X_2 be two Vector Spaces over the same field $\mathbb{K}$. The Cartesian product $X = X_1 \times X_2$ is a Vector Space under the following two algebraic operations

PS 1: $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$,

PS 2: $\alpha\,(x_1, x_2) = (\alpha\, x_1, \alpha\, x_2)$,

for all $x_1, y_1 \in X_1$, $x_1, y_2 \in X_2$ and $\alpha \in \mathbb{K}$.
The Vector Space X is called the product space of X_1 and X_2.

3.2.3 Quotient Spaces

Let W be a linear subspace of a Vector Space V.

Definition 3.6

The coset of an element $x \in V$ with respect to W is defined by the set

$$x + W = \{x + w \mid w \in W\}.$$

The distinct cosets form a partition of V. The Quotient Space or Factor Space is written by

$$V/W = \{x + W \mid x \in V\}.$$

Definition 3.7

The linear operations on V/W are defined by

QS 1: $(\mathrm{x}+\mathrm{W})+(\mathrm{y}+\mathrm{W})=(\mathrm{x}+\mathrm{y})+\mathrm{W}$,

QS 2: $\alpha(\mathrm{x}+\mathrm{W})=\alpha\,\mathrm{x}+\mathrm{W}$,

for all $x, y \in V$ and $\alpha \in \mathbb{K}$.

It is easily verified that the Quotient Space V/W, with the defined addition and the scalar multiplication, is a linear Vector Space over $\mathbb{K}$.

Remark 3.2

Working with cosets:

a. $x + W$ and $v + W$ are equal $\Leftrightarrow$ $(x - v) \in W$,
b. $(x - v) \in W \Leftrightarrow v \in x + W$,
c. the zero in V/W is W, also written as $0 + W$,
d. $-(x + W) = (-x) + W$ for every $x \in V$,
e. the cosets $x + W$ and $v + W$ are either equal or disjoint.

The sets that are elements of V/W partition V into equivalence classes.

Remark 3.3

The sum of two cosets is just the algebraic sum of two sets, as defined in Remark 3.1.

The product of a scalar $\alpha \neq 0$ with a coset $x + W$ is just equal to the product of α with the set $x + W$, as defined in Remark 3.1.
But be careful with $\alpha = 0$:

$$0\,(x + W) = \begin{cases} 0\,(\mathrm{x}+\mathrm{W}) = \mathrm{W} & \text{in the sense of Remark 3.2,} \\ \{0\} & \text{in the sense of Remark 3.1.} \end{cases}$$

Prevent confusion to describe in what context the expression $0\,(x + W)$ is meant, so an operation on the set $x + W$ or an operation at the coset $x + W$.

Example 3.1

Consider the Vector Space $\mathbb{R}^2$. Let M be a one-dimensional subspace of $\mathbb{R}^2$, so M is a straight line through the origin. A coset of M is a translation of M by a vector in $\mathbb{R}^2$.
The result of such a translation of M has not to be a subspace of $\mathbb{R}^2$. And there are infinitely many choices of translations that give the same coset. So there are some particular settings:

a. Two cosets of M are either identical or entirely disjoint.
b. The union of the cosets is all of $\mathbb{R}^2$.
c. The set of distinct cosets is a partition of $\mathbb{R}^2$.

Example 3.2

For the space of the continuous functions, see **Section 5.1.2** and for the space of the polynomials, see **Section 5.1.1**.
But much of the details given in the sections above are not of direct importance in this example.
Let $C(\mathbb{R})$ the space of continuous functions on $\mathbb{R}$ and let $\mathbb{P}(\mathbb{R})$ the subspace of $C(\mathbb{R})$ containing the polynomials. Given $f \in C(\mathbb{R})$, the coset determined by f is

$$f + \mathbb{P} = \{f + p \mid p \in \mathbb{P}(\mathbb{R})\}.$$

Further, $f + \mathbb{P} = g + \mathbb{P}$ if and only if $f - g$ is a polynomial. $f + \mathbb{P}$ is the equivalence class obtained by identifying functions which differ by a polynomial,
"f modulo the polynomials".

Example 3.3

Take $M = \{(x,0) \mid x \in \mathbb{R}\}$ in **Example 3.1**, then

$$\mathbb{R}^2/M = \{y + M \mid y \in \mathbb{R}\} = \{(x,0) + M \mid x \in \mathbb{R}\},$$

so $\mathbb{R}^2/M$ is the set of all horizontal lines in $\mathbb{R}^2$. Note that $\mathbb{R}^2/M$ is a $1-1$ correspondence with the set of distinct heights, so there is a natural bijection of $\mathbb{R}^2/M$ onto $\mathbb{R}$.
A equivalence class can be seen as "collapsing information modulo M".

There are slightly different viewpoints of those cosets or quotient sets, but they are mutually equivalent. In books and lecture notes most of the time one of the following approaches is chosen.

Definition 3.8

Let X be a non-empty set. Mutually equivalent definitions of quotient sets.

a. The *quotient set* $\pi(X)$ associated to a surjective function $\pi : X \to Y$ onto a non-empty set Y is defined to be $\pi(X) = Y$.
b. The *quotient set* $X/\sim$ associated to an equivalence relation $\sim$ on X is the set of equivalence classes: $(X/\sim) = \{[x] \mid x \in X\}$ with $[x] = \{x' \in X \mid x' \sim x\}$.
c. The *quotient set* $X/\mathcal{P}$ associated to a partition $\mathcal{P} = \{\mathcal{P}_i \mid i \in \mathcal{I}\}$ of X is defined as $X/\mathcal{P} = \mathcal{I}$.

These three notions coincide, here is the way how it can be done:

($\pi \Rightarrow \sim$) Let $\pi : X \to Y$ be a surjective function and define the relation $\sim$ as

$$x_1 \sim x_2 \quad \text{if} \quad \pi(x_1) = \pi(x_2).$$

($\sim \Rightarrow \mathcal{P}$) Given an equivalence relation $\sim$ on X, the partition $\mathcal{P}$ on X is defined by

$$\mathcal{P} = \{[x] \mid x \in X\},$$

where $[x]$ is as **Definition 3.8**, **part b**, $[x]$ is thought as an equivalence class and as an element out of the partition $\mathcal{P}$.

($\mathcal{P} \Rightarrow \pi$) Let $\mathcal{P} = \{\mathcal{P}_i \mid i \in \mathcal{I}\}$ be a partition of X with $\mathcal{P}_i \neq \emptyset$ for all $i \in \mathcal{I}$. The function $\pi : X \to Y$, by setting $Y = \mathcal{I}$ and

$$\pi(x) = i \quad \text{if} \quad x \in \mathcal{P}_i.$$

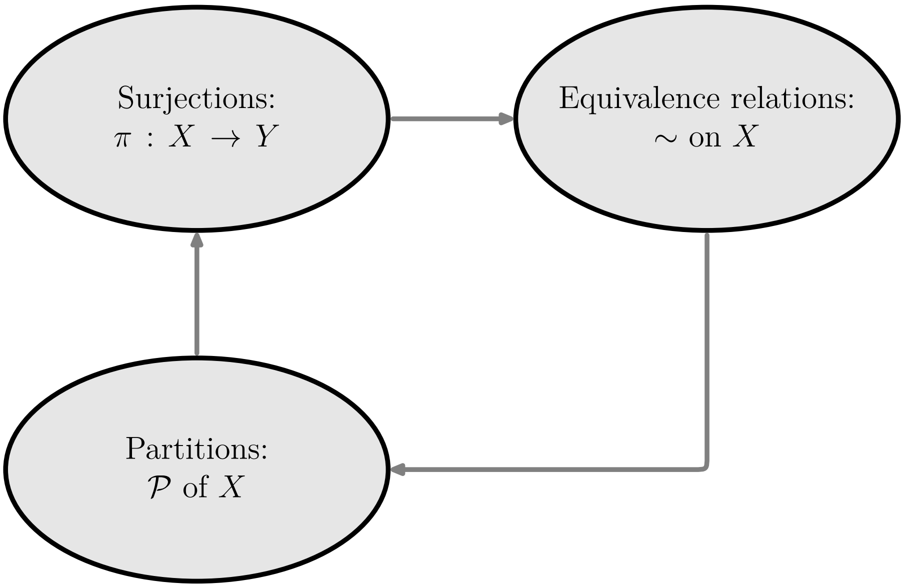


Figure 3.2 Commutative diagram, each map is a bijection.

3.2.4 Bases

Let X be a Vector Space and given some set $\{x_1, \cdots, x_p\}$ of p vectors or elements out of X. Let $x \in X$, the question becomes if x can be described on a unique way by that given set of p elements out of X? Problems are for instance if some of these p elements are just summations of each other of scalar multiplications, are they linear independent? Another problem is if these p elements are enough to describe x, the dimension of such set of vectors or the Vector Space X?

Definition 3.9

Let X be a Vector Space. A system of p vectors $\{x_1, \cdots, x_p\} \subset X$ is called linear independent, if the following equation gives that

$$\sum_{j=1}^{p} \alpha_j\, x_j \;=\; 0 \Rightarrow \alpha_1 \;=\; \cdots \;=\; \alpha_p \;=\; 0 \tag{3.1}$$

is the only solution.
If there is just one $\alpha_i \neq 0$ then the system $\{x_1, \cdots, x_p\}$ is called linear dependent.

If the system has infinitely many vectors $\{x_1, \cdots, x_p, \cdots\}$ then this system is called linear independent, if is it linear independent for every finite part of the given system, so

$$\forall N \in \mathbb{N} : \sum_{j=1}^{N} \alpha_j\, x_j \;=\; 0 \Rightarrow \alpha_1 \;=\; \cdots \;=\; \alpha_N \;=\; 0$$

is the only solution.
There can be looked at all possible finite linear combinations of the vectors out of the system $\{x_1, \cdots, x_p, \cdots\}$. All possible finite linear combinations of $\{x_1, \cdots, x_p, \cdots\}$ is called the span of $\{x_1, \cdots, x_p, \cdots\}$.

Definition 3.10

The span of the system $\{x_1, \cdots, x_p, \cdots\}$ is defined and denoted by

$$\text{span}(x_1, \cdots, x_p, \cdots) \;= < x_1, \cdots, x_p, \cdots >= \{\sum_{j=1}^{N} \alpha_j\, x_j \mid N \in \mathbb{N}, \alpha_1,\, \alpha_2, \cdots,\, \alpha_N \in \mathbb{K}\},$$

so all finite linear combinations of the system $\{x_1, \cdots, x_p, \cdots\}$.

If every $x \in X$ can be expressed as a unique linear combination of the elements out of the system $\{x_1, \cdots, x_p\}$ then that system is called a basis of X.

Definition 3.11

The system $\{x_1, \cdots, x_p\}$ is called a basis of X if:

B 1: the elements out of the given system are linear independent

B 2: and $< x_1, \cdots, x_p >= X$.

The number of elements, needed to describe a Vector Space X, is called the dimension of X, abbreviated by dim X.

Definition 3.12

Let X be a Vector Space. If $X = \{0\}$ then dim $X = 0$ and if X has a basis $\{x_1, \cdots, x_p\}$ then dim $X = p$. If $X \neq \{0\}$ has no finite basis then dim $X = \infty$, or if for every $p \in \mathbb{N}$ there exist a linear independent system $\{x_1, \cdots, x_p\} \subset X$ then dim $X = \infty$.

3.2.5 Finite dimensional Vector Space X

The Vector Space X is finite dimensional, in this case dim $X = n$, then a system of n linear independent vectors is a basis for X, or a basis in X. If the vectors $\{x_1, \cdots, x_n\}$ are linear independent, then every $x \in X$ can be written in an unique way as a linear combination of these vectors, so

$$x = \alpha_1 x_1 + \cdots + \alpha_n x_n$$

and the numbers $\alpha_1, \cdots, \alpha_n$ are unique.
The element x can also be given by the sequence $(\alpha_1, \alpha_2, \cdots, \alpha_n)$ and $\alpha_i, 1 \leq i \leq n$ are called the coordinates of x with respect to the basis $\alpha = \{x_1, \cdots, x_n\}$, denoted by $x_\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$. The sequence $x_\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ can be seen as an element out of the sequence space $\mathbb{R}^n$, see section **5.2.8**.
Such a sequence x_α can be written as

$$
\begin{aligned}
x_\alpha \;=\; & \alpha_1\,(\,1,\,0,\,0,\cdots,\,0) \,+ \\
& \alpha_2\,(\,0,\,1,\,0,\cdots,\,0) \,+ \\
& \cdots \\
& \alpha_n\,(\,0,\,0,\cdots,\,0,\,1),
\end{aligned}
$$

which is a linear combination of the elements out of the canonical basis for $\mathbb{R}^n$. The canonical basis for $\mathbb{R}^n$ is defined by

$$
\begin{aligned}
e_1 \;&=\; (\,1,\,0,\,0,\cdots,\,0) \\
e_2 \;&=\; (\,0,\,1,\,0,\cdots,\,0) \\
\cdots \;& \;\;\;\cdots \\
e_n \;&=\; (\,\underbrace{0,\,0,\cdots,\,0}_{(n-1)},\,1).
\end{aligned}
$$

It is important to note that, in the case of a finite dimensional Vector Space, there is only made use of algebraïc operations by defining a basis. Such a basis is also called an algebraiïc basis , or Hamel basis .

3.2.6 Infinite dimensional Vector Space X

There are some very hard problems in the case that the dimension of a Vector Space X is infinite. Look for instance to the definition 3.10 of a span. There are taken only finite summations and that in combination with an infinite dimensional space? Another problem is that, in the finite dimensional case, the number of basis vectors are countable, question becomes if that is also in the infinite dimensional case?
In comparison with a finite dimensional Vector Space there is also a problem with the norms, because there exist norms which are not equivalent. This means that different norms can generate quite different topologies on the same Vector Space X. So in the infinite dimensional case are several problems. Like, if there exists some set which is dense in X (see section 2.7) and if this set is countable (see section 2.6)?
The price is that infinite sums have to be defined. Besides the algebraïc calculations, the analysis becomes of importance (norms, convergence, etc.).

Just an ordinary basis, without the use of a topology, is difficult to construct, sometimes impossible to construct and in certain sense never used.

Example 3.4

Here an example to illustrate the above mentioned problems.
Look at the set of rows

$$S = \{(1, \alpha, \alpha^2, \alpha^3, \cdots) \mid \mid \alpha \mid < 1, \alpha \in \mathbb{R}\}.$$

It is not difficult to see that $S \subset \ell^2$, for the defintion of ℓ^2, see section 5.2.4.
All the elements out of S are linear independent, in the sense of section 3.2.4.
The set S is a linear independent uncountable subset of ℓ^2.

An index set is an abstract set to label different elements, such set can be uncountable.

Example 3.5

Define the set of functions $\text{Id}_r : \mathbb{R} \to \{0, 1\}$ by

$$\text{Id}_r(x) = \begin{cases} 1 & \text{if } x = r, \\ 0 & \text{if } x \neq r. \end{cases}$$

The set of all the Id_r functions is an uncountable set, which is indexed by $\mathbb{R}$.

The definition of a Hamel basis in some Vector Space $X \neq 0$.

Definition 3.13

A Hamel basis is a set H such that every element of the Vector Space $X \neq 0$ is a unique finite linear combination of elements in H.

Let X be some Vector Space of sequences, for instance ℓ^2, see section 5.2.4.
Let $A = \{e_1, e_2, e_3, \cdots\}$ with $e_i = (\delta_{i1}, \delta_{i2}, \cdots, \delta_{ij}, \cdots)$ and δ_{ij} is the Krönecker symbol,

$$\delta_{ij} = \begin{cases} i = j & \text{then } 1, \\ i \neq j & \text{then } 0. \end{cases}$$

The sequences e_i are linear independent, but A is not a Hamel basis of ℓ^2, since there are only finite linear combinations allowed. The sequence $x = (1, \frac{1}{2}, \frac{1}{3}, \cdots) \in \ell^2$ cannot be written as a finite linear combination of elements out of A.

Theorem 3.2

Every Vector Space $X \neq 0$ has a Hamel basis H.

Proof of Theorem 3.2

A proof will not be given here, but only an outline of how this theorem can be proved. It dependents on the fact, if you accept the Axiom of Choice, see **wiki-axiom-choice**. In Functional Analysis is used the lemma of Zorn, see **wiki-lemma-Zorn**. Mentioned the Axiom of Choice and the lemma of Zorn it is also worth to mention the Well-ordering Theorem, see **wiki-well-order-th**.
The mentioned Axiom, Lemma and Theorem are in certain sense equivalent, not accepting one of these makes the mathematics very hard and difficult.

The idea behind the proof is that there is started with some set H that is too small, so some element of X can not be written as a finite linear combination of elements out of H. Then you add that element to H, so H becomes a little bit larger. This larger H still violates that any finite linear combination of its elements is unique.

The set inclusion is used to define a partial ordering on the set of all possible linearly independent subsets of X. See **wiki-partial-order** for definition of a partial ordening.

By adding more and more elements, you reach some maximal set H, that can not be made larger. For a good definition of a maximal set, see **wiki-maximal**. The existence of such a maximal H is guaranteed by the lemma of Zorn.

Be careful by the idea of adding elements to H. It looks as if the elements are countable but look at the indices k of the set $H = \{v_\alpha\}_{\alpha \in A}$. The index set A is not necessarily $\mathbb{N}$, it is may be uncountable, see the examples 3.4 and 3.5.

Let H be maximal. Let $Y = \text{span}(H)$, then is Y a linear subspace of X and $Y = X$. If not, then $H' = H \cup \{z\}$ with $z \in X, z \notin Y$ would be a linear independent set, with H as a proper subset. That is contrary to the fact that H is maximal. □

In the section about Normed Spaces, the definition of an infinite sequence is given, see definition 3.26. An infinite sequence will be seen as the limit of finite sequences, if possible.

3.3 Topological Spaces

A nice overview of Topological Spaces is given in an article written by (Moller,). A lot of information is also given in the book written by (Taylor, 1958), but that is not only about Topological Spaces. The book of (Taylor, 1958) is also a nice introduction to the functional analysis.

A Topological Space is a set with a collection of subsets. The union or the intersection of these subsets is again a subset of the given collection.

Definition 3.14

A Topological Space $TS = \{A, \Psi\}$ consists of a non-empty set A together with a fixed collection Ψ of subsets of A satisfying

TS 1: $A, \varnothing \in \Psi$,

TS 2: the intersection of a finite collection of sets Ψ is again in Ψ,

TS 3: the union of any collection of sets in Ψ is again in Ψ.

The collection Ψ is called a topology of A and members of Ψ are called *open sets* of TS. Ψ is a subset of the power set of A.

The power set of A is denoted by $\mathcal{P}(A)$ and is the collection of all subsets of A.
For a nice paper about topological spaces, written by J.P. Möller, with at the end of it a scheme with relations between topological spaces, see paper-top-moller.

3.3.1 T_i Spaces, $i = 0, \cdots, 4$

There are no separation axioms so far. There are several types of separation. Here follow the different definitions of the T_i-spaces, $i = 0, \cdots, 4$.

Definition 3.15

Let X be a topological space. X is called:

1. T_0-space if and only if given any two distinct points $x \neq y \in X$ there is an open set containing one but not the other;
2. T_1-space if and only if given any two distinct points $x \neq y \in X$ there are open sets U and V such that $x \in U$, $y \in V$ but $x \notin V$, $y \notin U$;
3. T_2-space or Hausdorff space if and only if for any two distinct points $x_1 \neq x_2 \in X$, there exist open sets U, V with $x_1 \in U$ and $x_2 \in V$ and $U \cap V = \varnothing$;
4. T_3-space or regular if and only if X is T_1 and for every $x \in X$ and closed set C such that $x \notin C$, there are disjoint open sets U and V such that $x \in U$ and $C \subseteq V$;
5. $T_{3\frac{1}{2}}$-space or Tychonoff if and only if X is T_1 and for every $x \in X$ and closed set C such that $x \notin C$, there is a continuous function $f : X \to [0, 1]$ such that $f(x) = 0$ and $f(y) = 1$ for all $y \in C$;
6. T_4-space or normal if and only if for every pair disjoint sets C and D, there are disjoint sets U and V such that $C \subseteq U$ and $D \subseteq V$.

3.4 Topological Vector Spaces

Definition 3.16

A Topological Vector Space space $TVS = \{VS, \Psi\}$ consists of a non-empty vectorspace VS together with a topology Ψ.

Definition 3.17

Let $V_1 \; =, \{X_1, \psi_1\}$ and $V_2 \; =, \{X_2, \psi_2\}$ be topological vector spaces. Let $X_1 \times X_2$ be the cartesian product of X_1 and X_2, see **definition 3.5**.
The product topology ψ is the topology with basis
$B = \{U_1 \times U_2 \,|\, U_1 \in \psi_1,\, U_2 \in \psi_2\}$.

3.5 Metric Spaces

If $x, y \in X$ then the distance between these points is denoted by $d(x, y)$. The function $d(\cdot, \cdot) : X \times X \to \mathbb{R}$ has to satisfy several conditions before the function d is called a distance function on X or a metric on X.

Definition 3.18

A Metric Space MS is a pair (X, d). X is a Topological Space and the topology on X is defined by a distance function d, called the *metric* on X. The distance function d is defined on XxX and satisfies, for all $x, y, z \in X$,

M 1: $d(x, y) \in \mathbb{R}$ and $0 \leq d(x, y) < \infty$,

M 2: $d(x, y) = 0 \iff x = y$,

M 3: $d(x, y) = d(y, x)$ (Symmetry),

M 4: $d(x, y) \leq d(x, z) + d(z, y)$, (Triangle inequality).

Here follow some examples of metrics.

Example 3.6

If two points $x = (x_1, \cdots, x_n)$ and $y = (y_1, \cdots, y_n)$ are given in the $\mathbb{R}^n$ then the Euclidean metric is defined by

$$d(x, y) = \sqrt{\sum_{i=1}^{n}(y_i - x_i)^2}$$

and the so-called taxicab metric is defined by

$$d(x, y) = \sum_{i=1}^{n} \mid y_i - x_i \mid .$$

Example 3.7

The examples given in Example 3.6 are both special cases of the Minkowsky metric. Given two points $x = (x_1, \cdots, x_n)$ and $y = (y_1, \cdots, y_n)$ in $\mathbb{R}^n$ then the Minkowsky distance between x and y is defined by

$$d(x,y) = \left(\sum_{i=1}^{n} (y_i - x_i)^p \right)^{\frac{1}{p}}, \ 0 < p \in \mathbb{R},$$

if $p \to \infty$ then the Chebychev metric is obtained

$$d(x,y) = \lim_{p \to \infty} \left(\sum_{i=1}^{n} (y_i - x_i)^p \right)^{\frac{1}{p}} = \max_{i \in \{1,\cdots,n\}} \mid y_i - x_i \mid .$$

Example 3.8

Given a set X the discrete metric d on X is defined by

$$d(x,y) = \begin{cases} 1 & \text{if} \quad x \neq y \\ 0 & \text{if} \quad x = y \end{cases}$$

The definition of an open and a closed ball in the Metric Space (X, d).

Definition 3.19

The set $\{x \,|\, x \in X, \, d(x, x_0) < r\}$ is called an open ball of radius r around the point x_0 and denoted by $B_r(x_0, d)$.
A closed ball of radius r around the point x_0 is defined and denoted by $\overline{B}_r(x_0, d) = \{x \,|\, x \in X, \, d(x, x_0) \leq r\}$.
A sphere of radius r around the point x_0 is defined and denoted by $S_r(x_0, d) = \{x \,|\, x \in X, \, d(x, x_0) = r\}$.

The definitions immediately implies that

$$S_r(x_0, d) = \overline{B}_r(x_0, d) - B_r(x_0, d).$$

Remark 3.4

Be aware of the fact, that the closed ball $\overline{B}_r(x_0, d)$ has not always to be equal to the closure of the open ball $B_r(x_0, d)$, denoted by $\overline{B_r(x_0, d)}$.
Take for the metric d the discrete metric, defined in 3.8, take $r = 1$ and see the difference.

The definition of an interior point and the interior of some subset G of the Metric Space (X, d).

Definition 3.20

Let G be some subset of X. $x \in G$ is called an interior point of G, if there exists some $r > 0$, such that $B_r(x_0, d) \subset G$.
The set of all interior points of G is called the interior of G and is denoted by G°.

Theorem 3.3

The distance function $d(\cdot, \cdot) : X \times X \to \mathbb{R}$ is continuous.

Proof of Theorem 3.3

Let $\epsilon > 0$ be given and x_0 and y_0 are two arbitrary points of X. For every $x \in X$ with $d(x, x_0) < \frac{\epsilon}{2}$ and for every $y \in X$ with $d(x, x_0) < \frac{\epsilon}{2}$, it is easily seen that

$$d(x, y) \leq d(x, x_0) + d(x_0, y_0) + d(y_0, y) < d(x_0, y_0) + \epsilon$$

and

$$d(x_0, y_0) \leq d(x_0, x) + d(x, y) + d(y, y_0) < d(x, y) + \epsilon$$

such that

$$|\ d(x, y) - d(x_0, y_0)\ | < \epsilon.$$

The points x_0 and y_0 are arbitrary chosen so the function d is continuous in X. □

The distance function d is used to define the distance between a point and a set, the distance between two sets and the diameter of a set.

Definition 3.21

Let (X, d) be a metric space.

a. The distance between a point $x \in X$ and a set $A \subset X$ is denoted and defined by

$$\text{dist}(\,x,\ A) = \inf\{\,d(x, y)\,|\,y \in A\}.$$

b. The distance between the sets $A \subset X$ and $B \subset X$ is denoted and defined by

$$\text{dist}(\,A,\ B) = \inf\{\,d(x, y)\,|\,x \in A,\ y \in B\}.$$

c. The diameter of $A \subset X$ is denoted and defined by

$$\text{diam}(\,A) = \sup\{\,d(x, y)\,|\,x \in A,\ y \in A\}.$$

The sets A and B are non-empty sets of X and $x \in X$.

Remark 3.5

The distance function $\text{dist}(\,\cdot,\ A)$ is most of the time denoted by $d(\,\cdot,\ A)$.

Theorem 3.4

The distance function $d(\cdot, A) : X \to \mathbb{R}$, defined in 3.21 is continuous.

Proof of Theorem 3.4

Let $x, y \in X$ then for each $a \in A$

$$d(x, a) \leq d(x, y) + d(y, a).$$

So that

$$d(x, A) \leq d(x, y) + d(y, a),$$

for each $a \in A$, so that

$$d(x, A) \leq d(x, y) + d(y, A),$$

which shows that

$$d(x, A) - d(y, A) \leq d(x, y).$$

Interchanging the names of the variables x and y and the result is

$$| \ d(x, A) - d(y, A) \ | \leq d(x, y),$$

which gives the continuity of $d(\cdot, A)$. □

Theorem 3.5

Let $\{X, d\}$ be a Metric Space. Let the sequence $\{x_n\}_{n \in \mathbb{N}}$ be a Cauchy sequence in X with a convergent subsequence $\{x_{n_k}\}_{k \in \mathbb{N}}$,

$$\text{if} \quad \lim_{k \to \infty} x_{n_k} = x \quad \text{then} \quad \lim_{n \to \infty} x_n = x.$$

Proof of Theorem 3.5

Let $\epsilon > 0$ be given. The subsequence $\{x_{n_k}\}_{k \in \mathbb{N}}$ has a limit x, so there exists some $K(\epsilon) \in \mathbb{N}$ such that for every $k > K(\epsilon)$: $d(x_{n_k}, x) < \frac{\epsilon}{2}$. The sequence

$\{x_n\}_{n \in \mathbb{N}}$ is a Cauchy sequence, so there exists some $N(\epsilon) \in \mathbb{N}$ such that for every $n, m > N(\epsilon): \ d(x_n, x_m) < \frac{\epsilon}{2}$. Let $n > \max\{ n_{K(\epsilon)}, N(\epsilon)\}$ and let $k > K(\epsilon)$ then

$$d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

The number $n > \max\{n_{K(\epsilon)}, N(\epsilon)\}$ is arbitrary chosen, so the limit of the sequence $\{x_n\}_{n \in \mathbb{N}}$ exists and is equal to x. □

3.5.1 Urysohn's Lemma

Theorem 3.6

Let $\{X, d\}$ be a Metric Space and let A, B be non-empty closed subsets of X, such that $A \cap B = \varnothing$. Then there exists a continuous function $g : X \to [0, 1]$ such that

$$g(x) = \begin{cases} 1 & \forall x \in A, \\ 0 & \forall x \in B. \end{cases}$$

Proof of Theorem 3.6

The definition of the distance function $\mathrm{dist}(\cdot, \cdot)$ is given in **definition 3.21**. The distance function $\mathrm{dist}(\cdot, Y)$ is denoted by $d(\cdot, Y)$ for any non-empty set $Y \subseteq X$. There is proved in **theorem 3.3** that the distance function is continuous, it is even uniform continuous. If the set Y is closed, then $d(x, Y) = 0 \Leftrightarrow x \in Y$. Given are the closed sets A and B, define for every $x \in X$

$$g(x) = \frac{d(x, B)}{d(x, A) + d(x, B)}.$$

The function g is continuous on X and satisfies the desired properties. □

3.6 Complete Metric Spaces

Definition 3.22

If every Cauchy row in a Metric Space MS_1 converges to an element of that same space MS_1 then the space MS_1 is called complete.
The space MS_1 is called a Complete Metric Space.

Theorem 3.7

If M is a subspace of a Complete Metric Space MS_1 then
M is complete *if and only if* M is closed in MS_1.

Proof of Theorem 3.7

($\Rightarrow$) Take some $x \in \overline{M}$. Then there exists a convergent sequence $\{x_n\}$ to x, see theorem 2.1. The sequence $\{x_n\}$ is a Cauchy sequence, see section 2.3 and since M is complete the sequence $\{x_n\}$ converges to an unique element $x \in M$. Hence $\overline{M} \subseteq M$.

($\Leftarrow$) Take a Cauchy sequence $\{x_n\}$ in the closed subspace M. The Cauchy sequence converges in MS_1, since MS_1 is a Complete Metric Space, this implies that $x_n \rightarrow x \in MS_1$, so $x \in \overline{M}$. M is closed, so $M = \overline{M}$ and this means that $x \in M$. Hence the Cauchy sequence $\{x_n\}$ converges in M, so M is complete. □

□

Theorem 3.8

For $1 \leq p \leq \infty$, the metric space ℓ^p is complete.

Proof of Theorem 3.8

1. Let $1 \leq p < \infty$. Consider a Cauchy sequence $\{x_n\}$ in ℓ^p.
   Given $\epsilon > 0$, then there exists a $N(\epsilon)$ such that for all $m, n > N(\epsilon)$ $d_p(x_n, x_m) < \epsilon$, with the metric d_p, defined by

$$d_p(x, y) = (\sum_{i=1}^{\infty} \mid x_i - y_i \mid^p)^{\frac{1}{p}}.$$

For $n,\, m > N(\epsilon)$ and for $i = 1, 2, \cdots$

$$\mid (x_n)_i - (x_m)_i \mid \leq d_p(x_n,\, x_m) \leq \epsilon.$$

For each fixed $i \in \{1, 2, \cdots\}$, the sequence $\{(x_n)_i\}$ is a Cauchy sequence in $\mathbb{K}$.
$\mathbb{K}$ is complete, so $(x_n)_i \rightarrow x_i$ in $\mathbb{K}$ for $n \rightarrow \infty$.
Define $x = (x_1,\, x_2, \cdots)$, there has to be shown that $x \in \ell^p$ and $x_n \rightarrow x$ in ℓ^p, for $n \rightarrow \infty$.
For all $n,\, m > N(\epsilon)$

$$\sum_{i=1}^{k} \mid (x_n)_i - (x_m)_i \mid^p < \epsilon^p$$

for $k = 1, 2, \cdots$. Let $m \rightarrow \infty$ then for $n > N(\epsilon)$

$$\sum_{i=1}^{k} \mid (x_n)_i - x_i \mid^p \leq \epsilon^p$$

for $k = 1, 2, \cdots$. Now letting $k \rightarrow \infty$ and the result is that

$$d_p(x_n, x) \leq \epsilon \tag{3.2}$$

for $n > N(\epsilon)$, so $(x_n - x) \in \ell^p$. Using the Minkowski inequality ?? ii.b, there follows that $x = x_n + (x - x_n) \in \ell^p$.
Inequality 3.2 implies that $x_n \rightarrow x$ for $n \rightarrow \infty$.
The sequence $\{x_n\}$ was an arbitrary chosen Cauchy sequence in ℓ^p, so ℓ^p is complete for $1 \leq p < \infty$.

2. For $p = \infty$, the proof is going almost on the same way as for $1 \leq p < \infty$, only with the metric d_∞, defined by

$$d_\infty(x, y) = \sup_{i \in \mathbb{N}} | x_i - y_i |$$

for every $x, y \in \ell^\infty$. □

□

3.7 Normed Spaces

Definition 3.23

A Normed Space NS is a pair $(X, \| \,.\, \|)$. X is a topological vector space, the topology of X is defined by the *norm* $\| \,.\, \|$. The *norm* is a real-valued function on X and satisfies for all $x, y \in X$ and $\alpha \in \mathbb{R}$ or $\mathbb{C}$,

N 1: $\| x \| \geq 0$, (positive)

N 2: $\| x \| = 0 \Longleftrightarrow x = 0$,

N 3: $\| \alpha x \| = | \alpha | \| x \|$, (homogeneous)

N 4: $\| x + y \| \leq \| x \| + \| y \|$, (Triangle inequality).

A normed space is also a metric space. A metric d induced by the norm is given by

$$d(x, y) = \| x - y \| . \tag{3.3}$$

A mapping $p : X \to \mathbb{R}$, that is almost a norm, is called a seminorm or a pseudonorm .

Definition 3.24

Let X be a Vector Space. A mapping $p : X \to \mathbb{R}$ is called a seminorm or pseudonorm if it satisfies the conditions (N 1), (N 3) and (N 4), given in definition 3.23.

Remark 3.6

If p is a seminorm on the Vector Space X and if $p(x) = 0$ implies that $x = 0$ then p is a norm.
A seminorm p satisfies:

$$p(0) = 0,$$
$$|\, p(x) - p(y) \,| \leq p(x - y).$$

Besides the triangle inequality given by (N 4), there is also the so-called inverse triangle inequality

$$|\, \| x \| - \| y \| \,| \leq \| x - y \| . \tag{3.4}$$

The inverse triangle inequality is also true in Metric Spaces

$$|\ d(x, y) - d(y, z)\ | \leq d(x, z).$$

With these triangle inequalities lower and upper bounds can be given of $\| x - y \|$ or
$\| x + y \|$.

Theorem 3.9

Given is a Normed Space $(X, \| \cdot \|)$. The map

$$\| \cdot \| : X \to [0, \infty)$$

is continuous in $x = x_0$, for every $x_0 \in X$.

Proof of Theorem 3.9

Let $\epsilon > 0$ be given. Take $\delta = \epsilon$ then is obtained, that for every $x \in X$ with $\| x - x_0 \| < \delta$ that $| \| x \| - \| x_0 \| | \leq \| x - x_0 \| < \delta = \epsilon$. □

There is also said that the norm is continuous in its own topology on X.

On a Vector Space X there can be defined an infinitely number of different norms. Between some of these different norms there is almost no difference in the topology they generate on the Vector Space X. If some different norms are not to be distinguished of each other, these norms are called equivalent norms.

Definition 3.25

Let X be a Vector Space with norms $\| \cdot \|_0$ and $\| \cdot \|_1$. The norms $\| \cdot \|_0$ and $\| \cdot \|_1$ are said to be equivalent if there exist numbers $m > 0$ and $M > 0$ such that for every $x \in X$

$$m \| x \|_0 \leq \| x \|_1 \leq M \| x \|_0 .$$

The constants m and M are independent of x!

In Linear Algebra there is used, most of the time, only one norm and that is the Euclidean norm: $\| \cdot \|_2$, if $x = (x_1, \cdots, x_N) \in \mathbb{R}^N$ then $\| x \|_2 = \sqrt{\sum_{i=1}^{N} | x_i |^2}$. Here beneath the reason why!

Theorem 3.10

All norms on a finite-dimensional Vector Space X (over $\mathbb{R}$ or $\mathbb{C}$) are equivalent.

Proof of Theorem 3.10

Let $\| \cdot \|$ be a norm on X and let $\{x_1, x_2, \cdots, x_N\}$ be a basis for X, the dimension of X is N. Define another norm $\| \cdot \|_2$ on X by

$$\| \sum_{i=1}^{N} \alpha_i x_i \|_2 = (\sum_{i=1}^{N} | \alpha_i |^2)^{\frac{1}{2}}.$$

If the norms $\| \cdot \|$ and $\| \cdot \|_2$ are equivalent then all the norms on X are equivalent. Define $M = (\sum_{i=1}^{N} \| x_i \|^2)^{\frac{1}{2}}$, M is positive because $\{x_1, x_2, \cdots, x_N\}$ is a basis for X. Let $x \in X$ with $x = \sum_{i=1}^{N} \alpha_i x_i$, using the triangle inequality and the inequality of Cauchy-Schwarz, see theorem 5.42, gives

$$\begin{aligned}
\| x \| = \| \sum_{i=1}^{N} \alpha_i x_i \| &\leq \sum_{i=1}^{N} \| \alpha_i x_i \| \\
&= \sum_{i=1}^{N} | \alpha_i | \| x_i \| \\
&\leq (\sum_{i=1}^{N} | \alpha_i |^2)^{\frac{1}{2}} (\sum_{i=1}^{N} \| x_i \|^2)^{\frac{1}{2}} \\
&= M \| \sum_{i=1}^{N} \alpha_i x_i \|_2 = M \| x \|_2
\end{aligned}$$

Define the function $f : \mathbb{K}^N \to \mathbb{K}$, with $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ by

$$f(\alpha_1, \alpha_2, \cdots, \alpha_N) = \| \sum_{i=1}^{N} \alpha_i x_i \| .$$

The function f is continuous in the $\| \cdot \|_2$-norm, because

$$| f(\alpha_1, \cdots, \alpha_N) - f(\beta_1, \cdots, \beta_N) | \leq \| \sum_{i=1}^{N} (\alpha_i - \beta_i) x_i \|$$

$$\leq M (\sum_{i=1}^{N} | \alpha_i - \beta_i |^2)^{\frac{1}{2}} \; (= M \| \sum_{i=1}^{N} (\alpha_i - \beta_i) x_i \|_2).$$

Above are used the continuity of the norm $\| \cdot \|$ and the inequality of Cauchy-Schwarz.
The set

$$S_1 = \{(\gamma_1, \cdots, \gamma_N) \in \mathbb{K}^N \mid \sum_{i=1}^{N} | \gamma_i |^2 = 1\}$$

is a compact set, the function f is continuous in the $\| \cdot \|_2$-norm, so there exists a point $(\theta_1, \cdots, \theta_N) \in S_1$ such that

$$m = f(\theta_1, \cdots, \theta_N) \leq f(\alpha_1, \cdots, \alpha_N)$$

for all $(\alpha_1, \cdots, \alpha_N) \in S_1$.
If $m = 0$ then $\| \sum_{i=1}^{N} \theta_i x_i \| = 0$, so $\sum_{i=1}^{N} \theta_i x_i = 0$ and there follows that $\theta_i = 0$ for all $1 < i < N$, because $\{x_1, x_2, \cdots, x_N\}$ is basis of X, but this contradicts the fact that $(\theta_1, \cdots, \theta_N) \in S_1$.
Hence $m > 0$.
The result is that, if $\| \sum_{i=1}^{N} \alpha_i x_i \|_2 = 1$ then $f(\alpha_1, \cdots, \alpha_N) = \| \sum_{i=1}^{N} \alpha_i x_i \| \geq m$.
For every $x \in X$, with $x \neq 0$, is $\| \frac{x}{\| x \|_2} \|_2 = 1$, so $\| \frac{x}{\| x \|_2} \| \geq m$ and this results in

$$\| x \| \geq m \| x \|_2,$$

which is also valid for $x = 0$. The norms $\| \cdot \|$ and $\| \cdot \|_2$ are equivalent

$$m \| x \|_2 \leq \| x \| \leq M \| x \|_2 . \tag{3.5}$$

If $\| \cdot \|_1$ should be another norm on X, then with the same reasoning as above, there can be found constants $m_1 > 0$ and $M_1 > 0$, such that

$$m_1 \| x \|_2 \leq \| x \|_1 \leq M_1 \| x \|_2 . \tag{3.6}$$

and combining the results of 3.5 and 3.6 results in

$$\frac{m}{M_1} \| x \|_1 \leq \| x \| \leq \frac{M}{m_1} \| x \|_1$$

so the norms $\| \cdot \|_1$ and $\| \cdot \|$ are equivalent. □

3.7.1 Hamel and Schauder bases

In section 3.2, about Vector Spaces, there is made some remark about problems by defining infinite sums, see section 3.2.6. In a normed space, the norm can be used to overcome some problems.
Every Vector Space has a Hamel basis, see Theorem 3.2, but in the case of infinite dimensional Vector Spaces it is difficult to find the right form of it. It should be very helpfull to get some basis, where elements x out of the normed space X can be approximated by limits of finite sums. If such a basis exists it is called a Schauder basis .

Definition 3.26

Let X be a Vector Space over the field $\mathbb{K}$. If the Normed Space $(X, \| \cdot \|)$ has a countable sequence $\{e_n\}_{n \in \mathbb{N}}$ with the property that for every $x \in X$ there exists an unique sequence $\{\alpha_n\}_{n \in \mathbb{N}} \subset \mathbb{K}$ such that

$$\lim_{n \to \infty} \| x - \sum_{i=1}^{n} \alpha_i \, e_i \| = 0.$$

then $\{e_n\}$ is called a Schauder basis of X.

Some textbooks will define Schauder bases for Banach Spaces, see section 3.8, and not for Normed Spaces. Having a Schauder basis $\{e_n\}_{n \in \mathbb{N}}$, it is now possible to look to all possible linear combinations of these basis vectors $\{e_n\}_{n \in \mathbb{N}}$. To be careful, it is may be better to look to all possible Cauchy sequences, which can be constructed with these basis vectors $\{e_n\}_{n \in \mathbb{N}}$.
The Normed Space X united with all the limits of these Cauchy sequences in X, is denoted by $\hat{X}$ and in most cases it will be greater then the original Normed Space X. The space $(\hat{X}, \| \cdot \|_1)$ is called the completion of the normed space $(X, \| \cdot \|)$ and is complete, so a Banach Space.
May be it is useful to read how the real numbers ($\mathbb{R}$) can be constructed out of the rational numbers ($\mathbb{Q}$), with the use of Cauchy sequences, see **wiki-constr-real**. Keep in mind that, in general, elements of a Normed Space can not be multiplied with each other. There is defined a scalar multiplication on such a Normed Space. Further there is, in general, no ordening relation between elements of a Normed Space. These two facts are the great differences between the completion of the rational numbers and the completion of an arbitrary Normed Space, but further the construction of such a completion is almost the same.

Theorem 3.11

Every Normed Space $(X, \| \cdot \|)$ has a completion $(\hat{X}, \| \cdot \|_1)$.

Proof of Theorem 3.11

Here is not given a proof, but here is given the construction of a completion.
There has to overcome a problem with the norm $\| \cdot \|$. If some element $y \in \hat{X}$ but $y \notin X$, then $\| y \|$ has no meaning. That is also the reason of the index 1 to the

norm on the Vector Space $\hat{X}$.

The problem is easily fixed by looking at equivalence classes of Cauchy sequences. More information about equivalence classes can be found in **wiki-equi-class**. Important is the equivalence relation, denoted by $\sim$. If $\{x_n\}_{n \in \mathbb{N}}$ and $\{y_n\}_{n \in \mathbb{N}}$ are two Cauchy sequences in X then an equivalence relation $\sim$ is defined by

$$\{x_n\} \sim \{y_n\} \iff \lim_{n \to \infty} \| x_n - y_n \| = 0.$$

An equivalence class is denoted by $\tilde{x} = [\{x_n\}]$ and equivalence classes can be added, or multiplied by a scalar, such that $\hat{X}$ is a Vector Space. The norm $\| \cdot \|_1$ is defined by

$$\| \tilde{x} \|_1 = \lim_{n \to \infty} \| x_n \|$$

with $\{x_n\}$ a sequence out of the equivalence class $\tilde{x}$.

To complete the proof of the theorem several things have to be done, such as to proof that

1. there exists a norm preserving map of X onto a subspace W of X, with W dense in $\hat{X}$,
2. the constructed space $(\hat{X}, \| \cdot \|_1)$ is complete,
3. the space $\hat{X}$ is unique, except for isometric isomorphisms[3].

It is not difficult to prove these facts but it is lengthy.

See **section 3.11.4** for a proof, but then for a Metric Space.

It becomes clear, that is easier to define a Schauder basis for a Banach Space then for a Normed Space, the problems of a completion are circumvented.

Next are given some nice examples of a space with a Hamel basis and set of linear independent elements, which is a Schauder basis, but not a Hamel basis.

[3] For isometric isomorphisms, see page **121**

Example 3.9

Look at the space c_{00} out of section **5.2.7**, the space of sequences with only a finite number of coefficients not equal to zero. c_{00} is a linear subspace of ℓ^{∞} and equipped with the norm $\| \cdot \|_{\infty}$-norm, see section **5.2.1**.
The canonical base of c_{00} is defined by

$$\begin{aligned} e_1 &= (1, 0, 0, \cdots), \\ e_2 &= (0, 1, 0, \cdots), \\ \cdots & \quad \cdots \\ e_k &= (\underbrace{0, \cdots, 0}_{(k-1)}, 1, 0, \cdots), \\ \cdots & \end{aligned}$$

and is a Hamel basis of c_{00}.

Explanation of Example 3.9

Take an arbitrary $x \in c_{00}$ then $x = (\alpha_1, \alpha_2, \cdots, \alpha_n, 0, 0, \cdots)$ with $\alpha_i = 0$ for $i > n$ and $n \in \mathbb{N}$. So x can be written by a finite sum of the basisvectors out of the given canonical basis:

$$x = \sum_{i=1}^{n} \alpha_i e_i,$$

and the canonical basis is a Hamel basis of c_{00}. □

Example 3.10

Look at the space c_{00}, see example 3.9. Let's define a set of sequences

$$\begin{aligned} b_1 &= (1, \frac{1}{2}, 0, \cdots) \\ b_2 &= (0, \frac{1}{2}, \frac{1}{3}, 0, \cdots) \\ \cdots & \quad \cdots \\ b_k &= (\underbrace{0, \cdots, 0}_{(k-1)}, \frac{1}{k}, \frac{1}{k+1}, 0, \cdots), \\ \cdots & \end{aligned}$$

The system $\{b_1, b_2, b_3, \cdots\}$ is a Schauder basis of c_{00} but it is not a Hamel basis of c_{00}.

Explanation of Example 3.10

If the set given set of sequences $\{b_n\}_{n \in \mathbb{N}}$ is a basis of c_{00} then it is easy to see that

$$e_1 = \lim_{N \to \infty} \sum_{j=1}^{N} (-1)^{(j-1)} b_j,$$

and because of the fact that

$$\| b_k \|_\infty = \frac{1}{k}$$

for every $k \in \mathbb{N}$, it follows that:

$$\| e_1 - \sum_{j=1}^{N} (-1)^{(j-1)} b_j \|_\infty \leq \frac{1}{N+1}.$$

Realize that $(e_1 - \sum_{j=1}^{N} (-1)^{(j-1)} b_j) \in c_{00}$ for every $N \in \mathbb{N}$, so there are no problems by calculating the norm.
This means that e_1 is a summation of an infinite number of elements out of the set $\{b_n\}_{n \in \mathbb{N}}$, so this set can not be a Hamel basis.
Take a finite linear combination of elements out of $\{b_n\}_{n \in \mathbb{N}}$ and solve

$$\sum_{j=1}^{N} \gamma_i \, b_j = (0, 0, \cdots, 0, 0, \cdots),$$

this gives $\gamma_j = 0$ for every $1 \leq j \leq N$, with $N \in \mathbb{N}$ arbitrary chosen. This means that the set of sequences $\{b_n\}_{n \in \mathbb{N}}$ is linear independent in the sense of section 3.2.4. Take now an arbitrary $x \in c_{00}$ then $x = (\alpha_1, \alpha_2, \cdots, \alpha_n, 0, 0, \cdots)$ with $\alpha_i = 0$ for $i > n$ and $n \in \mathbb{N}$. To find, is a sequence $(\gamma_1, \gamma_2, \cdots)$ such that

$$x = \sum_{j=1}^{\infty} \gamma_j \, b_j. \tag{3.7}$$

Equation 3.7 gives the following set of linear equations

$$\begin{aligned}
\alpha_1 &= \gamma_1, \\
\alpha_2 &= \frac{1}{2}\gamma_1 + \frac{1}{2}\gamma_2, \\
\cdots & \quad \cdots \\
\alpha_n &= \frac{1}{n}\gamma_{n-1} + \frac{1}{n}\gamma_n, \\
0 &= \frac{1}{n+1}\gamma_n + \frac{1}{n+1}\gamma_{n+1}, \\
\cdots & \quad \cdots,
\end{aligned}$$

which is solvable. Since γ_1 is known, all the values of γ_i with $2 \leq i \leq n$ are known. Remarkable is that $\gamma_{k+1} = -\gamma_k$ for $k \geq n$ and because of the fact that γ_n is known all the next coeffcients are also known.
One thing has to be done! Take $N \in \mathbb{N}$ great enough and calculate

$$\| \, x - \sum_{j=1}^{N} \gamma_j \, b_j \, \|_\infty = \| \, (\underbrace{0, \cdots, 0}_{N}, \gamma_N, -\gamma_N, \cdots) \, \|_\infty \leq | \, \gamma_N \, | \, \| \, e_{N+1} \, \|_\infty = \frac{| \, \gamma_N \, |}{(N+1)}$$

So $\lim_{N \to \infty} \| \, x - \sum_{j=1}^{N} \gamma_j \, b_j \, \|_\infty = 0$ and the conclusion becomes that the system $\{b_n\}_{n \in \mathbb{N}}$ is a Schauder basis of c_{00}. □

Sometimes there is also spoken about a total set or fundamental set.

Definition 3.27

A total set (or fundamental set) in a Normed Space X is a subset $M \subset X$ whose span 3.10 is dense in X.

Remark 3.7

According the definition:

$$M \text{ is total in } X \text{ if and only if } \overline{\operatorname{span} M} = X.$$

Be careful: a complete set is total, but the converse need not hold in infinite-dimensional spaces.

3.8 Banach Spaces

Definition 3.28

If every Cauchy row in a Normed Space $(X, \|\cdot\|)$ converges to an element of that same space X then that Normed Space $(X, \|\cdot\|)$ is called complete in the metric induced by the norm.

A complete Normed Space $(X, \|\cdot\|)$ is called a Banach Space.

Theorem 3.12

Let Y be a subspace of a Banach Space $(X, \|\cdot\|)$. Then,
Y is closed *if and only if* Y is complete.

Proof of Theorem 3.12

($\Rightarrow$) Let $\{x_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence in Y, then it is also in BS. BS is complete, so there exists some $x \in BS$ such that $x_n \rightarrow x$. Every neighbourhood of x contains points out of Y, take $x_n \neq x$, with n great enough. This means that x is an accumulation point of Y, see section **2.5**. Y is closed, so $x \in Y$ and there is proved that Y is complete.

($\Leftarrow$) Let x be a limitpoint of Y. So there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subset Y$, such that $x_n \rightarrow x$ for $n \rightarrow x\infty$. A convergent sequence is a Cauchy sequence. Y is complete, so the sequence $\{x_n\}_{n\in\mathbb{N}}$ converges in Y. It follows that $x \in Y$, so Y is closed.

□

3.9 Inner Product Spaces

The norm of an Inner Product Space can be expressed as an inner product and so the inner product defines a topology on such a space. An Inner Product gives also information about the position of two elements with respect to each other.

Definition 3.29

An Inner Product Space IPS is a pair $(X, (.,.))$. X is a topological vector space, the topology on X is defined by the norm induced by the *inner product* $(.,.)$. The *inner product* $(.,.)$ is a real or complex valued function on XxX and satisfies for all $x, y, z \in X$ and $\alpha \in \mathbb{R}$ or $\mathbb{C}$

IP 1: $0 \leq (x, x) \in \mathbb{R}$ and $(x, x) = 0 \Longleftrightarrow x = 0,$

IP 2: $(x, y) = \overline{(y, x)},$

IP 3: $(\alpha x, y) = \alpha(x, y),$

IP 4: $(x + y, z) = (x, z) + (y, z),$

with $\overline{(y, x)}$ is meant, the complex conjugate[4] of the value (y, x).

The inner product $(.,.)$ defines a norm $\| \,.\, \|$ on X

$$\| x \| = \sqrt{(x, x)} \tag{3.8}$$

and this norm induces a metric d on X by

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

in the same way as formula (3.3).
An Inner Product Space is also called a pre-Hilbert space.

3.9.1 Inequality of Cauchy-Schwarz (general)

The inequality of Cauchy-Schwarz is valid for every inner product.

Theorem 3.13

Let X be an Inner Product Space with inner product $(\cdot, \cdot)$, for every $x, y \in X$ holds that

$$|\,(x, y)\,| \leq \|\, x \,\|\, \|\, y \,\|\,. \tag{3.9}$$

Proof of Theorem 3.13

Condition IP 1 and definition 3.8 gives that

$$0 \leq (x - \alpha\, y, x - \alpha\, y) = \|\, x - \alpha\, y \,\|^2$$

for every $x, y \in X$ and $\alpha \in \mathbb{K}$, with $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.
This gives

$$\begin{aligned} 0 &\leq (x, x) - (x, \alpha\, y) - (\alpha\, y, x) + (\alpha\, y, \alpha\, y) \\ &= (x, x) - \overline{\alpha}\,(x, y) - \alpha(y, x) + \overline{\alpha}\,\alpha\,(y, y). \end{aligned} \tag{3.10}$$

If $(y, y) = 0$ then $y = 0$ (see condition IP 1) and there is no problem. Assume $y \neq 0$, in the sense that $(y, y) \neq 0$, and take

$$\alpha = \frac{(x, y)}{(y, y)}.$$

Put α in inequality 3.10 and use that

$$(x, y) = \overline{(y, x)},$$

see condition IP 2. Writing out and some calculations gives the inequality of Cauchy-Schwarz. □

Theorem 3.14

If $(X, (.,.))$ is an Inner Product Space, then is the inner product $(.,.) : X \times X \to \mathbb{K}$ continuous. This means that if

$$x_n \to x \text{ and } y_n \to y \text{ then } (x_n, y_n) \to (x, y) \text{ for } n \to \infty.$$

Proof of Theorem 3.14

With the triangle inequality and the inequality of Cauchy-Schwarz is obtained

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

since $\|\,x_n - x\,\| \to 0$ and $\|\,y_n - y\,\| \to 0$ for $n \to \infty$. □

So the norm and the inner product are continuous, see theorem **3.9** and theorem **3.14**.

3.9.2 Parallelogram Identity and Polarization Identity

An important equality is the parallelogram equality, see figure **3.3**.

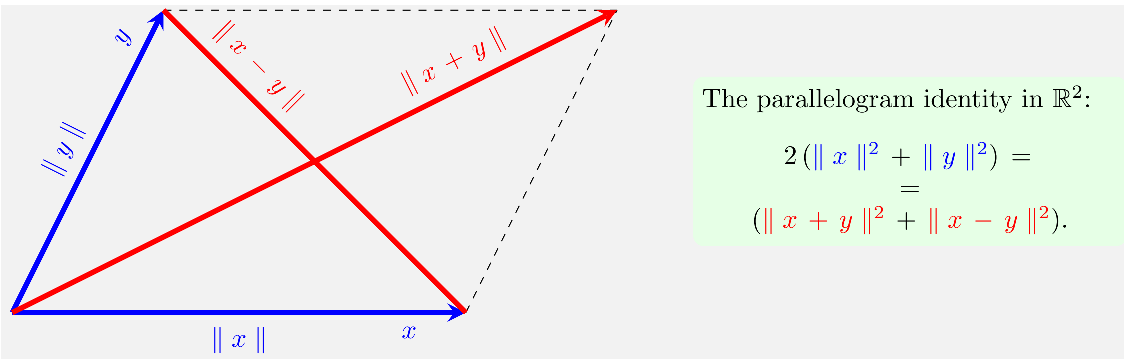


Figure 3.3 Parallelogram Identity

If it is not sure, if the used norm $\|\cdot\|$ is induced by an inner product, the check of the parallelogram identity will be very useful. If the norm $\|\cdot\|$ satisfies the parallelogram identity then the inner product $(\cdot,\cdot)$ can be recovered by the norm, using the so-called polarization identity.

Theorem 3.15

An inner product $(\cdot,\cdot)$ can be recovered by the norm $\|\cdot\|$ on a Vector Space X *if and only if* the norm $\|\cdot\|$ satisfies the parallelogram identity

$$2\,(\| x \|^2 + \| y \|^2) = (\| x + y \|^2 + \| x - y \|^2). \tag{3.11}$$

The inner product is given by the polarization identity

$$(x, y) = \frac{1}{4}\Big\{(\| x + y \|^2 - \| x - y \|^2) + i\,(\| x + i\,y \|^2 - \| x - i\,y \|^2)\Big\}. \tag{3.12}$$

Proof of Theorem 3.15

($\Rightarrow$) If the inner product can be recovered by the norm $\| x \|$ then $(x,\, x) = \| x \|^2$ and

$$\| x + y \|^2 = (x + y,\, x + y)$$
$$= \| x \|^2 + (x,\, y) + (y,\, x) + \| y \|^2 = \| x \|^2 + (x,\, y) + \overline{(x,\, y)} + \| y \|^2,$$

where with $\overline{(x,\, y)}$ is meant the complex conjugate of $(x,\, y)$.
Replace y by $(-y)$ and there is obtained

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

Adding the obtainded formulas together gives the parallelogram identity 3.11.

($\Leftarrow$) Here the question becomes if the right-hand site of formula 3.12 is an inner product? The first two conditions, IP1 and IP2 are relative easy. The conditions IP3 and IP4 require more attention. Conditon IP4 is used in the proof of the scalar multiplication, condition IP3. The parallelogram identity is used in the proof of IP4.

IP 1: The inner product $(\cdot,\cdot)$ induces the norm $\|\cdot\|$:

$$(x,x) = \frac{1}{4}\left\{(\| x + x \|^2 - \| x - x \|^2) + i\,(\| x + i\,x \|^2 - \| x - i\,x \|^2)\right\}$$
$$= \frac{1}{4}\left\{4 \| x \|^2 + i\,(|\,(1 + i\,|^2 - |\,(1 - i\,|^2)\ \| x \|^2)\right\}$$
$$= \| x \|^2 .$$

IP 2:

$$\overline{(y,x)} = \frac{1}{4}\left\{(\| y + x \|^2 - \| y - x \|^2) - i\,(\| y + i\,x \|^2 - \| y - i\,x \|^2)\right\}$$
$$= \frac{1}{4}\left\{(\| x + y \|^2 - \| x - y \|^2) - i\,(|\,-i\,|^2 \| y + i\,x \|^2 - |\,i\,|^2 \| y - i\,x \|^2)\right\}$$
$$= \frac{1}{4}\left\{(\| x + y \|^2 - \| x - y \|^2) - i\,(\| -i\,y + x \|^2 - \| i\,y + x \|^2)\right\}$$
$$= \frac{1}{4}\left\{(\| x + y \|^2 - \| x - y \|^2) + i\,(\| x + i\,y \|^2 - \| x - i\,y \|^2)\right\} = (x,y)$$

IP 3: Take first notice of the result of IP4. The consequence of **3.16** is that by a trivial induction can be proved that

$$(n\,x,\, y) = n\,(x,\, y) \tag{3.13}$$

and hence $(x,\, y) = (n\,\frac{x}{n},\, y) = n\,(\frac{x}{n},\, y)$, such that

$$(\frac{x}{n},\, y) = \frac{1}{n}\,(x,\, y), \tag{3.14}$$

for every positive integer n. The above obtained expressions **3.13** and **3.14** imply that

$$(q\,x,\, y) = q\,(x,\, y),$$

for every rational number q, and $(0,y) = 0$ by the polarization identity.
The polarization identity also ensures that

$$(\,-x,y) = (-1)\,(x,y).$$

Every real number can b e approximated by a row of rational numbers, $\mathbb{Q}$ is dense in $\mathbb{R}$. Take an arbitrary $\alpha \in \mathbb{R}$ and there exists a sequence $\{q_n\}_{n \in \mathbb{N}}$ such that q_n converges in $\mathbb{R}$ to α for $n \to \infty$, this together with

$$-(\alpha\, x,\, y) \;=\; (-\alpha\, x, y)$$

gives that

$$\mid (q_n\, x, y) \;-\; (\alpha\, x,\, y) \mid = \mid ((q_n \;-\; \alpha)\, x,\, y) \mid .$$

The polarization identity and the continuity of the norm ensures that $\mid ((q_n \;-\; \alpha)\, x,\, y) \mid \rightarrow 0$ for $n \rightarrow \infty$. This all here results in

$$(\alpha\, x, y) \;=\; \lim_{n\to\infty} (q_n x,\, y) \;=\; \lim_{n\to\infty} q_n(\, x,\, y) \;=\; \alpha\, (x, y).$$

The polarization identity ensures that $i\, (x, y) \;=\; (\, i\, x,\, y)$ for every $x,\, y \in X$. Take $\lambda = \alpha \,+\, i\, \beta \in \mathbb{C}$ and $(\lambda\, x,\, y) \;=\; (\,(\alpha \,+\, i\, \beta)\, x,\, y) \;=\; (\alpha\, x,\, y) \,+\, (i\, \beta\, x,\, y) \;=\; (\alpha \,+\, i\, \beta)\, (x,\, y) \;=\; \lambda\, (x,\, y)$, conclusion

$$(\, \lambda\, x,\, y) \;=\; \lambda(\, x,\, y)$$

for every $\lambda \in \mathbb{C}$ and for all $x, y \in X$.

IP 4: The parallelogram identity is used. First $(x \,+\, z)$ and $(y \,+\, z)$ are rewritten

$$\begin{aligned} x \,+\, z &= (\frac{x \,+\, y}{2} \,+\, z) \,+\, \frac{x \,-\, y}{2}, \\ y \,+\, z &= (\frac{x \,+\, y}{2} \,+\, z) \,-\, \frac{x \,-\, y}{2}. \end{aligned}$$

The parallelogram identity is used, such that

$$\| \, x \,+\, z \,\|^2 \,+\, \| \, y \,+\, z \,\|^2 = 2\, (\| \, \frac{x \,+\, y}{2} \,+\, z \,\|^2 \,+\, \| \, \frac{x \,-\, y}{2} \, \|^2).$$

Hence

$$\begin{aligned} &(x, z) \,+\, (y, z) \;=\; \frac{1}{4}\Big\{(\| \, x \,+\, z \,\|^2 \,+\, \| \, y \,+\, z \,\|^2) \,-\, (\| \, x \,-\, z \,\|^2 \,+\, \| \, y \,-\, z \,\|^2) \\ &+ i\, (\| \, x \,+\, i\, z \,\|^2 \,+\, \| \, y \,+\, i\, z \,\|^2) \,-\, i\, (\| \, x \,-\, i\, z \,\|^2 \,+\, \| \, y \,-\, i\, z \,\|^2)\Big\} \\ &= \frac{1}{2}\Big\{(\| \, \frac{x \,+\, y}{2} \,+\, z \,\|^2 \,+\, \| \, \frac{x \,-\, y}{2} \,\|^2) \,-\, (\| \, \frac{x \,+\, y}{2} \,-\, z \,\|^2 \,+\, \| \, \frac{x \,-\, y}{2} \,\|^2) \\ &+ i\, (\| \, \frac{x \,+\, y}{2} \,+\, i\, z \,\|^2 \,+\, \| \, \frac{x \,-\, y}{2} \,\|^2) \,-\, i\, (\| \, \frac{x \,+\, y}{2} \,-\, i\, z \,\|^2 \,+\, \| \, \frac{x \,-\, y}{2} \,\|^2)\Big\} \\ &= 2\, (\frac{x \,+\, y}{2}, z) \end{aligned}$$

for every $x, y, z \in X$, so also for $y = 0$ and that gives

$$(x, z) = 2\,(\frac{x}{2}, z) \tag{3.15}$$

for every $x, z \in X$. The consequence of **3.15** is that

$$(x, z) + (y, z) = (x + y, z) \tag{3.16}$$

for every $x, y, z \in X$.

□

3.9.3 Orthogonality

In an Inner Product Space $(X, (.,.))$, there can be get information about the position of two vectors x and y with respect to each other. With the geometrical definition of an inner product the angle can be calculated between two elements x and y.

Definition 3.30

Let $(X, (.,.))$ be an Inner Product Space, the geometrical definition of the inner product $(.,.)$ is

$$(x, y) = \| x \| \| y \| \cos(\angle x, y),$$

for every $x, y \in X$, with $\angle x, y$ is denoted the angle between the elements $x, y \in X$.

An important property is if elements in an Inner Product Space are perpendicular or not.

Definition 3.31

Let $(X, (.,.))$ be an Inner Product Space. A vector $0 \neq x \in X$ is said to be orthogonal to the vector $0 \neq y \in X$ if

$$(x, y) = 0,$$

x and y are called orhogonal vectors, denoted by $x \perp y$.

If $A, B \subset X$ are non-empty subsets of X then

a. $x \perp A$, if $(x, y) = 0$ for each $y \in A$,
b. $A \perp B$, if $(x, y) = 0$ if $x \perp y$ for each $x \in A$ and $y \in B$.

If $A, B \subset X$ are non-empty subspaces of X and $A \perp B$ then is $A + B$, see 3.3, called the orthogonal sum of A and B.
All the elements of X, which stay orthogonal to some non-empty subset $A \subset X$ is called the orthoplement of A.

Definition 3.32

Let $(X, (.,.))$ be an Inner Product Space and let A be an non-empty subset of X, then

$$A^{\perp} = \{x \in X \mid (x, y) = 0 \text{ for every } y \in A\}$$

is called the orthoplement of A.

Theorem 3.16

Let A, B be non-empty subsets of some Inner Product Space $(X, (.,.))$.

a. If A be a subset of X then is the set $A^{\perp}$ a closed subspace of X.
b. $A \cap A^{\perp}$ is empty or $A \cap A^{\perp} = \{0\}$.
c. If A be a subset of X then $A \subset A^{\perp\perp}$.
d. If A, B are subsets of X and $A \subset B$, then $A^{\perp} \supset B^{\perp}$.
e. $A^{\perp} = (\text{span}(A))^{\perp} = (\overline{\text{span(A)}})^{\perp}$.

Proof of Theorem 3.16

a. Let $x, y \in A^{\perp}$ and $\alpha \in \mathbb{K}$, then

$$(x + \alpha\, y, z) = (x, z) + \alpha\, (y, z) = 0$$

for every $z \in A$. Hence $A^{\perp}$ is a linear subspace of X.

Remains to prove: $A^{\perp} = \overline{A^{\perp}}$.

($\Rightarrow$) The set $\overline{A^{\perp}}$ is equal to $A^{\perp}$ unified with all its accumulation points, so $A^{\perp} \subseteq \overline{A^{\perp}}$.

($\Leftarrow$) Let $x \in \overline{A^{\perp}}$ then there exist a sequence $\{x_n\}$ in $A^{\perp}$ such that $\lim_{n\to\infty} x_n = x$. Hence

$$(x, z) = \lim_{n\to\infty} (x_n, z) = 0,$$

for every $z \in A$. (Inner product is continuous.) So $x \in A^{\perp}$ and $\overline{A^{\perp}} \subseteq A^{\perp}$.

b. If $x \in A \cap A^{\perp} \neq \varnothing$ then $x \perp x$, so $x = 0$.

c. If $x \in A$, and $x \perp A^{\perp}$ means that $x \in (A^{\perp})^{\perp}$, so $A \subset A^{\perp\perp}$.

d. If $x \in B^{\perp}$ then $(x, y) = 0$ for each $y \in B$ and in particular for every $x \in A \subset B$. So $x \in A^{\perp}$, this gives $B^{\perp} \subset A^{\perp}$.

e. ($\Leftarrow$) Since $A \subset \mathrm{span}(A) \subset \overline{\mathrm{span(A)}}$, from (d.) follows that $(\overline{\mathrm{span(A)}})^{\perp} \subset (\mathrm{span}(A))^{\perp} \subset A^{\perp}$.

($\Rightarrow$) If $x \in A^{\perp}$ then $(x, y) = 0$ for all $y \in A$. Since $\mathrm{span}(A)$ are finite linear combinations of elements out of A, $(x, y) = 0$ for all $y \in \mathrm{span}(A)$ as well. If $t \in \overline{\mathrm{span(A)}}$, then there exist a sequence $\{t_n\}_{n\in\mathbb{N}} \subset \mathrm{span}(A)$ such that $\lim_{n\to\infty} t_n = t$. The inner product is continuous so

$$(x, t) = \lim_{n\to\infty} (x, t_n) = 0,$$

so $x \in (\overline{\mathrm{span(A)}})^{\perp}$.

□

3.9.4 Orthogonal and orthonormal systems

Important systems in Inner Product spaces are the orthogonal and orthonormal systems. Orthonormal sequences are often used as basis for an Inner Product Space, see for bases: section 3.2.4.

Definition 3.33

Let $(X, (.,.))$ be an Inner Product Space and $S \subset X$ is a system, with $0 \notin S$.

1. The system S is called orthogonal if for every $x, y \in S$:

$$x \neq y \Rightarrow x \perp y.$$

2. The system S is called orthonormal if the system S is orthogonal and

$$\| x \| = 1.$$

3. The system S is called an orthonormal sequence, if $S = \{x_n\}_{n \in \mathbb{I}}$, and

$$(x_n, x_m) = \delta_{nm} = \begin{cases} 0, & \text{if } n \neq m, \\ 1, & \text{if } n = m. \end{cases}$$

with mostly $\mathbb{I} = \mathbb{N}$ or $\mathbb{I} = \mathbb{Z}$.

Remark 3.8

From an orthogonal system $S = \{\, x_i \mid 0 \neq x_i \in S,\, i \in \mathbb{N} \,\}$ can simply be made an orthonormal system $S_1 = \{\, e_i = \dfrac{x_i}{\| x_i \|} \mid x_i \in S,\, i \in \mathbb{N} \,\}$. Divide the elements through by their own length.

Theorem 3.17

Orthogonal systems are linear independent systems.

Proof of Theorem 3.17

The system S is linear independent if every finite subsystem of S is linear independent. Let S be an orthogonal system. Assume that

$$\sum_{i=1}^{N} \alpha_i x_i = 0,$$

with $x_i \in S$, then $x_i \neq 0$ and $(x_i, x_j) = 0$, if $i \neq j$. Take a k, with $1 \leq k \leq N$, then

$$0 = (0, x_k) = (\sum_{i=1}^{N} \alpha_i x_i, x_k) = \alpha_k \parallel x_k \parallel^2 .$$

Hence $\alpha_k = 0$, k was arbitrary chosen, so $\alpha_k = 0$ for every $k \in \{1, \cdots, N\}$. Further N was arbitrary chosen so the system S is linear independent. $\square$

Theorem 3.18

Let $(X, (.,.))$ be an Inner Product Space.

1. Let $S = \{x_i \mid 1 \leq i \leq N\}$ be an orthogonal set in X, then

$$\parallel \sum_{i=1}^{N} x_i \parallel^2 = \sum_{i=1}^{N} \parallel x_i \parallel^2,$$

the theorem of Pythagoras.

2. Let $S = \{x_i \mid 1 \leq i \leq N\}$ be an orthonormal set in X, and $0 \notin S$ then

$$\parallel x - y \parallel = \sqrt{2}$$

for every $x \neq y$ in S.

Proof of Theorem 3.18

1. If $x_i,\, x_j \in S$ with $i \neq j$ then $(\,x_i,\, x_j) = 0$, such that

$$\| \sum_{i=1}^{N} x_i \|^2 = (\sum_{i=1}^{N} x_i, \sum_{i=1}^{N} x_i) = \sum_{i=1}^{N} \sum_{j=1}^{N} (\,x_i,\, x_j) = \sum_{i=1}^{N} (\,x_i,\, x_i) = \sum_{i=1}^{N} \| x_i \|^2 .$$

2. S is orhonormal, then for $x \neq y$

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

$x \neq 0$ and $y \neq 0$, because $0 \notin S$.

□

The following inequality can be used to give certain bounds for approximation errors or it can be used to prove the convergence of certain series. It is called the inequality of Bessel (or Bessel's inequality).

Theorem 3.19

(Inequality of Bessel) Let $\{e_n\}_{n \in \mathbb{N}}$ be an orthonormal sequence in an Inner Product Space $(X, (.,.))$, then

$$\sum_{i \in \mathbb{N}} | (x, e_i) |^2 \leq \| x \|^2,$$

for every $x \in X$. (Instead of $\mathbb{N}$ there may also be chosen another countable index set.)

Proof of Theorem 3.19

The proof exists out of several parts.

1. For arbitrary chosen complex numbers α_i holds

$$\| x - \sum_{i=1}^{N} \alpha_i e_i \|^2 = \| x \|^2 - \sum_{i=1}^{N} | (x, e_i) |^2 + \sum_{i=1}^{N} | (x, e_i) - \alpha_i |^2 . \qquad (3.17)$$

Take $\alpha_i = (x, e_i)$ and

$$\| x - \sum_{i=1}^{N} \alpha_i \, e_i \|^2 = \| x \|^2 - \sum_{i=1}^{N} | (x, e_i) |^2 .$$

2. The left-hand site of 3.17 is non-negative, so

$$\sum_{i=1}^{N} | (x, e_i) |^2 \leq \| x \|^2 .$$

3. Take the limit for $N \rightarrow \infty$. The limit exists because the series is monotone increasing and bounded above.

□

If there is given some countable linear indenpendent set of elements in an Inner Product Spaces $(X, (.,.))$, there can be constructed an orthonormal set of elements with the same span as the original set of elements. The method to construct such an orthonormal set of elements is known as the Gram-Schmidt proces. In fact is the orthogonalisation of the set of linear independent elements the most important part of the Gram-Schmidt proces, see Remark 3.8.

Theorem 3.20

Let the elements of the set $S = \{x_i \mid i \in \mathbb{N}\}$ be a linear independent set of the Inner Product Spaces $(X, (.,.))$. Then there exists an orthonormal set $ONS = \{e_i \mid i \in \mathbb{N}\}$ of the Inner Product Spaces $(X, (.,.))$, such that

$$\operatorname{span}(x_1, x_2, \cdots, x_n) = \operatorname{span}(e_1, e_2, \cdots, e_n),$$

for every $n \in \mathbb{N}$.

Proof of Theorem 3.20

Let $n \in \mathbb{N}$ be given. Let's first construct an orthogonal set of elements $OGS = \{y_i \mid i \in \mathbb{N}\}$.
The first choice is the easiest one. Let $y_1 = x_1$, $y_1 \neq 0$ because $x_1 \neq 0$ and $\operatorname{span}(x_1) = \operatorname{span}(y_1)$. The direction y_1 will not be changed anymore, the only thing that will be changed, of y_1, is it's length.
The second element y_2 has to be constructed out of y_1 and x_2. Let's take $y_2 =$

$x_2 - \alpha\, y_1$, the element y_2 has to be orthogonal to the element y_1. That means that the constant α has to be chosen such that $(y_2, y_1) = 0$, that gives

$$(y_2, y_1) = (x_2 - \alpha\, y_1, y_1) = 0 \Rightarrow \alpha = \frac{(x_2, y_1)}{(y_1, y_1)}.$$

The result is that

$$y_2 = x_2 - \frac{(x_2, y_1)}{(y_1, y_1)}\, y_1.$$

It is easy to see that

$$\mathrm{span}(y_1,\, y_2) = \mathrm{span}(x_1,\, x_2),$$

because y_1 and y_2 are linear combinations of x_1 and x_2.
Let's assume that there is constructed an orthogonal set of element $\{y_1, \cdots, y_{(n-1)}\}$, with the property $\mathrm{span}(y_1, \cdots, y_{(n-1)}) = \mathrm{span}(x_1, \cdots, x_{(n-1)})$. How to construct y_n?
The easiest way to do is to subtract from x_n a linear combination of the elements y_1 to $y_{(n-1)}$, in formula form,

$$y_n = x_n - \left(\alpha_1\, y_1 + \alpha_2\, y_2 \cdots + \alpha_{(n-1)}\, y_{(n-1)}\right),$$

such that y_n becomes perpendicular to the elements y_1 to $y_{(n-1)}$. That means that

$$\left((y_n, y_i) = 0 \Rightarrow \alpha_i = \frac{(x_n, y_i)}{(y_i, y_i)}\right) \text{ for } 1 \leq i \leq (n-1).$$

It is easily seen that y_n is a linear combination of x_n and the elements $y_1, \cdots, y_{(n-1)}$, so $\mathrm{span}(y_1, \cdots, y_n) = \mathrm{span}(y_1, \cdots, y_{(n-1)}, x_n) = \mathrm{span}(x_1, \cdots, x_{(n-1)}, x_n)$.
Since n is arbitrary chosen, this set of orthogonal elements $OGS = \{y_i \mid 1 \leq i \leq n\}$ can be constructed for every $n \in \mathbb{N}$. The set of orthonormal elements is easily constructed by $ONS = \{\frac{y_i}{\| y_i \|} = e_i \mid 1 \leq i \leq n\}$. □

3.10 Hilbert Spaces

Definition 3.34

A Hilbert space H is a complete Inner Product Space, complete in the metric induced by the inner product.

A Hilbert Space can also be seen as a Banach Space with a norm, which is induced by an inner product. Further the term pre-Hilbert space is mentioned at page 75. The next theorem makes clear why the word *pre-* is written before Hilbert. For the definition of an isometric isomorphism see page 121.

Theorem 3.21

If X is an Inner Product Space, then there exists a Hilbert Space H and an isometric isomorphism $T : X \rightarrow W$, where W is a dense subspace of H. The Hilbert Space H is unique except for isometric isomorphisms.

Proof of Theorem 3.21

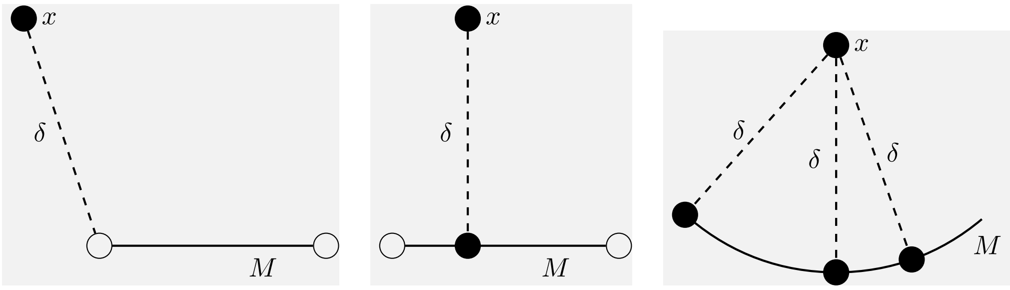
The Inner Product Space with its inner product is a Normed Space. So there exists a Banach Space H and an isometry $T : X \rightarrow W$ onto a subspace of H, which is dense in H, see theorem 3.11 and the proof of the mentioned theorem.
The problem is the inner product. But with the help of the continuity of the inner product, see theorem 3.14, there can be defined an inner product on H by

$$(\tilde{x}, \hat{y}) = \lim_{n \rightarrow \infty} (x_n, y_n)$$

for every $\tilde{x}, \tilde{y} \in H$. The sequences $\{x_n\}_{n \in \mathbb{N}}$ and $\{y_n\}_{n \in \mathbb{N}}$ represent the equivalence classes $\tilde{x}$ and $\tilde{y}$, see also theorem 3.11. The norms on X and W satisfy the parallelogram identity, see theorem 3.11, such that T becomes an isometric isomorphism between Inner Product Spaces. Theorem 3.11 guarantees that the completion is unique execept for isometric isomorphisms. □

3.10.1 Minimal distance and orthogonal projection

The definition of the distance of a point x to a set A is given in **3.21**.
Let M be subset of a Hilbert Space H and $x \in H$, then it is sometimes important to know if there exists some $y \in M$ such that $\text{dist}(x, M) = \| x - y \|$. And if there exists such a $y \in M$, the question becomes if this y is unique? See the figures **3.4** for several complications which can occur.



No solution. | An unique solution. | Infinitely many solutions.

Figure 3.4 Minimal distance δ to some subset $M \subset X$.

To avoid several of these problems it is of importance to assume that M is a closed subset of H and also that M is a convex set.

Definition 3.35

A subset A of a Vector Space X is said to be convex if

$$\alpha x + (1 - \alpha) y \in A$$

for every $x, y \in A$ and for every α with $0 \leq \alpha \leq 1$.

Any subspace of a Vector Space is obviously convex and intersections of convex subsets are also convex.

Theorem 3.22

Let X be an Inner Product Space and $M \neq \varnothing$ is a convex subset of X. M is complete in the metric induced by the inner product on X. Then for every $x \in X$, there exists an unique $y_0 \in M$ such that

$$\mathrm{dist}(x, M) = \| x - y_0 \| .$$

Proof of Theorem 3.22

Just write

$$\lambda = \mathrm{dist}(x, M) = \inf\{ d(x, y) \,|\, y \in M\},$$

then there is a sequence $\{y_n\}$ in M such that

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

If the sequence $\{y_n\}$ is a Cauchy sequence, the completeness of M can be used to prove the existence of such $y_0 \in M$(!).
Write

$$\lambda_n = \| y_n - x \|$$

so that $\lambda_n \to \lambda$, as $n \to \infty$.
The norm is induced by an inner product such that the parallelogram identity can be used in the calculation of

$$\begin{aligned} &\| y_n - y_m \|^2 = \| (y_n - x) - (y_m - x) \|^2 \\ &= 2 (\| (y_n - x) \|^2 + \| (y_m - x) \|^2) - 2 \| \frac{(y_n + y_m)}{2} - x \|^2 \\ &\leq 2 (\lambda_n^2 + \lambda_m^2) - \lambda^2, \end{aligned}$$

because $\frac{(y_n + y_m)}{2} \in M$ and $\| \frac{(y_n + y_m)}{2} - x \| \geq \lambda$.
This shows that $\{y_n\}$ is a Cauchy sequence, since $\lambda_n \to \lambda$, as $n \to \infty$. M is complete, so $y_n \to y_0 \in M$, as $n \to \infty$ and

$$\lim_{n \to \infty} \| x - y_n \| = \| x - y_0 \| = \lambda.$$

Is y_0 unique? Assume that there is some $y_1 \in M, y_1 \neq y_0$ with $\| x - y_1 \| = \lambda = \| x - y_0 \|$. The parallelogram identity is used again and also the fact that M is convex

$$\begin{aligned}
&\| y_0 - y_1 \|^2 = \| (y_0 - x) - (y_1 - x) \|^2 \\
&= 2 (\| y_0 - x \|^2 + \| y_1 - x \|^2) - \| (y_0 - x) + (y_1 - x) \|^2 \\
&= 2 (\| y_0 - x \|^2 + \| y_1 - x \|^2) - 4 \, \| \frac{(y_0 + y_1)}{2} - x \|^2 \\
&\leq 2 (\lambda^2 + \lambda^2) - 4 \lambda^2 = 0.
\end{aligned}$$

Hence $y_1 = y_0$. □

Theorem 3.23

See theorem **3.22**, but now within a real Inner Product Space. The point $y_0 \in M$ can be characterised by

$$(x - y_0, z - y_0) \leq 0$$

for every $z \in M$. The angle between $x - y_0$ and $z - y_0$ is obtuse for every $z \in M$.

Proof of Theorem 3.23

Step 1: If the inequality is valid then

$$\begin{aligned}
&\| x - y_0 \|^2 - \| x - z \|^2 \\
&= 2 (x - y_0, z - y_0) - \| z - y_0 \|^2 \leq 0.
\end{aligned}$$

Hence for every $z \in M$: $\| x - y_0 \| \leq \| x - z \|$.

Step 2: The question is if the inequality is true for the closest point y_0? Since M is convex, $\lambda z + (1 - \lambda) y_0 \in M$ for every $0 < \lambda < 1$.
About y_0 is known that

$$\| x - y_0 \|^2 \leq \| x - \lambda z - (1 - \lambda) y_0 \|^2 \quad (3.18)$$
$$= \| (x - y_0) - \lambda (z - y_0) \|^2 . \quad (3.19)$$

Because X is a real Inner Product Space, inequality 3.18 becomes

$$\| x - y_0 \|^2$$
$$\leq \| (x - y_0) \|^2 - 2 \lambda (x - y_0, z - y_0) + \lambda^2 \| z - y_0 \|^2 .$$

and this leads to the inequality

$$(x - y_0, z - y_0) \leq \frac{\lambda}{2} \| z - y_0 \|^2$$

for every $z \in M$. Take the limit of $\lambda \to 0$ and the desired result is obtained.

□

Theorem 3.23 can also be read as that it is possible to construct a hyperplane through y_0, such that x lies on a side of that plane and that M lies on the opposite site of that plane, see figure 3.5. Several possibilities of such a hyperplane are drawn.

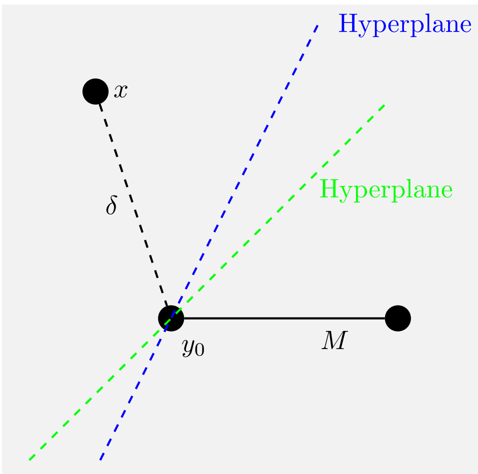


Figure 3.5 Some hyperplanes through y_0.

If there is only an unique hyperplane than the direction of $(x - y_0)$ is perpendicular to that plane, see figure 3.6.

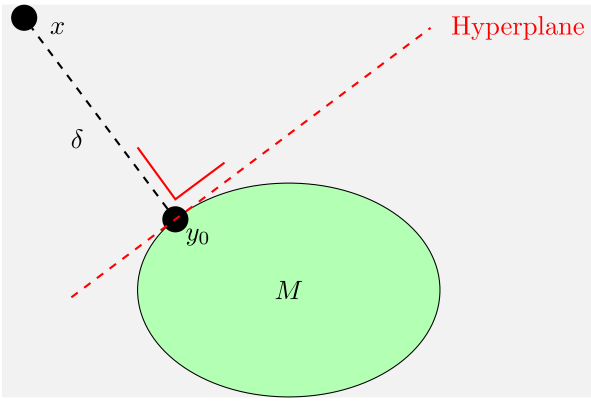


Figure 3.6 Unique hyperplane through y_0.

Given a fixed point x and certain plane M, the shortest distance of x to the plane is found by dropping a perpendicular line through x on M. With the point of intersection of this perpendicular line with M and the point x, the shortest distance can be calculated. The next theorem generalizes the above mentioned fact. Read theorem 3.22 very well, there is spoken about a non-empty convex subset, in the next theorem is spoken about a linear subspace.

Theorem 3.24

See theorem 3.22, but now with M a complete subspace of X, then $z = x - y_0$ is orthogonal to M.

Proof of Theorem 3.24

A subspace is convex, that is easy to verify. So theorem 3.22 gives the existence of an element $y_0 \in M$, such that $\mathrm{dist}(x, M) = \| x - y_0 \| = \delta$.
If $z = x - y_0$ is not orthogonal to M then there exists an element $y_1 \in M$ such that

$$(z, y_1) = \beta \neq 0. \tag{3.20}$$

It is clear that $y_1 \neq 0$ otherwise $(z, y_1) = 0$. For any γ

$$\begin{aligned} \| z - \gamma y_1 \|^2 &= (z - \gamma y_1, z - \gamma y_1) \\ &= (z, z) - \overline{\gamma}(z, y_1) - \gamma (y_1, z) + | \gamma |^2 (y_1, y_1) \end{aligned}$$

If $\overline{\gamma}$ is chosen equal to

$$\overline{\gamma} = \frac{\overline{\beta}}{\| y_1 \|^2}$$

then

$$\| z - \gamma y_1 \|^2 = \| z \|^2 - \frac{| \beta |^2}{\| y_1 \|^2} < \delta^2.$$

This means that $\| z - \gamma y_1 \| = \| x - y_0 - \gamma y_1 \| < \delta$, but by definition $\| z - \gamma y_1 \| > \delta$, if $\gamma \neq 0$.

Hence 3.20 cannot hold, so $z = x - y_0$ is orthogonal to M. □

From theorem 3.24, it is easily seen that $x = y_0 + z$ with $y_0 \in M$ and $z \in M^{\perp}$. In a Hilbert Space this representation is very important and useful.

Theorem 3.25

If M is closed subspace of a Hilbert Space H. Then

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

Proof of Theorem 3.25

Since M is a closed subspace of H, M is also a complete subspace of H, see theorem 3.7. Let $x \in H$, theorem 3.24 gives the existence of a $y_0 \in M$ and a $z \in M^{\perp}$ such that $x = y_0 + z$.
Assume that $x = y_0 + z = y_1 + z_1$ with $y_0, y_1 \in M$ and $z, z_1 \in M^{\perp}$. Then $y_0 - y_1 = z - z_1$, since $M \cap M^{\perp} = \{0\}$ this implies that $y_1 = y_0$ and $z = z_1$. So y_0 and z are unique. □

In section 3.7.1 is spoken about total subset M of a Normed Space X, i.e. $\overline{\text{span}(M)} = X$. How to characterize such a set in a Hilbert Space H?

Theorem 3.26

Let M be a non-empty subset of a Hilbert Space H.
M is total in H *if and only if* $x \perp M \Longrightarrow x = 0$ (or $M^{\perp} = \{0\}$).

Proof of Theorem 3.26

($\Rightarrow$) Take $x \in M^{\perp}$. M is total in H, so $\overline{\text{span}(M)} = H$. This means that for $x \in H (M^{\perp} \subset H)$, there exists a sequence $\{x_n\}_{n \in \mathbb{N}}$ in $\text{span}(M)$ such that $x_n \rightarrow x$. Since $x \in M^{\perp}$ and $M^{\perp} \perp \text{span}(M)$, $(x_n, x) = 0$. The continuity of the inner product implies that $(x_n, x) \rightarrow (x, x)$, so $(x, x) = \| x \|^2 = 0$ and this means that $x = 0$. $x \in M^{\perp}$ was arbitrary chosen, hence $M^{\perp} = \{0\}$.

($\Leftarrow$) Given is that $M^{\perp} = \{0\}$. If $x \perp \text{span}(M)$ then $x \in M^{\perp}$ and $x = 0$. Hence $\text{span}(M)^{\perp} = \{0\}$. The $\text{span}(M)$ is a subspace of H. With theorem 3.25 is obtained that $\overline{\text{span}(M)} = H$, so M is total in H.

□

Remark 3.9

In Inner Product Spaces theorem 3.26 is true from right to the left. If X is an Inner Product Space then: "If M is total in X *then* $x \perp M \Longrightarrow x = 0$."
The completeness of the Inner Product Space X is of importance for the opposite!

Lemma 3.1

If S is a subspace of a Hilbert space H then $\overline{S} = S^{\perp\perp}$,
so $S^{\perp\perp}$ is the smallest closed subspace containing S.

Proof of Theorem 3.1

If $x \in S$ then $x \perp y$ for all $y \in S^{\perp}$, so $x \in S^{\perp\perp}$ and therefore $S \subseteq S^{\perp\perp}$.
Since $(S^{\perp})^{\perp}$ is closed, so is obtained one direction of the containment

$$\overline{S} \subseteq S^{\perp\perp}.$$

Suppose that $S^{\perp\perp}$ is strictly larger than $\overline{S}$. Then there is some $y \in S^{\perp\perp}$ not lying in $\overline{S}$. $S^{\perp\perp}$ is a Hilbert space in its own and $\overline{S}$ is a closed subset, so the orthogonal complement of $\overline{S}$ in $S^{\perp\perp}$ contains an element $z \neq 0$. But then $z \in S^{\perp}$ and $z \in S^{\perp\perp}$, this contradicts the fact that

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

See for the comment about the smallest closed subspace, Theorem 2.4. □

3.10.2 Orthogonal base, Fourier expansion and Parseval's relation

The main problem will be to show that sums can be defined in a reasonable way. It should be nice to prove that orhogonal bases of H are countable.

Definition 3.36

An orthogonal set M of a Hilbert Space H is called an orthogonal base of H, if no orthonormal set of H contains M as a proper subset.

Remark 3.10

An orthogonal base is sometimes also called a complete orthogonal system . Be careful, the word "complete" has <u>nothing</u> to do with the topological concept: completeness.

Theorem 3.27

A Hilbert Space $H\,(0 \neq x \in H)$ has at least one orthonormal base. If M is any orthogonal set in H, there exists an orthonormal base containing M as subset.

Proof of Theorem 3.27

There exists a $x \neq 0$ in H. The set, which contains only $\frac{x}{\| x \|}$ is orthonormal. So there exists an orthonormal set in H.
Look to the totality $\{S\}$ of orthonormal sets which contain M as subset. $\{S\}$ is partially ordered. The partial order is written by $S_1 \prec S_2$ what means that $S_1 \subseteq S_2$. $\{S^{'}\}$ is the linear ordered subset of $\{S\}$. $\cup_{S^{'} \in \{S^{'}\}}$ is an orthonormal set and an upper bound of $\{S^{'}\}$. Thus by Zorn's Lemma, there exists a maximal element S_0 of $\{S\}$. $S \subseteq S_0$ and because of it's maximality, S_0 is an orthogonal base of H. □

There exists an orthonormal base S_0 of a Hilbert Space H. This orthogonal base S_0 can be used to represent elements $f \in H$, the so-called Fourier expansion of f. With the help of the Fourier expansion the norm of an element $f \in H$ can be calculated by Parseval's relation .

Theorem 3.28

Let $S_0 = \{e_\alpha \mid \alpha \in \Lambda\}$ be an orthonormal base of a Hilbert Space H. For any $f \in H$ the Fourier-coefficients, with respect to S_0, are defined by

$$f_\alpha = (f, e_\alpha)$$

and

$$f = \sum_{\alpha \in \Lambda} f_\alpha \, e_\alpha,$$

which is called the Fourier expansion of f. Further

$$\| f \|^2 = \sum_{\alpha \in \Lambda} | f_\alpha |^2,$$

for any $f \in H$, which is called Parseval's relation.

Proof of Theorem 3.28

The proof is splitted up into several steps.

1. First will be proved the inequality of Bessel. In the proof given in theorem 3.19 there was given a countable orthonormal sequence. Here is given an orthonormal base S_0. If this base is countable, that is till this moment not known. Let's take a finite system $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ out of Λ.
For arbitrary chosen complex numbers c_{α_i} holds

$$\| f - \sum_{i=1}^{n} c_{\alpha_i} e_{\alpha_i} \|^2 = \| f \|^2 - \sum_{i=1}^{n} | (f, e_{\alpha_i}) |^2 + \sum_{i=1}^{n} | (f, e_{\alpha_i}) - c_{\alpha_i} |^2 . \tag{3.21}$$

For fixed $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$, the minimum of $\| f - \sum_{i=1}^{n} c_{\alpha_i} e_{\alpha_i} \|^2$ is attained when $c_{\alpha_i} = f_{\alpha_i}$. Hence

$$\sum_{i=1}^{n} | f_{\alpha_i} |^2 \leq \| f \|^2$$

2. Define

$$E_j = \{e_\alpha \mid | (f, e_\alpha) | \geq \frac{\| f \|}{j}, \, e_\alpha \in S_0\}$$

for $j = 1, 2, \cdots$. Suppose that E_j contains the distinct elements $\{e_{\alpha_1}, e_{\alpha_1}, \cdots, e_{\alpha_m}\}$ then by Bessel's inequality,

$$\sum_{i=1}^{m} (\frac{\| f \|}{j})^2 \leq \sum_{\alpha_i} | (f, e_{\alpha_i}) |^2 \leq \| f \|^2 .$$

This shows that $m \leq j^2$, so E_j contains at most j^2 elements.
Let

$$E_f = \{e_\alpha \mid (f, e_\alpha) \neq 0, \, e_\alpha \in S_0\}.$$

E_f is the union of all E_j's, $j = 1, 2, \cdots$, so E_f is a countable set.

3. Also if E_f is denumerable then

$$\sum_{i=1}^{\infty} | f_{\alpha_i} |^2 \leq \| f \|^2 < \infty,$$

such that the term $f_{\alpha_i} = (f, e_{\alpha_i})$ of that convergent series tends to zero if $i \rightarrow \infty$.
Also important to mention

$$\sum_{\alpha \in \Lambda} \mid f_{\alpha} \mid^2 = \sum_{i=1}^{\infty} \mid f_{\alpha_i} \mid^2 \leq \parallel f \parallel^2 < \infty,$$

so Bessel's inequality is true.

4. The sequence $\{\sum_{i=1}^{n} f_{\alpha_i} e_{\alpha_i}\}_{n \in \mathbb{N}}$ is a Cauchy sequence, since, using the orthonormality of $\{e_{\alpha}\}$,

$$\parallel \sum_{i=1}^{n} f_{\alpha_i} e_{\alpha_i} - \sum_{i=1}^{m} f_{\alpha_i} e_{\alpha_i} \parallel^2 = \sum_{i=m+1}^{n} \mid f_{\alpha_i} \mid^2$$

which tends to zero if $n, m \rightarrow \infty$, $(n > m)$. The Cauchy sequence converges in the Hilbert Space H, so $\lim_{n \rightarrow \infty} \sum_{i=1}^{n} f_{\alpha_i} e_{\alpha_i} = g \in H$.
By the continuity of the inner product

$$(f - g, e_{\alpha_k}) = \lim_{n \rightarrow \infty} (f - \sum_{i=1}^{n} f_{\alpha_i} e_{\alpha_i}, e_{\alpha_k}) = f_{\alpha_k} - f_{\alpha_k} = 0,$$

and when $\alpha \neq \alpha_j$ with $j = 1, 2, \cdots$ then

$$(f - g, e_{\alpha}) = \lim_{n \rightarrow \infty} (f - \sum_{i=1}^{n} f_{\alpha_i} e_{\alpha_i}, e_{\alpha}) = 0 - 0 = 0.$$

The system S_0 is an orthonormal base of H, so $(f - g) = 0$.

5. By the continuity of the norm and formula 3.21 follows that

$$0 = \lim_{n \rightarrow \infty} \parallel f - \sum_{i=1}^{n} f_{\alpha_i} e_{\alpha_i} \parallel^2 = \parallel f \parallel^2 - \lim_{n \rightarrow \infty} \sum_{i=1}^{n} \mid f_{\alpha_i} \mid^2 = \parallel f \parallel^2 - \sum_{\alpha \in \Lambda} \mid f_{\alpha} \mid^2 .$$

□

3.10.3 Representation of bounded linear functionals

In the chapter about Dual Spaces, see chapter 4, there is something written about the representation of bounded linear functionals. Linear functionals are in certain sense nothing else then linear operators on a vectorspace and their range lies in the field $\mathbb{K}$ with $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. About their representation is also spoken, for the finite dimensional case, see 4.4.1 and for the vectorspace ℓ^1 see 4.6.1. The essence is that these linear functionals can be represented by an inner product. The same can be done for bounded linear functionals on a Hilbert Space H.

Remark 3.11

Be careful:
The ℓ^1 space is not an Inner Product space, the representation can be read as an inner product.

The representation theorem of Riesz (functionals).

Theorem 3.29

Let H be a Hilbert Space and f is a bounded linear functional on H, so $f : H \rightarrow \mathbb{K}$ and there is some $M > 0$ such that $\mid f(x) \mid \leq M \parallel x \parallel$ then there is an unique $a \in H$ such that

$$f(x) = (x, a)$$

for every $x \in H$ and

$$\parallel f \parallel = \parallel a \parallel .$$

Proof of Theorem 3.29

The proof is splitted up in several steps.

1. First the existence of such an $a \in H$.
If $f = 0$ then satisfies $a = 0$. Assume that there is some $z \neq 0$ such that $f(z) \neq 0$, $(z \in H)$. The nullspace of f, $N(f) = \{x \in H | f(x) = 0\}$ is a closed linear subspace of H, hence $N(f) \oplus N(f)^{\perp} = H$. So z can be written as $z = z_0 + z_1$ with $z_0 \in N(f)$ and $z_1 \in N(f)^{\perp}$ and $z_1 \neq 0$. Take now $x \in H$ and write x as follows $x = (x - \dfrac{f(x)}{f(z_1)} z_1) + \dfrac{f(x)}{f(z_1)} z_1 = x_0 + x_1$. It is easily to check that $f(x_0) = 0$, so $x_1 \in N(f)^{\perp}$ and that means that

$(x - \frac{f(x)}{f(z_1)} z_1) \perp z_1$. Hence, $(x, z_1) = \frac{f(x)}{f(z_1)} (z_1, z_1) = f(x) \frac{\| z_1 \|^2}{f(z_1)}$. Take $a = \frac{\overline{f(z_1)}}{\| z_1 \|^2} z_1$ and for every $x \in H : f(x) = (x, a)$.

2. Is a unique?
If there is some $b \in H$ such that $(x, b) = (x, a)$ for every $x \in H$ then $(x, b-a) = 0$ for every $x \in H$. Take $x = b - a$ then $\| b - a \|^2 = 0$ then $(b - a) = 0$, hence $b = a$.

3. The norm of f?
Using Cauchy-Schwarz gives $|f(x)| = |(x, a)| \leq \| x \| \| a \|$, so $\| f \| \leq \| a \|$. Further $f(a) = \| a \|^2$, there is no other possibility then $\| f \| = \| a \|$.

□

3.10.4 Representation of bounded sesquilinear forms

In the paragraphs before is, without knowing it, already worked with sesquilinear forms , because inner products are sesquilinear forms. Sesquilinear forms are also called sesquilinear functionals .

Definition 3.37

Let X and Y be two Vector Spaces over the same field $\mathbb{K}$. A mapping

$$h : X \times Y \to \mathbb{K}$$

is called a sesquilinear form, if for all $x_1, x_2 \in X$ and $y_1, y_2 \in Y$ and $\alpha \in \mathbb{K}$

SQL 1: $h(x_1 + x_2, y_1) = h(x_1, y_1) + h(x_2, y_1)$,

SQL 2: $h(x_1, y_1 + y_2) = h(x_1, y_1) + h(x_1, y_2)$,

SQL 3: $h(\alpha\, x_1, y_1) = \alpha\, h(x_1, y_1)$,

SQL 4: $h(x_1, \alpha\, y_1) = \overline{\alpha}\, h(x_1, y_1)$.

In short h is *linear* it the first argument and *conjugate linear* in the second argument.

Inner products are bounded sesquilinear forms. The definition of the norm of a sesquilinear form is almost the same as the definition of the norm of a linear functional or a linear operator.

Definition 3.38

If X and Y are Normed Spaces, the sesquilinear form is *bounded* if there exists some positive number $c \in \mathbb{R}$ such that

$$|h(x, y)| \leq c \parallel x \parallel \parallel y \parallel$$

for all $x \in X$ and $y \in Y$.
The norm of h is defined by

$$\parallel h \parallel = \sup_{\begin{cases} 0 \neq x \in X, \\ 0 \neq y \in Y \end{cases}} \frac{|h(x, y)|}{\parallel x \parallel \parallel y \parallel} = \sup_{\begin{cases} \parallel x \parallel = 1, \\ \parallel y \parallel = 1 \end{cases}} |h(x, y)|.$$

When the Normed Spaces X and Y are Hilbert Spaces then the representation of a sesquilinear form can be done by an inner product and the help of a bounded linear operator, the so-called Riesz representation .

Theorem 3.30

Let H_1 and H_2 be Hibert Spaces over the field $\mathbb{K}$ and

$$h : H_1 \times H_2 \to \mathbb{K}$$

is a bounded sesquilinear form. Let $(\cdot, \cdot)_{H_1}$ be the inner product in H_1 and let $(\cdot, \cdot)_{H_2}$ be the inner product in H_2. Then h has a representation

$$h(x, y) = (S(x), y)_{H_2}$$

where $S : H_1 \to H_2$ is a uniquely determined bounded linear operator and

$$\| S \| = \| h \| .$$

Proof of Theorem 3.30

The proof is splitted up in several steps.

1. The inner product?
   Let $x \in H_1$ be fixed and look at $\overline{h(x, y)}$. $\overline{h(x, y)}$ is linear in y because there is taken the complex conjugate of $h(x, y)$. Then using Theorem **3.29** gives the existence of an unique $z \in H_2$, such that

$$\overline{h(x, y)} = (y, z)_{H_2},$$

   therefore

$$h(x, y) = (z, y)_{H_2}. \tag{3.22}$$

2. The operator S?
   $z \in H_2$ is unique, but depends on the fixed $x \in H_1$, so equation **3.22** defines an operator $S : H_1 \to H_2$ given by

$$z = S(x).$$

3. Is S linear?
   For $x_1, x_2 \in H_1$ and $\alpha \in \mathbb{K}$:

$$\begin{aligned}(S(x_1 + x_2), y)_{H_2} &= h((x_1 + x_2), y) = h(x_1, y) + h(x_2, y) \\ &= (S(x_1), y)_{H_2} + (S(x_2), y)_{H_2} = ((S(x_1) + S(x_2)), y)_{H_2}\end{aligned}$$

for every $y \in H_2$. Hence, $S(x_1 + x_2) = S(x_1) + S(x_2)$.
On the same way, using the linearity in the first argument of h:
$S(\alpha x_1) = \alpha S(x_1)$.

4. Is S bounded?

$$\| h \| = \sup_{\begin{cases} 0 \neq x \in H_1, \\ 0 \neq y \in H_2 \end{cases}} \frac{(S(x), y)_{H_2}}{\| x \|_{H_1} \| y \|_{H_2}}$$

$$\geq \sup_{\begin{cases} 0 \neq x \in H_1, \\ 0 \neq S(x) \in H_2 \end{cases}} \frac{(S(x), S(x))_{H_2}}{\| x \|_{H_1} \| S(x) \|_{H_2}} = \| S \|,$$

so the linear operator S is bounded.

5. The norm of S?

$$\| h \| = \sup_{\begin{cases} 0 \neq x \in H_1, \\ 0 \neq y \in H_2 \end{cases}} \frac{(S(x), y)_{H_2}}{\| x \|_{H_1} \| y \|_{H_2}}$$

$$\leq \sup_{\begin{cases} 0 \neq x \in H_1, \\ 0 \neq y \in H_2 \end{cases}} \frac{\| S(x) \|_{H_2} \| y \|_{H_2}}{\| x \|_{H_1} \| y \|_{H_2}} = \| S \|$$

using the Cauchy-Schwarz-inequality. Hence, $\| S \| = \| T \|$.

6. Is S unique?
If there is another linear operator $T : H_1 \to H_2$ such that

$$h(x, y) = (T(x), y)_{H_2} = (S(x), y)_{H_2}$$

for every $x \in H_1$ and $y \in H_2$, then

$$(T(x) - S(x), y) = 0$$

for every $x \in H_1$ and $y \in H_2$. Hence, $T(x) = S(x)$ for every $x \in H_1$, so $S = T$.

3.11 Quotient Spaces

See Section 3.2.3 for the definition of a Quotient Space and its linear operations. The book of (Megginson, 1998) is used to the following overview of the properties of Quotient Spaces.
Suppose that W is a linear subspace of a normed space $(V, \| \cdot \|)$. With a norm there can be easily defined a metric, see formula 3.3.
With the help of that metric, there can be defined a distance between cosets and with that distance function, there will be defined a norm at the Quotient Space V/W.

3.11.1 Metric and Norm on V/W

The natural way to define a distance between the cosets $x + W$ and $y + W$, is to think as if the cosets were sets. The distance between the sets $x + W$ and $y + W$ in a Metric Space is defined by

$$d(x + W, y + W) = \inf\{\| s - t \| \mid s \in x + W, t \in y + W\}, \tag{3.23}$$

used is definition 3.21. Since

$$\{s - t \mid s \in x + W, t \in y + W\} = \{(x + z_1) - (y + z_2) \mid z_1, z_2 \in W\} =$$
$$\{x - (y - z_1 + z_2) \mid z_1, z_2 \in W\} = \{x - (y + z) \mid z \in W\} =$$
$$\{x - w \mid w \in y + W\}$$

$d(x + W, y + W) = d(x, y + W)$, whenever $x, y \in V$.
If $x \in \overline{W} \setminus W$ then $0 \leq d(x + W, 0 + W) = d(x, W) = 0$, but $x + W \neq 0 + M$, so formula 3.23 is not a metric at the Quotient Space V/W. But if W is closed, there are no problems anymore, because $\overline{W} \setminus W = \emptyset$.

If W is closed, formula 3.23 defines a (quotient) metric on the Quotient Space V/W. That will also be the reason that most of the time the linear space W is assumed to be closed.
Since the metric is induced by a norm, it also possible to define a norm at the Quotient Space V/W, that will be the distance of a coset to the origin of V/W.

Definition 3.39

Let W be a closed linear subspace of the Normed Space $(V, \| \cdot \|)$. The Quotient Norm of the Quotient Space V/W is given by

$$\| x + W \| = d(x + W, 0 + W) = \inf\{\| x + y \| \mid y \in W\}. \tag{3.24}$$

Let's look, if the conditions in definition 3.23 are satisfied.
It's clear that $\| x + W \| \geq 0$, so condition 1 is satisfied.
If $\| x + W \| = 0$, there exists a sequence $\{w_n\}_{n \in \mathbb{N}} \subset W$, such that $x + w_n \to 0$, as $n \to \infty$, so $w_n \to (-x)$. W is closed, so that $(-x) \in W$ and that means that $x + W = x + (-x) + W = 0 + W$ in V/W and condition 2 is fulfilled.
Let's now look at condition 3, for $x \in V$ and $0 \neq k \in \mathbb{K}$,

$$\begin{aligned}
&\| k\,(x + W) \| = \| k\,x + W \| = \inf\{\| k\,x + y \| \mid y \in W\} = \\
&\inf\{| k \| \| x + \frac{y}{k} \| \mid y \in W\} = | k | \inf\{\| x + y \| \mid y \in W\} = \\
&| k \| \| x + W \| .
\end{aligned}$$

And now the triangle-inequality, let $x_1, x_2 \in V$. The infimum is the greatest lower bound

$$\begin{aligned}
&\| (x_1 + W) + (x_2 + W) \| = \| (x_1 + x_2) + W \| = \\
&\inf\{\| (x_1 + x_2) + y \| \mid y \in W\} \\
&= \inf\{\| (x_1 + y_1) + (x_2 + y_2) \| \mid y_1, y_2 \in W\} \leq \\
&\inf\{(\| (x_1 + y_1) \| + \| (x_2 + y_2) \|) \mid y_1, y_2 \in W\} = \\
&\| x_1 + W \| + \| x_2 + W \|
\end{aligned}$$

Since $x_1, x_2 \in V$ were arbitrary chosen, so condition 4 is also fulfilled. It follows that the expression 3.24 is a norm on the Quotient Space V/W.

Theorem 3.31

Let W be a closed subspace of the Normed Space $(V, \| \cdot \|)$.

a. If $x \in X$ then $\| x \| \geq \| x + W \|$.
b. If $x \in X$ and $\epsilon > 0$, then there exists an $x_0 \in V$ such that $x_0 + W = x + W$ and $\| x_0 \| < \| x + W \| + \epsilon$.

Proof of Theorem 3.31

a. $\| x \| = \| x - 0 \| \geq d(x, 0 + W) = \| x + W \|$.
b. Suppose that $x \in V$ and $\epsilon > 0$. There holds $d(x, W) \leq \| x - y \|$ for every $y \in W$. Let y be an element of W such that $\| x - y \| < d(x, W) + \epsilon = \| x + W \| + \epsilon$.
So take $x_0 = x - y$.

□

To do certain estimations, it is of importance to know about the existence of certain elements in some subspace. Let $(V, \| \cdot \|)$ be a Normed Space and W a closed subspace of V. Suppose that $x, y \in V$ and $\| x - y + W \| < \delta$ then there exists a sequence $\{z_i\}_{i \in \mathbb{N}} \subset W$ such that

$$\lim_{i \to \infty} \| x - y + z_i \| = \| x - y + W \| = \inf\{\| x - y + z \| \mid z \in W\}$$

and $\| x - y + W \| \leq \| x - y + z_i \|$ for every $i \in \mathbb{N}$. If $\| x - y + W \| < \delta$ then there is some $z_{i_0} (\in W)$ such that $\| x - y + z_{i_0} \| < \delta$ and $(x - y) + z_{i_0} + W = (x - y) + W$. So there exists some $y_0 = y - z_{i_0} \in V$ such that $x - y_{i_0} + W = x - y + W$ and $\| x - y_{i_0} \| < \delta$.

Theorem 3.32

Let W be a closed subspace of the Normed Space $(V, \| \cdot \|)$.
The sequence $\{x_n + W\}_{n \in \mathbb{N}}$ converges to $x + W$ in V/W if and only if there is a sequence $\{y_n\}_{n \in \mathbb{N}} \subset W$ such that the sequence $\{(x_n + y_n)\}_{n \in \mathbb{N}}$ converges to $x \in V$.

Proof of Theorem 3.32

(⇒) Assume that $(x_n + W) \rightarrow (x + W)$ in V/W. Since

$$\| (x_n + W) - (x + W) \| = \inf\{\| x_n - x + y \| \mid y \in W\},$$

choose $y_n \in W$ such that

$$\| x_n - x + y_n \| < \| (x_n + W) - (x + W) \| + \frac{1}{n}$$

for $n = 1, 2, \cdots$, see **Theorem 3.31 b**. There follows that $(x_n - x + y_n) \rightarrow 0$ in V, so $(x_n + y_n) \rightarrow x$ in V, as $n \rightarrow \infty$.

(⇐) Let $\{x_n + W\}_{n \in \mathbb{N}}$ be a sequence in V/W. If $\{y_n\}_{n \in \mathbb{N}}$ is a sequence in W such that $(x_n + y_n) \rightarrow x$ in V, then

$$\| (x_n + W) - (x + W) \| = \| (x_n - x) + W \| \leq \| (x_n - x + y_n) \|$$

for every n, so that $(x_n + W) \rightarrow (x + W)$ in V/W.

□

3.11.2 Completeness is three-space property

It would be nice to deduce a property of some space on the basis of some other facts that are known about that space. If W is a closed subspace of a normed space $(V, \| \cdot \|)$ and there is known that the quotient space V/W is complete. Does this fact imply that the space $(V, \| \cdot \|)$ is complete or not? Here it is a question about completeness, but there are also other properties of spaces, where this question can be asked.

Definition 3.40

Let P be a property defined for Normed Spaces. Suppose that $(V, \| \cdot \|)$ is a Normed Space with a closed subspace W such that two of the spaces V, W, V/W have the property P, then the third must also have it. Then P is called a three-space property.

Theorem 3.33

If W is a closed subspace of a Normed Space $(V, \| \cdot \|)$.
Completeness is a three-space property.
The normed space $(V, \| \cdot \|)$ is complete if and only if W and V/W are complete.

Proof of Theorem 3.33

($\Rightarrow$) The normed space $(V, \| \cdot \|)$ is complete, so it is a Banach Space. So W is a closed linear subspace of the Banach Space $(V, \| \cdot \|)$, so W is a Banach Space, see Theorem 3.12.
The Quotient Space V/W is a Normed linear Space, see Definition 3.39.
Suppose that $\{(x_n + W\}_{n \in \mathbb{N}}$ is a Cauchy sequence in V/W. If some subsequence of $\{(x_n + W)\}_{n \in \mathbb{N}}$ has a limit, the entire sequence will converge to the same limit. (Idea: $(x_n - x) = (x_n - x_{n_k}) + (x_{n_k} - x)$.)
There is a subsequence $\{(x_{n_k} + W)\}_{k \in \mathbb{N}}$ with $\| (x_{n_k} - x_{n_{k+1}}) + W \| < 2^{-k}$.
Hence there exists a sequence $\{y_k\}_{k \in \mathbb{N}} \subset W$ such that $\| (x_{n_k} - x_{n_{k+1}} - y_k) \| < 2^{-k}$. Write $y_k = w_{k+1} - w_k$ with $w_1 = 0$ and $w_k \in W$, $k = 2, 3, \cdots$. So the sequence $\{(x_{n_k} - w_k)\}_{k \in \mathbb{N}}$ is a Cauchy sequence in V, since V is complete, it converges to a limit $x \in V$. With Theorem 3.32 follows that $(x_{n_k} + W) \rightarrow (x + W)$, hence the Quotient Space V/W is complete.

($\Leftarrow$) Suppose that W and V/W are complete. Let $\{x_n\}_{n \in \mathbb{N}}$ be Cauchy sequence in V. Since $\| (x_n - x_m) + W \| \leq \| x_n - x_m \|$ for all $n, m \in \mathbb{N}$, the sequence $\{(x_n + W)\}_{n \in \mathbb{N}}$ is a Cauchy sequence in V/W and so converges to some $(z + W) \in V/W$.
With Theorem 3.31 b follows that there exists a sequence $\{y_n\}_{n \in \mathbb{N}} \subset W$ such that $\| (x_n + y_n) - z \| \rightarrow 0$ in V.

Since $y_n - y_m = y_n + x_n - z - x_n + x_m - x_m - y_m + z$ and $\| y_n - y_m \| \leq \| y_n + x_n - z \| + \| -x_n + x_m \| + \| -x_m - y_m + z \|$ for $n, m = 1, 2, \cdots$, if follows that the sequence $\{y_n\}_{n \in \mathbb{N}}$ is a Cauchy sequence in W. W is complete, so $y_n \to y$ in W and $x_n = (x_n + y_n) - y_n \to z - y$ in V. This shows that V is complete.

□

3.11.3 Quotient Map

Let $(V, \| \cdot \|)$ be a Normed Space and W a closed subspace of V. The equivalence classes are the members of the Quotient Space V/W. Quite often, use is made of a *projection* from V onto V/W.

Definition 3.41

Let $(V, \| \cdot \|)$ be a Normed Space and W a closed subspace of V. The quotient map from V onto V/W is the function π defined by the formula
$\pi(x) = x + W$. $(\pi : V \to V/W)$
The addition and scalar multiplication are defined by

$$\pi(x + y) = \pi(x) + \pi(y), \quad \pi(\alpha\, x) = \alpha\, \pi(x),$$

with $x, y \in V$ and $\alpha \in \mathbb{K}$.
Be careful: if $\alpha = 0$ then $\alpha\, \pi(x) = 0 + W$.

Lemma 3.2

If W is a closed subspace of a normed space $(V, \| \cdot \|)$ and π is the quotient map of V onto V/W then the image of the open unit ball in V is the open unit ball of V/M.

Proof of Theorem 3.2

($\pi(U_V) \subseteq U_{V/W}$):
Let U_V be the open unit ball of V and let $U_{V/W}$ be the open unit ball of V/W. If $x \in U_V$ then $\| \pi(x) \| = \| x + W \| \leq \| x \| < 1$, so . Here is used Theorem 3.31 a.

($U_{V/W} \subseteq \pi(U_V)$):
If $y + W \in U_{V/W}$ then $\| y + W \| < 1$, so there exists some $\epsilon > 0$, such that $\| y + W \| + \epsilon < 1$. Theorem 3.31 b gives that there exists a $z \in V$, such that $z + W = y + W$ and such that $\| z \| < \| y + W \| + \epsilon < 1$.

This proves $\pi(U_V) = U_{V/W}$.
□

3.11.4 Important Construction: Completion

If (X, d) is a Metric Space, which is not complete, then it is always possible to construct a larger space, which is complete. This larger space contains just enough elements such that every Cauchy sequence in X has a limit in that larger space. It is an important construction, which is often used.
New points are adjoined to the space (X, d) and d has to be extended to all these new points. And Cauchy sequences, which had first no limit, find a limit among those new points. Those new points become limits of sequences in X.

Definition 3.42

Let (X, d) be a metric space. The set of all the Cauchy sequences with respect to the metric d is defined by

$$\mathrm{cs}(X, d) = \{\mathbf{x} = \{x_n\}_{n \in \mathbb{N}} \mid \mathbf{x} \,\text{Cauchy sequence in}\, X\}.$$

Cauchy sequences $\mathbf{x} = \{x_n\}_{n \in \mathbb{N}}$ and $\mathbf{y} = \{y_n\}_{n \in \mathbb{N}}$ are said to be equivalent if

$$\lim_{n\to\infty} d(x_n, y_n) = 0$$

and then is written $\mathbf{x} \sim \mathbf{y}$. It is fairly obvious that $\sim$ is indeed an equivalence relation, see section 2.14.

Reflexivity: $[x_n] \sim [x_n]$, since $d(x_n, x_n) = 0$ for every n and so $\lim_{n\to\infty} d(x_n, x_n) = 0$.

Symmetry: If $[x_n] \sim [y_n]$, then $\lim_{n\to\infty} d(x_n, y_n) = 0$ and since $d(x_n, y_n) = d(y_n, x_n)$ for every n, $\lim_{n\to\infty} d(y_n, x_n) = 0$, so that $[y_n] \sim [x_n]$.

Transitivity: If $[x_n] \sim [y_n]$ and $[y_n] \sim [z_n]$ then $\lim_{n\to\infty} d(x_n, y_n) = 0$ and $\lim_{n\to\infty} d(y_n, z_n) = 0$. Since $0 \leq d(x_n, z_n) \leq d(x_n, y_n) + d(y_n, z_n)$ for all n, it follows that $0 \leq \lim_{n\to\infty} d(x_n, z_n) \leq \lim_{n\to\infty} d(x_n, y_n) + \lim_{n\to\infty} d(y_n, z_n) = 0$, so that $[x_n] \sim [z_n]$.

With $\mathrm{cs}(X, d)$ and $\sim$, there is defined the quotient space

$$\widetilde{X} = \mathrm{cs}(X, d)/\sim .$$

For an element $\mathbf{x} \in \widetilde{X}$, its equivalence class is denoted by $\widetilde{\mathbf{x}}$.

For a point $x \in X$, there is defined $\langle x \rangle \in \widetilde{X}$, to be the equivalence class of the constant sequence x. So $\langle x \rangle = \{x, x, x, \cdots\}$, which of course is a Cauchy sequence.

Remark 3.12

Let (X, d) be a Metric Space and let $\mathbf{x} = \{x_n\}_{n\in\mathbb{N}}$ and $\mathbf{y} = \{y_n\}_{n\in\mathbb{N}}$ be Cauchy sequences in X. The sequence of real numbers $\{d(x_n, y_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence, since for any n, m:

$$\begin{aligned}
&\mid d(x_m, y_m) - d(x_n, y_n) \mid \leq \\
&\mid d(x_m, y_m) - d(x_n, y_m) \mid + \mid d(x_n, y_m) - d(x_n, y_n) \mid \leq \\
&d(x_m, x_n) + d(y_m, y_n).
\end{aligned}$$

Every Cauchy sequence in $\mathbb{R}$ is convergent, so the sequence of numbers $(d(x_n, y_n)_{n\in\mathbb{N}}$ converges. This can be used to define a metric at $\widetilde{X}$.

Define the map $\delta : \mathrm{cs}(X, d) \times \mathrm{cs}(X, d) \to [0, \infty)$ by

$$\delta(\mathbf{x}, \mathbf{y}) = \lim_{n\to\infty} d(x_n, y_n).$$

Theorem 3.34

Let (X, d) be a Metric Space.

A. The map $\delta : \mathrm{cs}(X, d) \times \mathrm{cs}(X, d) \to [0, \infty)$ has the following properties:

i. $\delta(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{y}, \mathbf{x}), \quad \forall \mathbf{x}, \mathbf{y} \in \mathrm{cs}(X, d)$;

ii. $\delta(\mathbf{x}, \mathbf{y}) \leq \delta(\mathbf{x}, \mathbf{z}) + \delta(\mathbf{z}, \mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathrm{cs}(X, d)$;

iii. $\delta(\mathbf{x}, \mathbf{y}) = 0 \Rightarrow \mathbf{x} \sim \mathbf{y}$;

iv. If $\mathbf{x}, \mathbf{x}', \mathbf{y}, \mathbf{y}' \in \mathrm{cs}(X, d)$ are such that $\mathbf{x} \sim \mathbf{x}'$ and $\mathbf{y} \sim \mathbf{y}'$, then $\delta(\mathbf{y}, \mathbf{x}) = \delta(\mathbf{x}', \mathbf{y}')$.

B. The map $\widetilde{d} : \widetilde{X} \times \widetilde{X} \to [0, \infty)$, correctly defined by

$$\widetilde{d}(\widetilde{\mathbf{x}}, \widetilde{\mathbf{y}}) = \delta(\mathbf{x}, \mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in \mathrm{cs}(X, d),$$

is a metric on $\widetilde{X}$.

C. The map $X \ni x \longmapsto \langle x \rangle \in \widetilde{X}$ is isometric, in the sense that

$$\widetilde{d}(\langle x \rangle, \langle y \rangle) = d(x, y), \quad \forall x, y \in X.$$

Proof of Theorem 3.34

A. The properties i, ii and iii are obvious. See the reflexivity, symmetry and the transitivity of the equivalence relation $\sim$, beneath **Definition 3.42**.
To prove property iv, let $\mathbf{x} = \{\mathbf{x_n}\}_{\mathbf{n} \in \mathbb{N}}$, $\mathbf{x}' = \{\mathbf{x'_n}\}_{\mathbf{n} \in \mathbb{N}}$, $\mathbf{y} = \{\mathbf{y_n}\}_{\mathbf{n} \in \mathbb{N}}$ and $\mathbf{y}' = \{\mathbf{y'_n}\}_{\mathbf{n} \in \mathbb{N}} \in \mathrm{cs}(X, d)$. The next inequality

$$d(x'_n, y'_n) \leq d(x'_n, x_n) + d(x_n, y_n) + d(y_n, y'_n),$$

together with the fact that $\lim_{n \to \infty} d(x'_n, x_n) = 0$ and $\lim_{n \to \infty} d(y_n, y'_n) = 0$ immediately gives

$$\delta(\mathbf{x}', \mathbf{y}') = \lim_{n \to \infty} d(x'_n, y'_n) \leq \lim_{n \to \infty} d(x_n, y_n) = \delta(\mathbf{x}, \mathbf{y}).$$

By symmetry: $\delta(\mathbf{x}, \mathbf{y}) \leq \delta(\mathbf{x}', \mathbf{y}')$.

B. This follows immediately from **A**.

C. This follows from the definition.

$\square$

Theorem 3.35

Let (X, d) be a Metric Space.

i. For any Cauchy sequence $\mathbf{x} = \{x_n\}_{n\in\mathbb{N}}$ in X, there exists a limit in $\widetilde{X}$, so

$$\lim_{n\to\infty} \langle x_n \rangle = \widetilde{\mathbf{x}} \in \widetilde{X}.$$

ii. The metric space $(\widetilde{X}, \widetilde{d})$ is complete.

Proof of Theorem 3.35

i. For every $n \geq 1$, there holds that

$$\widetilde{d}(\langle x_n \rangle, \widetilde{x}) = \delta(x_n, x_m) = \lim_{n\to\infty} d(x_n, x_m).$$

If $\epsilon > 0$ is given, there exists a $N(\epsilon)$ such that

$$d(x_n, x_m) < \epsilon \quad \text{for all } n, m \geq N(\epsilon),$$

and this shows that

$$\widetilde{d}(\langle x_n \rangle, \widetilde{x}) \leq \epsilon \quad \text{for all } n \geq N(\epsilon).$$

The result is that

$$\lim_{n\to\infty} \widetilde{d}(\langle x_n \rangle, \widetilde{x}) = 0.$$

ii. Let $\{\widetilde{p}_n\}_{n \in \mathbb{N}}$ be a Cauchy sequence in $\widetilde{X}$.
For each n is $\widetilde{p}_n$ an equivalence class in $\widetilde{X}$, containing Cauchy sequences in X, converging to $\widetilde{p}_n$, see **part i**. So for each $n \geq 1$, there is some element $x_n \in X$ such that

$$\widetilde{d}(\langle x_n \rangle, \widetilde{p}_n) \leq \frac{1}{2^n}.$$

The sequence $\mathbf{x} = \{x_n\}_{n \in \mathbb{N}}$ is a Cauchy sequence in X.

For $i \geq j \geq 1$:

$$\begin{aligned} d(x_i, x_j) &= \widetilde{d}(\langle x_i \rangle, \langle x_j \rangle) \leq \widetilde{d}(\langle x_i \rangle, \widetilde{p}_i) + \widetilde{d}(\widetilde{p}_i, \widetilde{p}_j) + \widetilde{d}(\widetilde{p}_j, \langle x_j \rangle) \\ &\leq \widetilde{d}(\widetilde{p}_i, \widetilde{p}_j) + \frac{2}{2^j}. \end{aligned}$$

So $\mathbf{x} = \{x_n\}_{n \in \mathbb{N}}$ is a Cauchy sequence in X.

There holds that $\lim_{n \to \infty} \widetilde{p}_n = \widetilde{\mathbf{x}} \in \widetilde{X}$.

First of all, for $i \geq j \geq 1$, there is the inequality

$$\widetilde{d}(\widetilde{p}_j, \langle x_i \rangle) \leq \widetilde{d}(\widetilde{p}_j, \langle x_j \rangle) + \widetilde{d}(\langle x_j \rangle, \langle x_i \rangle) \leq \frac{1}{2^j} + d(x_j, x_i). \tag{3.25}$$

If $\epsilon > 0$ is given, there exists a $N(\epsilon)$ such that

$$d(x_j, x_i) < \epsilon \quad \text{for all } i, j \geq N(\epsilon),$$

and inequality (**3.25**) becomes

$$\widetilde{d}(\widetilde{p}_j, \langle x_i \rangle) \leq \frac{1}{2^j} + \epsilon \quad \text{for all } j \geq N(\epsilon).$$

Keep $j \geq N(\epsilon)$ fixed and let $i \to \infty$, together with **part i**, this gives:

$$\widetilde{d}(\widetilde{p}_j, \widetilde{\mathbf{x}}) = \lim_{j \to \infty} \widetilde{d}(\widetilde{p}_j, \langle x_i \rangle) \leq \frac{1}{2^j} + \epsilon \quad \text{for all } j \geq N(\epsilon).$$

This altogether proves that

$$\lim_{j \to \infty} \widetilde{d}(\widetilde{p}_j, \widetilde{\mathbf{x}}) = 0,$$

so the Cauchy sequence $\{\widetilde{p}_n\}_{n \in \mathbb{N}}$ converges to $\widetilde{\mathbf{x}} \in \widetilde{X}$. □

Definition 3.43

The metric space $(\widetilde{X}, \widetilde{d})$ is called the completion of (X, d).

Theorem 3.36

Let (X, d) be a Metric Space and let $(\widetilde{X}, \widetilde{d})$ be its completion.
If (Y, ρ) is a complete Metric Space and if $T : X \rightarrow Y$ is a map, which is Lipschitz continuous, see section 2.10, then there exists an unique Lipschitz continuous extension $\widetilde{T} : \widetilde{X} \rightarrow Y$ of the map T, such that

$$\widetilde{T}(\langle x \rangle) = T(x) \quad \text{for every } x \in X.$$

Moreover, the extension $\widetilde{T}$ and T have the same Lipschitz constant $L > 0$.

Proof of Theorem 3.36

There will be started with some Cauchy sequence $\mathbf{x} = \{x_n\}_{n \in \mathbb{N}}$ in X. Since the map T is Lipschitz continuous, it follows that

$$\rho(T(x_m), T(x_n)) \leq L\, d(x_m, x_n) \quad \text{for all } m, n \in \mathbb{N}.$$

So the sequence $\{T(x_n\}_{n \in \mathbb{N}}$ is a Cauchy sequence in Y. The Metric Space Y is complete , so the sequence $\{T(x_n\}_{n \in \mathbb{N}}$ converges in Y and there can be constructed the map

$$\phi(\mathbf{x}) = \lim_{n \rightarrow \infty} T(x_n)$$

and $\phi : \mathrm{cs}(X, d) \rightarrow Y$.
If $\mathbf{x} \sim \mathbf{x}'$ then $\phi(\mathbf{x}) = \phi(\mathbf{x}')$.
If $\mathbf{x} = \{x_n\}_{n \in \mathbb{N}}$ and $\mathbf{x}' = \{x'_n\}_{n \in \mathbb{N}}$ then the Lipschitz continuity gives

$$\rho(T(x_n, T(x'_n)) \leq L\, d(x_n, x'_n) \quad \text{for all } n \in \mathbb{N}.$$

Since $\lim_{n \rightarrow \infty} d(x_n, x'_n) = 0$, the result becomes that

$$\lim_{n\in\mathbb{N}} \rho(T(x_n), T(x'_n)) = 0$$

and that means that

$$\lim_{n\in\mathbb{N}} T(x_n) \;=\; \lim_{n\in\mathbb{N}} T(x'_n),$$

so $\phi(\mathbf{x}) = \phi(\mathbf{x}')$.

The extended map $\widetilde{T} : \widetilde{X} \to Y$ is correctly defined, with the property that

$$\widetilde{T}(\widetilde{\mathbf{x}}) \;=\; \phi(\mathbf{x}) \quad \text{ for all } \mathbf{x} \in \mathrm{cs}(X, d)$$

and the equality

$$\widetilde{T}(\langle x \rangle) \;=\; T(x) \quad \text{ for all } x \in X$$

is also satisfied.

Two things have to be checked,

the Lipschitz continuity and the uniqueness of $\widetilde{T}$.

The Lipschitz continuity.

Take two arbitrary elements $\widetilde{p}$ and $\widetilde{q}$ out of $\widetilde{X}$, represented as $\widetilde{p} = \widetilde{\mathbf{x}}$ and $\widetilde{q} = \widetilde{\mathbf{y}}$, for two Cauchy sequences $\widetilde{\mathbf{x}} = \{x_n\}_{n\in\mathbb{N}}$ and $\widetilde{\mathbf{y}} = \{y_n\}_{n\in\mathbb{N}}$ in X. Using the definition of $\widetilde{T}$ gives

$$\widetilde{T}(\widetilde{p}) \;=\; \lim_{n\to\infty} T(x_n) \,\text{and}\, \widetilde{T}(\widetilde{q}) \;=\; \lim_{n\to\infty} T(y_n)$$

and

$$\rho(\widetilde{T}(\widetilde{p}), \widetilde{T}(\widetilde{q})) \;=\; \lim_{n\to\infty} \rho(T(x_n), T(y_n)).$$

For every $n \in \mathbb{N}$ holds

$$\rho(T(x_n), T(y_n)) \;\leq\; L\, d(x_n, y_n),$$

taking the limit yields

$$\rho(\widetilde{T}(\widetilde{p}), \widetilde{T}(\widetilde{q})) \;=\; \lim_{n\to\infty} \rho(T(x_n), T(y_n)) \;\leq\; L \lim_{n\to\infty} d(x_n, y_n) \;=\; L\, d(\widetilde{p}, \widetilde{q}).$$

The uniqueness of $\widetilde{T}$.

A map, which is Lipschitz continuous, is also continuous.

Let $F : \widetilde{X} \to Y$ be another continuous map with $F(\langle x \rangle) = T(x)$ for all

$x \in X$. Take an arbitrary point $p \in \widetilde{X}$, represented as $p = \mathbf{x}$, for some Cauchy sequence $\mathbf{x} = \{x_n\}_{n\in\mathbb{N}} \in X$. Since $\lim_{n\to\infty}\langle x_n\rangle = p$ in $\widetilde{X}$ and by the use of **Remark 2.1**, there follows that

$$F(p) \;=\; \lim_{n\to\infty} F(\langle x_n\rangle) \;=\; \lim_{n\to\infty} T(x_n) \;=\; \phi(\mathbf{x}) \;=\; \widetilde{T}(p).$$

$\square$

4 Dual Spaces

Working with a dual space, it means that there is a vector space X. A dual space is not difficult, if the vector space X has an finite dimension, for instance $\dim X = n$. If first instance the vector space X is kept finite dimensional.
To make clear, what the differences are between finite and infinite dimensional vector spaces there will be given several examples with infinite dimensional vector spaces. The sequence spaces ℓ^1, ℓ^∞ and c_0, out of section **5.2**, are used.
Working with dual spaces, there becomes sometimes the question: "If the vector space X is equal to the dual space of X or if these spaces are really different from each other." Two spaces can be different in appearance but, with the help of a mapping, they can be "essential identical".
The scalars of the Vector Space X are taken out of some field $\mathbb{K}$, most of the time the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$.

4.1 Spaces X and $\tilde{X}$ are "essential identical"

To make clear that the spaces X and $\tilde{X}$ are "essential identical", there is needed a bijective mapping T between the spaces X and $\tilde{X}$.
If $T : X \rightarrow \tilde{X}$, then T has to be onto and one to one, such that T^{-1} exists.
But in some cases, T also satisfies some other conditions. T is called a isomorphism if it also preserves the structure on the space X and there are several possibilities. For more information about an isomorphism, see **wiki-homomorphism**.
Using the following abbvreviations, VS for a vector space (see section **3.2**), MS for a metric space (see section **3.5**), NS for a normed vector space (see section **3.7**),
several possibilities are given in the following scheme:

VS: An isomorphism T between vector spaces X and $\tilde{X}$, i.e.
T is a bijective mapping, but it also preserves the linearity

$$\begin{cases} T(x+y) = T(x) + T(y) \\ T(\beta x) = \beta T(x) \end{cases} \tag{4.1}$$

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

MS: An isomorphism T between the metric space (X, d) and $(\tilde{X}, \tilde{d})$. Besides that T is a bijective mapping, it also preserves the distance

$$\tilde{d}(T(x), T(y)) = d(x, y) \tag{4.2}$$

for all $x, y \in X$, also called an distance-preserving isomorphism.

Remark 4.1

If a map T satisfies (4.2) then T is necessarily injective. Because out of $\tilde{d}(T(x), T(y)) = 0$ follows that $d(x, y) = 0$, so $x = y$. If $\tilde{X} = T(X)$ then the map $T : X \to \tilde{X}$ is also bijective. There is said that $(\tilde{X}, \tilde{d})$ is an isometric copy of (X, d) and T is called an isometry. $T^{-1} : T(X) \Rightarrow X$ is also an isometry.

NS: An isomorphism T between Normed Spaces X and $\tilde{X}$. Besides that T is an isomorphism between vector spaces, it also preserves the norm

$$\| T(x) \| = \| x \| \tag{4.3}$$

for all $x \in X$, also called an isometric isomorphism.

4.2 Linear functional and sublinear functional

Definition 4.1

If X is a Vector Space over $\mathbb{K}$, with $\mathbb{K}$ the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$, then a linear functional is a function $f : X \rightarrow \mathbb{K}$, which is linear

LF 1: $f(x + y) = f(x) + f(y)$,
LF 2: $f(\alpha x) = \alpha f(x)$,
for all $x, y \in X$ and for all $\beta \in \mathbb{K}$.

Sometimes linear functionals are just defined on a subspace Y of some Vector Space X. To extend such functionals on the entire space X, the boundedness properties are defined in terms of sublinear functionals.

Definition 4.2

Let X be a Vector Space over the field $\mathbb{K}$. A mapping $p : X \rightarrow \mathbb{R}$ is called a sublinear functional on X if

SLF 1: $p(x + y) \leq p(x) + p(y)$,
SLF 2: $p(\alpha x) = \alpha p(x)$,
for all $x \in X$ and for all $0 \leq \alpha \in \mathbb{R}$

Example 4.1

The *norm* on a Normed Space is an example of a sublinear functional.

Example 4.2

If the elements of $\underline{x} \in \mathbb{R}^N$ are represented by columns

$$\underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

and there is given a row $\underline{a}$, with N known real numbers

$$\underline{a} = \begin{bmatrix} a_1 & \cdots & a_N \end{bmatrix}$$

then the matrix product

$$f(\underline{x}) = \begin{bmatrix} a_1 & \cdots & a_N \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \tag{4.4}$$

defines a linear functional f on $\mathbb{R}^N$.
If all the linear functionals g, on $\mathbb{R}^N$, have the same representation as given in (4.4), then each functional g can be identified by a column $\underline{b} \in \mathbb{R}^N$

$$\underline{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_N \end{bmatrix}.$$

In that case each linear functional g can be written as an inner product between the known element $\underline{b} \in \mathbb{R}^N$ and the unknown element $\underline{x} \in \mathbb{R}^N$

$$\underline{g(\underline{x})} = \underline{b} \bullet \underline{x},$$

for the notation, see (5.39).

4.3 Algebraïc dual space of X, denoted by X^*

Let X be a Vector Space and take the set of all linear functionals $f : X \rightarrow \mathbb{K}$. This set of all these linear functionals is made a Vector Space by defining an addition and a scalar multiplication. If f_1, f_2 are linear functionals on X and β is a scalar, then the addition and scalar multiplication are defined by

$$\begin{cases} (f_1 + f_2)(x) = f_1(x) + f_2(x) \\ f_1(\beta\, x) = \beta\, f_1(x) \end{cases} \tag{4.5}$$

for all $x \in X$ and for all $\beta \in \mathbb{K}$.
The set of all linear functionals on X, together with the above defined addition and scalar multiplication, see (**4.5**), is a Vector Space and is called the algebraïc dual space of X and is denoted by X^* .
In short there is spoken about the the dual space X^*, the space of all the linear functionals on X. X^* becomes a Vector Space, if the addition and scalar multiplication is defined as in (**4.5**).

4.4 Vector space X, $\dim X = n$

Let X be a finite dimensional vector space, $\dim X = n$. Then there exists a basis $\{e_1, ..., e_n\}$ of X. Every $x \in X$ can be written in the form

$$x = \alpha_1 e_1 + ... + \alpha_n e_n \tag{4.6}$$

and the coefficients α_i, with $1 = i \leq n$, are unique.

4.4.1 Unique representation of linear functionals

Let f be a linear functional on X, the image of x is

$$f(x) \in \mathbb{K}.$$

Theorem 4.1

The functional f is uniquely determined if the images of the $y_k = f(e_k)$ of the basis vectors $\{e_1, \cdots, e_n\}$ are prescribed.

Proof of Theorem 4.1

Choose a basis $\{e_1, \cdots, e_n\}$ then every $x \in X$ has an unique representation

$$x = \sum_{i=1}^{n} \alpha_i \, e_i. \tag{4.7}$$

The functional f is linear and x has as image

$$f(x) = f(\sum_{i=1}^{n} \alpha_i \, e_i) = \sum_{i=1}^{n} \alpha_i \, f(e_i).$$

Since 4.7 is unique, the result is obtained. □

4.4.2 Unique representation of linear operators between finite dimensional spaces

Let T be a linear operator between the finite dimensional Vector Spaces X and Y

$$T : X \rightarrow Y.$$

Theorem 4.2

The operator T is uniquely determined if the images of the $y_k = T(e_k)$ of the basis vectors $\{e_1, \cdots, e_n\}$ of X are prescribed.

Proof of Theorem 4.2

Take the basis $\{e_1, \cdots, e_n\}$ of X then x has an unique representation

$$x = \sum_{i=1}^{n} \alpha_i \, e_i. \tag{4.8}$$

The operator T is linear and x has as image

$$T(x) = T(\sum_{i=1}^{n} \alpha_i \, e_i) = \sum_{i=1}^{n} \alpha_i \, T(e_i).$$

Since 4.8 is unique, the result is obtained. □

Let $\{b_1, \cdots, b_k\}$ be a basis of Y.

Theorem 4.3

The image of $y = T(x) = \sum_{i=1}^{k} \beta_i \, b_i$ of $x = \sum_{i=1}^{n} \alpha_i \, e_i$ can be obtained with

$$\beta_j = \sum_{i=1}^{n} \tau_{ij} \, \alpha_i$$

for $1 \leq j \leq k$. (See formula 4.10 for τ_{ij}.)

Proof of Theorem 4.3

Since $y = T(x)$ and $y_k = T(e_k)$ are elements of Y they have an unique representation with respect tot the basis $\{b_1, \cdots, b_k\}$,

$$y = \sum_{i=1}^{k} \beta_i \, b_i, \tag{4.9}$$

$$T(e_j) = \sum_{i=1}^{k} \tau_{jk} \, b_i, \tag{4.10}$$

Substituting the formulas of 4.9 and 4.10 together gives

$$T(x) = \sum_{j=1}^{k} \beta_j \, b_j = \sum_{i=1}^{n} \alpha_i \, T(e_i) = \sum_{i=1}^{n} \alpha_i \, (\sum_{j=1}^{k} \tau_{ij} \, b_j) = \sum_{j=1}^{k} (\sum_{i=1}^{n} \alpha_i \, \tau_{ij}) \, b_j. \tag{4.11}$$

Since $\{b_1, \cdots, b_k\}$ is basis of Y, the coefficients

$$\beta_j = \sum_{i=1}^{n} \alpha_i \, \tau_{ij}$$

for $1 \leq j \leq k$. □

4.4.3 Dual basis $\{f_1, f_2, ..., f_n\}$ of $\{e_1, ..., e_n\}$

Going back to the space X with $\dim X = n$, with its base $\{e_1, ..., e_n\}$ and the linear functionals f on X.
Given a linear functional f on X and $x \in X$.
Then x can be written in the following form $x = \sum_{i=1}^{n} \alpha_i e_i$. Since f is a linear functional on X, $f(x)$ can be written in the form

$$f(x) = f(\sum_{i=1}^{n} \alpha_i e_i) = \sum_{i=1}^{n} \alpha_i f(e_i) = \sum_{i=1}^{n} \alpha_i \gamma_i,$$

with $\gamma_i = f(e_i), i = 1, ..., n$.
The linear functional f is uniquely determined by the values $\gamma_i, i = 1, ..., n$, at the basis vectors $e_i, i = 1, ..., n$, of X.
Given n values of scalars $\gamma_1, ..., \gamma_n$, and a linear functional is determined on X, see in section 4.4.1, and see also example 4.2.
Look at the following n-tuples:

$$\begin{array}{c} (1, 0, ..., 0), \\ (0, 1, 0, ..., 0), \\ ...., \\ (0, ...0, 1, 0, .., 0), \\ ...., \\ (0, ..., 0, 1), \end{array}$$

these define n linear functionals $f_1, ..., f_n$ on X by

$$f_k(e_j) = \delta_{jk} = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

The defined set $\{f_1, f_2, ..., f_n\}$ is called the dual basis of the basis $\{e_1, e_2, ..., e_n\}$ for X. To prove that these functionals $\{f_1, f_2, ..., f_n\}$ are linear independent, the following equation has to be solved

$$\sum_{k=1}^{n} \beta_k f_k = 0.$$

Let the functional $\sum_{k=1}^{n} \beta_k f_k$ work on e_j and it follows that $\beta_j = 0$, because $f_j(e_j) = 1$ and $f_j(e_k) = 0$, if $j \neq k$.
Every functional $f \in X^*$ can be written as a linear combination of $\{f_1, f_2, ..., f_n\}$. Write the functional $f = \gamma_1 f_1 + \gamma_2 f_2 + ... + \gamma_n f_n$ and realize that when $x = \alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n$ that $f_j(x) = f_j(\alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n) = \alpha_j$, so $f(x) = f(\alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n) = \alpha_1 \gamma_1 + ... + \alpha_n \gamma_n$.
It is interesting to note that: $\dim X^* = \dim X = n$.

Theorem 4.4

Let X be a finite dimensional vector space, $\dim X = n$. If $x_0 \in X$ has the property that $f(x_0) = 0$ for all $f \in X^*$ then $x_0 = 0$.

Proof of Theorem 4.4

Let $\{e_1, ..., e_n\}$ be a basis of X and $x_0 = \sum_{i=1}^{n} \alpha_i e_i$, then

$$f(x_0) = \sum_{i=1}^{n} \alpha_i \gamma_i = 0,$$

for every $f \in X^*$, so for every choice of $\gamma_1, \cdots, \gamma_n$. This can only be the case if $\alpha_j = 0$ for $1 \leq j \leq n$. □

4.4.4 Second algebraïc dual space of X, denoted by X^{**}

Let X be a finite dimensional with $\dim X = n$.
An element $g \in X^{**}$, which is a linear functional on X^*, can be obtained by

$$g(f) = g_x(f) = f(x),$$

so $x \in X$ is fixed and $f \in X^*$ variable. In short X^{**} is called the second dual space of X. It is easily seen that

$$g_x(\alpha f_1 + \beta f_2) = (\alpha f_1 + \beta f_2)(x) = \alpha f_1(x) + \beta f_2(x) = \alpha g_x(f_1) + \beta g_x(f_2)$$

for all $\alpha, \beta \in \mathbb{K}$ and for all $f_1, f_2 \in X^*$. Hence g_x is an element of X^{**}.
To each $x \in X$ there corresponds a $g_x \in X^{**}$.
This defines the canonical mapping C of X into X^{**},

$$C : X \to X^{**},$$
$$C : x \to g_x$$

The mapping C is linear, because

$$(C(\alpha x + \beta y))(f) = g_{(\alpha x + \beta y)}(f) = f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) =$$
$$\alpha g_x(f) + \beta g_y(f) = \alpha (C(x))(f) + \beta (C(y))(f)$$

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

Theorem 4.5

The canonical mapping C is injective.

Proof of Theorem 4.5

If $C(x) = C(y)$ then $f(x) = f(y)$ for all $f \in X^*$. f is a linear functional, so $f(x - y) = 0$ for all $f \in X^*$. Using theorem 4.4 gives that $x = y$. □

Result so far is that C is a (vector space) isomorphism of X onto its range $R(C) \subset X^{**}$. The range $R(C)$ is a linear vectorspace of X^{**}, because C is a linear mapping on X. Also is said that X is embeddable in X^{**}.
The question becomes if C is surjective, is C onto? $(R(C) = X^{**}?)$

Theorem 4.6

The canonical mapping C is surjective.

Proof of Theorem 4.6

The domain of C is finite dimensional. C is injective from C to $R(C)$, so the inverse mapping of C, from $R(C)$ to C, exists. The dimension of $R(C)$ and the dimension of the domain of C have to be equal, this gives that $\dim R(C) = \dim X$. Further is know that $\dim (X^*)^* = \dim X^* (= \dim X)$ and the conclusion becomes that $\dim R(C) = \dim X^{**}$. The mapping C is onto the space X^{**}. □

C is vector isomorphism, so far it preserves only the linearity, about the preservation of other structures is not spoken. There is only looked at the perservation of the algebraic operations.
The result is that X and X^{**} look "algebraic identical". So speaking about X or X^{**}, it doesn't matter, but be careful: $\dim X = n < \infty$.

Definition 4.3

A Vector Space X is called algebraic reflexive if $R(C) = X^{**}$.

Important to note is that the canonical mapping C defined at the beginning of this section, is also called a natural embedding of X into X^{**}. There are examples of Banach spaces $(X, \|\cdot\|)$, which are isometric isomorph with $(X^{**}, \|\cdot\|)$, but not reflexive. For reflexivity, you need the natural embedding.

4.5 The dual space X' of a Normed Space X

In section 4.4 the dimension of the Normed Space X is finite.
In the finite dimensional case the linear functionals are always bounded. If a Normed Space is infinite dimensional that is not the case anymore. There is

a distinction between bounded linear functionals and unbounded linear functional. The set of all the linear functionals of a space X is often denoted by X^* and the set of bounded linear functionals by $X^{'}$.
In this section there will be looked at Normed Space in general, so they may also be infinite dimensional. There will be looked in the main to the bounded linear functionals.
Let X be a Normed Space, with the norm $\| \cdot \|$. This norm is needed to speak about a norm of a linear functional on X.

Definition 4.4

The norm of a linear functional f is defined by

$$\| f \| = \sup_{\left\{ \begin{array}{c} x \in X \\ x \neq 0 \end{array} \right\}} \frac{| f(x) |}{\| x \|} = \sup_{\left\{ \begin{array}{c} x \in X \\ \| x \| = 1 \end{array} \right\}} | f(x) | \tag{4.12}$$

If the Normed Space X is finite dimensional then the linear functionals of the Normed Space X are always bounded. But if X is infinite dimensional there are also unbounded linear functionals.

Definition 4.5

A functional f is bounded if there exists a number A such that

$$| f(x) | \leq A \| x \| \tag{4.13}$$

for all x in the Normed Space X.

The two definitions of a norm of a linear functional are equivalent because of the fact that

$$\frac{| f(x) |}{\| x \|} = | f(\frac{x}{\| x \|}) |$$

for all $0 \neq x \in X$. Interesting to note is, that the dual space $X^{'}$ of a Normed Space X is always a Banach space, because $BL(X, \mathbb{K})$ is a Banach Space, see Theorem 7.8 with $Y = \mathbb{K}$, $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ both are Banach Spaces.
Working with linear functionals, there is no difference between bounded or continuous functionals. Keep in mind that a linear functional f is nothing else

as a special linear operator $f : X \rightarrow \mathbb{K}$. Results derived for linear operators are also applicable to linear functionals.

Theorem 4.7

A linear functional, on a Normed Space, is bounded *if and only if* it is continuous.

Proof of Theorem 4.7

The proof exists out of two parts.

($\Rightarrow$) Suppose f is linear and bounded, then there is a positive constant A such that $\mid f(x) \mid \leq A \parallel x \parallel$ for all x.
If $\epsilon > 0$ is given, take $\delta = \frac{\epsilon}{A}$ and for all y with $\parallel x - y \parallel \leq \delta$

$$\mid f(x) - f(y) \mid = \mid f(x - y) \mid \leq A \parallel x - y \parallel \leq A\,\delta = A\,\frac{\epsilon}{A} = \epsilon.$$

So the functional f is continuous in x.
If $A = 0$, then $f(x) = 0$ for all x and f is trivally continuous.

($\Leftarrow$) The linear functional is continous, so continuous in $x = 0$.
Take $\epsilon = 1$ then there exists a $\delta > 0$ such that

$$\mid f(x) \mid < 1 \text{ for } \parallel x \parallel < \delta.$$

For some arbirary y, in the Normed Space, it follows that

$$\mid f(y) \mid = \frac{2 \parallel y \parallel}{\delta} f(\frac{\delta}{2 \parallel y \parallel} y) < \frac{2}{\delta} \parallel y \parallel,$$

since $\parallel \frac{\delta}{2 \parallel y \parallel} y \parallel = \frac{\delta}{2} < \delta$. Take $A = \frac{2}{\delta}$ in formula 4.13, this positive constant A is independent of y, the functional f is bounded.

□

4.6 Difference between finite and infinite dimensional Normed Spaces

If X is a finite dimensional Vector Space then there is in certain sense no difference between the space X^{**} and the space X, as seen in section 4.4.4. Be careful if X is an infinite dimensional Normed Space.

Theorem 4.8

$(\ell^1)^{'} = \ell^\infty$ and $(c_0)^{'} = \ell^1$

Proof of Theorem 4.8

See the sections 4.6.1 and 4.6.2. □

Theorem 4.8 gives that $((c_0)^{'})^{'} = (\ell^1)^{'} = \ell^\infty$. One thing can always be said and that is that $X \subseteq X''$, see theorem 4.14. So $c_0 \subseteq (c_0)'' = \ell^\infty$. c_0 is a separable Normed Space and ℓ^∞ is a non-separable Normed Space, so $c_0 \neq (c_0)''$ but $c_0 \subset \ell^\infty (= (c_0)'')$.
So, be careful in generalising results obtained in finite dimensional spaces to infinite dimensional Vector Spaces.

4.6.1 Dual space of ℓ^1 is ℓ^∞, ($(\ell^1)' = \ell^\infty$)

With the dual space of ℓ^1 is meant $(\ell^1)^{'}$, the space of bounded linear functionals of ℓ^1. The spaces ℓ^1 and ℓ^∞ have a norm and in this case there seems to be an isomorphism between two normed vector spaces, which are bove infinitely dimensional.
For ℓ^1 there is a basis $(e_k)_{k \in \mathbb{N}}$ and $e_k = \delta_{kj}$, so every $x \in \ell^1$ can be written as

$$x = \sum_{k=1}^{\infty} \alpha_k e_k.$$

The norm of $x \in \ell^1$ is

$$\| x \|_1 = \sum_{k=1}^{\infty} | \alpha_k | \; (< \infty)$$

and the norm of $x \in \ell^\infty$ is

$$\| x \|_\infty = \sup_{k \in \mathbb{N}} | \alpha_k | (< \infty).$$

A bounded linear functional f of ℓ^1, $(f : \ell^1 \to \mathbb{R})$ can be written in the form

$$f(x) = f(\sum_{k=1}^{\infty} \alpha_k e_k) = \sum_{k=1}^{\infty} \alpha_k \gamma_k,$$

with $f(e_k) = \gamma_k$.
Take a look at the row $(\gamma_k)_{k \in \mathbb{N}}$, realize that $\| e_k \|_1 = 1$ and

$$| \gamma_k | = | f(e_k) | \leq \| f \|_1 \| e_k \|_1 = \| f \|_1$$

for all $k \in \mathbb{N}$. Such that $(\gamma_k)_{k \in \mathbb{N}} \in \ell^\infty$, since

$$\sup_{k \in \mathbb{N}} | \gamma_k | \leq \| f \|_1 .$$

Given a linear functional $f \in (\ell^1)^{'}$ there is constructed a row $(\gamma_k)_{k \in \mathbb{N}} \in \ell^\infty$.
Now the otherway around, given an element of ℓ^∞, can there be constructed a bounded linear functionals in $(\ell^1)^{'}$?
An element $(\gamma_k)_{k \in \mathbb{N}} \in \ell^\infty$ is given and it is not difficult to construct the following linear functional f on ℓ^1

$$f(x) = \sum_{k=1}^{\infty} \alpha_k \gamma_k,$$

with $x = \sum_{k=1}^{\infty} \alpha_k e_k \in \ell^1$.
Linearity is no problem, but the boundedness of the linear functional g is more difficult to proof

$$| f(x) | \leq \sum_{k=1}^{\infty} | \alpha_k \gamma_k | \leq \sup_{k \in \mathbb{N}} | \gamma_k | \sum_{k=1}^{\infty} | \alpha_k | \leq \sup_{k \in \mathbb{N}} | \gamma_k | \, \| x \|_1 = \| (\gamma_k)_{k \in \mathbb{N}} \|_\infty \| x \|_1 .$$

The result is, that the functional f is linear and bounded on ℓ^1, so $f \in (\ell^1)^1$.
Looking at an isomorphism between two normed vector spaces, it is also of importance that the norm is preserved.
In this case, it is almost done, because

$$| f(x) | = | \sum_{k=1}^{\infty} \alpha_k \gamma_k | \leq \sup_{k \in \mathbb{N}} | \gamma_k | \sum_{k=1}^{\infty} | \alpha_k | \leq \sup_{k \in \mathbb{N}} | \gamma_k | \, \| x \|_1 = \| (\gamma_k) \|_\infty \| x \|_1 .$$

Take now the supremum over all the $x \in \ell^1$ with $\| x \|_1 = 1$ and the result is

$$\| f \|_1 \leq \sup_{k \in \mathbb{N}} | \gamma_k | = \| (\gamma_k)_{k \in \mathbb{N}} \|_\infty,$$

above the result was

$$\| (\gamma_k)_{k \in \mathbb{N}} \|_\infty = \sup_{k \in \mathbb{N}} | \gamma_k | \leq \| f \|_1,$$

taking these two inequalities together and there is proved that the norm is preserved,

$$\| f \|_1 = \| (\gamma_k)_{k \in \mathbb{N}} \|_\infty$$

The isometric isorphism between the two given Normed Spaces $(\ell^1)^{'}$ and ℓ^∞ is a fact.
So taking a element out of $(\ell^1)^{'}$ is in certain sense the same as speaking about an element out of ℓ^∞.

4.6.2 Dual space of c_0 is ℓ^1, ($(c_0)^{‘} = \ell^1$)

Be careful the difference between finite and infinite plays an important rule in this proof.
Take an arbitrary $x \in c_0$ then

$$x = \sum_{k=1}^{\infty} \lambda_k e_k \text{ with } \lim_{k \to \infty} \lambda_k = 0,$$

see the definition of c_0 in section 5.2.6.
Taking finite sums, there is contructed the following approximation of x

$$s_n = \sum_{k=1}^{n} \lambda_k e_k,$$

because of the $\| \, . \, \|_\infty$-norm

$$\lim_{n \to \infty} \| s_n - x \|_\infty = 0.$$

If f is a bounded functional on c_0, it means that f is continuous on c_0 (see theorem 4.7), so if $s_n \to x$ then $f(s_n) \to f(x)$ as $n \to \infty$.
Known is that

$$f(x) = \sum_{k=1}^{\infty} \lambda_k f(e_k) \; = \sum_{k=1}^{\infty} \lambda_k \gamma_k.$$

Look at the row $(\gamma_k)_{k \in \mathbb{N}}$, the question becomes if $(\gamma_k)_{k \in \mathbb{N}} \in \ell^1$?
Speaking about f in $(c_0)'$ should become the same as speaking about the row $(\gamma_k)_{k \in \mathbb{N}} \in \ell^1$.
With γ_k, $k \in \mathbb{N}$, is defined a new symbol

$$\lambda_k^0 = \begin{cases} \dfrac{\gamma_k}{\mid \gamma_k \mid} & \text{if } \gamma_k \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Now it easy to define new sequences $x_0^n = (\eta_k^0)_{k \in \mathbb{N}} \in c_0$, with

$$\eta_k^0 = \begin{cases} \lambda_k^0 & \text{if } 1 \leq k \leq n, \\ 0 & n < k, \end{cases}$$

and for all $n \in \mathbb{N}$.
It is clear that $\| x_0^n \|_\infty = 1$ and

$$\mid f(x_0^n) \mid = \mid \sum_{k=1}^{n} \eta_k^0 \gamma_k \mid = \sum_{k=1}^{n} \mid \gamma_k \mid = \sum_{k=1}^{n} \mid f(e_k) \mid \leq \| f \|_\infty \| x_0^n \|_\infty \leq \| f \|_\infty, \tag{4.14}$$

so $\sum_{k=1}^{n} \mid f(e_k) \mid = \sum_{k=1}^{n} \mid \gamma_k \mid < \infty$, and that is for every $n \in \mathbb{N}$ and $\| f \|_\infty$ is independent of n.
Out of the last inequalities, for instance inequality 4.14, follows that

$$\sum_{k=1}^{n} \mid \gamma_k \mid = \sum_{k=1}^{\infty} \mid f(e_k) \mid \leq \| f \|_\infty . \tag{4.15}$$

This means that $(\gamma_k)_{k \in \mathbb{N}} \in \ell^1$!
That the norm is preserved is not so difficult. It is easily seen that

$$\mid f(x) \mid \leq \sum_{k=1}^{\infty} \mid \lambda_k \mid \mid \gamma_k \mid \leq \| x \|_\infty \sum_{k=1}^{\infty} \mid \lambda_k \mid \leq \| x \|_\infty \sum_{k=1}^{\infty} \mid f(e_k) \mid,$$

and this means that

$$\frac{\mid f(x) \mid}{\| x \|_\infty} \leq \sum_{k=1}^{\infty} \mid f(e_k) \mid,$$

together inequalty 4.15, gives that $\| (\gamma_k)_{k \in \mathbb{N}} \|_1 = \| f \|_\infty$.
Known some $f \in (c_0)$‘ gives us an element in ℓ^1.

Is that mapping also onto?
Take some $(\alpha_k)_{k \in \mathbb{N}} \in \ell^1$ and an arbirtrary $x = (\lambda_k)_{k \in \mathbb{N}} \in c_0$ and define the linear functional $f(x) = \sum_{k=1}^{\infty} \lambda_k \, \alpha_k$. The series $\sum_{k=1}^{\infty} \lambda_k \, \alpha_k$ is absolute convergent and

$$\frac{| f(x) |}{\| x \|_\infty} \leq \sum_{k=1}^{\infty} | \alpha_k | \leq \| (\alpha_k)_{k \in \mathbb{N}} \|_1 .$$

The constructed linear functional f is bounded (and continuous) on c_0.
The isometric isorphism between the two given Normed Spaces $(c_0)^{'}$ and ℓ^1 is a fact.

4.7 The extension of functionals, the Hahn-Banach theorem

In section 3.10.1 is spoken about the minimal distance of a point x to some convex subset M of an Inner Product Space X. Theorem 3.23 could be read as that it is possible to construct hyperplanes through y_0, which separate x form the subset M, see figures 3.5 and 3.6. Hyperplanes can be seen as level surfaces of functionals. The inner products are of importantance because these results were obtained in Hilbert Spaces.
But a Normed Space has not to be a Hilbert Space and so the question becomes if it is possible to separate points of subsets with the use of linear functionals? Not anymore in an Inner Product Space, but in a Normed Space.
Let X be a Normed Space and M be some proper linear subspace of X and let $x_0 \in X$ such that $d(x_0, M) = d > 0$ with $d(\cdot, M)$ as defined in definition 3.21. The question is if there exists some bounded linear functional $g \in X^{'}$ such that

$$g(x_0) = 1, \ g|_M = 0, \text{ and may be } \| g \| = \frac{1}{d}? \tag{4.16}$$

This are conditions of a certain functional g on a certain subspace M of X and in a certain point $x_0 \in X$. Can this functional g be extended to the entire Normed Space X, preserving the conditions as given? The theorem of Hahn-Banach will prove the existence of such an extended functional.

Remark 4.2

Be careful! Above is given that $d(x_0, M) = d > 0$. If not, if for instance is given some proper linear subspace M and $x_0 \in X \backslash M$, it can happpen that $d(x_0, M) = 0$, for instance if $x_0 \in \overline{M} \backslash M$.
But if M is closed and $x_0 \in X \backslash M$ then $d(x_0, M) = d > 0$. A closed linear subspace M gives no problems, if nothing is known about $d(x_0, M)$.

Proving the theorem of Hahn-Banach is a lot of work and the lemma of Zorn is used, see theorem 9.1. Difference with section 3.10 is, that there can not be made use of an inner product, there can not be made use of orthogonality. To construct a bounded linear functional g, which satisfies the conditions as given in formula 4.16 is not difficult. Let $x = m + \alpha x_0$, with $m \in M$ and $\alpha \in \mathbb{R}$, define the bounded linear functional g on the linear subspace $\widehat{M} = \{m + \alpha x_0 | m \in M \text{ and } \alpha \in \mathbb{R}\}$ by

$$g(m + \alpha x_0) = \alpha.$$

It is easily seen that $g(m) = 0$ and $g(m + x_0) = 1$, for every $m \in M$.
The functional g is linear on $\widehat{M}$

$$\begin{cases} g((m_1 + m_2) + (\alpha_1 + \alpha_2) x_0) = (\alpha_1 + \alpha_2) = g(m_1 + \alpha_1 x_0) + g(m_2 + \alpha_2 x_0) \\ g(\gamma(m_1 + \alpha_1 x_0)) = \gamma \alpha_1 = \gamma g(m_1 + \alpha_1 x_0). \end{cases}$$

Further, $\alpha \neq 0$,

$$\| m + \alpha x_0 \| = |\alpha| \, \| \frac{m}{\alpha} + x_0 \| \geq |\alpha| \, d(x_0, M) = |\alpha| \, d,$$

since $\frac{m}{\alpha} \in M$, so

$$\frac{|g(m + \alpha x_0)|}{\| m + \alpha x_0 \|} \leq \frac{|\alpha|}{|\alpha| \, d} = \frac{1}{d}, \tag{4.17}$$

so the linear functional g is bounded on $\widehat{M}$ and $\| g \| \leq \frac{1}{d}$.
The distance of x_0 to the linear subspace M is defined as an infimum, what means that there exists a sequence $\{m_k\}_{k \in \mathbb{N}}$ such that $\lim_{k \to \infty} \| x_0 - m_k \| = d$. Using the definition and the boundedness of the linear functional g

$$g(-m_k + x_0) = 1 \leq \| g \| \, \| -m_k + x_0 \|,$$

let $k \to \infty$ and it follows that

$$\| g \| \geq \frac{1}{d} \tag{4.18}$$

on $\widehat{M}$. With the inequalities 4.18 and 4.17 it follows that $\| g \| = \frac{1}{d}$ on $\widehat{M}$ and there is constructed a $g \in \widehat{M}'$, which satisfies the conditions given in 4.16. The problem is to extended g to the entire Normed Space X.

First will be proved the Lemma of Hahn-Banach and after that the Theorem of Hahn-Banach . In the Lemma of Hahn-Banach is spoken about a sublinear functional, see **definition 4.2**. If $f \in X'$ then is an example of a sublinear functional p given by

$$p(x) = \| f \| \| x \|, \tag{4.19}$$

for every $x \in X$. If the bounded linear functional f is only defined on some linear subspace M of the Normed Space X, then can also be taken the norm of f on that linear subspace M in definition 4.19 of the sublinear functional p. The conditions **SLF ii: 1** and **SLF ii: 2** are easy to check. First will be proved the Lemma of Hahn-Banach .

Theorem 4.9

Let X be real linear space and let p be a sublinear functional on X. If f is a linear functional on a linear subspace M of X which satisfies

$$f(x) \leq p(x),$$

for every $x \in M$, then there exists a real linear functional $f_{\mathcal{E}}$ on X such that

$$f_{\mathcal{E}}|M = f \text{ and } f_{\mathcal{E}}(x) \leq p(x),$$

for every $x \in X$.

Proof of Theorem 4.9

The proof is splitted up in several steps.

1. First will be looked at the set of all possible extensions of (M, f) and the question will be if there exists some maximal extension? See **Step ii: 1**.
2. If there exists some maximal extension, the question will be if that is equal to $(X, f_{\mathcal{E}})$? See **Step ii: 2**.

Step 1: An idea to do is to enlarge M with one extra dimension, a little bit as the idea written in the beginning of this section 4.7 and then to keep doing that until the entire space X is reached. The problem is to find a good argument that indeed the entire space X is reached.
To prove the existence of a maximal extension the lemma of Zorn will be used, see section 9.3. To use that lemma there has to be defined some order $\preceq$, see section 2.13.
The order will be defined on the set $\mathcal{P}$ of all possible linear extensions (M_α, f_α) of (M, f), satisfying the condition that

$$f_\alpha(x) \leq p(x),$$

for every $x \in M_\alpha$, so

$$\mathcal{P} = \{(M_\alpha, f_\alpha) \,|\, M_\alpha \text{ a linear subspace of } X \text{ and } M \subset M_\alpha,$$
$$f_\alpha|M = f \text{ and } f_\alpha(x) \leq p(x) \text{ for every } x \in M_\alpha\}.$$

The order $\preceq$ on $\mathcal{P}$ is defined by

$$(M_\alpha, f_\alpha) \preceq (M_\beta, f_\beta) \Longleftrightarrow M_\alpha \subset M_\beta$$

and $f_\beta \,|\, M_\alpha = f_\alpha$, so f_β is an extension of f_α.
It is easy to check that the defined order $\preceq$ is a partial order on $\mathcal{P}$, see definition 2.4. Hence, $(\mathcal{P}, \preceq)$ is a partial ordered set.
Let $\mathcal{Q}$ be a total ordered subset of $\mathcal{P}$ and let

$$\widehat{M} = \bigcup \{M_\gamma \,|\, (M_\gamma, f_\gamma) \in \mathcal{Q}\}.$$

$\widehat{M}$ is a linear subspace, because of the total ordering of $\mathcal{Q}$.
Define $\widehat{f} : \widehat{M} \to \mathbb{R}$ by

$$\widehat{f}(x) = f_\gamma(x) \text{ if } x \in M_\gamma.$$

It is clear, that $\widehat{f}$ is a linear functional on the linear subspace $\widehat{M}$ and

$$\widehat{f} \,|\, M = f \text{ and } \widehat{f}(x) \leq p(x)$$

for every $x \in \widehat{M}$. Further is $(\widehat{M}, \widehat{f})$ an upper bound of $\mathcal{Q}$, because

$$M_\gamma \subset \widehat{M} \text{ and } \widehat{f} \,|\, M_\gamma = f_\gamma.$$

Hence, $(M_\gamma, f_\gamma) \preceq (\widehat{M}, \widehat{f})$.
Since $\mathcal{Q}$ is an arbitrary total ordered subset of $\mathcal{P}$, Zorn's lemma implies that $\mathcal{P}$ possesses at least one maximal element (M_ϵ, f_ϵ).

Step 2: The problem is to prove that $M_\epsilon = X$ and $f_\epsilon = f_\mathcal{E}$. It is clear that when is proved that $M_\epsilon = X$ that $f_\epsilon = f_\mathcal{E}$ and the proof of the theorem is completed.
Assume that $M_\epsilon \neq X$, then there is some $y_1 \in (X \setminus M_\epsilon)$ and $y_1 \neq 0$, since $0 \in M_\epsilon$. look at the subspace $\widehat{M_\epsilon}$ spanned by M_ϵ and y_1. Elements are of the form $z + \alpha\, y_1$ with $z \in M_\epsilon$ and $\alpha \in \mathbb{R}$. If $z_1 + \alpha_1\, y_1 = z_2 + \alpha_2\, y_1$ then $z_1 - z_2 = (\alpha_2 - \alpha_1)\, y_1$, the only possible solution is $z_1 = z_2$ and $\alpha_1 = \alpha_2$, so the representation of elements out of $\widehat{M_\epsilon}$ is unique.
A linear functional h on $\widehat{M_\epsilon}$ is easily defined by

$$h(z + \alpha\, y_1) = f_\epsilon(z) + \alpha\, C$$

with a constant $C \in \mathbb{R}$. h is an extension of f_ϵ, if there exists some constant C such that

$$h(z + \alpha\, y_1) \leq p(z + \alpha\, y_1) \tag{4.20}$$

for all elements out of $\widehat{M_\epsilon}$. The existence of such a C is proved in Step **ii: 3**. If $\alpha = 0$ then $h(z) = f_\epsilon(z)$, further $M_\epsilon \subset \widehat{M_\epsilon}$, so $(M_\epsilon, f_\epsilon) \preceq (\widehat{M_\epsilon}, h)$, but this fact is in contradiction with the maximality of (M_ϵ, f_ϵ), so

$$M_\epsilon = X.$$

Step 3: It remains to choose C on such a way that

$$h(z + \alpha\, y_1) = f_\epsilon(z) + \alpha\, C \leq p(z + \alpha\, y_1) \tag{4.21}$$

for all $z \in M_\epsilon$ and $\alpha \in \mathbb{R} \setminus \{0\}$. Replace z by $\alpha\, z$ and divide both sides of **formula 4.21** by $|\alpha|$. That gives two conditions

$$\begin{cases} h(z) + C \leq p(z + y_1) & \text{if } z \in M_\epsilon \text{ and } \alpha > 0, \\ -h(z) - C \leq p(-z - y_1) & \text{if } z \in M_\epsilon \text{ and } \alpha < 0. \end{cases}$$

So the constant C has to be chosen such that

$$-\,h(v) - p(-v - y_1) \leq C \leq -\,h(w) + p(w + y_1)$$

for all $v, w \in M_\epsilon$. The condition, which C has to satisfy, is now known, but not if such a constant C also exists.
For any $v, w \in M_\epsilon$

$$h(w) - h(v) = h(w - v) \leq p(w - v)$$
$$= p(w + y_1 - v - y_1) \leq p(w + y_1) + p(-v - y_1),$$

and therefore

$$-h(v) - p(-v - y_1) \leq -h(w) + p(w + y_1).$$

Hence, there exists a real constant C such that

$$\sup_{v \in M_\epsilon} (-h(v) - p(-v - y_1)) \leq C \leq \inf_{w \in M_\epsilon} (-h(w) + p(w + y_1)). \tag{4.22}$$

With the choice of a real constant C, which satisfies inequality **4.22**, the extended functional h can be constructed, as used in **Step ii: 2**.

□

In the Lemma of Hahn-Banach, see **theorem 4.9**, is spoken about some sublinear functional p. In the Theorem of Hahn-Banach this sublinear functional is more specific given. The Theorem of Hahn-Banach gives the existence of an extended linear functional g of f on a Normed Space X, which preserves the norm of the functional f on some linear subspace M of X. In first instance only for real linear vectorspaces $(X, \mathbb{R})$ and after that the complex case.

Theorem 4.10

Let M be a linear subspace of the Normed Space X over the field $\mathbb{K}$, and let f be a bounded functional on M. Then there exists a norm-preserving extension g of f to X, so

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

Proof of Theorem 4.10

The proof is splitted up in two cases.

1. The real case $\mathbb{K} = \mathbb{R}$, see **Case ii: 1**.
2. The complex case $\mathbb{K} = \mathbb{C}$, see **Case ii: 2**.

Case 1: Set $p(x) = \| f \| \| x \|$, p is a sublinear functional on X and by the Lemma of Hahn-Banach, see **theorem 4.9**, there exists a real linear functional g on X such that

$$g \,|\, M \;=\; f \text{ and } g(x) \;\leq \| \; f \; \| \, \| \; x \; \|,$$

for every $x \in X$. Then

$$|\, g(x) \,| \;=\; \pm\, g(x) \;=\; g(\pm\, x) \;\leq\; p(\pm\, x) \;\leq \| \; f \; \| \, \| \pm x \, \| = \| \; f \; \| \, \| \; x \; \| \,.$$

Hence, g is bounded and

$$\| \; g \; \| \leq \| \; f \; \| \,. \tag{4.23}$$

Take some $y \in M$ then

$$\| \; g \; \| \geq \frac{|\, g(y) \,|}{\| \; y \; \|} \;=\; \frac{|\, f(y) \,|}{\| \; y \; \|}.$$

Hence,

$$\| \; g \; \| \geq \| \; f \; \| \,. \tag{4.24}$$

The inequalities 4.23 and 4.24 give that $\| \; g \; \| = \| \; f \; \|$ and complete the proof.

Case 2: Let X be a complex Vector Space and M a complex linear subspace. Set $p(x) \;= \| \; f \; \| \, \| \; x \; \|$, p is a sublinear functional on X. The functional f is complex-valued and the functional f can be written as

$$f(x) \;=\; u(x) \;+\; \imath\, v(x)$$

with u, v real-valued. Regard, for a moment, X and M as real Vector Spaces, denoted by X_r and M_r, just the scalar multiplication is restricted to real numbers. Since f is linear on M, u and v are linear functionals on M_r. Further

$$u(x) \;\leq\; |\, f(x) \,| \;\leq\; p(x)$$

for all $x \in M_r$. Using the result of theorem 4.9, there exists a linear extension $\widehat{u}$ of u from M_r to X_r, such that

$$\widehat{u}(x) \;\leq\; p(x)$$

for all $x \in X_r$.
Return to X, for every $x \in M$ yields

$$\imath(u(x) \;+\; \imath\, v(x)) \;=\; \imath\, f(x) \;=\; f(\imath\, x) \;=\; u(\imath\, x) \;+\; \imath\, v(\imath\, x),$$

so $v(x) \;=\; -\, u(\imath\, x)$ for every $x \in M$.
Define

$$g(x) = \widehat{u}(x) - \imath\,\widehat{u}(\imath\,x) \tag{4.25}$$

for all $x \in X$, $g(x) = f(x)$ for all $x \in M$, so g is an extension of f from M to X.
Is the extension g linear on X?
The summation is no problem. Using **formula 4.25** and the linearity of u on X_r, it is easily seen that

$$\begin{aligned} g((a + \imath\,b)\,x) &= \widehat{u}((a + \imath\,b)\,x) - \imath\,\widehat{u}((a + \imath\,b)\,\imath\,x) \\ &= a\,\widehat{u}(x) + b\,\widehat{u}(\imath\,x) - \imath\,(a\,\widehat{u}(\imath\,x) - b\,\widehat{u}(x)) \\ &= (a + \imath\,b)\,(\widehat{u}(x) - \imath\,\widehat{u}(\imath\,x)) = (a + \imath\,b)\,g(x), \end{aligned}$$

for all $a, b \in \mathbb{R}$. Hence, g is linear on X.
Is the extension g norm-preserving on M?
Since g is an extension of f, this implies that

$$\| g \| \geq \| f \| . \tag{4.26}$$

Let $x \in X$ then there is some real number ϕ such that

$$g(x) = |\,g(x)\,| \exp(\imath\,\phi).$$

Then

$$\begin{aligned} |\,g(x)\,| &= \exp(-\imath\,\phi)\,g(x) \\ &= \mathrm{Re}(\exp(-\imath\,\phi)\,g(x)) = \mathrm{Re}(g(\exp(-\imath\,\phi)\,x)) \\ &= \widehat{u}(\exp(-\imath\,\phi)\,x) \leq \| f \| \| \exp(-\imath\,\phi)\,x \| \\ &= \| f \| \| x \| . \end{aligned}$$

This shows that g is bounded and $\| g \| \leq \| f \|$, together with **inequality 4.26** it completes the proof.

☐

At the begin of this section, the problem was the existence of a bounded linear functional g on X, such that

$$g(x_0) = 1,\; g|_M = 0, \text{ and may be } \| g \| = \frac{1}{d}, \tag{4.27}$$

with $x_0 \in X$ such that $d(x_0, M) = d > 0$.
Before the **Lemma of Hahn-Banach, theorem 4.9**, there was constructed a bounded linear functional g on $\widehat{M}$, the span of M and x_0, which satisfied

the condition given in 4.27. The last question was if this constructed g could be extended to the entire space X?
With the help of the **Hahn-Banach theorem, theorem 4.10**, the constructed bounded linear functional g on $\widehat{M}$ can be extended to the entire space X and the existence of a $g \in X^{'}$, which satisfies all the conditions, given 4.27, is a fact.
The result of the question in 4.16 can be summarized into the following theorem:

Theorem 4.11

Let X be a Normed Space over some field $\mathbb{K}$ and M some linear subspace of X. Let $x_0 \in X$ be such that $d(x_0, M) > 0$. Then there exists a linear functional $g \in X^{'}$ such that

i. $g(x_0) = 1,$
ii. $g(M) = 0,$
iii. $\| g \| = \dfrac{1}{d}.$

Proof of Theorem 4.11

Read this section 4.7. □

Remark 4.3

With the result of **theorem 4.11** can be generated all kind of other results, for instance there is easily made another functional $h \in X^{'}$, by $h(x) = d \cdot g(x)$, such that

i. $h(x_0) = d$,
ii. $h(M) = 0$,
iii. $\| h \| = 1$.

And also that there exist a functional $k \in X^{'}$, such that

i. $k(x_0) \neq 0$,
ii. $k(M) = 0$,

of k is known, that $\| k \|$ is bounded, because $k \in X^{'}$.
Be careful with the choice of x_0, see **remark 4.2**.

4.7.1 Useful results with Hahn-Banach

There are enough bounded linear functionals on a Normed Space X to distinguish between the points of X.

Theorem 4.12

Let X be a Normed Space over the field $\mathbb{K}$ and let $0 \neq x_0 \in X$, then there exists a bounded linear functional $g \in X^{'}$ such that

i. $g(x_0) = \| x_0 \|$
ii. $\| g \| = 1$.

Proof of Theorem 4.12

Consider the linear subspace M spannend by x_0, $M = \{x \in X \,|\, x = \alpha\, x_0$ with $\alpha \in \mathbb{K}\}$ and define $f : M \rightarrow \mathbb{K}$ by

$$f(x) = f(\alpha x_0) = \alpha \parallel x_0 \parallel,$$

with $\alpha \in \mathbb{K}$. f is a linear functional on M and

$$|f(x)| = |f(\alpha x_0)| = |\alpha| \parallel x_0 \parallel = \parallel x \parallel$$

for every $x \in M$. Hence, f is bounded and $\parallel f \parallel = 1$.
By the theorem of Hahn-Banach, theorem 4.10, there exists a functional $g \in X^{'}$, such that $g|M = f$ and $\parallel g \parallel = \parallel f \parallel$. Hence, $g(x_0) = f(x_0) = \parallel x_0 \parallel$, and $\parallel g \parallel = 1$. □

Theorem 4.13

Let X be a Normed Space over the field $\mathbb{K}$ and $x \in X$, then

$$\parallel x \parallel = \sup\{\frac{|f(x)|}{\parallel f \parallel} \,|\, f \in X^{'} \text{ and } f \neq 0\}.$$

Proof of Theorem 4.13

The case that $x = 0$ is trivial.
Let $0 \neq x \in X$. With theorem 4.12 there exists a $g \in X'$, such that $g(x) = \parallel x \parallel$, and $\parallel g \parallel = 1$. Hence,

$$\sup\{\frac{|f(x)|}{\parallel f \parallel} \,|\, f \in X^{'} \text{ and } f \neq 0\} \geq \frac{|g(x)|}{\parallel g \parallel} = \parallel x \parallel . \tag{4.28}$$

Further,

$$|f(x)| \leq \parallel f \parallel \parallel x \parallel,$$

for every $f \in X^{'}$, therefore

$$\sup\{\frac{|f(x)|}{\parallel f \parallel} \,|\, f \in X^{'} \text{ and } f \neq 0\} \leq \parallel x \parallel . \tag{4.29}$$

The inequalities 4.28 and 4.29 complete the proof. □

4.8 The dual space X'' of a Normed Space X

The dual space X' has its own dual space X'', the second dual space of X, it is also called the bidual space of X. If the Vector Space X is finite dimensional then $R(C) = X^{**}$, where $R(C)$ is the range of the canonical mapping C of X to X^{**}.
In the infinite dimensional case, there can be proved that the canonical mapping C is onto some subspace of X''. In general $R(C) = C(X) \subseteq X''$ for every Normed Space X. The second dual space X'' is always complete, see **theorem 7.8**. So completeness of the space X is essential for the Normed Space X to be reflexive ($C(X) = X''$), but not enough. Completenes of the space X is a neccessary condition to be reflexive, but not sufficient.
It is clear that when X is not a Banach Space then X is non-reflexive, $C(X) \neq X''$.
With the theorem of Hahn-Banach, **theorem 4.10**, is derived that the dual space X' of a normed space X has enough bounded linear functionals to make a distinguish between points of X. A result that is necessary to prove that the canonical mapping C is unique.
To prove reflexivity, the canonical mapping is needed. There are examples of spaces X and X'', which are isometrically isomorphic with another mapping then the canonical mapping, but with X non-reflexive.

Theorem 4.14

Let X be a Normed Space over the field $\mathbb{K}$. Given $x \in X$ en let

$$g_x(f) = f(x),$$

for every $f \in X'$. Then g_x is a bounded linear functional on X', so $g_x \in X''$. The mapping $C : x \to g_x$ is an isometry of X onto the subspace $Y = \{g_x \,|\, x \in X\}$ of X''.

Proof of Theorem 4.14

The proof is splitted up in several steps.

1. Several steps are already done in section 4.4.4. The linearity of $g_x : X^{'} \rightarrow X^{''}$ and $C : X \rightarrow X^{''}$ that is not a problem.
The functional g_x is bounded, since

$$|g_x(f)| = |f(x)| \leq \| f \| \| x \|,$$

for every $f \in X^{'}$, so $g_x \in X^{''}$.

2. To every $x \in X$ there is an unique g_x. Suppose that $g_x(f) = g_y(f)$ for every $f \in X^{'}$ then $f(x - y) = 0$ for every $f \in X^{'}$. Hence, $x = y$, see **theorem 4.13**. Be careful the normed space X is may be not finite dimensional anymore, so **theorem 4.4** cannot be used. Hence, the mapping C is injective.
3. The mapping C preserves the norm, because

$$\| C(x) \| = \| g_x \| = \sup\{\frac{|g_x(f)|}{\| f \|} \,|\, f \in X^{'} \text{ and } f \neq 0\} = \| x \|,$$

see **theorem 4.13**.

Hence, C is an isometric isomorphism of X onto the subspace Y $(= C(X))$ of $X^{''}$. □

Some other terms are for instance for the canonical mapping: the natural embedding and for the functional $g_x \in X^{''}$: the functional induced by the vector x. The functional g_x is an induced functional . With the canonical mapping it is allowed to regard X as a part of $X^{''}$ without altering its structure as a Normed Space.

Theorem 4.15

Let $(X, \| \cdot \|)$ be a Normed Space. If $X^{'}$ is separable *then* X is separable.

Proof of Theorem 4.15

The proof is splitted up in several steps.

1. First is searched for a countable set S of elements in X, such that possible $\overline{S} = X$, see **Step ii: 1**.
2. Secondly there is proved, by a contradiction, that $\overline{S} = X$, see **Step ii: 2**.

Step 1: Because $X^{'}$ is separable, there is a countable set $M = \{f_n \in X^{'} | n \in \mathbb{N}\}$ which is dense in $X^{'}$, $\overline{M} = X^{'}$. By **definition 4.4**, $\| f_n \| = \sup_{\|x\|=1} |f_n(x)|$, so there exist a $x \in X$, with $\| x \| = 1$, such that for small $\epsilon > 0$

$$\| f_n \| - \epsilon \| f_n \| \leq |f_n(x)|,$$

with $n \in \mathbb{N}$. Take $\epsilon = \frac{1}{2}$ and let $\{v_n\}$ be sequence such that

$$\| v_n \| = 1 \text{ and } \frac{1}{2} \| f_n \| \leq |f_n(v_n)|.$$

Let S be the subspace of X generated by the sequence $\{v_n\}$,

$$S = \text{span}\{v_n | n \in \mathbb{N}\}.$$

Step 2: Assume that $\overline{S} \neq X$, then there exists a $w \in X$ and $w \notin \overline{S}$. An immediate consequence of the **formulas 4.27** is that there exists a functional $g \in X^{'}$ such that

$$\begin{aligned} g(w) &\neq 0, \\ g(\overline{S}) &= 0, \\ \| g \| &= 1. \end{aligned} \quad (4.30)$$

In particular $g(v_n) = 0$ for all $n \in \mathbb{N}$ and

$$\begin{aligned} \frac{1}{2} \| f_n \| \leq |f_n(v_n)| &= |f_n(v_n) - g(v_n) + g(v_n)| \\ &\leq |f_n(v_n) - g(v_n)| + |g(v_n)|. \end{aligned}$$

Since $\| v_n \| = 1$ and $g(v_n) = 0$ for all $n \in \mathbb{N}$, it follows that

$$\frac{1}{2} \| f_n \| \leq \| f_n - g \| . \quad (4.31)$$

Since M is dense in $X^{'}$, choose f_n such that

$$\lim_{n \to \infty} \| f_n - g \| = 0. \quad (4.32)$$

Using the formulas 4.30, 4.30 and 4.32, the result becomes that

$$\begin{aligned} 1 &= \| g \| = \| g - f_n + f_n \| \\ &\leq \| g - f_n \| + \| f_n \| \\ &\leq \| g - f_n \| + 2 \| g - f_n \|, \end{aligned}$$

such that

$$1 = \| g \| = 0.$$

Hence, the assumption is false and $\overline{S} = X$.

□

There is already known that the canonical mapping C is an isometric isomorphism of X onto the some subspace Y $(= C(X))$ of $X^{''}$, see theorem 4.14 and $X^{''}$ is a Banach Space.

Theorem 4.16

A Normed Space X is isometrically isomorphic to a dense subset of a Banach Space.

Proof of Theorem 4.16

The proof is not difficult.
X is a Normed Space and C is the canonical mapping $C : X \rightarrow X^{''}$.
The spaces $C(X)$ and X are isometrically isomorphic, and $C(X)$ is dense in $\overline{C(X)}$. $\overline{C(X)}$ is a closed subspace of the Banach Space $X^{''}$, so $\overline{C(X)}$ is a Banach Space, see theorem 3.12. Hence, X is isometrically isomorphic with to the dense subspace $C(X)$ of the Banach Space $\overline{C(X)}$. □

An nice use of theorem 4.15 is the following theorem.

Theorem 4.17

Let $(X, \| \cdot \|)$ be a separable Normed Space. If $X^{'}$ is non-separable *then* X is non-reflexive.

Proof of Theorem 4.17

The proof will be done by a contradiction.
Assume that X is reflexive. Then is $X^{''}$ isometrically isomorphic to X under the canonical mapping $C : X \to X^{''}$. X is separable, so $X^{''}$ is separable and with the use of **theorem 4.15**, the space $X^{'}$ is separable. But that contradicts the hypothesis that $X^{'}$ is non-separable. □

4.9 Weak and Weak* Convergence

Sometimes the convergence of sequences in the norm of a Normed Space X is too strong. Here will be introduced new modes of convergence of sequences in a Normed Space X and in its dual space $X^{'}$. In general, they are not as strong as the norm convergence, more freely available and useful.

Definition 4.6

Let X be a normed linear space with norm $\| \cdot \|$. Let $\{x_n\}_{n \in \mathbb{N}}$ be a sequence in X and $x \in X$, the sequence $\{x_n\}_{n \in \mathbb{N}}$ converges strongly or converges in norm to x, written as $x_n \to x$, if

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

Definition 4.7

Let X be a linear normed space with norm $\| \cdot \|$. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence in X and $x \in X$, the sequence $\{x_n\}_{n\in\mathbb{N}}$ converges weakly to x, written as $x_n \xrightarrow{w} x$, if

$$\lim_{n\to\infty} \mu(x_n) = \mu(x)$$

for every $\mu \in X'$.

Definition 4.8

Let X be a normed linear space with norm $\| \cdot \|$. Let $\{\mu_n\}_{n\in\mathbb{N}}$ be a sequence in X' and $\mu \in X'$, the sequence $\{\mu_n\}_{n\in\mathbb{N}}$ converges weak* to μ, written as $\mu_n \xrightarrow{w^*} x$, if

$$\lim_{n\to\infty} \mu_n(x) = \mu(x)$$

for every $x \in X$.

So, weak* convergence is just pointwise convergence of the operators μ_n.

Remark 4.4

Weak* convergence makes only sense for a sequence that lies in the dual space $X^{'}$ of X. If there is a sequence $\{\mu_n\}_{n \in \mathbb{N}}$ in $X^{'}$, there can be looked to three types of convergence of μ_n to μ. These are:

i. strong:

$$\mu_n \rightarrow \mu \iff \lim_{n \rightarrow \infty} \| \mu_n - \mu \| = 0,$$

with $\| \cdot \|$, the norm used in the dual space $X^{'}$,

ii. weak:

$$\mu_n \xrightarrow{w} \mu \iff \lim_{n \rightarrow \infty} T(\mu_n) = T(\mu)$$

for every $T \in X^{''}$,

iii. weak*:

$$\mu_n \xrightarrow{w^*} \mu \iff \lim_{n \rightarrow \infty} \mu_n(x) = \mu(x)$$

for every $x \in X$,

4.9.1 Schur's property and the Radon-Riesz or Kadets-Klee property

Definition 4.9

A Normed Space $(X, \| \cdot \|)$ has Schur's property if

$$x_n \xrightarrow{w} x \Rightarrow \lim_{n \rightarrow \infty} \| x_n - x \| = 0.$$

Definition 4.10

A Normed Space $(X, \|\cdot\|)$ has the Radon-Riesz or the Kadets-Klee property if

$$x_n \xrightarrow{w} x \text{ and } \lim_{n\to\infty} \| x_n \| = \| x \| = 0 \Rightarrow \lim_{n\to\infty} \| x_n - x \| = 0.$$

Theorem 4.18

The space ℓ^2 does not have Schur's property, see Definition 4.9, but has the Radon-Riesz property, see Definition 4.10.

Proof of Theorem 4.18

Let $\{e_n\}_{n\in\mathbb{N}}$ be the sequence of unit vectors in ℓ^2. The dual space of ℓ^2 can be identified by itself, see Theorem 5.15. It is clear that $x(e_n) \to 0$, $(n \to \infty)$ for each $x \in \ell^2$. This means that sequence $\{e_n\}_{n\in\mathbb{N}}$ converges weakly to 0. But $\| e_n \|_2 = 1$ for each $n \in \mathbb{N}$. So the sequence $\{e_n\}_{n\in\mathbb{N}}$ converges not in the norm to 0.
Thus the space ℓ^2 does not satisfy the Schur's property.
It is clear, that the weak topology and the norm topology, of ℓ^2, are different.

Let $\{x_n\}_{n\in\mathbb{N}}$ be sequence in ℓ^2 and $x \in \ell^2$ such that $x_n \xrightarrow{w} x$ and $\| x_n \|_2 \to \| x \|_2$, for $n \to \infty$. For each n, let $x_n = \{\alpha_n^k\}_{k\in\mathbb{N}}$ then

$$\begin{aligned}
&\| x_n - x \|_2^2 = \sum_{k=1}^{\infty} (\alpha_n^k - \alpha_n)\overline{(\alpha_n^k - \alpha_n)} = \\
&\sum_{k=1}^{\infty} | \alpha_n^k |^2 - \sum_{k=1}^{\infty} \alpha_n \overline{\alpha_n^k} - \sum_{k=1}^{\infty} \overline{\alpha_n} \alpha_n^k + \sum_{k=1}^{\infty} | \alpha_n |^2 = \\
&\| x_n \|_2^2 - x(\overline{x_n}) - x_n(x) + \| x \|_2^2 \to \\
&\| x \|_2^2 - x(\overline{x}) - \overline{x}(x) + \| x \|_2^2 = 0
\end{aligned}$$

for $n \rightarrow \infty$.

Thus, the space ℓ^2 satisfies the Radon-Riesz property.
□

Theorem 4.19

The space ℓ^1 satisfies Schur's property, see Definition 4.9.

Proof of Theorem 4.19

Assume that there exists a sequence $\{x_n\}_{n \in \mathbb{N}}$ in ℓ^1 such that $x_n \xrightarrow{w} 0$ but $x_n \nrightarrow 0$ in ℓ^1, so $\lim_{n \rightarrow \infty} \| x_n - 0 \|_1 \neq 0$.
If necessary, there can be looked to a subsequence of $\{x_n\}_{n \in \mathbb{N}}$, with an increasing sequence $n_1 < n_2 < \cdots$ such that

$$\| x_{n_j} \|_1 = \sum_{k=1}^{\infty} | x_{n_j}(k) | > \epsilon,$$

$x_{n_j}(k) = f_k(x_{n_j})$, with the $k-$th coordinate functional $f_k(x) = x(k)$, $x \in \ell^1$ and $k \in \mathbb{N}$.
Choose N_1 such that $\sum_{k=(N_1+1)}^{\infty} | x_{n_1}(k) | < \frac{1}{5}\epsilon$. This is possible since $x_{n_1} \in \ell^1$.
Then $\sum_{k=1}^{N_1} | x_{n_1}(k) | \geq \frac{4}{5}\epsilon$, this can also be written as
$\sum_{k=1}^{N_1} \epsilon_{n_1}(k) x_{n_1}(k) \geq \frac{4}{5}\epsilon$ with $\epsilon_{n_1}(k) = (\frac{x_{n_1}(k)}{| x_{n_1}(k) |})$ for $k = 1, \cdots, N_1$.
Choose an arbitrary sequence of signs $\{\epsilon_k = \pm 1\}$, but such that $\epsilon_k = \epsilon_{n_1}(k)$ for $k \leq N_1$, then

$$| \sum_{k=1}^{\infty} \epsilon_k x_{n_1}(k) | = | \sum_{k=1}^{N_1} \epsilon_{n_1}(k) x_{n_1}(k) + \sum_{k=(N_1+1)}^{\infty} \epsilon_k x_{n_1}(k) | \geq$$

$$| \sum_{k=1}^{N_1} \epsilon_{n_1}(k) x_{n_1}(k) | - \sum_{k=(N_1+1)}^{\infty} | x_{n_1}(k) | \geq \frac{4}{5}\epsilon - \frac{1}{5}\epsilon = \frac{3}{5}\epsilon.$$

Since $x_n \xrightarrow{w} 0$ for $n \to \infty$, then $f_k(x_n) \to 0$ for $n \to \infty$, so there exists a $n_{j_2} > n_1$ such that $\sum_{k=1}^{N_1} \mid x_{n_{j_2}}(k) \mid < \frac{1}{5}\epsilon$. Then choose $N_2 > N_1$ such that $\sum_{k=(N_2+1)}^{\infty} \mid x_{n_{j_2}}(k) \mid < \frac{1}{5}\epsilon$ and consequently $\sum_{k=1}^{N_2} \mid x_{n_{j_2}}(k) \mid \geq \frac{4}{5}\epsilon$.
Then for arbitrary choice of signs $\{\epsilon_k = \pm 1\}$ satisfying $\epsilon_k = \epsilon_{n_1}(k)$ for $k \leq N_1$ and $\epsilon_k = \epsilon_{n_{j_2}}(k)$ for $N_1 < k < N_2$ follows that

$$\mid \sum_{k=1}^{\infty} \epsilon_k x_{n_{j_2}}(k) \mid \geq \mid \sum_{k=(N_1+1)}^{N_2} \epsilon_k x_{n_{j_2}}(k) \mid - \sum_{k=1}^{N_1} \mid x_{n_{j_2}}(k) \mid - \sum_{k=(N_2+1)}^{\infty} \mid x_{n_{j_2}}(k) \mid \geq$$
$$\mid \sum_{k=(N_1+1)}^{N_2} \epsilon_k x_{n_{j_2}}(k) \mid - \frac{2}{5}\epsilon = \sum_{k=1}^{N_2} \mid x_{n_{j_2}}(k) \mid - \sum_{k=1}^{N_1} \mid x_{n_{j_2}}(k) \mid - \frac{2}{5}\epsilon \geq$$
$$\frac{4}{5}\epsilon - \frac{1}{5}\epsilon - \frac{2}{5}\epsilon = \frac{1}{5}\epsilon.$$

Repeating this process, there is an element $w = \{w_k\}_{k \in \mathbb{N}} \in \ell^{\infty}$ constructed, with $w_k = \epsilon_{n_{jm}}(k)$ for $N_{(m-1)} < k \leq N_m$. The dual space of ℓ^1 is equal to ℓ^{∞}, see **subsection 4.6.1**. The constructed element $w \in \ell^{\infty}$ has the property that

$$w(x^{n_{jm}}) > \frac{1}{5}\epsilon,$$

for all m, but that is in contradiction with $x_n \xrightarrow{w} 0$.

This proof has been found in the book of (Fabian, 2001).

□

5 Example Spaces

There are all kind of different spaces, which can be used as illustration for particular behaviour of convergence or otherwise.

5.1 Function Spaces

The function spaces are spaces, existing out of functions, which have a certain characteristic or characteristics. Characteristics are often described in terms of norms. Different norms can be given to a set of functions and so the same set of functions can get a different behaviour.
In first instance the functions are assumed to be real-valued. Most of the given spaces can also be defined for complex-valued functions.
Working with a Vector Space means that there is defined an addition and a scalar multiplication. Working with Function Spaces means that there has to be defined a summation between functions and a scalar multiplication of a function.
Let $\mathbb{K}$ be the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. Let I be an open interval (a, b), or a closed interval $[a, b]$ or may be $\mathbb{R}$ and look at the set of all functions $S = \{f \mid f : I \to \mathbb{K}\}$.

Definition 5.1

Let $f, g \in S$ and let $\alpha \in \mathbb{K}$.
The addition $(+)$ between the functions f and g and the scalar multiplication $(\cdot)$ of α with the function f are defined by:

addition $(+)$: $(f + g)$ means $(f + g)(x) := f(x) + g(x)$ for all $x \in I$,
scalar m. $(\cdot)$: $(\alpha \cdot f)$ means $(\alpha \cdot f)(x) := \alpha(f(x))$ for all $x \in I$.

The symbol $:=$ means: is defined by.

The quartet $(S, \mathbb{K}, (+), (\cdot))$, with the above defined addition and scalar multiplication, is a Vector Space.
The Vector Space $(S, \mathbb{K}, (+), (\cdot))$ is very big, it exists out of all the functions defined on the interval I and with their function values in $\mathbb{K}$. Most of the time

is looked at subsets of the Vector Space $(S, \mathbb{K}, (+), (\cdot))$. For instance there is looked at functions which are continuous on I, have a special form, or have certain characteristic described by integrals. If characteristics are given by certain integrals the continuity of such functions is often dropped.
To get an Inner Product Space or a Normed Space there has to be defined an inner product or a norm on the Vector Space, that is of interest on that moment.

5.1.1 Polynomials

A polynomial p of degree less or equal to n is written in the following form

$$p_n(t) \;=\; a_0 + a_1\, t + \cdots + a_n\, t^n \;=\; \sum_{i=0}^{n} a_i\, t^i.$$

If p_n is exactly of the degree n, it means that $a_n \neq 0$. A norm, which can be defined on this space of polynomials of degree less or equal to n is

$$\| \, p_n \, \| = \max_{i=0,\cdots,n} \mid a_i \mid . \tag{5.1}$$

Polynomials have always a finite degree, so $n < \infty$. Looking to these polynomials on a certain interval $[a, b]$, then another norm can be defined by

$$\| \, p_n \, \|_\infty = \sup_{a \leq t \leq b} \mid p_n(t) \mid,$$

the so-called sup-norm, on the interval $[a, b]$.
With $\mathbb{P}_N([a, b])$ is meant the set of all polynomial functions on the interval $[a, b]$, with a degree less or equal to N. The number $N \in \mathbb{N}$ is a fixed number.
With $\mathbb{P}([a, b])$ is meant the set of all polynomial functions on the interval $[a, b]$, which have a finite degree.

5.1.2 $C([a, b])$ with $\| \cdot \|_\infty$-norm

The normed space of all continuous function on the closed and bounded interval $[a,b]$. The norm is defined by

$$\| f \|_{\infty} = \sup_{x \in [a,b]} | f(x) | . \tag{5.2}$$

and is often called the sup-norm of the function f at the interval $[a,b]$.

Dense subspaces are of importance, also in the Normed Space $(C([a,b]), \| \cdot \|_{\infty})$. After that an useful formula is proved, it will be shown that the set $\mathbb{P}([a,b])$ is dense in $(C([a,b]), \| \cdot \|_{\infty})$. This spectacular result is know as the Weierstrass Approximation Theorem.

Theorem 5.1

Let $n \in \mathbb{N}$ and let t be a real parameter then

$$\sum_{k=0}^{n} (t - \frac{k}{n})^2 \binom{n}{k} t^k (1-t)^{(n-k)} = \frac{1}{n} t (1-t)$$

Proof of Theorem 5.1

First is defined the function $G(s)$ by

$$G(s) = (s t + (1-t))^n, \tag{5.3}$$

using the binomial formula, the function $G(s)$ can be rewritten as

$$G(s) = \sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{(n-k)} s^k. \tag{5.4}$$

Differentiating the formulas 5.3 and 5.4 to s results in

$$G^{'}(s) = n\, t\, (s t + (1-t))^{n-1} = \sum_{k=0}^{n} k \binom{n}{k} t^k (1-t)^{(n-k)} s^{k-1}$$

and

$$G^{''}(s) = n(n-1)\, t^2 \,(st + (1-t))^{n-1} = \sum_{k=0}^{n} k(k-1) \binom{n}{k} t^k (1-t)^{(n-k)}\, s^{k-2}.$$

Take $s = 1$ and the following functions values are obtained:

$$G(1) = 1 = \sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{(n-k)},$$

$$G^{'}(1) = n\, t = \sum_{k=0}^{n} k \binom{n}{k} t^k (1-t)^{(n-k)},$$

$$G^{''}(1) = n(n-1)\, t^2 = \sum_{k=0}^{n} k(k-1) \binom{n}{k} t^k (1-t)^{(n-k)}.$$

The following computation

$$\begin{aligned}
&\sum_{k=0}^{n} (t - \frac{k}{n})^2 \binom{n}{k} t^k (1-t)^{(n-k)} \\
&= \sum_{k=0}^{n} (t^2 - 2\,\frac{k}{n} t + (\frac{k}{n})^2\, t^2) \binom{n}{k} t^k (1-t)^{(n-k)} \\
&= t^2\, G(1) - \frac{2}{n}\, t\, G^{'}(1) + \frac{1}{n^2} G^{''}(1) + \frac{1}{n^2} G^{'}(1) \\
&= t^2 - \frac{2}{n} tnt + \frac{1}{n^2} n(n-1) t^2 + \frac{1}{n^2} nt \\
&= \frac{1}{n}\, t\,(1 - t),
\end{aligned}$$

completes the proof. □

If a and b are finite, the interval $[a, b]$ can always be rescaled to the interval $[0, 1]$, by $t = \frac{x - a}{b - a}$, $0 \leq t \leq 1$ if $x \in [a, b]$. Therefore will now be looked at the Normed Space $(C([0, 1]), \| \cdot \|_{\infty})$.
The Bernstein polynomials $p_n(f) : [0, 1] \rightarrow \mathbb{R}$ are defined by

$$p_n(f)(t) = \sum_{k=0}^{n} f(\frac{k}{n}) \binom{n}{k} t^k (1-t)^{(n-k)} \tag{5.5}$$

with $f \in C[0,1]$ and are used to proof the following theorem, also known as the Weierstrass Approximation Theorem.

Theorem 5.2

The Normed Space $(C([0,1]), \|\cdot\|_\infty)$ is the completion of the Normed Space $(\mathbb{P}([0,1]), \|\cdot\|_\infty)$.

Proof of Theorem 5.2

Let $\epsilon > 0$ be given and an arbitrary function $f \in C[0,1]$. f is continuous on the compact interval $[0,1]$, so f is uniformly continuous on $[0,1]$, see **theorem 2.10**. Further f is bounded on the compact interval $[0,1]$, see **theorem 2.9**, so let

$$\sup_{t \in [0,1]} |\, f(t)\,| = M.$$

Since f is uniformly continuous, there exists some $\delta > 0$ such that for every $t_1, t_2 \in [0,1]$ with $|\, t_1 - t_2\,| < \delta, |\, f(t_1) - f(t_2)\,| < \epsilon$. Important is that δ only depends on ϵ. Using

$$1 = (t + (1-t))^n = \sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{(n-k)}$$

the following computation can be done for some arbitrary $t \in [0,1]$

$$\begin{aligned} |\, f(t) - p_n(f)(t)\,| &= |\sum_{k=0}^{n} (f(t) - f(\frac{k}{n})) \binom{n}{k} t^k (1-t)^{(n-k)}| \\ &\leq \sum_{k=0}^{n} |(f(t) - f(\frac{k}{n}))| \binom{n}{k} t^k (1-t)^{(n-k)} \end{aligned}$$

The fact that δ depends only on ϵ makes it useful to split the summation into two parts, one part with $|\, t - \frac{k}{n}\,| < \delta$ and the other part with $|\, t - \frac{k}{n}\,| \geq \delta$. On the first part will be used the uniform continuity of f and on the other part will be used the boundedness of f, so

$$| f(t) - p_n(f)(t)| \leq \sum_{|t-\frac{k}{n}|<\delta} |(f(t) - f(\frac{k}{n}))| \binom{n}{k} t^k(1-t)^{(n-k)}$$

$$+ \sum_{|t-\frac{k}{n}|\geq\delta} |(f(t) - f(\frac{k}{n}))| \binom{n}{k} t^k(1-t)^{(n-k)}$$

$$\leq \sum_{k=0}^{n} \epsilon \binom{n}{k} t^k(1-t)^{(n-k)} + \sum_{|t-\frac{k}{n}|\geq\delta} 2\,M \binom{n}{k} t^k(1-t)^{(n-k)}.$$

The fact that $|t - \dfrac{k}{n}| \geq \delta$ means that

$$1 \leq \frac{|t - \frac{k}{n}|}{\delta} \leq \frac{(t - \frac{k}{n})^2}{\delta^2}. \tag{5.6}$$

Inequality **5.6** and the use of **theorem 5.1** results in

$$| f(t) - p_n(f)(t)| \leq \epsilon + \frac{2\,M}{\delta^2} \sum_{k=0}^{n} (t - \frac{k}{n})^2 \binom{n}{k} t^k(1-t)^{(n-k)}$$

$$= \epsilon + \frac{2\,M}{\delta^2} \frac{1}{n} t(1-t)$$

$$\leq \epsilon + \frac{2\,M}{\delta^2} \frac{1}{n} \frac{1}{4},$$

for all $t \in [0,1]$. The upper bound $(\epsilon + \dfrac{2\,M}{\delta^2} \dfrac{1}{n} \dfrac{1}{4})$ does not depend on t and for
$n > \dfrac{M}{2\,\delta^2\,\epsilon}$, this implies that

$$\| f(t) - p_n(f)(t) \|_\infty < 2\,\epsilon.$$

The consequence is that

$$p_n(f) \to f \text{ for } n \to \infty \text{ in } (C([0,1]), \|\cdot\|_\infty).$$

Since f was arbitrary, it follows that $\overline{\mathbb{P}([0,1])} = C([0,1])$, in the $\|\cdot\|_\infty$-norm, and the proof is complete. □

Theorem 5.3

The Normed Space $(C([a,b]), \|\cdot\|_\infty)$ is separable.

Proof of Theorem 5.3

According the Weierstrass Approximation Theorem, **theorem 5.2**, every continous function f on the bounded and closed interval $[a,b]$, can be approximated by a sequence of polynomials $\{p_n\}$ out of $(\mathbb{P}([a,b]), \|\cdot\|_\infty)$. The convergence is uniform, see **section 2.12**. The coeffcients of these polynomials can be approximated with rational coefficients, since $\mathbb{Q}$ is dense in $\mathbb{R}$ ($\overline{\mathbb{Q}} = \mathbb{R}$). So any polynomial can be uniformly approximated by a polynomial with rational coeficients.
The set $\mathbb{P}_{\mathbb{Q}}$ of all polynomials on $[a,b]$, with rational coefficients, is a countable set and $\overline{\mathbb{P}_{\mathbb{Q}}([a,b])} = C[a,b]$. □

Theorem 5.4

The Normed Space $(C([a,b]), \|\cdot\|_\infty)$ is a Banach Space.

Proof of Theorem 5.4

See **Section 2.12**. □

5.1.3 $C([a,b])$ with $\mathbb{L}_p$-norm and $1 \leq p < \infty$

The normed space of all continuous function on the closed and bounded interval $[a,b]$. The norm is defined by

$$\| f \|_p = \left(\int_a^b | f(x) |^p \; dx \right)^{\left(\frac{1}{p}\right)}. \tag{5.7}$$

and is often called the $\mathbb{L}_p$-norm of the function f at the interval $[a, b]$.

5.1.4 $C([a, b])$ with $\mathbb{L}_2$-inner product

The inner product space of all continuous function on the closed and bounded interval $[a, b]$. Let $f, g \in C([a, b])$ then it is easily to define the inner product between f and g by

$$(f, g) = \int_a^b f(x) \, g(x) \, dx \tag{5.8}$$

and it is often called the $\mathbb{L}_2$-inner product between the functions f and g at the interval $[a, b]$. With the above defined inner product the $\mathbb{L}_2$-norm can calculated by

$$\| f \|_2 = (f, f)^{\frac{1}{2}}. \tag{5.9}$$

When the functions are complex-valued then the inner product has to be defined by

$$(f, g) = \int_a^b f(x) \, \overline{g(x)} \, dx. \tag{5.10}$$

The value of $\overline{f(x)}$ is the complex conjugate of the value of $f(x)$.

5.1.5 $\mathbb{L}_p(a, b)$ with $1 \leq p < \infty$

In the section 5.1.3 and 5.1.4 there are taken functions which are continuous on the closed and bounded interval $[a, b]$. To work with more generalized functions, the continuity can be dropped and there can be looked at classes of functions on the open interval (a, b). The functions $f, g \in \mathbb{L}_p(a, b)$ belong to the same class in $\mathbb{L}_p(a, b)$ if and only if

$$\| f - g \|_p = 0.$$

The functions f and g belong to $\mathbb{L}_p(a, b)$, if $\| f \|_p < \infty$ and $\| g \|_p < \infty$. With the Lebesgue integration theory, the problems are taken away to calculate the given integrals. Using the theory of Riemann integration gives problems. For more information about these different integration techniques, see for instance Chen-2 and see section 5.1.6.
From the Lebesgue integration theory it is known that

$$\| f - g \|_p = 0 \Leftrightarrow f(x) = g(x) \textit{ almost everywhere.}$$

With *almost everywhere* is meant that the set $\{x \in (a, b) \mid f(x) \neq g(x)\}$ has measure 0, for more information see wiki-measure.

Example 5.1

An interesting example is the function $f \in \mathbb{L}_p(0, 1)$ defined by

$$f(x) = \begin{cases} 1 & \text{for } x \in \mathbb{Q} \\ 0 & \text{for } x \notin \mathbb{Q} \end{cases} \tag{5.11}$$

This function f is equal to the zero-function *almost everywhere*, because $\mathbb{Q}$ is countable.

Very often the expression $\mathbb{L}_p(a, b)$ is used, but sometimes is also written $\mathcal{L}_p(a, b)$. What is the difference between these two spaces? Let's assume that $1 \leq p < \infty$.
First of all, most of the time there will be written something like $\mathcal{L}_p(\Omega, \Sigma, \mu)$, instead of $\mathcal{L}_p$. In short, Ω is a subset out of some space. Σ is a collection of subsets of Ω and these subsets satisfy certain conditions. And μ is called a measure, with μ the elements of Σ can be given some number (they can be measured), for more detailed information about the triplet (Ω, Σ, μ), see page 270. In this simple case, $\Omega = (a, b)$, for Σ can be thought to the set of open subsets of (a, b) and for μ can be thought to the absolute value $| \cdot |$. Given

are very easy subsets of $\mathbb{R}$, but what to do in the case $\Omega = (\mathbb{R} \setminus \mathbb{Q}) \cap (a, b)$? How to measure the length of a subset? May be the function defined in 5.1 can be used in a proper manner.
A function $f \in \mathcal{L}_p(\Omega, \Sigma, \mu)$ satisfies

$$N_p(f) = (\int_\Omega | f |^p \, d\mu)^{\frac{1}{p}} < \infty. \tag{5.12}$$

Now is the case, that there exist functions $g \in \mathcal{L}_p(\Omega, \Sigma, \mu)$, which have almost the same look as the function f. There can be defined an equivalence relation between f and g,

$$f \sim g \quad \text{if} \quad N_p(f - g) = 0, \tag{5.13}$$

the functions f and g are said to be equal *almost everywhere*, see page 271. With the given equivalence relation, it is possible to define equivalence classes of functions.
Another way to define these equivalence classes of functions is to look at all those functions which are *almost everywhere* equal to the zero function

$$\ker(N_p) = \{f \in \mathcal{L}_p \mid N_p(f) = 0\}.$$

So be careful! If $N_p(f) = 0$, it does not mean that $f = 0$ everywhere, but it means, that the set $\{x \in \Omega \mid f(x) \neq 0\}$ has measure zero. So the expression N_p is not really a norm on the space $\mathcal{L}_p(\Omega, \Sigma, \mu)$, but a seminorm, see **definition 3.24**. The expression N_p becomes a norm, if the $\ker(N_p)$ is divided out of the space $\mathcal{L}_p(\Omega, \Sigma, \mu)$.
So it is possible to define the space $\mathbb{L}_p(\Omega, \Sigma, \mu)$ as the quotient space (see **section 3.11**) of $\mathcal{L}_p(\Omega, \Sigma, \mu)$ and $\ker(N_p)$

$$\mathbb{L}_p(\Omega, \Sigma, \mu) = \mathcal{L}_p(\Omega, \Sigma, \mu) \,/\, \ker(N_p).$$

The Normed Space $\mathbb{L}_p(\Omega, \Sigma, \mu)$ is a space of equivalence classes and the norm is given by the expression N_p in **5.12**. The equivalence relation is given by **5.13**.
Be still careful! $N_p(f) = 0$ means that in $\mathbb{L}_p(\Omega, \Sigma, \mu)$ the zero-function can be taken as représentant of all those functions with $N_p(f) = 0$, but f has not to be zero everywhere. The zero-function represents an unique class of functions in $\mathbb{L}_p(\Omega, \Sigma, \mu)$ with the property that $N_p(f) = 0$.
More interesting things can be found at the internet site **wiki-Lp-spaces** and see also (Bouziad and Clabrix, 1993, page 109).

5.1.6 Riemann integrals and Lebesgue integration

To calculate the following integral

$$\int_a^b f(x)\,dx,$$

with a nice and friendly function f, most of the time the the method of Riemann is used. That means that the domain of f is partitioned into pieces, for instance
$\{a = x_0 < x_1 < x_2 < \cdots x_0 < x_n = b\}$. On a small piece $x_{i-1} < x < x_i$ is taken some x and $c_i = f(x)$ is calculated, this for $i = 1, \cdots, n$. The elementary integral is then defined by,

$$\int_a^b f(x)\,dx \doteq \sum_{i=1}^{n} c_i\,(x_i - x_{i-1}). \tag{5.14}$$

With $\doteq$ is meant that the integral is approximated by the finite sum on the right-side of formula 5.14. For a positive function this means the area beneath the graphic of that function, see figure 5.1. How smaller the pieces $x_{i-1} < x < x_i$, how better the integral is approximated.
Step functions are very much used to approximate functions. An easy example of the step fuction is the function ψ with $\psi(x) = c_i$ at the interval $x_{i-1} < x < x_i$ then

$$\int_a^b \psi(x)\,dx = \sum_{i=1}^{n} c_i\,(x_i - x_{i-1}).$$

How smaller the pieces $x_{i-1} < x < x_i$, how better the function f is approximated.
Another way to calculate that area beneath the graphic of a function is to partition the range of a function and then to ask how much of the domain is mapped between some endpoints of the range of that function. Partitioning the range, instead of the domain, is called the method of Lebesgue .
Lebesgue integrals solves many problems left by the Riemann integrals.
To have a little idea about how Lebesgue integrals are calculated, the characteristic functions are needed. On some subset A, the characteristic function χ_A is defined by

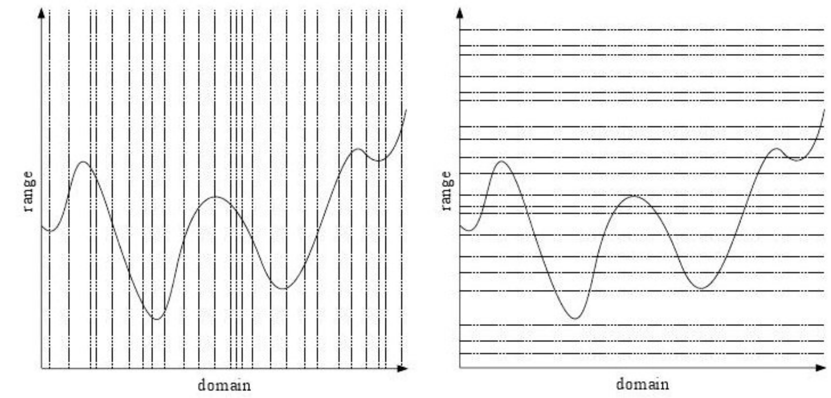


Figure 5.1 Left: Riemann-integral, right: Lebesgue-integral.

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \qquad (5.15) \\ 0 & \text{if } x \notin A. \end{cases} \tag{5.16}$$

As already mentioned the range of a function is partitioned in stead of it's domain. The range can be partitioned in a similar way as the domain is partitioned in the Riemann integral. The size of the intervals have not to be the same, every partition is permitted.

A simple example, let f be a positive function and continous. Consider the finite collection of subsets B defined by

$$B_{n,\alpha} = \{x \in [a, b] \,|\, \frac{\alpha - 1}{2^n} \leq f(x) < \frac{\alpha}{2^n}\}$$

for $\alpha = 1, 2, \cdots, 2^{2\,n}$, see figure 5.2,

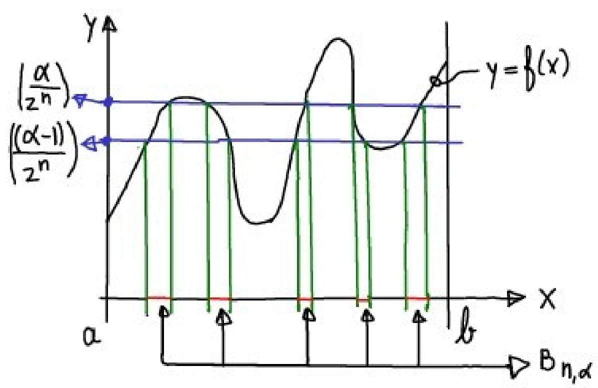


Figure 5.2 A subset $B_{n,\alpha}$.

and if $\alpha = 1 + 2^{2n}$

$$B_{n,1+2^{2n}} = \{x \in [a,b] \,|\, f(x) \geq 2^n\}.$$

Define the sequence $\{f_n\}$ of functions by

$$f_n = \sum_{\alpha=1}^{1+2^{2n}} \frac{(\alpha - 1)}{2^n} \chi_{B_{n,\alpha}}. \tag{5.17}$$

It is easily seen that the sequence $\{f_n\}$ converges (pointwise) to f at the interval $[a,b]$. The function f is approximated by step functions.
The sets $B_{n,\alpha}$, which have a certain length (have a certain measure), are important to calculate the integral. May be it is interesting to look at the internet site **wiki-measures**, for all kind of measures . Let's notate the measure of $B_{n,\alpha}$ by $m(B_{n,\alpha})$. In this particular case, the function f is continous on a closed and bounded interval, so f is bounded. Hence, only a limited part of $B_{n,\alpha}$ will have a measure not equal to zero.
The function f_n is a finite sum, so

$$\int_a^b f_n(x)\,dx = \sum_{\alpha=1}^{1+2^{2n}} \frac{(\alpha - 1)}{2^n} m(\chi_{B_{n,\alpha}}).$$

In this particular case,

$$\lim_{n\to\infty} \sum_{\alpha=1}^{1+2^{2n}} \frac{(\alpha - 1)}{2^n} m(\chi_{B_{n,\alpha}}) = \int_a^b f(x)\,dx, \tag{5.18}$$

but be careful in all kind of other situations, for instance if f is not continuous or if the interval $[a,b]$ is not bounded, etc.

5.1.7 Fundamental convergence theorems of integration

The following theorems are very important in the case that the question becomes if the limit and the integral sign may be changed or not. There will be tried to give an outline of the proofs of these theorems. Be not disturbed and try to read the outlines of the proofs. See for more information, for instance (Royden, 1988) or (Kolmogorv and Fomin, 1961).

The best to do, is first to give two equivalent definitions of the Lebesgue integral. With equivalent is meant that, when f satisfies the conditions given in the definitions, both integrals give the same value. To understand the definitions there can be thought to the Riemann integrals which, if possible, are approximated by lower- and under-sums. With the Lebesgue integration there is worked with simple functions, see for instance **formula 5.17** or **theorem 8.5**. There is looked at simple functions ψ with $\psi \geq f$ or at simple functions φ with $\varphi \leq f$.

Definition 5.2

Let f be a bounded measurable function defined on a measurable set E, with $\mu(E) < \infty$. The Lebesgue integral of f over E is defined by

$$\int_E f \, d\mu = \inf \left(\int_E \psi \, d\mu \right)$$

for all simple functions $\psi \geq f$.

Definition 5.3

Let f be a bounded measurable function defined on a measurable set E, with $\mu(E) < \infty$. The Lebesgue integral of f over E is defined by

$$\int_E f \, d\mu = \sup \left(\int_E \varphi \, d\mu \right)$$

for all simple functions $\varphi \leq f$.

Theorem 5.5

Lebesgue's Bounded Convergence Theorem

Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions defined on a set E of finite measure. Suppose that there exists a positive real number M such that $\mid f_n(x) \mid\leq M$ for every n and for all $x \in E$. If $\lim_{n\to\infty} f_n(x) = f(x)$ for every $x \in E$, then

$$\lim_{n\to\infty} \int_E f_n \, \mathbf{d}\mu = \int_E f \, \mathbf{d}\mu.$$

Proof of Theorem 5.5

Let $\epsilon > 0$ be given.

a. **In the case of uniform convergence**:
If the sequence $\{f_n\}_{n\in\mathbb{N}}$ should converge uniformly to f, then there would be no problem to change the integral sign and the limit. In that case there is some $N(\epsilon)$ such that for all $n > N(\epsilon)$ and for all $x \in E$, $\mid f_n(x) - f(x) \mid< \epsilon$. Thus

$$\mid \int_E f_n \mathrm{d}\mu - \int_E f \mathrm{d}\mu \mid\leq \int_E \mid f_n - f \mid \mathrm{d}\mu < \epsilon\mu(E).$$

b. **Pointwise convergence and uniformly bounded**:
The sequence $\{f_n\}_{n\in\mathbb{N}}$ converges pointwise and is uniformly bounded. Define the sets

$$H_n = \{x \in E \mid \mid f_n(x) - f(x) \mid\geq \epsilon\}$$

and let

$$G_N = \cup_{n=N}^{\infty} H_n.$$

It is easily seen $G_{N+1} \subset G_N$ and for each $x \in E$ there is some N such that $x \in G_N$, since $\lim_{n\to\infty} f_n(x) = f(x)$. Thus $\cap_{n=1}^{\infty} G_n = \varnothing$, so $\lim_{n\to\infty} \mu(G_n) = 0$.
Given some $\delta > 0$, there is some N such that $\mu(G_N) < \delta$.

c. **Difference between the integrals**:
Take $\delta = \frac{\epsilon}{4M}$, then there is some $N(\epsilon)$ such that $\mu(G_{N(\epsilon)}) < \delta$.

So there is a measurable set $A = G_{N(\epsilon)} \subset E$, with $\mu(A) < \frac{\epsilon}{4M}$, such that for $n > N(\epsilon)$ and $x \in E \backslash A$ $\mid f_n(x) - f(x) \mid < \epsilon$. Thus

$$\mid \int_E f_n \mathrm{d}\mu - \int_E f \mathrm{d}\mu \mid = \mid \int_E (f_n - f) \mathrm{d}\mu \mid$$
$$\leq \int_E \mid f_n - f \mid \mathrm{d}\mu$$
$$= \int_{E \backslash A} \mid f_n - f \mid \mathrm{d}\mu + \int_A \mid f_n - f \mid \mathrm{d}\mu$$
$$\epsilon\mu(E \backslash A) + 2M\,\mu(A) < \epsilon\mu(E) + \frac{\epsilon}{2}$$

Hence the proof is completed, since $\int_E f_n \mathrm{d}\mu \to \int_E f \mathrm{d}\mu$ for $n \to \infty$.

□

Remark 5.1

Theorem 5.5 is not valid for the Riemann integral,
see example 5.1. May be it is good to remark that the function given in example 5.1 is *nowhere* continuous.

Theorem 5.6

Fatou's Lemma

Let $\{f_n\}_{n \in \mathbb{N}}$ be a sequence of nonnegative measurable functions and $\lim_{n\to\infty} f_n(x) = f(x)$ *almost everywhere* on a set E. Then

$$\int_E f \,\mathrm{d}\mu \leq \liminf_{n\to\infty} \int_E f_n \,\mathrm{d}\mu \tag{5.19}$$

Proof of Theorem 5.6

If $\lim_{n\to\infty} f_n(x) = f(x)$ *almost everywhere* on a set E, it means that there exist a set $N \subset E$ with $\mu(N) = 0$ such that $\lim_{n\to\infty} f_n(x) = f(x)$ is *everywhere* on $E' = (E \backslash N)$. And integrals over sets with measure zero are zero. Let h be a measurable function, bounded by f and zero outside E', a set of finite measure. Define the sequence of functions $\{h_n\}_{n\in\mathbb{N}}$ by

$$h_n(x) = \min\left(h(x), f_n(x)\right)$$

Out of the definition of the functions h_n follows that, the functions h_n are bounded by the function h and are zero outside E', so

$$\int_E h \mathrm{d}\mu = \int_{E'} h \mathrm{d}\mu. \tag{5.20}$$

The functions h_n are bounded, because the function h is bounded by f, so $\lim_{n\to\infty} h_n(x) = h(x)$ for each $x \in E'$. This means that

$$\int_{E'} h \mathrm{d}\mu = \lim_{n\to\infty} \int_{E'} h_n \mathrm{d}\mu \tag{5.21}$$

and since $h_n \leq f_n$, follows with **theorem 5.5** that

$$\lim_{n\to\infty} \int_{E'} h_n \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_E f_n \mathrm{d}\mu. \tag{5.22}$$

Put the results of (5.20),(5.21) and (5.22) behind each other and the following inequality is obtained

$$\int_E h \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_E f_n \mathrm{d}\mu.$$

Take the supremum over h and with **definition 5.3**, the result (5.19) is obtained. □

Theorem 5.7

Monotone Convergence Theorem

Let $\{f_n\}_{n\in\mathbb{N}}$ be a non-decreasing sequence of nonnegative measurable functions, which means that $0 \leq f_1 \leq f_2 \leq \cdots$ and let $\lim_{n\to\infty} f_n(x) = f(x)$ *almost everywhere*. Then

$$\lim_{n\to\infty} \int_E f_n \, \mathbf{d}\mu = \int_E f \, \mathbf{d}\mu.$$

Proof of Theorem 5.7

With **theorem 5.6** it follows that

$$\int_E f \, d\mu \le \liminf_{n\to\infty} \int_E f_n \, d\mu.$$

For each n is $f_n \le f$, so $\int_E f_n d\mu \le \int_E f d\mu$ and this means that

$$\limsup_{n\to\infty} \int_E f_n \, d\mu \le \int_E f d\mu.$$

Hence

$$\lim_{n\to\infty} \int_E f_n \, d\mu = \int_E f \, d\mu.$$

The theorem is proved. □

Theorem 5.8

Lebesgue's Dominated Convergence Theorem

Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions such that for *almost every* x **on** E**,** $\lim_{n\to\infty} f_n(x) = f(x)$**. If there exists an integrable function** g**, which dominates the functions** f_n **on** E**, i.e.** $\mid f_n(x) \mid \le g(x)$**. Then**

$$\lim_{n\to\infty} \int_E f_n \, \mathbf{d}\mu = \int_E f \, \mathbf{d}\mu.$$

Proof of Theorem 5.8

It is sufficient to prove this for nonnegative functions.
From the fact that $(g - f_n) \ge 0$ for all $x \in E$ and n follows with Fatou's Lemma, see **5.6**, that

$$\int_E (g - f) \, d\mu \le \liminf_{n\to\infty} \int_E (g - f_n) \, d\mu.$$

Subtract the integral of g and use the fact that

$$\liminf_{n\to\infty} \int_E (-f_n)\, \mathrm{d}\mu = -\limsup_{n\to\infty} \int_E f_n\, \mathrm{d}\mu.$$

Thus

$$\limsup_{n\to\infty} \int_E f_n\, \mathrm{d}\mu \leq \int_E f\, \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_E f_n\, \mathrm{d}\mu.$$

The proof is completed. □

5.1.8 Inequality of Cauchy-Schwarz (functions)

The exactly value of an inner product is not always needed. But it is nice to have an idea about maximum value of the absolute value of an inner product. The inequality of Cauchy-Schwarz is valid for every inner product, here is given the theorem for functions out of $\mathbb{L}_2(a,b)$.

Theorem 5.9

Let $f, g \in \mathbb{L}_2(a,b)$ and let the inner product be defined by

$$(f, g) = \int_a^b f(x)\,\overline{g(x)}\,dx.$$

then

$$|\,(f,g)\,| \leq \|\, f\,\|_2 \,\|\, g\,\|_2, \tag{5.23}$$

with $\|\cdot\|_2$ defined as in 5.9.

Proof of Theorem 5.9

See the proof of theorem 3.9.1. Replace x by f and y by g. See section 5.1.5 about what is meant by $\|\, g\,\|_2 = 0$. □

5.1.9 $B(\Omega)$ with $\| \cdot \|_\infty$-norm

Let Ω be a set and with $B(\Omega)$ is meant the space of all real-valued bounded functions $f : \Omega \rightarrow \mathbb{R}$, the norm is defined by

$$\| f \|_\infty = \sup_{x \in \Omega} | f(x) | . \tag{5.24}$$

It is easily to verify that $B(\Omega)$, with the defined norm, is a Normed Linear Space. (If the the functions are complex-valued, it becomes a complex Normed Linear Space.)

Theorem 5.10

The Normed Space $(B(\Omega), \| \cdot \|_\infty)$ is a Banach Space.

Proof of Theorem 5.10

Let $\epsilon > 0$ be given and let $\{f_n\}_{n \in \mathbb{N}}$ be a Cauchy row in $B(\Omega)$. Then there exists a $N(\epsilon)$ such that for every $n, m > N(\epsilon)$ and for every $x \in \Omega$, $|f_n(x) - f_m(x)| < \epsilon$. For a fixed x is $\{f_n(x)\}_{n \in \mathbb{N}}$ a Cauchy row in $\mathbb{R}$. The real numbers are complete, so there exists some limit $g(x) \in \mathbb{R}$. x is arbitrary chosen, so there is constructed a new function g.
If x is fixed then there exists a $M(x, \epsilon)$ such that for every $n > M(x, \epsilon)$, $|f_n(x) - g(x)| < \epsilon$.
It is easily seen that $|g(x) - f_n(x)| \leq |g(x) - f_m(x)| + |f_m(x) - f_n(x)| < 2\epsilon$ for $m > M(x, \epsilon)$ and $n > N(\epsilon)$. The result is that $\| g - f_n \|_\infty < 2\epsilon$ for $n > N(\epsilon)$ and this means that the convergence is uniform.
The inequality $\| g \| \leq \| g - f_n \| + \| f_n \|$ gives that, for an appropriate choice of n. The constructed function g is bounded, so $g \in B(\Omega)$. □

5.1.10 The functions spaces $C(\mathbb{R})$, $C_c(\mathbb{R})$ and $C_0(\mathbb{R})$

Most of the time are those place of interest, where some function f is not equal to zero. The support of the function f is the smallest closed set outside f is equal to zero.

Definition 5.4

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be some function, then

$$\mathrm{supp}(f) = \overline{\{x \in \mathbb{R} \,|\, f(x) \neq 0\}},$$

the set $\mathrm{supp}(f)$ is called the support of f.

Definition 5.5

The continous functions $f : \mathbb{R} \rightarrow \mathbb{R}$ are denoted by $C(\mathbb{R})$, continuous in the $\| \cdot \|_\infty$-norm. So if $f \in C(\mathbb{R})$ then $\| f \|_\infty = \sup_{x \in \mathbb{R}} | f(x) |$.

In the integration theory are often used continuous functions with a compact support.

Definition 5.6

The continuous functions $f : \mathbb{R} \rightarrow \mathbb{R}$, with a compact support, are denoted by $C_c(\mathbb{R})$

$$C_c(\mathbb{R}) \;=]\, \{f : \mathbb{R} \rightarrow \mathbb{R} \,|\, f \in C(\mathbb{R}) \,\mathrm{and}\, \mathrm{supp}(f) \,\mathrm{is\ compact}\}.$$

Sometimes $C_c(\mathbb{R})$ is also denoted by $C_{00}(\mathbb{R})$.

And then there also functions with the characteristic that they vanish at infinity.

Definition 5.7

The continuous functions $f : \mathbb{R} \to \mathbb{R}$, which vanish at infinity, are denoted by $C_0(\mathbb{R})$

$$C_0(\mathbb{R}) \;=]\, \{f : \mathbb{R} \to \mathbb{R} \,|\, f \in C(\mathbb{R}) \,\text{and}\, \lim_{|x| \to \infty} \mid f(x) \mid = 0\}.$$

5.2 Sequence Spaces

The sequence spaces are most of the time normed spaces, existing out of rows of numbers $\underline{\xi} = (\xi_1, \xi_2, \xi_3, \ldots)$, which have a certain characteristic or characteristics.
The indices of the elements out of those rows are most of the time natural numbers, so out of $\mathbb{N}$. Sometimes the indices are be taken out of $\mathbb{Z}$, for instance if calculations have to be done with complex numbers.
Working with a Vector Space means that there is defined an addition and a scalar multiplication. Working with Sequence Spaces means that there has to be defined a summation between sequences and a scalar multiplication of a sequence.
Let $\mathbb{K}$ be the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$ and look to the set of functions $\mathbb{K}^{\mathbb{N}} = \{f \mid f : \mathbb{N} \to \mathbb{K}\}$[5]. The easiest way to describe such an element out of $\mathbb{K}^{\mathbb{N}}$ is by a row of numbers, notated by $\underline{x}$. If $\underline{x} \in \mathbb{K}^{\mathbb{N}}$ then $\underline{x} = (x_1, x_2, \cdots, x_i, \cdots)$, with $x_i = f(i)$. A row of numbers out of $\mathbb{K}$ described by some function f. (The set $\mathbb{K}^{\mathbb{Z}}$ can be defined on the same way.)

[5] Important: The sequence spaces are also function spaces, only their domain is most of the time $\mathbb{N}$ or $\mathbb{Z}$.

Definition 5.8

Let $\underline{x}$, $\underline{y} \in \mathbb{K}^{\mathbb{N}}$ and let $\alpha \in \mathbb{K}$.
The addition $(+)$ between the sequences $\underline{x}$ and $\underline{y}$ and the scalar multiplication $(\cdot)$ of α with the sequence $\underline{x}$ are defined by:

addition $(+)$: $(\underline{x} + \underline{y})$ means $(\underline{x} + \underline{y})_i := x_i + y_i$ for all $i \in \mathbb{N}$,
scalar m. $(\cdot)$: $(\alpha \cdot \underline{x})$ means $(\alpha \cdot \underline{x})_i := \alpha\, x_i$ for all $i \in \mathbb{N}$.

The quartet $(\mathbb{K}^{\mathbb{N}}, \mathbb{K}, (+), (\cdot))$, with the above defined addition and scalar multiplication, is a Vector Space. The Vector Space $(\mathbb{K}^{\mathbb{N}}, \mathbb{K}, (+), (\cdot))$ is very big, it exists out of all possible sequences. Most of the time is looked at subsets of the Vector Space $(\mathbb{K}^{\mathbb{N}}, \mathbb{K}, (+), (\cdot))$, there is looked at the behaviour of the row $(x_1, x_2, \cdots, x_i, \cdots)$ for $i \to \infty$. That behaviour can be described by just looking at the single elements x_i for all $i > N$, with $N \in \mathbb{N}$ finite. But often the behaviour is described in terms of series, like $\lim_{N \to \infty} \sum_1^N \mid x_i \mid$, which have to be bounded for instance.
To get an Inner Product Space or a Normed Space there have to be defined an inner product or a norm on the Vector Space, that is of interest on that moment.

5.2.1 ℓ^∞ with $\| \,.\, \|_\infty$-norm

The norm used in this space is the $\| \,.\, \|_\infty$-norm, which is defined by

$$\| \underline{\xi} \|_\infty = \sup_{i \in \mathbb{N}} \mid \xi_i \mid \tag{5.25}$$

and $\underline{\xi} \in \ell^\infty$, if $\| \underline{\xi} \|_\infty < \infty$.
The Normed Space $(\ell^\infty, \| \cdot \|_\infty)$ is complete.

Theorem 5.11

The space ℓ^∞ is not separable.

Proof of Theorem 5.11

Let $S = \{ x \in \ell^\infty \,|\, x(j) = 0$ or 1, for $j = 1, 2, \cdots \}$ and $y = (\eta_1, \eta_2, \eta_3, \cdots) \in S$. y can be seen as a binary representation of of a number $\gamma = \sum_{i=1}^\infty \frac{\eta_i}{2^i} \in [0, 1]$. The interval $[0, 1]$ is uncountable. If $x, y \in S$ and $x \neq y$ then $\| x - y \|_\infty = 1$, so there are uncountable many sequences of zeros and ones. Let each sequence be a center of ball with radius $\frac{1}{4}$, these balls don't intersect and there are uncountable many of them.
Let M be a dense subset in ℓ^∞. Each of these non-intersecting balls must contain an element of M. There are uncountable many of these balls. Hence, M cannot be countable. M was an arbitrary dense set, so ℓ^∞ cannot have dense subsets, which are countable. Hence, ℓ^∞ is not separable. □

Theorem 5.12

The dual space $(\ell^\infty)' = ba(\mathcal{P}(\mathbb{N}))$.

Proof of Theorem 5.12

This will become a difficult proof[6].

1. $\mathcal{P}(\mathbb{N})$ that is the power set of $\mathbb{N}$, the set of all subsets of $\mathbb{N}$. There exists a bijective map between $\mathcal{P}(\mathbb{N})$ and the real numbers $\mathbb{R}$, for more information, see **Section 8.2**.
-This part is finished.

2. What is $ba(\mathcal{P}(\mathbb{N}))$? At this moment, not really an answer to the question, but may be "bounded additive functions on $\mathcal{P}(\mathbb{N})$". See **Step ii: 2** of **Section 8.5** for more information.
-This part is finished.

3. An additive function f preserves the addition operation:

$$f(x + y) = f(x) + f(y),$$

[6] At the moment of writing, no idea if it will become a succesful proof.

for all x, y out of the domain of f.
-This part gives some information.

4. It is important to realize that ℓ^∞ is a non-separable Banach Space. It means that ℓ^∞ has no countable dense subset. Hence, this space has no Schauder basis. There is no set $\{z_i\}_{i \in \mathbb{N}}$ of sequences in ℓ^∞, such that every $x \in \ell^\infty$ can be written as

$$x = \lim_{N \to \infty} \sum_{i=1}^{N} \alpha_i \, z_i \text{ in the sense that } \lim_{N \to \infty} \| \, x - \sum_{i=1}^{N} \alpha_i \, z_i \, \|_\infty = 0,$$

for suitable $\alpha_i \in \mathbb{R}, i \in \mathbb{N}$.
Every element $x \in \ell^\infty$ is just a bounded sequence of numbers, bounded in the $\| \cdot \|_\infty$-norm.
See also **Theorem 5.2.1**.
-This part gives some information.

5. $\ell^1 \subset (\ell^\infty)'$, because of the fact that $(\ell^1)' = \ell^\infty$. ($C(\ell^1) \subset (\ell^1)''$ with C the canonical mapping.) For an example of a linear functional $L \in (\ell^\infty)'$, not necessarily in ℓ^1, see the Banach Limits, **theorem 5.13**.
-This part gives some information.

6. In the literature (Aliprantis, 2006) can be found that

$$(\ell^\infty)' = \ell^1 \oplus \ell^1_d = ca \oplus pa,$$

with ca the countably additive measures and pa the pure finitely additive charges[7].
It seems that $\ell^1 = ca$ and $\ell^1_d = pa$. Further is written that every countably additive finite signed measure on $\mathbb{N}$ corresponds to exactly one sequence in ℓ^1. And every purely additive finite signed charge corresponds to exactly one extension of a scalar multiple of the limit functional on c, that is ℓ^1_d ?
-This part gives some information. The information given is not completely clear to me. Countable additivity is no problem anymore, see **Definition 8.6**, but these charges?

[7] At the moment of writing, no idea what this means!

7. Reading the literature, there is much spoken about σ-algebras and measures, for more information about these subjects, see section 8.3.
-This part gives some information.

8. In the literature, see (Morrison, 2001, page 50), can be read a way to prove theorem 5.12. For more information, see section 8.5.
-This part gives a way to a proof of Theorem 5.12, it uses a lot of information out of the steps made above.

Theorem 5.12 is proved yet, see ii.8!!

It was a lot of hard work. To search through literature, which is not readable in first instance and then there are still questions, such as these charges in ii.6. So in certain sense not everything is proved. Still is not understood that $(\ell^{\infty})' = \ell^1 \oplus \ell^1_d = ca \oplus pa$, so far nothing found in literature. But as ever, written the last sentence and may be some useful literature is found, see (Rao and Rao, 1983). □

Linear functionals of the type described in theorem 5.13 are called Banach Limits.

Theorem 5.13

There is a bounded linear functional $L : \ell^{\infty} \to \mathbb{R}$ such that

a. $\| L \| = 1$.
b. If $x \in c$ then $L(x) = \lim_{n \to \infty} x_n$.
c. If $x \in \ell^{\infty}$ and $x_n \geq 0$ for all $n \in \mathbb{N}$ then $L(x) \geq 0$.
d. If $x \in \ell^{\infty}$ then $L(x) = L(\sigma(x))$, where $\sigma : \ell^{\infty} \to \ell^{\infty}$ is the shift-operator,
defined by $(\sigma(x))_n = x_{n+1}$.

Proof of Theorem 5.13

The proof is splitted up in several parts and steps.
First the parts ii.a and ii.d. Here Hahn-Banach, theorem 4.11, will be used:

1. Define $M = \{v - \sigma(v) \,|\, v \in \ell^\infty\}$. It is easy to verify that M is a linear subspace of ℓ^∞. Further $e = (1, 1, 1, \cdots) \in \ell^\infty$ and $e \notin M$.
2. Since $0 \in M$, $\mathrm{dist}(e, M) \leq 1$.
   If $(x - \sigma(x))_n \leq 0$ for all $n \in \mathbb{N}$, then $\| e - (x - \sigma(x)) \|_\infty \geq |1 - (x - \sigma(x))_n| \geq 1$.
   If $(x - \sigma(x))_n \geq 0$ for all $n \in \mathbb{N}$, then $x_{n+1} \geq x_n$ for all $n \in \mathbb{N}$. The sequence $\{x_n\}_{n\in\mathbb{N}}$ is nonincreasing and bounded, because $x \in \ell^\infty$, so $\lim_{n\to\infty} x_n$ exists. Thus $\lim_{n\to\infty}(x - \sigma(x))_n = 0$ and $\| e - (x - \sigma(x)) \|_\infty \geq 1$.
   This proves that $\mathrm{dist}(e, M) = 1$.
3. By **theorem 4.11** there is linear functional $L : \ell^\infty \to \mathbb{R}$ such that $\| L \| = 1$, $L(e) = 1$ and $L(M) = 0$. The bounded functional L satisfies **ii.a** and **ii.d** of the theorem.
   $(L(x - \sigma(x)) = 0$, L is linear, so $L(x) = L(\sigma(x)).)$

Part **ii.b**:

1. Let $\epsilon > 0$ be given. Take some $x \in c_0$, then there is a $N(\epsilon)$ such that for every $n \geq N(\epsilon)$, $|\, x_n \,| < \epsilon$.
2. If the sequence x is shifted several times then the norm of the shifted sequences become less then ϵ after some while. Since $L(x) = L(\sigma(x))$, see **ii.d**, also $L(x) = L(\sigma(x)) = L(\sigma(\sigma(x))) = \cdots = L(\sigma^{(n)}(x))$. Hence, $|\, L(\sigma^{(n)}(x)) \,| \leq \| \sigma^{(n)}(x) \|_\infty < \epsilon$ for all $n > N(\epsilon)$. The result becomes that
   $$|\, L(x)| = |\, L(\sigma^{(n)}(x)) \,| < \epsilon. \tag{5.26}$$
   Since $\epsilon > 0$ is arbitrary chosen, **inequality 5.26** gives that $L(x) = 0$. That means that $x \in \mathcal{N}(L)$ (the kernel of L), so $c_0 \subset \mathcal{N}(L)$.
3. Take $x \in c$, then there is some $\alpha \in \mathbb{R}$ such that $\lim_{n\to\infty} x_n = \alpha$. Then $x = \alpha\, e + (x - \alpha\, e)$ with $(x - \alpha\, e) \in c_0$ and
   $$L(x) = L(\alpha\, e + (x - \alpha\, e)) = L(\alpha\, e) + L(x - \alpha\, e) = \alpha = \lim_{n\to\infty} x_n.$$

Part **ii.c**:

1. Suppose that $v \in \ell^\infty$, with $v_n \geq 0$ for all $n \in \mathbb{N}$, but $L(v) < 0$.
2. $v \neq 0$ can be scaled. Let $w = \dfrac{v}{\| v \|_\infty}$, then $0 \leq w_n \leq 1$ and since L is linear, $\dfrac{1}{\| v \|_\infty} L(v) = L(\dfrac{v}{\| v \|_\infty}) = L(w) < 0$. Further is $\| e - w \|_\infty \leq 1$ and $L(e - w) = 1 - L(w) > 1$. Hence,

$$\frac{L(e - w)}{\| e - w \|_\infty} > 1$$

but this contradicts with **ii.a**, so $L(v) \geq 0$.

Theorem 5.13, about the Banach Limits, is proved. □

Example 5.2

Here an example of a non-convergent sequence, which has a unique Banach limit. If $x = (1, 0, 1, 0, 1, 0, \cdots)$ then $x + \sigma(x) = (1, 1, 1, 1, \ldots)$ and $2\,L(x) = L(x) + L(x) = L(x) + L(\sigma(x)) = L(x + \sigma(x)) = 1$. So, for the Banach limit, this sequence has limit $\frac{1}{2}$.

5.2.2 ℓ^1 with $\| \,.\, \|_1$-norm

The norm used in this space is the $\| \,.\, \|_1$-norm, which is defined by

$$\| \underline{\xi} \|_1 = \sum_{i=1}^{\infty} | \xi_i | \tag{5.27}$$

and $\underline{\xi} \in \ell^1$, if $\| \underline{\xi} \|_1 < \infty$.

The Normed Space $(\ell^1, \| \cdot \|_1)$ is complete.

The space ℓ^1 is separable, see ℓ^p with $p = 1$ in **section 5.2.3**.

5.2.3 ℓ^p with $\| \,.\, \|_p$-norm and $1 \leq p < \infty$

The norm used in this space is the $\| \,.\, \|_p$-norm, which is defined by

$$\| \underline{\xi} \|_p = (\sum_{i=1}^{\infty} | \xi_i |^p)^{\frac{1}{p}} \tag{5.28}$$

and $\underline{\xi} \in \ell^p$, if $\| \underline{\xi} \|_p < \infty$.
The Normed Space $(\ell^p, \| \cdot \|_p)$ is complete.

Theorem 5.14

The space ℓ^p is separable.

Proof of Theorem 5.14

The set $S = \{\underline{y} = (\eta_1, \eta_2, \cdots, \eta_n, 0, 0, \cdots) \,|\, \eta_i \in \mathbb{Q}, 1 \leq i \leq n, \mathbb{N}\}$ is a countable subset of ℓ^p.
Given $\epsilon > 0$ and $\underline{x} = (\xi_1, \xi_2, \xi_3, \cdots) \in \ell^p$ then there exists a $N(\epsilon)$ such that

$$\sum_{j=N(\epsilon)+1}^{\infty} | \xi_j |^p < \frac{\epsilon^p}{2}.$$

$\mathbb{Q}$ is dense in $\mathbb{R}$, so there is a $\underline{y} \in S$ such that

$$\sum_{j=1}^{N(\epsilon)} | \eta_j - \xi_j |^p < \frac{\epsilon^p}{2}.$$

Hence, $\| \underline{x} - \underline{y} \|_p = (\sum_{j=1}^{N(\epsilon)} | \eta_j - \xi_j |^p + \sum_{j=N(\epsilon)+1}^{\infty} | \xi_j |^p)^{\frac{1}{p}} < \epsilon$,
so $\overline{S} = \ell^p$. □

Theorem 5.15

The dual space of ℓ^p is ℓ^q, with $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$.

Proof of Theorem 5.15

The proof is done with almost the same steps as done SubSection 4.8.

Every $x \in \ell^p$ has a unique representation

$$\underline{x} = \sum_{i=1}^{\infty} \xi_i \, e_i$$

with respect to a Schauder basis $(e_i)_{i \in \mathbb{N}}$ for ℓ^p.
Let $f \in (\ell^p)'$, where $(\ell^p)'$ is the dual space of ℓ^p, f is linear and bounded, so

$$f(x) = \sum_{i=1}^{\infty} \xi_i \, \gamma_i, \text{ with } \gamma_i = f(e_i). \tag{5.29}$$

Let q be such that $\frac{1}{p} + \frac{1}{q} = 1$, the same as $q = p(q-1)$.
p and q are often called conjugate exponents .
Consider $x_n = \{\xi_i^n\}_{i \in \mathbb{N}}$, defined by

$$\xi_i^n = \begin{cases} \dfrac{\mid \gamma_i \mid^q}{\gamma_i} & \text{if } k \leq n \text{ and } \gamma_i \neq 0, \\ 0 & \text{if } k > n \text{ or } \gamma_i = 0. \end{cases} \tag{5.30}$$

So is obtained that

$$f(x_n) = \sum_{i=1}^{\infty} \xi_i^n \, \gamma_i = \sum_{i=1}^{\infty} \mid \gamma_i \mid^q .$$

Since $f \in (\ell^p)'$ and with the use of **5.30**, there follows that

$$f(x_n) \leq \| f \| \| x_n \| = \| f \| \, (\sum_{i=1}^{n} \mid \xi_i^n \mid^p)^{(\frac{1}{p})} =$$

$$\| f \| \, (\sum_{i=1}^{n} \mid \gamma_i^n \mid^{p(q-1)})^{(\frac{1}{p})} = \| f \| \, (\sum_{i=1}^{n} \mid \gamma_i^n \mid^q)^{(\frac{1}{p})}.$$

All together gives that

$$f(x_n) = \sum_{i=1}^{n} \mid \gamma_i^n \mid^q \leq \| f \| \, (\sum_{i=1}^{n} \mid \gamma_i^n \mid^q)^{(\frac{1}{p})},$$

divide by the last factor and using that $1 - \frac{1}{p} = \frac{1}{q}$ gives

$$(\sum_{i=1}^{n} \mid \gamma_i^n \mid^q)^{(1 - \frac{1}{p})} = (\sum_{i=1}^{n} \mid \gamma_i^n \mid^q)^{(\frac{1}{q})} \leq \| f \| .$$

Take the limit $n \to \infty$ and there is obtained that

$$\sum_{i=1}^{\infty} \mid \gamma_i^n \mid^q)^{(\frac{1}{q})} \leq \parallel f \parallel, \tag{5.31}$$

so $\{\gamma_i\} \in \ell q$.

Let $b = \{\beta_i\}_{i \in \mathbb{N}} \in \ell^q$ and there can be constructed a bounded linear functional g on ℓ^p. The definition of g on ℓ^p is given by

$$g(x) = \sum_{i=1}^{\infty} \xi_i \, \beta_i,$$

where $x = \{\xi_i\} \in \ell^p$. g is linear and bounded, use the Hölder-inequality, see **Theorem ii.a**. Hence $g \in (\ell^p)'$.

With these two steps of above, there is proven that there is a bijective map between the spaces ℓ^p and $(\ell^p)'$.
Let $c = \{\gamma_i\}_{i \in \mathbb{N}}$ and $\gamma_i = f(e_i)$. The mapping ψ of $(\ell^p)'$ onto ℓ^p defined by $\psi : f \mapsto c$ is linear and bijective. Further has to be proven that the map ψ is norm preserving.
From 5.29 and the Hölder inequality, ii.a, there follows that

$$\mid f(x) \mid = \mid \sum_{i=1}^{\infty} \xi_i \, \gamma_i \leq$$

$$(\sum_{i=1}^{\infty} \mid \xi_i \mid^p)^{(\frac{1}{p})} \, ((\sum_{i=1}^{\infty} \mid \gamma_i \mid^q))^{(\frac{1}{q})} = \parallel x \parallel \; ((\sum_{i=1}^{\infty} \mid \gamma_i \mid^q))^{(\frac{1}{q})},$$

taking the supremum over all x with $\parallel x \parallel = 1$, gives as result

$$\parallel f \parallel \leq ((\sum_{i=1}^{\infty} \mid \gamma_i \mid^q))^{(\frac{1}{q})}. \tag{5.32}$$

From 5.31 and 5.32 follows that $\parallel f \parallel = \parallel c \parallel_q$, so the map ψ is norm preserving and so it is an isometric isomorphism.
□

5.2.4 ℓ^2 with $\parallel . \parallel_2$-norm

The norm used in this space is the $\| \,.\, \|_2$-norm, which is defined by

$$\| \underline{\xi} \|_2 = \sqrt{\sum_{i=1}^{\infty} | \xi_i |^2} \tag{5.33}$$

and $\underline{\xi} \in \ell^2$, if $\| \underline{\xi} \|_2 < \infty$.
The Normed Space $(\ell^2, \| \cdot \|_2)$ is complete.
The space ℓ^2 is separable, see ℓ^p with $p = 2$ in **section 5.2.3**.

Theorem 5.16

If $x \in \ell^2$ and $y \in \ell^2$ then $(x + y) \in \ell^2$.

Proof of Theorem 5.16

Let $x = (x_1, x_2, \cdots)$ and $y = (y_1, y_2, \cdots)$ then $(x+y) = (x_1+y_1, x_2+y_2+\cdots)$.
Question: $\lim_{N\to\infty}(\sum_{i=1}^{N} | x_i + y_i |^2)^{\frac{1}{2}} < \infty$?
Take always finite sums and afterwards the limit of $N \to \infty$, so

$$\sum_{i=1}^{N} | x_i + y_i |^2 = \sum_{i=1}^{N} | x_i |^2 + \sum_{i=1}^{N} | y_i |^2 + 2 \sum_{i=1}^{N} | x_i || y_i | .$$

Use the inequality of Cauchy-Schwarz, see **3.13**, to get

$$\sum_{i=1}^{N} | x_i + y_i |^2 \leq \sum_{i=1}^{N} | x_i |^2 + \sum_{i=1}^{N} | y_i |^2 + 2 (\sum_{i=1}^{N} | x_i |^2)^{\frac{1}{2}} (\sum_{i=1}^{N} | y_i |^2)^{\frac{1}{2}}$$

$$= ((\sum_{i=1}^{N} | x_i |^2)^{\frac{1}{2}} + (\sum_{i=1}^{N} | y_i |^2)^{\frac{1}{2}})^2.$$

On such way there is achieved that

$$\lim_{N\to\infty}(\sum_{i=1}^{N} \mid x_i + y_i \mid^2)^{\frac{1}{2}} \leq \lim_{N\to\infty}((\sum_{i=1}^{N} \mid x_i \mid^2)^{\frac{1}{2}} + (\sum_{i=1}^{N} \mid y_i \mid^2)^{\frac{1}{2}})$$

$$= (\sum_{i=1}^{\infty} \mid x_i \mid^2)^{\frac{1}{2}} + (\sum_{i=1}^{\infty} \mid y_i \mid^2)^{\frac{1}{2}} < \infty.$$

□

5.2.5 $c \subseteq \ell^{\infty}$

The norm of the normed space ℓ^{∞} is used and for every element $\underline{\xi} \in c$ holds that $\lim_{i\to\infty} \xi_i$ exists.

The Normed Space $(c, \| \cdot \|_{\infty})$ is complete.

The space c is separable.

5.2.6 $c_0 \subseteq c$

The norm of the normed space ℓ^{∞} is used and for every element $\underline{\xi} \in c_0$ holds that $\lim_{i\to\infty} \xi_i = 0$.

The Normed Space $(c_0, \| \cdot \|_{\infty})$ is complete.

The space c_0 is separable.

Theorem 5.17

The mapping $T : c \rightarrow c_0$ is defined by

$$T(x_1, x_2, x_3, \cdots) = (x_\infty, x_1 - x_\infty, x_2 - x_\infty, x_3 - x_\infty, \cdots),$$

with $x_\infty = \lim_{i \rightarrow \infty} x_i$.
T is

a. bijective,

b. continuous,

c. and the inverse map T^{-1} is continuous,

in short: T is a homeomorphism .

T is also linear, so T is a linear homeomorphism .

Proof of Theorem 5.17

It is easy to verify that T is linear, one-to-one and surjective.
The spaces c and c_0 are Banach spaces.
If $x = (x_1, x_2, x_3, \cdots) \in c$, then

$$|x_i - x_\infty| \leq |x_i| - |x_\infty| \leq 2 \parallel x \parallel_\infty \tag{5.34}$$

and

$$|x_i| \leq |x_i| - |x_\infty| + |x_\infty| \leq 2 \parallel T(x) \parallel_\infty . \tag{5.35}$$

With the inequalities 5.34 and 5.35, it follows that

$$\frac{1}{2} \parallel x \parallel_\infty \leq \parallel T(x) \parallel_\infty \leq 2 \parallel x \parallel_\infty . \tag{5.36}$$

T is continuous, because T is linear and bounded. Further is T bijective and bounded from below. With theorem 7.10, it follows that T^{-1} is continuous.
The bounds given in 5.36 are sharp

$$\parallel T(1, -1, -1, -1, \cdots) \parallel_\infty = \parallel (-1, 2, 0, 0, \cdots) \parallel_\infty = 2 \parallel (1, -1, -1, -1, \cdots) \parallel_\infty,$$

$$\parallel T(1, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \cdots) \parallel_\infty = \parallel (\frac{1}{2}, \frac{1}{2}, 0, 0, \cdots) \parallel_\infty = \frac{1}{2} \parallel (1, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \cdots) \parallel_\infty .$$

□

Remark 5.2

1. T is not an isometry,
$\| T(1,-1,-1,-1,\cdots) \|_\infty = 2 \; \| (1,-1,-1,-1,\cdots) \|_\infty \neq \| (1,-1,-1,-1,\cdots) \|_\infty$.
2. Define the set of sequences $\{e = (1,1,1,\cdots),\cdots,e_j,\cdots\}$ with $e_i = (0,\cdots,0,\delta_{ij},0,\cdots)$. If $x \in c$ and $x_\infty = \lim_{i\to\infty} x_i$ then

$$x = x_\infty \, e + \sum_{i=1}^{\infty} (x_i - x_\infty) \, e_i.$$

The sequence $\{e, e_1, e_2, \cdots\}$ is a Schauder basis for c.

5.2.7 $c_{00} \subseteq c_0$

The norm of the normed space ℓ^∞ is used. For every element $\underline{\xi} \in c_{00}$ holds that only a finite number of the coordinates ξ_i are not equal to zero.
If $\underline{\xi} \in c_{00}$ then there exists some $N \in \mathbb{N}$, such that $\xi_i = 0$ for every $i > N$. (N depends on $\underline{\xi}$.)

The Normed Space $(c_{00}, \| \cdot \|_\infty)$ is not complete.

5.2.8 $\mathbb{R}^N$ or $\mathbb{C}^N$

The spaces $\mathbb{R}^N$ or $\mathbb{C}^N$, with a fixed number $N \in \mathbb{N}$, are relative simple in comparison with the above defined Sequence Spaces. The sequences in the mentioned spaces are of finite length

$$\mathbb{R}^N = \{\underline{x} \mid \underline{x} = (x_1, \cdots, x_N), \, x_i \in \mathbb{R}, \, 1 \leq i \leq N\}, \tag{5.37}$$

replace $\mathbb{R}$ by $\mathbb{C}$ and we have the definition of $\mathbb{C}^N$.
An inner product is given by

$$(\underline{x}, \underline{y}) = \sum_{i=1}^{N} x_i \, \overline{y_i}, \tag{5.38}$$

with $\overline{y_i}$ the complex conjugate of y_i. The complex conjugate is only of interest in the space $\mathbb{C}^N$, in $\mathbb{R}^N$ it can be suppressed.
Some other notations for the inner product are

$$(\underline{x}, \underline{y}) = \underline{x} \bullet \underline{y} = < \underline{x}, \underline{y} > \tag{5.39}$$

Often the elements out of $\mathbb{R}^N$ or $\mathbb{C}^N$ are presented by columns, i.e.

$$\underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \tag{5.40}$$

If the elements of $\mathbb{R}^N$ or $\mathbb{C}^N$ are represented by columns then the inner product can be calculated by a matrix multiplication

$$(\underline{x}, \underline{y}) = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}^T \begin{bmatrix} \overline{y_1} \\ \vdots \\ \overline{y_N} \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_N \end{bmatrix} \begin{bmatrix} \overline{y_1} \\ \vdots \\ \overline{y_N} \end{bmatrix} \tag{5.41}$$

5.2.9 Inequality of Cauchy-Schwarz (vectors)

The exactly value of an inner product is not always needed. But it is nice to have an idea about maximum value of the absolute value of an inner product. The inequality of Cauchy-Schwarz is valid for every inner product, here is given the theorem for sets of sequences of the form $(x_1, \cdots, x_N)$, with $N \in \mathbb{N}$ finite.

Theorem 5.18

Let $\underline{x} = (x_1, \cdots, x_N)$ and $\underline{y} = (y_1, \cdots, y_N)$ with $x_i, y_i \in \mathbb{C}^N$ for $1 \leq i \leq N$, with $N \in \mathbb{N}$, then

$$| (\underline{x}, \underline{y}) | \leq \| \underline{x} \|_2 \| \underline{y} \|_2 . \tag{5.42}$$

With $\| \cdot \|_2$ is meant expression **5.33**, but not the length of a vector. Nothing is known about how the coördinates are chosen.

Proof of Theorem 3.13

It is known that

$$0 \leq (\underline{x} - \alpha \underline{y}, \underline{x} - \alpha \underline{y}) = \| \underline{x} - \alpha \underline{y} \|_2^2$$

for every $\underline{x}, \underline{y} \in \mathbb{C}^N$ and for every $\alpha \in \mathbb{C}$, see formula **3.8**. This gives

$$\begin{aligned} 0 &\leq (\underline{x}, \underline{x}) - (\underline{x}, \alpha \underline{y}) - (\alpha \underline{y}, \underline{x}) + (\alpha \underline{y}, \alpha \underline{y}) \\ &= (\underline{x}, \underline{x}) - \overline{\alpha} (\underline{x}, \underline{y}) - \alpha (\underline{y}, \underline{x}) + \overline{\alpha} \alpha (\underline{y}, \underline{y}) \end{aligned} \tag{5.43}$$

If $(\underline{y}, \underline{y}) = 0$ then $y_i = 0$ for $1 \leq i \leq N$ and there is no problem. Assume $\underline{y} \neq \underline{0}$ and take

$$\alpha = \frac{(\underline{x}, \underline{y})}{(\underline{y}, \underline{y})}.$$

Put α in inequality **5.43** and use that

$$(x, y) = \overline{(y, x)},$$

see definition **3.29**. Writing out, and some calculations, gives the inequality of Cauchy-Schwarz. □

5.2.10 Inequalities of Hölder, Minkowski and Jensen (vectors)

The inequality of Hölder and Minkowski are generalizations of Cauchy-Schwarz and the triangle-inequality. They are most of the time used in the Sequence Spaces ℓ^p with $1 < p < \infty$, be careful with $p = 1$ and $p = \infty$. Hölder's inequality is used in the proof of Minkowski's inequality. With Jensen's inequality it is easy to see that $\ell^p \subset \ell^r$ if $1 \leq p < r < \infty$.

Theorem 5.19

Let $a_j, b_j \in \mathbb{K}$, $j = 1, \cdots, n$, $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.
For $1 < p < \infty$, let q satisfy $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

a. **Hölder's inequality**, for $1 < p < \infty$:
$\sum_{i=1}^{n} \mid a_i\, b_i \mid \leq (\sum_{i=1}^{n} \mid a_i \mid^p)^{\frac{1}{p}} (\sum_{i=1}^{n} \mid b_i \mid^q)^{\frac{1}{q}}$.

If $a = \{a_j\} \in \ell^p$ and $b = \{b_j\} \in \ell^q$ then $\sum_{i=1}^{\infty} \mid a_i\, b_i \mid \leq \| a \|_p \| b \|_q$.

b. **Minkowski's inequality**, for $1 \leq p < \infty$:
$\sum_{i=1}^{n} \mid a_i + b_i \mid^{\frac{1}{p}} \leq (\sum_{i=1}^{n} \mid a_i \mid^p)^{\frac{1}{p}} + (\sum_{i=1}^{n} \mid b_i \mid^p)^{\frac{1}{p}}$.

If $a = \{a_j\} \in \ell^p$ and $b = \{b_j\} \in \ell^p$ then $\| a + b \|_p \leq \| a \|_p + \| b \|_p$.

c. **Jensen's inequality**, for $1 \leq p < r < \infty$:
$(\sum_{i=1}^{n} \mid a_i \mid^r)^{\frac{1}{r}} \leq (\sum_{i=1}^{n} \mid a_i \mid^p)^{\frac{1}{p}}$.

$\| a \|_r \leq \| a \|_p$ for every $a \in \ell^p$.

Proof of Theorem ??

a. If $a \geq 0$ and $b \geq 0$ then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \tag{5.44}$$

If $b = 0$, the inequality 5.44 is obvious, so let $b > 0$. Look at the function $f(t) = \dfrac{1}{q} + \dfrac{t}{p} - t^{\frac{1}{p}}$ with $t > 0$. The function f is a decreasing function for $0 < t < 1$ and an increasing function for $t > 1$, look to the sign of $\frac{df}{dt}(t) = \frac{1}{p}(1 - t^{-\frac{1}{q}})$. $f(0) = \dfrac{1}{q} > 0$ and $f(1) = 0$, so $x(t) \geq 0$ for $t \geq 0$. The result is that

$$t^{\frac{1}{p}} \leq \frac{1}{q} + \frac{t}{p}, \; t \geq 0. \tag{5.45}$$

Take $t = \dfrac{a^p}{b^q}$ and fill in formula 5.45, multiply the inequality by b^q and inequality 5.44 is obtained. Realize that $q - \frac{q}{p} = 1$.
Define

$$\alpha = (\sum_{i=1}^{n} \mid a_i \mid^p)^{\frac{1}{p}} \text{ and } \beta = (\sum_{i=1}^{n} \mid b_i \mid^q)^{\frac{1}{q}}$$

and assume $\alpha > 0$ and $\beta > 0$. The cases that $\alpha = 0$ or $\beta = 0$, the Holder's inequality is true. Take $a = \dfrac{a_j}{\alpha}$ and $b = \dfrac{b_j}{\alpha}$ and fill in in formula 5.44, $j = 1, \cdots, n$. Hence

$$\sum_{i=1}^{n} \frac{\mid a_i \, b_i \mid}{\alpha \, \beta} \leq (\frac{1}{p \, \alpha^p} \sum_{i=1}^{n} \mid a_i \mid^p + \frac{1}{q \, \beta^q} \sum_{i=1}^{n} \mid b_i \mid^q) = 1,$$

and Hölder's inequality is obtained.
The case $p = 2$ is the inequality of Cauchy, see 3.13.

b. The case $p = 1$ is just the triangle-inequality. Assume that $1 < p < \infty$.
With the help of Hölder's inequality

$$\begin{aligned}
&\sum_{i=1}^{n}(\mid a_i \mid + \mid b_i \mid)^p \\
&= \sum_{i=1}^{n} \mid a_i \mid (\mid a_i \mid + \mid b_i \mid)^{p-1} + \sum_{i=1}^{n} \mid b_i \mid (\mid a_i \mid + \mid b_i \mid)^{p-1} \\
&\leq (\sum_{i=1}^{n} \mid a_i \mid^p)^{\frac{1}{p}} (\sum_{i=1}^{n}(\mid a_i \mid + \mid b_i \mid)^{(p-1)q})^{\frac{1}{q}} \\
&+ (\sum_{i=1}^{n} \mid b_i \mid^p)^{\frac{1}{p}} (\sum_{i=1}^{n}(\mid a_i \mid + \mid b_i \mid)^{(p-1)q})^{\frac{1}{q}} \\
&= (\sum_{i=1}^{n} \mid a_i \mid^p)^{\frac{1}{p}} + \sum_{i=1}^{n} \mid b_i \mid^p)^{\frac{1}{p}}) (\sum_{i=1}^{n}(\mid a_i \mid + \mid b_i \mid)^p)^{\frac{1}{q}}
\end{aligned}$$

because $(p-1)\,q = p$, further $1 - \frac{1}{q} = \frac{1}{p}$.

c. Take $x \in \ell^p$ with $\| x \|_p \leq 1$, then $\mid x_i \mid \leq 1$ and hence $\mid x_i \mid^r \leq \mid x_i \mid^p$, so $\| x \|_r \leq 1$.
Take $0 \neq x \in \ell^p$ and consider $\dfrac{x}{\| x \|_p}$ then it follows that $\| x \|_r \leq \| x \|_p$ for $1 \leq p < r < \infty$.

□

Remark 5.3

Jensen's inequality ?? ii.c implies that $\ell^p \subset \ell^r$ and if $x_n \to x$ in ℓ^p then $x_n \to x$ in ℓ^r.

6 Types of Spaces

There are different types of spaces. The spaces in **Chapter 3** can have all kind of extra conditions on for instance the topology of such a space.
In **Chapter 3** are already discussed certain different type of spaces, see the **flow chart of spaces 3.1**. But here are discussed type of space, which are not so easily to put in a nice flowchart.

6.1 PreCompact\Compact Metric Spaces

In this section will be looked at compact and precompact metric spaces. The question to be answered is if these spaces are equivalent?
This question arose of the question whether a bounded sequence in a metric space has a convergent subsequence, like the classical **Bolzano-Weierstrass Theorem 6.1**. The **Arzela-Ascoli Theorem 6.8** is also proved as a kind of application of the theory.

6.1.1 Bolzano-Weierstrass

If there are bounded sequences and the convergence of some subsequence is of importance, the name of Bolzano-Weierstrass is used very much.
There is also spoken about the Bolzano-Weierstrass Property.

Definition 6.1

A metric space X has the Bolzano-Weierstrass Property if every infinite subset S in X has an accumulation point, also called a limit point.

What has the Bolzano-Weierstrass Property to do with the compactness of a subset in a metric space. Are these two properties equivalent in a metric space?

6.1.1.1 Theorems of Bolzano-Weierstrass

Definition 6.2

A set S in $\mathbb{R}$ is said to be bounded if it lies in an interval $[-a,\, a]$ for some $0 < a \in \mathbb{R}$.

Theorem 6.1

The classical theorem of Bolzano-Weierstrass:
If a bounded set S in $\mathbb{R}$ contains infinitely many points, then there is at least one point in $\mathbb{R}$ which is an accumulation point of S.

Proof of Theorem 6.1

Since S is bounded it lies in some interval $[-a,\, a]$ for some $0 < a \in \mathbb{R}$. At least one of the intervals $[-a, 0]$ or $[0, a]$ contains an infinite subset of S. Give one of these intervals the name $[a_1,\, b_1]$. Bisect the interval $[a_1,\, b_1]$ and obtain a subinterval $[a_2,\, b_2]$, which contains infinitely many points of S. Continue this process, such that a countable collection of intevals is obtained of which the length of the nth interval $[a_n,\, b_n]$ is equal to $b_n - a_n = \dfrac{a}{2^{(n-1)}}$.
The sup of the left endpoint a_n and the inf of the right endpoint b_n must be equal, say to x. The point x is an accumulation point of S. If r is any positive number, the interval $[a_n,\, b_n]$ will be contained in the interval $(x - r, x + r)$ as soon as n is large enough so that $b_n - a_n = \dfrac{r}{2}$. The interval $(x - r, x + r)$ contains a point of S distinct from x, so x is an accumulation point of S. (This accumulation point x may or may not belong to S.) $\square$

Theorem 6.2

The more general theorem of Bolzano-Weierstrass:
Any sequence $\{a_n\}_{n=1}^{\infty}$ in a compact metric space X has a convergent subsequence $\{a_{n_j}\}_{j=1}^{\infty}$.

6.1.2 Lebesgue-Number

Being involved with compactness in metric spaces, most of the time, there will also be spoken about a Lebesgue Number.

Definition 6.3

Let $M = (X, d)$ be metric space and let U be an open cover of M. A fixed positive real number $0 < \lambda \in \mathbb{R}$ is called a Lebesgue Number for U if

$$\forall x \in M : \exists U(x) \in U \text{ such that } N_\lambda(x, d) \subseteq U(x),$$

where $N_\lambda(x, d)$ is the λ-neighbourhood of x in M.

The λ-neighbourhood of x in some metric space $M = (X, d)$ is defined by

$$N_\lambda(x, d) = \{y \in M \mid d(y, x) < \lambda\}.$$

Example 6.1

Not every open cover has a Lebesgue Number.

Explanation of Example 6.1

Take the metric space $M = (X, d)$, with $X = (0, 1) \subset \mathbb{R}$ and $d(x, y) =| x - y |$, with the open cover $U = \{(\frac{1}{n}, 1) \mid n \geq 2\}$.
Let $0 < \lambda \in \mathbb{R}$ be a Lebesgues Number for U. Take some $n \in \mathbb{N}$ such that $\frac{1}{n} < \lambda$ and take $x = \frac{1}{n}$ then $N_\lambda(x, d) = N_\lambda(\frac{1}{n}, d) = (0, \frac{1}{n} + \lambda)$. But there is no $(\frac{1}{m}, 1) \in U$ such that $N_\lambda(x, d) \subseteq (\frac{1}{m}, 1)$. So λ is not a Lebesgue Number for U. □

Here follows the Lebesgue's Number Lemma.

Theorem 6.3

Let $M = (X, d)$ be a metric space. Let M be sequentially compact. Then there exists a Lebesgue Number for every open cover of M.

Proof of Theorem 6.3

Let's try to prove by contradiction.
Suppose that U is an open cover of M, which has no Lebesgue Number.
Then for any $n \in \mathbb{N}$ there exists some $x_n \in M$ such that $N_{\frac{1}{n}}(x_n, d) \subseteq U$ is false for every $U \in U$. Otherwise $\frac{1}{n}$ would be a Lebesgue Number for U.
So there is constructed a sequence (x_n). M is sequentially compact and that means that the sequence (x_n) has a subsequence $(x_{n(r)})$ which converges to some $x \in M$.
U covers M, so there is some $U_0 \in U$ such that $x \in U_0$. U_0 is open, so there is some $m \in \mathbb{N}$ such that $N_{\frac{2}{m}}(x, d) \subseteq U_0$.
Further there exists some $R \in \mathbb{N}$ such that $x_{n(r)} \in N_{\frac{1}{n}}(x, d)$ for $r \geq R$.
Choose some $r \geq R$ such that $n(r) \geq m$ and define $s = n(r)$, then

$$N_{\frac{1}{s}}(x_s, d) \subseteq N_{\frac{2}{m}}(x, d),$$

since

$$d(x_s, y) < \frac{1}{s} \Rightarrow d(x, y) \leq d(x, x_s) + d(x_s, y) < \frac{1}{m} + \frac{1}{s} \leq \frac{2}{m}.$$

So $N_{\frac{1}{s}}(xs, d) \subseteq U_0$ but this contradicts the choice of x_s!

The conclusion is that there has to be a Lebesgue Number for U. □

6.1.3 Totally-bounded or precompact

In a metric space $M = (X, d)$ the boundedness of a non-empty subset A is defined by

$$\varnothing \neq A \subseteq M \text{ is bounded if } \operatorname{diam}(A) < \infty, \tag{6.1}$$

so there exists some constant $K > 0$, such that $d(x, y) < K$ for all $x, y \in A$. This is a little bit different from what is used in definition (**2.6**). In (**6.1**) is used the metric d of the metric space M and not a metric induced by a norm.

Definition 6.4

The metric space $M = (X, d)$ is said to be totally bounded or precompact if for any $\lambda > 0$, there exists a finite cover of X by sets of diameter less than λ.

Precompact is also called relatively compact.

Theorem 6.4

A metric space $M = (X, d)$ is totally bounded *if and only if* every sequence in M has a Cauchy subsequence.

Proof of Theorem 6.4

The proof exists out of two parts. The shortest part will be done first and the difficult part as second.

($\Leftarrow$) Let's try to prove by contradiction.
Assume that $M = (X, d)$ is not totally bounded. Then there exists a

$\lambda_0 > 0$ such that X can not be covered by finitely many balls of radius λ_0.
Let $x_1 \in X$, then $B_{\lambda_0}(x_1, d) \neq X$. So there can be chose some $x_2 \in X \setminus B_{\lambda_0}(x_1, d)$ and go so on. So for each $n \in \mathbb{N}$, there can be chosen some $x_{n+1} \in X \setminus \cup_{i=1}^{n} B_{\lambda_0}(x_i, d)$. If $m > n$ then $x_m \notin B_{\lambda_0}(x_n, d)$ and thus $d(x_m, x_n) \geq \lambda_0$. So there is constructed a sequence (x_n) without a Cauchy subsequence, which contradicts the assumption.

($\Rightarrow$) Assume that $M = (X, d)$ is totally bounded and let (x_n) be a sequence in $M = (X, d)$. By a so-called diagonal argument, there will be constructed a Cauchy subsequence of (x_n).
There will be used an inductive construction.
Set $B^0 = X$. There exist a finite number of sets $A^{11}, A^{12}, \cdots, A^{1n_1} \subseteq X$ such that

$$\text{diam}(A^{1i}) < 1 \text{ with } i \in \{1, 2, \cdots, n_1\} \text{ such that } \bigcup_{i=1}^{n_1} A^{1i} = X.$$

At least one of these A^{1i}-sets must contain infinitely many terms of the squence (x_n), give it the name B^1. Let $(x_{11}, x_{12}, x_{13}, \cdots)$ be a subsequence of (x_n) which lies entirely in B^1. $B^1 \subseteq X$ and so B^1 is also totally bounded. There exist a a finite number of sets $A^{21}, A^{22}, \cdots, A^{2n_2} \subseteq B^1$ such that

$$\text{diam}(A^{2i}) < \frac{1}{2} \text{ with } i \in \{1, 2, \cdots, n_2\} \text{ such that } \bigcup_{i=1}^{n_2} A^{2i} = B^1.$$

At least one of these A^{2i}-sets must contain infinitely many terms of the squence (x_{1n}), give it the name B^2. Let $(x_{21}, x_{22}, x_{23}, \cdots)$ be a subsequence of (x_{1n}) which lies entirely in B^2. $B^2 \subseteq B^1$ and so B^2 is also totally bounded. There exist a a finite number of sets $A^{31}, A^{32}, \cdots, A^{3n_3} \subseteq B^2$ such that

$$\text{diam}(A^{3i}) < \frac{1}{3} \text{ with } i \in \{1, 2, \cdots, n_3\} \text{ such that } \bigcup_{i=1}^{n_3} A^{3i} = B^2.$$

And go so on.
So there can be constructed a sequence of sets (B^i) with
$B^i \subseteq B^{i-1}$ and $\text{diam}(B^i) < \frac{1}{i}$ for all $i \in \mathbb{N}$. And there can be choosen a subsequence $(x_{i1}, x_{i2}, x_{i3}, \cdots)$ of the sequence $(x_{(i-1)1}, x_{(i-1)2}, x_{(i-1)3}, \cdots)$ which lies entirely in B^i for all $i \in \mathbb{N}$. Those subsets $B^i \subseteq X$ are also totally bounded for all $i \in \mathbb{N}$.
The elements x_{ii} are taken out of each subsequence and so the sequence $(x_{11}, x_{22}, x_{33}, \cdots) = (x_{nn})$ is constructed. The question becomes if the

sequence (x_{nn}) is a Cauchy subsequence of the sequence (x_n).
Let $\epsilon > 0$ be given and choose some $N \in \mathbb{N}$ such that $\frac{1}{N} < \epsilon$. The construction of the sequence (x_{nn}) guarantees that the index of the constructed subsequence strictly increases. Let $m, n \in \mathbb{N}$ be such that $m, n \geq N$, then $x_{mm} \in B^m \subseteq B^N$ and $x_{nn} \in B^n \subseteq B^N$. So

$$d(x_{mm}, x_{nn}) \leq \text{diam}(B^N) < \frac{1}{N} < \epsilon$$

and there follows that (x_{nn}) is a Cauchy subsequence of (x_n). □

Theorem 6.5

A metric space $M = (X, d)$ is sequentially compact *if and only if* M is complete and totally bounded.

Proof of Theorem 6.5

($\Rightarrow$) M is sequentially compact. M is also totally bounded by Theorem 6.4. A convergent sequence is also a Cauchy sequence.
Let (x_n) be a Cauchy sequence in M. Since M is sequentially compact, the sequence (x_n) has a convergent subsequence, for instance $(x_{n(r)})$. If $\lim_{r \to \infty} x_{n(r)} = L$ then

$$| x_n - L | \leq | x_n - x_{n(r)} | + | x_{n(r)} - L |,$$

if the indices n and r are taken great enough, the right-hand side can be made as small as desired. So it follows that the whole sequence (x_n) converges and the conclusion is that M is complete.

($\Leftarrow$) The assumption is that M is totally bounded and complete. Let (x_n) be some sequence in X. By the use of Theorem 6.4, it follows that the sequence (x_n) has a Cauchy subsequence $(x_{n(r)})$. M is complete, which means that the Cauchy subsequence $(x_{n(r)})$ converges in M. Hence, the sequence (x_n) has a convergent subsequence and there follows that M is sequentially compact.

□

But first an example to show that the conditions, as given in the theorem of Heine-Borel (see **theorem 2.6**) are neccessary, but not sufficient for compactness.

Example 6.2

A closed and bounded set, that is not compact, is given by $\overline{B_1(0)} = \{f \in C[0,1] \mid \| f \|_\infty \leq 1\}$.

Explanation of Example 6.2

The metric d at $C[0,1]$ is defined by $d(f,g) = \| f - g \|_\infty$, with $f, g \in C[0,1]$. $\overline{B_1(0)}$ is a subset of the metric space $(C[0,1], d)$. $\overline{B_1(0)}$ is also closed (see **theorem 2.11**) and is bounded. See further **example 2.4**, the limit function in the mentioned example is clearly not continuous.
This shows that there exists a sequence in $\overline{B_1(0)}$, which has no subsequence, which converges in $\overline{B_1(0)}$. That means that the unit ball in the metric space $(C[0,1], d)$ is not a compact set. □

The question, in subsection **6.1.1**, about the possible equivalence between the Bolzano-Weierstrass Property and compactness, in a metric space, is answered in the following Theorem.

Theorem 6.6

In a metric space $M = (X, d)$ are the following statements equivalent:

a. M has the Bolzano-Weierstrass Property;

b. M is sequentially compact;

c. M is complete and totally bounded;

d. M is compact;

Proof of Theorem 6.6

The proof exists out of several parts. One part is the result of the foregoing **Theorem 6.5**.

(**ii.a** $\Rightarrow$ **ii.b**)

Assume that M has the Bolzano-Weierstrass Property and let (x_n) be a sequence in X. There are two possibilities:

Case I: The set $S = \{x_n | n \in \mathbb{N}\}$ is finite. Then there is an element $y \in X$, such that $x_n = y$ for inifinitely many n's. Let $V = \{n \in \mathbb{N} \mid x_n = y\}$ and $n_1 = \min(V)$ and let $n_k = \min(V \setminus \{n_1, \cdots, n_{(k-1)}\})$ for $k \geq 2$. Then $(x_{n(k)})$ is a constant subsequence of (x_n) and this subsequence is convergent.

Case II: The set $S = \{x_n | n \in \mathbb{N}\}$ is infinite and by assumption S has an accumulation point $x \in X$. So for each $n \in \mathbb{N}$, $S_n = B_{\frac{1}{n}}(x, d) \cap (S \setminus \{x\}) \neq \varnothing$
and let $V_n = \{n \in \mathbb{N} \mid x_n \in S_n\}$. The set V_n is an infinite set for each $n \in \mathbb{N}$.
Let $n_1 = \min(V_1)$ and $n_k = \min(V_k \setminus \{n_1, \cdots n_{(k-1)}\})$, for $k \geq 2$. Then $(x_{n(k)})$ is a subsequence of (x_n) such that $d(x_{n(k)}, x) < \frac{1}{k}$ for each $k \in \mathbb{N}$ and this subsequence converges to x.

(**ii.b** $\Leftrightarrow$ **ii.c**)

This is already proved in **Theorem 6.5**.

(**ii.b** $\Rightarrow$ **ii.d**)

The assumption is that the metric space M is sequentially compact. Let C be an open cover of M. By **Theorem 6.3** the cover C has a Lebesgue number $\lambda > 0$.
From **Theorem 6.5** is known that M is totally bounded. So there exist a finite number of subsets $A^1, \cdots, A^n \subseteq X$ such that $\bigcup_{i=1}^{n} A^i = X$ and $\operatorname{diam}(A^i) \leq \lambda$ for each $i \in \{1, \cdots, n\}$. For each $i \in \{1, \cdots, n\}$ there exists a $C^i \in C$ such that $A^i \subseteq C^i$ and $X = \bigcup_{i=1}^{n} C^i$. Hence the cover C has a finite subcover and this means that M is compact.

(**ii.d** $\Rightarrow$ **ii.a**)

Let's try to prove it by contradiction. The assumption is that M is a compact metric space and let S be an infinite subset of X. Suppose that S has no accumulation point. Hence for every $x \in X$, there exists an open neighbourhood V_x such that $V_x \cap (S \setminus x) = \varnothing$. All these open neighbourhoods V_x together $\{V_x \mid x \in X\} = C$ are an open cover of X. Since M is compact, there exists a finite subcover $\{V_{x_1}, \cdots, V_{x_n}\}$ of C. Each element V_{x_i} must contain at least one element of S. So $\bigcup_{i=1}^{n} V_{x_i} = X$ contains finitely many points of S and S has to be finite, which contradicts the assumption. Hence S has an accumulation point and the Bolzano-Weierstrass Property is satisfied.

□

6.1.4 Equicontinuity

Most of the time is worked with a set of maps and sometimes these maps have the same kind of behaviour at some single point or in every point of some set, where these maps are used. One of these behaviours is for instance the variation of such a family of maps over a neighbourhood of some point.

Definition 6.5

Let (X, d_1) and (Y, d_2) be two metric spaces and $\mathcal{F}$ a family of maps between X and Y.

a. The family $\mathcal{F}$ is equicontinuous in a point $x_0 \in X$, if for every $\epsilon > 0$ there exists some $\delta(\epsilon) > 0$, such that $d_2(f(x), f(x_0)) < \epsilon$ for all $f \in \mathcal{F}$ and for every $x \in X$ with $d_1(x, x_0) < \delta(\epsilon)$.

b. The family $\mathcal{F}$ is uniformly equicontinuous, if for every $\epsilon > 0$ there exists some $\delta(\epsilon) > 0$, such that $d_2(f(x_1), f(x_0)) < \epsilon$ for all $f \in \mathcal{F}$ and for every x_1, $x_0 \in X$ with $d_1(x_1, x_0) < \delta(\epsilon)$.

c. The family $\mathcal{F}$ is said to be equicontinuous if it is equicontinuous at every point $x \in X$.

In definition 6.5-ii.a, the δ may depend on ϵ and x_0, but is independent of f. In definition 6.5-ii.b, the δ may depend on ϵ, but is independent of f, x_0 and x_1.

Theorem 6.7

Let (X, d_1) and (Y, d_2) be two metric spaces. Assume that X is compact and $\mathcal{F} \subset C(X, Y)$. Then the following statements are equivalent:

a. $\mathcal{F}$ is equicontinuous.

b. $\mathcal{F}$ is uniform equicontinuous.

Proof of Theorem 6.7

(**ii.b** $\Rightarrow$ **ii.a**) It is clear that if $\mathcal{F}$ is uniform equicontinuous then it is also equicontinuous.

(**ii.a** $\Rightarrow$ **ii.b**) Assume that $\mathcal{F}$ is equicontinuous and let $\epsilon > 0$ be given. Then there exists some $\delta(\epsilon, x) > 0$ such that $f(B_{\delta(\epsilon,x)}(x) \subset B_\epsilon(f(x))$ for all $f \in \mathcal{F}$, with $B_{\delta(\epsilon,x)} = \{z \in X | d_1(z, x) < \delta(\epsilon, x)\}$ and $B_\epsilon(f(x)) = \{z \in Y | d_2(z, f(x)) < \epsilon\}$. The collection $\mathcal{O} = \{B_{\delta(\epsilon,x)} | x \in X\}$ forms an open covering of X. Since X is compact there exists a **Lebesgue Number 6.3** of the open cover $\mathcal{O}$. So there is some $\lambda > 0$ such that whenever $A \subset X$ and $\text{diam}(A) < \lambda$, that A is contained in some element of $\mathcal{O}$. This λ is independent of x, so if $x, y \in X$ and $d_1(x, y) < \lambda$ then $d_2(f(x), f(y)) < \epsilon$ for all $f \in \mathcal{F}$. So $\mathcal{F}$ is uniform equicontinuous.

□

6.1.5 Arzelà Ascoli theorem

Let X be a compact metric space, which means that the topology on X has the compactness property. Let $C(X)$ be the space of all continuous functions on X with values in $\mathbb{C}$. In $C(X)$ the metric *dist* is defined by

$$dist(f,g) = \max\{| f(x) - g(x) |: x \in X\}.$$

The space $C(X)$ with the given metric *dist* makes the space complete.
A subset $\mathcal{F}$ of $C(X)$ is bounded if there is positive constant $M < \infty$ such that $| f(x) |< M$ for each $x \in X$ and each $f \in \mathcal{F}$. M is independent of x and independent of f.
Since X is compact, a equicontinuous subset of functions $\mathcal{F}$ of $C(x)$ is also uniform equicontinuous, see **theorem 6.7**. This means that for every $\epsilon > 0$ there exists a $\delta(\epsilon)$ such that for every $x, y \in X$ with

$$d(x,y) < \delta \Rightarrow dist(f,g) < \epsilon \quad \text{for all} \quad f \in \mathcal{F},$$

with d the metric on X.

Theorem 6.8

The theorem of Arzelà-Ascoli:
If a sequence $\{f_n\}_{n=1}^{\infty}$ in $C(X)$, with X a compact metric space, is bounded and equicontinuous then it has a uniform convergent subsequence.

Proof of Theorem 6.8

The proof exists out of a several steps.

Step 1: The compact metric space X has a countable dense subset S, so the compact metric space is separable.
Given some $n \in \mathbb{N}$ and a point $x \in X$ than is

$$B(x, \frac{1}{n}) = \{y \in X \mid d(y,x) < \frac{1}{n}\}$$

an open ball centered at x with radius $\frac{1}{n}$. For given $n \in \mathbb{N}$, the collection of these balls as all $x \in X$ are taken, forms an open cover of X. Since X is compact there is also a finite subcover that covers X. Let's call this finite subset S_n. Each point $x \in X$ lies within a distance $\frac{1}{n}$ of a point of S_n. The union S of all the sets S_n is countable and dense in X.

Step 2: Let's find a subsequence of the bounded sequence $\{f_n\}_{n=1}^{\infty}$ that pointwise converges on S. This will be done by a so-called diagonal argument.

Let's make a list $\{x_1,\, x_2,\, \cdots\}$ of the countable many center points of the elemenst out of S. Look at the sequence of numbers $\{f_n(x_1)\}_{n=1}^{\infty}$, which is bounded and by the theorem of Bolzano-Weierstrass has a convergent subsequence, which is written by $\{f_{n,1}(x_1)\}_{n=1}^{\infty}$. The sequence $\{f_{n,1}(x_2)\}_{n=1}^{\infty}$ is also bounded and has a convergent subsequence $\{f_{n,2}(x_2)\}_{n=1}^{\infty}$. The sequence of functions $\{f_{n,2}\}_{n=1}^{\infty}$ converges in x_1 and x_2. Repeating this proces there is obtained a collection of subsequences of the original sequence:

$$\begin{array}{cccc} f_{1,1} & f_{1,2} & f_{1,3} & \cdots \\ f_{2,1} & f_{2,2} & f_{2,3} & \cdots \\ f_{3,1} & f_{3,2} & f_{3,3} & \cdots \\ . & . & . & \cdots \\ . & . & . & \cdots \\ . & . & . & \cdots \end{array}$$

where the n-th column converges at the points $x_1,\, \cdots,\, x_n$ and each column is a subsequence of the one left of it. Thus the diagonal sequence $\{f_{n,n}\}_{n=1}^{\infty}$ is a subsequence of the original sequence $\{f_n\}_{n=1}^{\infty}$ that converges at each point of S. Let's call this diagonal subsequence $\{h_n\}_{n=1}^{\infty}$.

Step 3: The produced sequence $\{h_n\}_{n=1}^{\infty}$ converges at each point of the dense set S. Let $\epsilon > 0$ be given, by compactness of X and the equicontinuity of the original sequence there exists a $\delta(\epsilon) > 0$ such that for every $x, y \in X$ with $d(x,y) < \delta(\epsilon) \quad \mid h_n(x) - h_n(y) \mid < \frac{\epsilon}{3}$ and for each $n \in \mathbb{N}$. Take $M > \frac{1}{\delta(\epsilon)}$, so that the set S_M, as produced in Step 1, is dense in X. The sequence $\{h_n\}_{n=1}^{\infty}$ converges at each point of S_M, so there exists a $N > 0$ such that

$$\text{for all} \quad n, m > N \Rightarrow \mid h_n(s) - h_m(s) \mid < \frac{\epsilon}{3} \, \text{for all} \, s \in S_M.$$

Take an arbitrary $x \in X$, then x lies within distance less than $\delta(\epsilon)$ of some $s \in S_M$ and so for all $n, m > N$

$$\mid h_n(x) - h_m(x) \mid \leq$$

$$\mid h_n(x) - h_n(s) \mid + \mid h_n(s) - h_m(s) \mid + \mid h_m(s) - h_m(x) \mid < 3\,(\frac{\epsilon}{3}),$$

because of the equicontinuity of the original sequence. Thus on X is the subsequence $\{h_n\}_{n=1}^{\infty}$, of $\{f_n\}_{n=1}^{\infty}$, a Cauchy sequence and therefore uniform convergent. That is the result, which completes the proof.

□

Theorem 6.9

If X is a compact metric space and $\{f_n\}_{n=1}^{\infty}$ a sequence of functions in $C(X)$, that converges uniformly, then the collection $\{f_n\}_{n=1}^{\infty}$ is equicontinuous.

Proof of Theorem 6.9

Let f be the limit of the uniform convergent sequence $\{f_n\}_{n=1}^{\infty}$. So given $\epsilon > 0$, there exists a N such that for all $n > N$ $|f_n(x) - f(x)| < \frac{\epsilon}{3}$ for all $x \in X$.
For each $j \leq N$ the function f_j is uniform continuous, so there exists a $\delta_j > 0$ such that for each $x, y \in X$ with $d(x, y) < \delta_j$ $|f_j(x) - f_j(y)| < \epsilon$. The limit function f is also uniform continuous, so there exists a $\delta_0 > 0$ such that for each $x, y \in X$ with $d(x, y) < \delta_0$ $|f(x) - f(y)| < \frac{\epsilon}{3}$.
Set $\delta_{\min} = \min(\delta_0, \min_{\{1 \leq j \leq N\}}(\delta_j)) > 0$. If $d(x, y) < \delta < \delta_{\min}$ then for $n > N$

$$|f_n(x) - f_n(y)| \leq |f_n(x) - f(x)| + |f(x) - f(y)| + |f(y) - f_n(y)| < \epsilon.$$

Thus this holds for all n, since $\delta \leq \delta_j$ for $j \leq N$ as well, so the collection of functions $\{f_n\}_{n=1}^{\infty}$ is equicontinuous. □

The **theorem of Arzelà-Ascoli 6.8** is the key to the following result.

Theorem 6.10

If X is a compact metric space then a subset $\mathcal{F}$ of $C(X)$ is compact if and only if it is closed, bounded and equicontinuous.

Proof of Theorem 6.10

The proof exists out of a several steps.

A continuous function on a compact metric space X is bounded, see **theorem 2.9**, so the function

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

is well-defined. $(C(X), d)$ is a metric space and convergence with respect to d is equivalent to **uniform convergence 2.12**. And if X is a compact metric space, the metric space $(C(X), d)$ is complete.
Because of the metric spaces and the compactness, equicontinuïty and uniform equicontinuïty are equivalent, see **theorem 6.7**.

($\Rightarrow$) The set $\mathcal{F}$ is a compact set in a metric space, so it is closed and bounded, see **theorem 2.5**. It remains to show that the set $\mathcal{F}$ is equicontinuous.
Equicontinuous means "uniform (in $f \in \mathcal{F}$) uniform (in the points of X) continuity". Suppose that the subset $\mathcal{F}$ is not equicontinuous. That means that there exists an $\epsilon > 0$ such that for each $\delta > 0$, there is a pair of points $x_0, y_0 \in X$ and a function $f_0 \in \mathcal{F}$ such that $d(x_0, y_0) < \delta$ and $|f_0(x) - f_0(y)| \geq \epsilon$.
So for each $n \in \mathbb{N}$, there is a pair of points $x_n, y_n \in X$ and a function $f_n \in \mathcal{F}$ such that $d(x_n, y_n) < \frac{1}{n}$ and $|f_n(x) - f_n(y)| \geq \epsilon$. This fixes a sequence of functions $\{f_n\}_{n=1}^{\infty}$ in $\mathcal{F}$, which has no equicontinuous subsequence.

This is in contradiction with **theorem 6.9**, because every sequence in $C(X)$ has a uniform convergent subsequence.
So the subset $\mathcal{F}$ is equicontinuous.

($\Leftarrow$) Let $\{f_n\}_{n=1}^{\infty}$ be a sequence in $\mathcal{F}$. Then the sequence $\{f_n\}_{n=1}^{\infty}$ is bounded and equicontinuous, X is compact, so by **theorem 6.8**, there exists a convergent subsequence. Since $\mathcal{F}$ is closed, the limit of the subsequence is an element of $\mathcal{F}$. Sequentially compactness and compactness are equivalent in a metric space, see **theorem 6.6**, so the subset $\mathcal{F}$ is compact.

□

7 Linear Maps

7.1 Linear Maps

In this chapter a special class of mappings will be discussed and that are linear maps .
In the literature is spoken about linear maps, linear operators and linear functionals. The distinction between linear maps and linear operators is not quite clear. Some people mean with a linear operator $T : X \to Y$, a linear map T that goes from some Vector Space into itself, so $Y = X$. Other people look to the fields of the Vector Spaces X and Y, if they are the same, then the linear map is called a linear operator.
If Y is another vectorspace then X, then the linear map can also be called a linear transformation.
About the linear functionals there is no confusion. A linear functional is a linear map from a Vector Space X to the field $\mathbb{K}$ of that Vector Space X.

Definition 7.1

Let X and Y be two Vector Spaces. A map $T : X \to Y$ is called a linear map if

LM 1: $T(x + y) = T(x) + T(y)$ for every $x, y \in X$ and

LM 2: $T(\alpha x) = \alpha T(x)$, for every $\alpha \in \mathbb{K}$ and for every $x \in X$.

If nothing is mentioned then the fields of the Vector Spaces X and Y are assumed to be the same. So there will be spoken about linear operators or linear functionals.
The definition for a linear functional is given in section 4.1.
Now there are repeated several notations, which are of importance, see section 2.1 and figure 7.1:

Domain : $\mathcal{D}(\mathrm{T}) \subset X$ is the domain of T;

Range : $\mathcal{R}$(T) $\subset Y$ is the range of T,
$\mathcal{R}(T) = \{y \in Y \mid \exists x \in X \text{with} T(x) = y\}$;

Nullspace : $\mathcal{N}$(T) $\subset \mathcal{D}(T)$ is the nullspace of T,
$\mathcal{N}(T) = \{x \in \mathcal{D}(T) \mid T(x) = 0\}$.
The nullspace of T is also called the kernel of T ;

New is the definition of the Graph of an operator:

Definition 7.2

Let $T : \mathcal{D}(T) \to Y$ be a linear operator, by $\mathcal{G}$(T) is defined the graph of T,

$$\mathcal{G}(T) = \{(x, y) \in X \times Y \mid x \in \mathcal{D}(T) \,\text{and}\, y = T(x) \in \mathcal{R}(T)\}.$$

Further: T is an operator *from* $\mathcal{D}(T)$ *onto* $\mathcal{R}(T)$, $T : \mathcal{D}(T) \to \mathcal{R}(T)$; T is an operator *from* $\mathcal{D}(T)$ *into* Y, $T : \mathcal{D}(T) \to Y$; if $\mathcal{D}(T) = X$ then $T : X \to Y$.
The $\mathcal{R}(T)$ is also called the image of $\mathcal{D}(T)$. If $V \subset \mathcal{D}(T)$ is some subspace of X then $T(V)$ is called the image of V. And if W is some subset of $\mathcal{R}(T)$ then $\{x \in X \mid T(x) \in W\}$ is called the inverse image of W, denoted by $T^{-1}(W)$.
The range and the nullspace of a linear operator have more structure then just an arbitrary mapping out of section 2.1.

Theorem 7.1

If X, Y are Vector Spaces and $T : X \to Y$ is a linear operator then:

a. $\mathcal{R}(T)$, the range of T, is a Vector Space,

b. $\mathcal{N}(T)$, the nullspace of T, is a Vector Space,

c. $\mathcal{G}(T)$, the graph of T, is a linear subspace of $X \times Y$.

Proof of Theorem 7.1

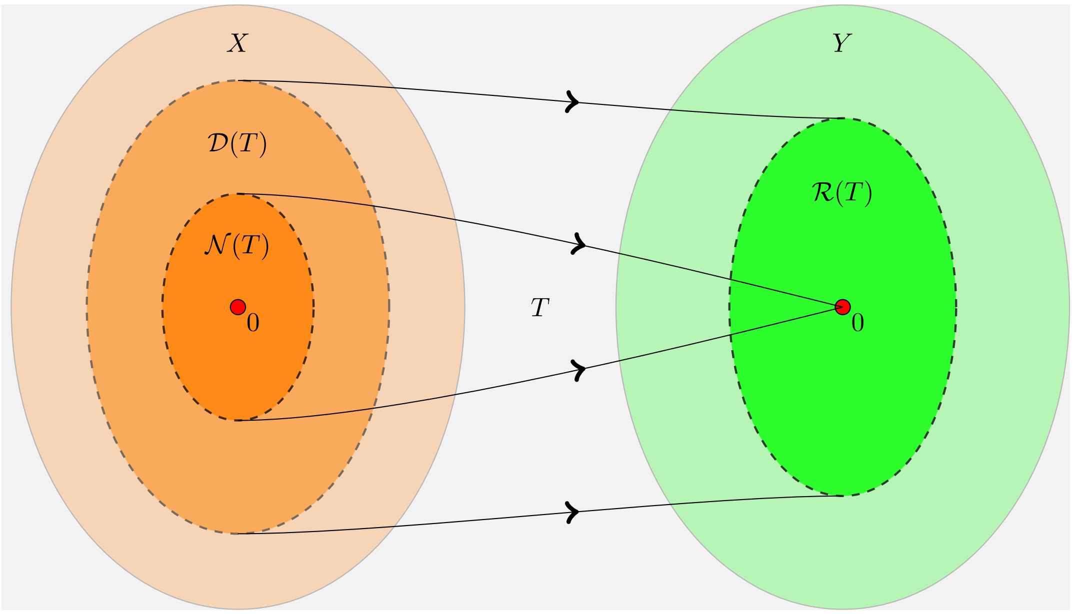


Figure 7.1 Domain, Range, Nullspace

a. Take $y_1,\, y_2 \in \mathcal{R}(T) \subseteq Y$, then there exist $x_1,\, x_2 \in \mathcal{D}(T) \subseteq X$ such that $T(x_1) = y_1$ and $T(x_2) = y_2$. Let $\alpha \in \mathbb{K}$ then $(y_1 + \alpha\, y_2) \in Y$, because Y is a Vector Space and

$$Y \ni y_1 + \alpha\, y_2 = T(x_1) + \alpha\, T(x_2) = T(x_1 + \alpha\, x_2).$$

This means that there exists an element $z_1 = (x_1 + \alpha\, x_2) \in \mathcal{D}(T)$, because $\mathcal{D}(T)$ is a Vector Space, such that $T(z_1) = y_1 + \alpha\, y_2$, so $(y_1 + \alpha\, y_2) \in \mathcal{R}(T) \subseteq Y$.

b. Take $x_1,\, x_2 \in \mathcal{D}(T) \subseteq X$ and let $\alpha \in \mathbb{K}$ then $(x_1 + \alpha\, x_2) \in \mathcal{D}(T)$ and

$$T(x_1 + \alpha\, x_2) = T(x_1) + \alpha\, T(x_2) = 0$$

The result is that $(x_1 + \alpha\, x_2) \in \mathcal{N}(T)$.

c. Take $(x_1, y_1) \in \mathcal{G}(T)$ and $(x_2, y_2) \in \mathcal{G}(T)$, this means that $y_1 = T(x_1)$ and $y_2 = T(x_2)$, then $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, T(x_1) + T(x_2)) = (\,(x_1 + x_2), T((x_1 + x_2))\,) \in \mathcal{G}(T)$. Let $\alpha \in \mathbb{K}$ then $\alpha\,(x_1, y_1) = (\alpha\, x_1,\, \alpha\, T(x_1)) = (\,(\alpha\, x_1),\, T((\alpha\, x_1))\,) \in \mathcal{G}(T)$.

Linear operators can be added together and multiplied by a scalar, the obvious way to do that is as follows.

Definition 7.3

If T, S : $X \rightarrow Y$ are linear operators and X, Y are Vector Spaces, then the addition and the scalar multiplication are defined by

LO 1: $(T + S)\,x = Tx + Sx$ and

LO 2: $(\alpha\, T)\,x = \alpha\,(Tx)$ for any scalar α
and for all $x \in X$.

The set of all the operators $X \rightarrow Y$ is a Vector Space, the zero-operator $\tilde{0} : X \rightarrow Y$ in that Vector Space maps every element of X to the zero element of Y.

Definition 7.4

If $(T - \lambda I)(x) = 0$ for some $x \neq 0$ then λ is called an *eigenvalue* of T. The vector x is called an *eigenvector* of T, or *eigenfunction* of T, $x \in \mathcal{N}(T - \lambda I)$.

It is also possible to define a product between linear operators .

Definition 7.5

Let X, Y and Z be Vector Spaces, if $T : X \rightarrow Y$ and $S : Y \rightarrow Z$ are linear operators then the product $S\,T : X \rightarrow Z$ of these linear operators is defined by

$$(S\,T)\,x = S(T\,x)$$

for every $x \in X$.

The product operator $ST : X \rightarrow Z$ is a linear operator,

1. $(ST)(x + y) = S(T(x + y)) = S(T(x) + T(y)) = S(T(x)) + (S(T(y)) = (ST)(x) + (ST)(y)$ and

2. $(ST)(\alpha\, x) = S(\alpha\, T(x)) = \alpha\, S(T(x)) = \alpha\,(ST)(x)$

for every x, $y \in X$ and $\alpha \in \mathbb{K}$.

7.2 Bounded and Continuous Linear Operators

An important subset of the linear operators are the bounded linear operators . Under quite general conditions the bounded linear operators are equivalent with the continuous linear operators.

Definition 7.6

Let X and Y be normed spaces and let $T : \mathcal{D}(T) \rightarrow Y$ be a linear operator, with $\mathcal{D}(T) \subset X$. The operator is bounded if there exists a positive real number M such that

$$\| T(x) \| \leq M \| x \|, \tag{7.1}$$

for every $x \in \mathcal{D}(T)$.

Read formula 7.1 carefully, on the left is used the norm on the Vector Space Y and on the right is used the norm on the Vector Space X. If necessary there are used indices to indicate that different norms are used. The constant M is independent of x.
If the linear operator $T : \mathcal{D}(T) \rightarrow Y$ is bounded then

$$\frac{\| T(x) \|}{\| x \|} \leq M, \text{ for all } x \in \mathcal{D}(T) \setminus \{0\},$$

so M is an upper bound, and the lowest upper bound is called the norm of the operator T, denoted by $\| T \|$.

Definition 7.7

Let T be a bounded linear operator between the normed spaces X and Y then

$$\| T \| = \sup_{x \in \mathcal{D}(T) \setminus \{0\}} \left(\frac{\| T(x) \|}{\| x \|} \right).$$

is called the norm of the operator.

Using the linearity of the operator T (see LM ii: 2) and the homogeneity of the norm $\| \cdot \|$ (see N 3), the norm of the operator T can also be defined by

$$\| T \| = \sup_{\substack{x \in \mathcal{D}(T), \\ \| x \| = 1}} \| T(x) \|,$$

because

$$\frac{\| T(x) \|}{\| x \|} = \| \frac{1}{\| x \|} T(x) \| = \| T(\frac{x}{\| x \|}) \| \quad \text{and} \quad \| \frac{x}{\| x \|} \| = 1$$

for all $x \in \mathcal{D}(T) \setminus \{0\}$.

A very nice property of linear operators is that boundedness and continuity are equivalent.

Theorem 7.2

Let $T : \mathcal{D}(T) \to Y$ be a linear operator, X and Y are normed spaces and $\mathcal{D}(T) \subset X$, then

a. T is continuous *if and only if* T is bounded,

b. if T is continuous in one point *then* T is continuous on $\mathcal{D}(T)$.

Proof of Theorem 7.2

Let $\epsilon > 0$ be given.

a. (⇒) T is continous in an arbitrary point $x \in \mathcal{D}(T)$. So there exists a $\delta > 0$ such that for every $y \in \mathcal{D}(T)$ with $\| x - y \| \leq \delta$, $\| T(x) - T(y) \| \leq \epsilon$. Take an arbitrary $z \in \mathcal{D}(T) \setminus \{0\}$ and construct $x_0 = x + \frac{\delta}{\| z \|} z$, then $(x_0 - x) = \frac{\delta}{\| z \|} z$ and $\| x_0 - x \| = \delta$. Using the continuity and the linearity of the operator T in x and using the homogenity of the norm gives that

$$\epsilon \geq \| T(x_0) - T(x) \| = \| T(x_0 - x) \| = \| T(\frac{\delta}{\| z \|} z) \| = \frac{\delta}{\| z \|} \| T(z) \| .$$

And the following inequality is obtained: $\frac{\delta}{\| z \|} \| T(z) \| \leq \epsilon$, rewritten it gives that the operator T is bounded

$$\| T(z) \| \leq \frac{\epsilon}{\delta} \| z \| .$$

The constant $\frac{\delta}{\epsilon}$ is independent of z, since $z \in \mathcal{D}(T)$ was arbitrary chosen.

(⇐) T is linear and bounded. Take an arbitrary $x \in \mathcal{D}(T)$. Let $\delta = \frac{\epsilon}{\| T \|}$ then for every $y \in \mathcal{D}(T)$ with $\| x - y \| < \delta$

$$\| T(x) - T(y) \| = \| T(x - y) \| \leq \| T \| \| x - y \| < \| T \| \; \delta = \epsilon.$$

The result is that T is continuous in x, x was arbitrary chosen, so T is continuous on $\mathcal{D}(T)$.

b. (⇒) If T is continuous in $x_0 \in \mathcal{D}(T)$ then is T bounded, see part a ((⇒)), so T is continuous, see Theorem 7.2 ii.a.

□

Theorem 7.3

Let $(X, \| \cdot \|_0)$ and $(Y, \| \cdot \|_1)$ be normed spaces and $T : X \to Y$ be a linear operator. If T is bounded on $B_r(\underline{0}, \| \cdot \|_0)$, for some $r > 0$ then

$$\| T(x) \|_1 \leq \alpha \; \| x \|_0 \quad \text{for all} \quad x \in X \quad \text{and some} \quad \alpha > 0.$$

Proof of Theorem 7.3

Let $\| F(x) \|_1 \leq \beta$, for all $x \in \overline{B_r(\underline{0}, \| \cdot \|_0)}, r > 0$. If $x = \underline{0}$ then $F(x) = \underline{0}$, and if $x \neq \underline{0}$, then since $r \dfrac{x}{\| x \|_0} \in \overline{B_r(\underline{0}, \| \cdot \|_0)}$, the result is that

$$\| F(x) \|_1 = \frac{\| x \|_0}{r} \, \| F(\frac{r\,x}{\| x \|_0}) \|_1 \leq \frac{\beta}{r} \, \| x \|_0 .$$

Take $\alpha = \dfrac{\beta}{r}$.

□

Theorem 7.4

Let $T : \mathcal{D}(T) \rightarrow Y$ be a bounded linear operator, with $\mathcal{D}(T) \subseteq X$ and X, Y are normed spaces *then* the nullspace $\mathcal{N}(T)$ is closed.

Proof of Theorem 7.4

Take a convergent sequence $\{x_n\}_{n \in \mathbb{N}}$ in $\mathcal{N}(T)$.
The sequence $\{x_n\}_{n \in \mathbb{N}}$ is convergent, so there exists some $x \in \mathcal{D}(T)$ such that $\| x - x_n \| \rightarrow 0$ if $n \rightarrow \infty$.
Using the linearity and the boundedness of the operator T gives that

$$\| T(x_n) - T(x) \| = \| T(x_n - x) \| \leq \| T \| \, \| x_n - x \| \rightarrow 0 \, (n \rightarrow \infty). \tag{7.2}$$

The sequence $\{x_n\}_{n \in \mathbb{N}}$ is a subset of $\mathcal{N}(T)$, so $T(x_n) = 0$ for every $n \in \mathbb{N}$.
By 7.2 follows that $T(x) = 0$, this means that $x \in \mathcal{N}(T)$, so $\mathcal{N}(T)$ is closed, see Theorem 2.2.

□

7.3 Space of bounded linear operators

Let X and Y be in first instance arbitrary Vector Spaces. Later on there can also be looked at Normed Spaces, Banach Spaces and other spaces, if necessary. Important is the space of linear operators from X to Y, denoted by $L(X, Y)$.

Definition 7.8

Let $L(X, Y)$ be the set of all the linear operators of X into Y. If $S, T \in L(X, Y)$ then the sum and the scalar multiplication are defined by

$$\begin{cases} (S + T)(x) = S(x) + T(x), \\ (\alpha S)(x) = \alpha (S(x)) \end{cases}$$

for all $x \in X$ and for all $\alpha \in \mathbb{K}$.

Theorem 7.5

The set $L(X, Y)$ is a Vector Space under the linear operations given in Definition 7.8.

Proof of Theorem 7.5

It is easy to check the conditons given in definition 3.1 of a Vector Space.

□

There will be looked at a special subset of $L(X, Y)$, but then it is of importance that X and Y are Normed Spaces. There will be looked at the bounded linear operators of the Normed Space X into the Normed Space Y, denoted by $BL(X, Y)$.

Theorem 7.6

Let $(X, \| \cdot \|_X)$ and $(Y, \| \cdot \|_Y)$ be Normed Spaces over the field $\mathbb{K}$.
The set $BL(X, Y)$ is a linear subspace of $L(X, Y)$.

Proof of Theorem 7.6

The set $BL(X,Y) \subset L(X,Y)$ and $BL(X,Y) \neq \varnothing$, for instance $0 \in BL(X,Y)$, the zero operator. For a linear subspace two conditions have to be checked, see definition 3.2. Let $S, T \in BL(X,Y)$, that means that there are positive constants C_1, C_2 such that

$$\begin{cases} \| S(x) \|_Y \leq C_1 \| x \|_X \\ \| T(x) \|_Y \leq C_2 \| x \|_X \end{cases}$$

for all $x \in X$. Hence,

1.

$$\| (S + T)(x) \|_Y \leq \| S(x) \|_Y + \| T(x) \|_Y \leq C_1 \| x \|_X + C_2 \| x \|_X \leq (C_1 + C_2) \| x \|_X,$$

2.

$$\| (\alpha S)(x) \|_Y = |\alpha| \, \| S(x) \|_Y \leq (|\alpha| \, C_1) \| x \|_X,$$

for all $x \in X$ and for all $\alpha \in \mathbb{K}$. The result is that $BL(X,Y)$ is a subspace of $L(X,Y)$.
□

The space $BL(X,Y)$ is more then just an ordinary Vector Space, if X and Y are Normed Spaces.

Theorem 7.7

If $(X, \| \cdot \|_X)$ and $(Y, \| \cdot \|_Y)$ are Normed Spaces, then $BL(X,Y)$ is a Normed Space, the norm is defined by

$$\| T \| = \sup_{0 \neq x \in X} \frac{\| T(x) \|_Y}{\| x \|_X} = \sup_{\begin{cases} x \in X \\ \| x \| = 1 \end{cases}} \| T(x) \|_Y$$

for every $T \in BL(X,Y)$.

Proof of Theorem 7.7

The norm of an operator is already defined in definition 7.7. It is not difficult to verify that the defined expression satisfies the conditions given in definition 3.23.

□

Remark 7.1

$\| T \|$ is the radius of the smallest ball in Y, around $0\,(\in Y)$, that contains all the images of the unit ball, $\{ x \in X \mid \| x \|_X = 1\}$ in X.

One special situation will be used very much and that is the case that Y is a Banach Space, for instance $Y = \mathbb{R}$ or $Y = \mathbb{C}$.

Theorem 7.8

If Y is a Banach Space, then $BL(X, Y)$ is a Banach Space.

Proof of Theorem 7.8

The proof will be split up in several steps.
First will be taken an Cauchy sequence $\{T_n\}_{n\in\mathbb{N}}$ of operators in $BL(X, Y)$. There will be constructed an operator T? Is T linear? Is T bounded? And after all the question if $T_n \to T$ for $n \to \infty$? The way of reasoning can be compared with the section about pointwise and uniform convergence, see section 2.12. Let's start!
Let $\epsilon > 0$ be given and let $\{T_n\}_{n\in\mathbb{N}}$ be an arbitrary Cauchy sequence of operators in $(BL(X, Y), \| \cdot \|)$.

1. Construct a new operator T:
   Let $x \in X$, then is $\{T_n(x)\}_{n\in\mathbb{N}}$ a Cauchy sequence in Y, since
   $$\| T_n(x) - T_m(x) \|_Y \leq \| T_n - T_m \| \| x \|_X .$$
   Y is complete, so the Cauchy sequence $\{T_n(x)\}_{n\in\mathbb{N}}$ converges in Y. Let $T_n(x) \to T(x)$ for $n \to \infty$. Hence, there is constructed an operator $T : X \to Y$, since $x \in X$ was arbitrary chosen.
2. Is the operator T linear?
   Let $x, y \in X$ and $\alpha \in \mathbb{K}$ then

$$T(x+y) \;=\; \lim_{n\to\infty} T_n(x+y) \;=\; \lim_{n\to\infty} T_n(x) + \lim_{n\to\infty} T_n(y) \;=\; T(x) + T(y)$$

and

$$T(\alpha\, x) \;=\; \lim_{n\to\infty} T_n(\alpha\, x) \;=\; \lim_{n\to\infty} \alpha\, T_n(x) \;=\; \alpha\, T(x)$$

Hence, T is linear.

3. Is T bounded?
The operators $T_n \in BL(X, Y)$ for every $n \in \mathbb{N}$, so

$$\| T_n(x) \|_Y \leq \| T_n \| \, \| x \|_X \, .$$

for every $x \in X$. Further is every Cauchy sequence in a Normed Space bounded. There exists some $N(\epsilon)$ such that $n, m > N(\epsilon)$, using the inverse triangle inequality gives

$$| \, \| T_n \| \, - \, \| T_m \| \, | \leq \| T_n \, - \, T_m \| < \epsilon$$

such that

$$-\epsilon + \, \| T_{N(\epsilon)} \| < \| T_n \| < \, \epsilon + \, \| T_{N(\epsilon)} \|,$$

for all $n > N(\epsilon)$. $N(\epsilon)$ is fixed, so $\{\| T_n \|\}_{n\in\mathbb{N}}$ is bounded. There exists some positive constant K, such that $\| T_n \| < \, K$ for all $n \in \mathbb{N}$. Hence,

$$\| T_n(x) \|_Y < \, K \, \| x \|_X$$

for all $x \in X$ and $n \in \mathbb{N}$. This results in

$$\| T(x) \|_Y \leq \| T(x) - T_n(x) \|_Y \, + \, \| T_n(x) \|_Y \leq \| T(x) - T_n(x) \|_Y \, + K \, \| x \|_X,$$

for all $x \in X$ and $n \in \mathbb{N}$. Be careful! Given some $x \in X$ and $n \to \infty$ then always

$$\| T(x) \|_Y \leq \, K \, \| x \|_X,$$

since $T_n(x) \to T(x)$, that means that $\| T_n(x) \, - \, T(x) \|_Y < \epsilon$ for all $n > N(\epsilon, x)$, since there is pointwise convergence.
Achieved is that the operator T is bounded, so $T \in BL(X, Y)$.

4. Finally, the question if $T_n \to T$ in $(BL(X, Y), \| \cdot \|)$?
The sequence $\{T_n(x)\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $BL(X, Y)$, so there is a $N(\epsilon)$ such that for all $n, m \, > \, N(\epsilon)$: $\quad \| T_n - T_m \| < \, \frac{\epsilon}{2}$.
Hence,

$$\| T_n(x) \, - \, T_m(x) \|_Y < \, \frac{\epsilon}{2} \, \| x \|_X$$

for every $x \in X$. Let $m \rightarrow \infty$ and use the continuity of the norm then

$$\| T_n(x) - T(x) \|_Y \leq \frac{\epsilon}{2} \| x \|_X$$

for every $n > N(\epsilon)$ and $x \in X$, this gives that

$$\frac{\| T_n(x) - T(x) \|_Y}{\| x \|_X} \leq \frac{\epsilon}{2},$$

for every $0 \neq x \in X$ and for every $n > N(\epsilon)$. The result is that

$$\| T_n - T \|_Y < \epsilon.$$

Hence, $T_n \rightarrow T$, for $n \rightarrow \infty$ in $(BL(X,Y), \| \cdot \|)$.

The last step completes the proof of the theorem.

□

7.4 Invertible Linear Operators

In section 2.1 are given the definitions of onto, see 2.5 and one-to-one, see 2.3 and 2.4, look also in the Index for the terms surjective (=onto) and injective (=one-to-one).

First the definition of the algebraic inverse of an operator .

Definition 7.9

Let $T : X \rightarrow Y$ be a linear operator and X and Y Vector Spaces. T is (algebraic) invertible, if there exists an operator $S : Y \rightarrow x$ such that $ST = I_X$ is the identity operator on X and $TS = I_Y$ is the identity operator on Y. S is called the algebraic inverse of T, denoted by $S = T^{-1}$.

Sometimes there is made a distinction between left and right inverse operators, for a nice example see **wiki-l-r-inverse**. It is of importance to know that this distiction can be made. In these lecture notes is spoken about the inverse of T. It can be of importance to restrict the operator to it's domain

$\mathcal{D}(T)$, see figure 7.2. The operator $T : \mathcal{D}(T) \to \mathcal{R}(T)$ is always onto, and the only thing to control if the inverse of T exists, that is to look if the operator is one-to-one.

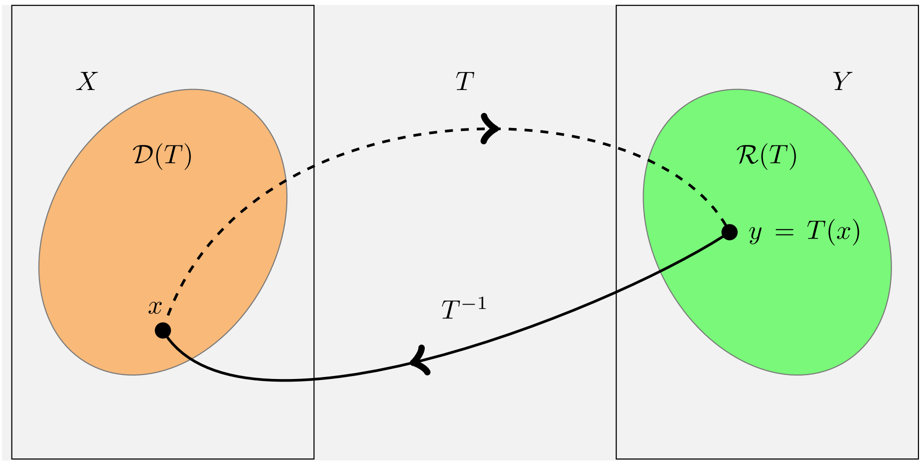


Figure 7.2 The inverse operator: T^{-1}

Theorem 7.9

Let X and Y be Vector Spaces and $T : \mathcal{D}(T) \to Y$ be a linear operator with $\mathcal{D}(T) \subseteq X$ and $\mathcal{R}(T) \subseteq Y$. Then

a. $T^{-1} : \mathcal{R}(T) \to \mathcal{D}(T)$ exists *if and only if*

$$T(x) = 0 \Rightarrow x = 0.$$

b. If T^{-1} exists *then* T^{-1} is a linear operator.

Proof of Theorem 7.9

a. ($\Rightarrow$) If T^{-1} exists, then is T injective and is obtained out of $T(x) = T(0) = 0$ that $x = 0$.

($\Leftarrow$) Let $T(x) = T(y)$, T is linear so $T(x - y) = 0$ and this implies that $x - y = 0$, using the hypothesis an that means $x = y$. T is onto $\mathcal{R}(T)$ and T is one-to-one, so T is invertibe.

b. The assumption is that T^{-1} exists. The domain of T^{-1} is $\mathcal{R}(T)$ and $\mathcal{R}(T)$ is a Vector Space, see Theorem 7.1 ii.a. Let $y_1,\, y_2 \in \mathcal{R}(T)$, so there exist $x_1,\, x_2 \in \mathcal{D}(T)$ with $T(x_1) = y_1$ and $T(x_2) = y_2$. T^{-1} exist, so $x_1 = T^{-1}(y_1)$ and $x_2 = T^{-1}(y_2)$. T is also a linear operator such that $T(x_1 + x_2) = (y_1 + y_2)$ and $T^{-1}(y_1 + y_2) = (x_1 + x_2) = T^{-1}(y_1) + T^{-1}(y_2)$. Evenso $T(\alpha\, x_1) = \alpha\, y_1$ and the result is that $T^{-1}(\alpha\, y_1) = \alpha\, x_1 = \alpha T^{-1}(y_1)$. The operator T^{-1} satisfies the conditions of linearity, see Definition 7.1. (α is some scalar.)

□

In this paragraph is so far only looked at Vector Spaces and not to Normed Spaces. The question could be if a norm can be used to see if an operator is invertible or not?
If the spaces X and Y are Normed Spaces, there is sometimes spoken about the topological inverse T^{-1} of some invertible operator T. In these lecture notes is still spoken about the inverse of some operator and no distinction will be made between the various types of inverses.

Example 7.1

Look to the operator $T : \ell^\infty \rightarrow \ell^\infty$ defined by

$$T(x) = y,\, x = \{\alpha_i\}_{i \in \mathbb{N}} \in \ell^\infty,\, y = \{\frac{\alpha_i}{i}\}_{i \in \mathbb{N}}.$$

The defined operator T is linear and bounded. The range $\mathcal{R}(T)$ is not closed.

The inverse operator $T^{-1} : \mathcal{R}(T) \rightarrow \ell^\infty$ exists and is unbounded.

Explanation of Example 7.1

The linearity of the operator T is no problem.
The operator is bounded because

$$\| \, T(x) \, \|_\infty = \sup_{i \in \mathbb{N}} | \, \frac{\alpha_i}{i} \, | \leq \sup_{i \in \mathbb{N}} | \, \frac{1}{i} \, | \, \sup_{i \in \mathbb{N}} | \, \alpha_i \, | = \| \, x \, \|_\infty \, . \tag{7.3}$$

The norm of T is easily calculated by the sequence $x = \{\, 1 \,\}_{i \in \mathbb{N}}$. The image of x becomes $T(x) = \{\, \frac{1}{i} \,\}_{i \in \mathbb{N}}$ with $\| \, T(x) \, \|_\infty = \| \, \{\, \frac{1}{i} \,\}_{i \in \mathbb{N}} \, \|_\infty = 1$, such

that $\| T(x) \|_{\infty} = \| x \|_{\infty}$. Inequality 7.3 and the just obtained result for the sequence x gives that $\| T \| = 1$.
The $\mathcal{R}(T)$ is a proper subset of ℓ^{∞}. There is no $x_0 \in \ell^{\infty}$ such that $T(x_0) = \{1\}_{i \in \mathbb{N}}$, because $\| x_0 \|_{\infty} = \| \{i\}_{i \in \mathbb{N}} \|_{\infty}$ is not bounded.
Look to the operator $T : \ell^{\infty} \to \mathcal{R}(T)$. If $T(x) = 0 \in \ell^{\infty}$ then $x = 0 \in \ell^{\infty}$, so T is one-to-one. T is always onto $\mathcal{R}(T)$. Onto and one-to-one gives that $T^{-1} : \mathcal{R}(T) \to \ell^{\infty}$ exists.
Look to the sequence $\{y_n\}_{n \in \mathbb{N}}$ with

$$y_n = (\underbrace{1, \frac{1}{\sqrt{2}}, \cdots, \frac{1}{\sqrt{n}}}_{n}, 0, \cdots)$$

and the element $y = (1, \frac{1}{\sqrt{2}}, \cdots, \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n+1}}, \cdots)$. It is easily seen that $y_n \in \ell^{\infty}$ for every $n \in \mathbb{N}$ and $y \in \ell^{\infty}$ and

$$\lim_{n \to \infty} \| y - y_n \|_{\infty} = \lim_{n \to \infty} \frac{1}{\sqrt{n+1}} = 0.$$

If $\mathcal{R}(T)$ is closed then there is an element $x \in \ell^{\infty}$ with $T(x) = y$.
Every y_n is an element out of the range of T, since there is an element $x_n \in \ell^{\infty}$ with $T(x_n) = y_n$,

$$x_n = (\underbrace{1, \sqrt{2}, \cdots, \sqrt{n}}_{n}, 0, \cdots).$$

with $\| x_n \|_{\infty} = \sqrt{n} < \infty$ for every $n \in \mathbb{N}$.
The sequence $\{x_n\}_{n \in \mathbb{N}}$ does not converge in ℓ^{∞}, since $\lim_{n \to \infty} \| x_n \|_{\infty} = \lim_{n \to \infty} \sqrt{n}$ not exists. The result is that there exists no element $x \in \ell^{\infty}$ such that $T(x) = y$,
the $\mathcal{R}(T)$ is not closed.
Another result is that the limit for $n \to \infty$ of

$$\frac{\| T^{-1}(y_n) \|_{\infty}}{\| y_n \|_{\infty}} = \frac{\sqrt{n}}{1}$$

does not exist. The inverse operator $T^{-1} : \mathcal{R}(T) \to \ell^{\infty}$ is not bounded. □

In example 7.1, the bounded linear operator T is defined between Normed Spaces and there exists an inverse operator T^{-1}. It is an example of an operator which is topological invertible, T^{-1} is called the topological inverse of T.

Definition 7.10

Let $T : X \to Y$ be a linear operator and X and Y Normed Spaces. T is (topological) invertible , if the algebraic inverse T^{-1} of T exists and also $\| T \|$ is bounded. T^{-1} is simply called the inverse of T.

Example 7.1 makes clear that the inverse of a bounded operator need not to be bounded. The inverse operator is sometimes bounded.

Theorem 7.10

Let $T : X \to Y$ be a linear and bounded operator from the Normed Spaces $(X, \| \cdot \|_1)$ onto the Normed Space $(Y, \| \cdot \|_2)$,
T^{-1} exists and is bounded *if and only if* there exists a constant $K > 0$ such that

$$\| T(x) \|_2 \geq K \ \| x \|_1 \tag{7.4}$$

for every $x \in X$. The operator T is called bounded from below.

Proof of Theorem 7.10

($\Rightarrow$) Suppose T^{-1} exists and is bounded, then there exists a constant $C_1 > 0$ such that $\| T^{-1}(y) \|_1 \leq C_1 \ \| y \|_2$ for every $y \in Y$. The operator T is onto Y that means that for every $y \in Y$ there is some $x \in X$ such that $y = T(x)$, x is unique because T^{-1} exists. Altogether

$$\| x \|_1 = \| T^{-1}(T(x)) \|_1 \leq C_1 \ \| T(x) \|_2 \Rightarrow \| T(x) \|_2 \geq \frac{1}{C_1} \ \| x \|_1 \tag{7.5}$$

Take $K = \dfrac{1}{C_1}$.

($\Leftarrow$) If $T(x) = 0$ then $\| T(x) \|_2 = 0$, using equality 7.4 gives that $\| x \|_1 = 0$ such that $x = 0$. The result is that T is one-to-one, together with the fact that T is onto, it follows that the inverse T^{-1} exists.

In Theorem 7.9 ii.b is proved that T^{-1} is linear.
Almost on the same way as in 7.5 there can be proved that T^{-1} is bounded,

$$\| T(T^{-1}(y)) \|_2 \geq K \| T^{-1}(y) \|_1 \Rightarrow \| T^{-1}(y) \|_1 \leq \frac{1}{K} \| y \|_2,$$

for every $y \in Y$, so T^{-1} is bounded.

□

The inverse of a composition of linear operators can be calculated, if the individual linear operators are bijective, see figure 7.3.

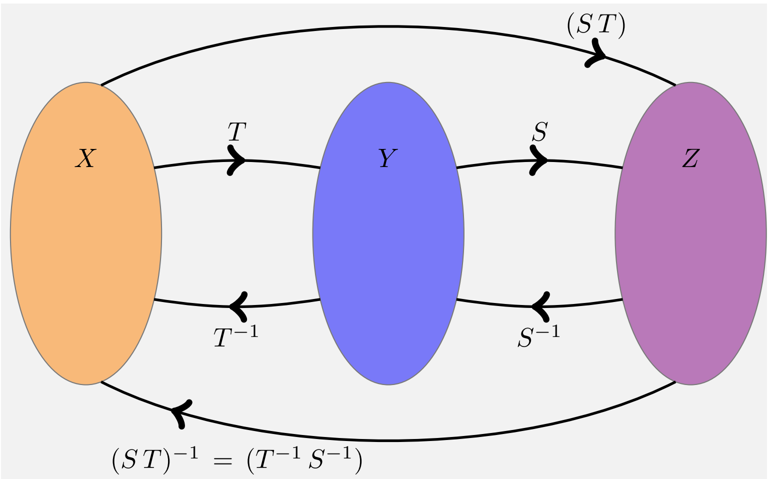


Figure 7.3 Inverse Composite Operator

Theorem 7.11

If $T : X \to Y$ and $S : Y \to Z$ are bijective linear operators, where X, Y and Z are Vector Spaces. Then the inverse $(S\,T)^{-1} : Z \to X$ exists and is given by

$$(S\,T)^{-1} = T^{-1}\,S^{-1}.$$

Proof of Theorem 7.11

The operator $(S\,T) : X \rightarrow Z$ is bijective, T^{-1} and S^{-1} exist such that

$(S\,T)^{-1}\,(ST) = (T^{-1}\,S^{-1})\,(S\,T) = T^{-1}\,(S^{-1}\,S)\,T = T^{-1}\,(I_Y\,T) = T^{-1}\,T = I_X$

with I_X and I_Y the identity operators on the spaces X and Y.

□

7.4.1 Power Series in $BL(X, X)$

Sometimes the inverse of an operator can be given by a Neumann series.

Theorem 7.12

Let $T \in BL(X, X)$, where $(X, \|\cdot\|)$ is a Banach Space and suppose that $\| I - T \| < 1$. Then T is invertible, where T^{-1} is given by the Neumann series

$$T^{-1} = \sum_{n=0}^{\infty} (I - T)^n. \tag{7.6}$$

The given series 7.6 converges in the operator norm and $T^{-1} \in BL(X, X)$

Proof of Theorem 7.12

The proof will be done in several steps:

i. Since I and T are bounded, so $(I - T)^n$ are bounded for every $n \in \mathbb{N} \cap \{0\}$.

ii. If $|x| < 1$ then $\dfrac{1}{1 - x} = \sum_{n=0}^{\infty} x^n$, replace $x = 1 - y$ and this leads to

$$\frac{1}{y} = \sum_{n=0}^{\infty} (1 - y)^n.$$

iii. There is given that $\| I - T \| = \alpha < 1$, so

$$\| x \| = \| (I - T)(x) + T(x) \| \leq$$
$$\| (I - T)(x) \| + \| T(x) \| \leq \alpha \| x \| + \| T(x) \| .$$

Therefore $\| T(x) \| \geq (1 - \alpha) \| x \|$, T is bounded from below and the result is that the operator T is invertible and $\| T^{-1} \| \leq (1 - \alpha)^{-1}$, see Theorem 7.10.

iv. Define the operator T_N by

$$T_N = \sum_{n=0}^{N} (I - T)^n,$$

$T_N \in L(X, X)$ for each N. Since X is a Banach Space, $BL(X, X)$ is a Banach Space, see Theorem 7.8. If $N > M$ then

$$\| T_N - T_M \| \leq \| \sum_{n=M+1}^{N} (I - T)^n \| \leq \sum_{n=M+1}^{N} \| (I - T)^n \| \leq$$
$$\sum_{n=M+1}^{N} \| (I - T) \|^n \leq \sum_{n=M+1}^{N} \alpha^n \rightarrow 0,$$

if $N, M \rightarrow \infty$. Therefore $\{T_N\}_{N \in \mathbb{N}}$ is a Cauchy sequence in $BL(X, X)$, so there exists some $S \in BL(X, X)$ such that $\| T_N - S \| \rightarrow 0$ for $N \rightarrow \infty$.

v. It has to be shown that $S = T^{-1}$. Let $y \in X$ and let $x = S(y)$. There has to be shown that $T(x) = y$, or equivalently $(I - T)(x) = (x - y)$.
Let's try to do:

$$(I-T)(x) = (I-T)S(y) = (I-T)(\lim_{N\to\infty} T_N)(y) =$$

$$(I-T)(\lim_{N\to\infty}\sum_{n=0}^{N}(I-T)^n)(y) = (\lim_{N\to\infty}\sum_{n=1}^{N}(I-T)^n)(y) =$$

$$(\lim_{N\to\infty}\sum_{n=0}^{N}(I-T)^n)(y) - I(y) = S(y) - y = (x-y).$$

Therefore, $S = T^{-1} \in BL(X,X)$.

□

Theorem 7.13

Let $(X, \| \cdot \|)$ be a Banach Space.

a. If $A \in BL(X,X)$ and invertible and $B \in BL(X,X)$, with $\| A^{-1}B \| < 1$, then $A+B$ invertible.

b. The set $IBL(X,X)$ of bounded invertible linear operators is open in $BL(X,X)$.

c. The inversion operator $INV : A \rightarrow A^{-1}$ is continuous on $BL(X,X) \cap IBL(X,X)$.

Proof of Theorem 7.13

The proofs of the different propositions.

a. Use Theorem 7.12 with $T = I + (A^{-1}B)$ then $(I-T) = -A^{-1}B : X \rightarrow X$ and $\| I - T \| = \| -A^{-1}B \| = \| A^{-1}B \| < 1$, so $T^{-1} = (I + (A^{-1}B))^{-1}$ exists and is given by

$$(I + (A^{-1}B))^{-1} = \sum_{n=0}^{\infty}(-1)^n(A^{-1}B)^n$$

and $(I + (A^{-1}B))^{-1}A^{-1} = (A(I + A^{-1}B))^{-1} = (A+B)^{-1}$.

b. Be careful: a bounded operator can be invertible, with a unbounded inverse operator, see Example 7.1.
Let $A \in IBL(X,X) \cap BL(X,X)$ then $A^{-1} \in BL(X,X)$. Take $\epsilon = \dfrac{1}{\| A^{-1} \|}$ and let $B \in B_{\epsilon}(A) \subset IBL(X,X)$ then

$$B = (B - A) + A = A(I + A^{-1}(B - A))$$

with $\| A^{-1}(B-A) \| \leq \| A^{-1} \| \| (B-A) \| < 1$, so B is invertible, see part ii.a.

c. Here to proof that the operator INV is continuous in A. The operator A is invertible and it's inverse A^{-1} is bounded. Given is some $0 < \epsilon < \dfrac{\| A^{-1} \|}{2}$.

i. First the wrong version!

$$\| INV(A) - INV(B) \| = \| A^{-1}(A - B)B^{-1} \| \leq$$
$$\| A^{-1} \| \| B^{-1} \| \| (A - B) \|$$

The problem is to find a $\delta(\epsilon) > 0$, independent of $\| B^{-1} \|$.

ii. Here the proper version. The parts ii.a and ii.b will be used. Take B in $B_{\delta}(A) \subset (IBL(X,X) \cap BL(X,X))$, with

$$\delta < \frac{\epsilon}{3 \| A^{-1} \|^2} \,(< \frac{1}{6 \| A^{-1} \|}),$$

then

$$\| A^{-1}(B - A) \| \leq \| A^{-1} \| \| (B - A) \| < \frac{\epsilon}{3 \| A^{-1} \|} < 1,$$

so $(I + A^{-1}(B-A))^{-1}$ exists and $\| INV(A) - INV(B) \| =$

$$\| A^{-1} - B^{-1} \| = \| A^{-1} - (B - A + A)^{-1} \| =$$
$$\| A^{-1} - (A(A^{-1}(B - A) + I))^{-1} \| =$$
$$\| A^{-1} - (I + A^{-1}(B - A))^{-1} A^{-1} \| =$$
$$\| (I - \sum_{n=0}^{\infty} (-1)^n (A^{-1}(B - A))^n) A^{-1} \| =$$

$$\| (\sum_{n=1}^{\infty}(-1)^n(A^{-1}(B-A))^n)A^{-1} \| \leq$$

$$\| A^{-1} \| \sum_{n=1}^{\infty} \| A^{-1} \|^n \| (B-A) \|^n \leq$$

$$\| A^{-1} \|^2 \frac{\epsilon}{3 \| A^{-1} \|^2} \frac{1}{(1 - \frac{\epsilon}{3 \| A^{-1} \|})} \leq \frac{\epsilon}{10} < \epsilon.$$

□

7.5 Projection operators

For the concept of a projection operator, see section 3.10.1.

Definition 7.11

See theorem 3.22, y_0 is called the projection of x on M, denoted by

$$P_M : x \rightarrow y_0, \text{ or } y_0 = P_M(x),$$

P_M is called the projection operator on M, $P_M : X \rightarrow M$.

But if M is just a proper subset and not a linear subspace of some Inner Product Space then the operator P_M, as defined in 7.11, is not linear. To get a linear projection operator M has to be a closed linear subspace of a Hilbert Space H.

Theorem 7.14

Projection Theorem

If M is a closed linear subspace of a Hilbert Space H then

$$H = M \oplus M^{\perp},$$
$$x = y + z.$$

Every $x \in H$ has a unique representation as the sum of $y \in M$ and $z \in M^{\perp}$, y and z are unique because of the direct sum of M and $M^{\perp}$.

Proof of Theorem 7.15

For the proof, see Theorem **3.25**. □

Definition 7.12

Let $T : X \to Y$ be a linear operator and X and Y Normed Spaces, the operator T is called idempotent, if $T^2 = T$, thus

$$T^2(x) = T(Tx) = T(x)$$

for every $x \in X$.

Remark 7.2

The projection operator P_M maps

a. X onto M and

b. M onto itself

c. $M^{\perp}$ onto $\{0\}$.

and is idempotent .

Remark 7.3

The projection operator P_M on M is idempotent, because $P_M(P_M(x)) = P_M(y_0) = y_0 = P_M(x)$, so $(P_M P_M)(x) = P_M(x)$.

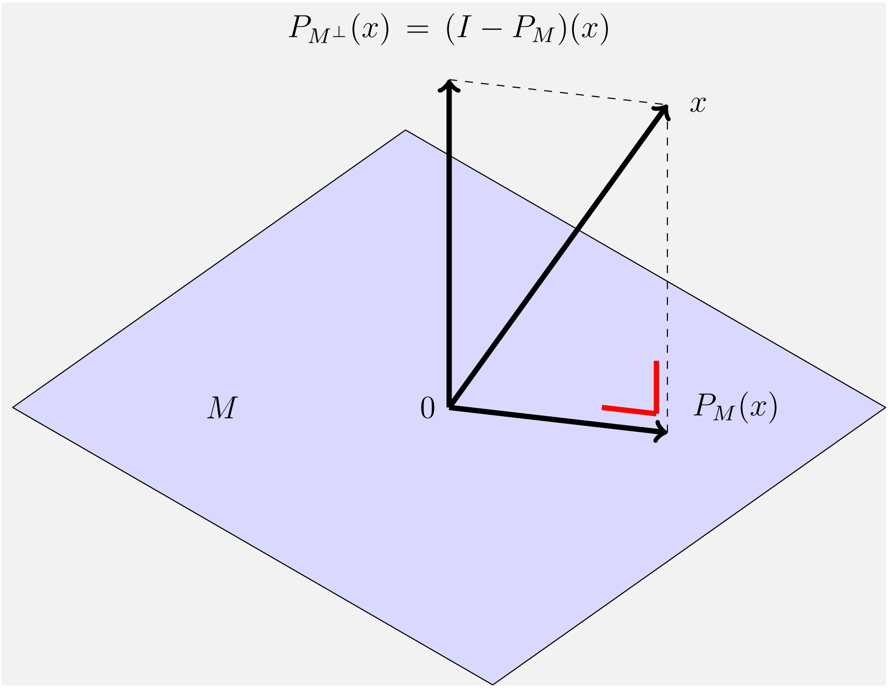


Figure 7.4 Orthogonal projection on a subspace M.

The projection operator P_M is called an orthogonal projection on M, see figure 7.4, because the nullspace of P_M is equal to $M^\perp$ (the orthogonal complement of M) and P_M is the identity operator on M. So every $x \in H$ can be written as

$$x = y + z = P_M(x) + P_{M^\perp}(x) = P_M(x) + (I - P_M)(x).$$

7.6 Adjoint operators

In first instance, it is the easiest way to introduce adjoint operators in the setting of Hilbert Spaces, see page 89. But the concept of the adjoint operator can also be defined in Normed Spaces.

Theorem 7.15

If $T : H \to H$ is a bounded linear operator on a Hilbert Space H, then there exists an unique operator $T^* : H \to H$ such that

$$(x, T^* y) \;=\; (Tx, y) \quad \text{for all} \quad x, y \in H.$$

The operator T^* is linear and bounded, $\| T^* \| = \| T \|$ and $(T^*)^* \;=\; T$. The operator T^* is called the adjoint of T.

Proof of Theorem 7.15

The proof exists out of several steps. First the existence of such an operator T^* and then the linearity, the uniqueness and all the other required properties.

a. Let $y \in H$ be fixed. Then the functional defined by $f(x) \;=\; (Tx, y)$ is linear, easy to prove. The functional f is also bounded since $| f(x) | = | (Tx, y) | \leq \| T \| \| x \| \| y \|$. The Riesz representation theorem, see theorem 3.29, gives that there exists an unique element $u \in H$ such that

$$(Tx, y) \;=\; (x, u) \quad \text{for all} \quad x \in H. \tag{7.7}$$

The element $y \in H$ is taken arbitrary. So there is a rule, given $y \in H$, which defines an element $u \in H$. This rule is called the operator T^*, such that $T^*(y) \;=\; u$, where u satisfies 7.7.

b. T^* satisfies $(x, T^* y) \;=\; (Tx, y)$, for all $x, y \in H$, by definition and that is used to prove the linearity of T^*. Take any $x, y, z \in H$ and any scalars $\alpha, \beta \in \mathbb{K}$ then

$$
\begin{aligned}
(x, T^*(\alpha y + \beta z)) &= (T(x), \alpha y + \beta z) \\
&= \overline{\alpha}(T(x), y) + \overline{\beta}(T(x), z) \\
&= \overline{\alpha}(x, T^*(y)) + \overline{\beta}(x, T^*(z)) \\
&= (x, \alpha T^*(y)) + (x, \beta T^*(z)) \\
&= (x, \alpha T^*(y) + \beta T^*(z))
\end{aligned}
$$

If $(x, u) = (x, v)$ for all $x \in H$ then $(x, u - v) = 0$ for all x and this implies that $u - v = 0 \in H$, or $u = v$. Using this result, together with the results of above, it is easily deduced that

$$T^*(\alpha y + \beta z) = \alpha T^*(y) + \beta T^*(z).$$

There is shown that T^* is a linear operator.

c. Let T_1^* and T_s^* be both adjoints of the same operator T. Then follows out of the definition that $(x, (T_1^* - T_2^*)y) = 0$ for all $x, y \in H$. This means that $(T_1^* - T_2^*)y = 0 \in H$ for all $y \in H$, so $T_1^* = T_2^*$ and the uniqueness is proved.

d. Since

$$(y, Tx) = (T^*(y), x) \quad \text{for all} \quad x, y \in H,$$

it follows that $(T^*)^* = T$. Used is the symmetry (or the conjugate symmmetry) of an inner product.

e. The last part of the proof is the boundedness and the norm of T^*.
The boundedness is easily achieved by

$$
\begin{aligned}
\| T^*(y) \|^2 &= (T^*(y), T^*(y)) \\
&= (T\,(T^*(y)), y) \\
&\leq \| T(T^*(y)) \| \| y \| \\
&\leq \| T \| \| T^*(y) \| \| y \| .
\end{aligned}
$$

So, if $\| T^*(y) \| \neq 0$ there is obtained that

$$\| T^*(y) \| \leq \| T \| \| y \|,$$

which is also true when $\| T^*(y) \| = 0$. Hence T^* is bounded

$$\| T^* \| \leq \| T \| . \tag{7.8}$$

Formula 7.8 is true for every operator, so also for the operator T^*, what means that $\| T^{**} \| \leq \| T^* \|$ and $T^{**} = T$. Combining the results of above results in $\| T^* \| = \| T \|$.

□

Lemma 7.1

If S is a subspace of a Hilbert space H then $S^{\perp}$ is closed.

Proof of Theorem 7.1

$S^{\perp}$ is a linear subspace of H. Let $\{t_n\}_{n \in \mathbb{N}}$ be a sequence in $S^{\perp}$ converging to t_0. An inner product is continuous, so for all $s \in S$,

$$(t_0, s) = \lim_{n \to \infty} (t_n, s) = 0,$$

so $t_0 \in S^{\perp}$. □

Theorem 7.16

If $T : H \to H$ is a bounded linear operator on a Hilbert Space H, and T^* its adjoint operator then:

a. $N(T) = (R(T^*))^{\perp}$,

b. $\overline{R(T)} = (N(T^*)^{\perp}$.

Proof of Theorem 7.16

The operators T and T^* are bounded, so the nullspaces $N(T)$ and $N(T^*)$ are closed, see Theorem 7.4.

a. If $x \in N(T)$ then $0 = (T(x), y) = (x, T^*(y))$ for every $y \in H$, so $x \in (R(T^*))^{\perp}$.
If $x \in (R(T^*))^{\perp}$ then $0 = (x, T^*(y)) = (T(x), y)$ for every $y \in H$, that means that $T(x) = 0$, so $x \in N(T)$.

b. If $y \in R(T)$ and $x \in H$ such that $y = T(x)$. Let $z^* \, in \, N(T^*)$ then

$$(y, z) = (T(x), z) = (x, T^*(z)) = 0,$$

so $R(T) \subset (N(T^*)^{\perp}$. Since $(N(T^*)^{\perp}$ is closed, see Lemma 7.1, there follows that

$$\overline{R(T)} \subset (N(T^*))^{\perp}.$$

If $z \in R(T)^{\perp}$ then for all $x \in H$

$$(x, T^*(z)) = (T(x), z) = 0,$$

so $T^*(z) = 0$. This means that $R(T)^{\perp} \subset N(T^*)$. Take on both sides the orthogonal complement and there follows that

$$(N(T^*))^{\perp} \subset (R(T))^{\perp\perp} = \overline{R(T)}.$$

□

Definition 7.13

If $T : H \to H$ is a bounded linear operator on a Hilbert Space H then T is said to be

a. *self-adjoint* if $T^* = T$,

b. *unitary* , if T is bijective and if $T^* = T^{-1}$,

c. *normal* if $T\,T^* = T^*\,T$.

Theorem 7.17

If $T : H \rightarrow H$ is a bounded self-adjoint linear operator on a Hilbert Space H then

a. the eigenvalues of T are real, if they exist, and

b. the eigenvectors of T corresponding to the eigenvalues λ, μ, with $\lambda \neq \mu$, are orthogonal,

for *eigenvalues* and *eigenvectors*, see **definition 7.4**.

Proof of Theorem 7.17

a. Let λ be an eigenvalue of T and x an corresponding eigenvector. Then $x \neq 0$ and $Tx = \lambda x$. The operator T is selfadjoint so

$$\lambda\,(x, x) = (\lambda\,x, x) = (Tx, x) = (x, T^{*}x)$$
$$= (x, Tx) = (x, \lambda x) = \overline{\lambda}\,(x, x).$$

Since $x \neq 0$ gives division by $\| x \|^2$ $(\neq 0)$ that $\lambda = \overline{\lambda}$. Hence λ is real.

b. T is self-adoint, so the eigenvalues λ an μ are real. If $Tx = \lambda\,x$ and $Ty = \mu\,y$, with $x \neq 0$ and $y \neq 0$, then

$$\lambda(x, y) = (\lambda x, y) = (Tx, y) =$$
$$(x, Ty) = (x, \mu y) = \mu(x, y).$$

Since $\lambda \neq \mu$, it follows that $(x, y) = 0$, which means that x and y are orthogonal.

□

7.7 Mapping Theorems

In this chapter are given important theorems, sometimes called the fundamental theorems for operators. Most of the time this will be operators on Banach Spaces, but the definition of certain kind of operators are given with respect to Normed Spaces.

The idea was to start with the Closed Graph Theorem, but to prove that theorem, the theorem of Baire's Category Theorem is needed, evenso to the proof of the Open Mapping Theorem. The Open Mapping Theorem and the Closed Graph Theorem are said to be equivalent to the so-called Bounded Inverse Theorem.

Another important theorem is the Banach-Steinhaus Theorem, also called the Uniform Boundedness Principle. These theorems are of great importance within the functionanalysis.

The Baire's Category Principle is also mentioned in **Section 9.3**. The proof of one of the variants of Baire's theorem will be given is this section. The theorem will be defined for complete Metric Spaces and that declares also the fact that the important theorems are often given with respect to Banach Spaces.

The T_i-spaces, $i = 0, \cdots, 4$, are also of importance, see for the definition of these spaces **Section 3.3.1**.

A lot of interesting material can be found in the book of (Kuttler, 2009).

But first are given the definitions of a closed linear operator and an open mapping.

Definition 7.14

Let $(X, \| \cdot \|_1)$ and $(Y, \| \cdot \|_2)$ be normed spaces. Then the linear operator $T : \mathcal{D}(T) \to Y$ is called a closed linear operator if its graph $\mathcal{G}(T)$, see **definition 7.2**, is closed in the normed space $X \times Y$. The norm on $X \times Y$ is defined by

$$\| (x, y) \| = \| x \|_1 + \| y \|_2 .$$

Let $T : X \to Y$ be a linear operator between Normed Spaces X and Y.

See **theorem 2.7** for the fact that:

T is continuous if and only if $x_n \to x$ implies that $T(x_n) \to T(x)$. Nothing is said about x and $T(x)$.

Theorem 7.18

Let $T : X \to Y$ be a linear operator between the normed spaces $(X, \|\cdot\|_1)$ and $(Y, \|\cdot\|_2)$. $\mathcal{G}(T)$ is closed if and only if the convergence of the sequences $\{x_n\} \subset X$ and $\{T(x_n)\} \subset Y$ implies $x_n \to x \in \mathcal{D}(T)$ and $T(x_n) \to y = T(x)$.

Proof of Theorem 7.18

$\mathcal{G}(T)$ is closed if and only if $(x, y) \in \overline{\mathcal{G}(\mathrm{T})} \Rightarrow (x, y) \in \mathcal{G}(T)$. With **theorem 2.2** $(x, y) = \overline{\mathcal{G}(\mathrm{T})}$ if and only if there exist $(x_n, T(x_n)) \in \mathcal{G}(T)$ such that $(x_n, T(x_n)) \to (x, y)$, hence

$$x_n \to x,\, T(x_n) \to y;$$

and $(x, y) \in \mathcal{G}(T)$ if and only if $x \in \mathcal{D}(T)$ and $T(x) = y$.

□

Example 7.2

Boundedness does not imply closedness:

Let $T : \mathcal{D}(T) \to \mathcal{D}(T) \subset X$ be the identity operator on $\mathcal{D}(T)$, where $\mathcal{D}(T)$ is a proper dense subspace of a Normed Space X. Then it is trivial that T is linear and bounded, but T is not closed. This follows from **Theorem 7.18**. Take $x \in X \setminus \mathcal{D}(T)$ and a sequence $\{x_n\} \subset \mathcal{D}(T)$ which converges to x.

Example 7.3

Closedness does not imply boundedness:

Let $X = C[0,1]$, with norm $\| x \| = \sup_{t\in[0,1]} |x(t)|$, and $T : \mathcal{D}(T) \to X$ and $T(x) = \dfrac{\mathrm{d}}{\mathrm{d}t}x$, with $\mathcal{D}(T)$ the subspace of functions $x \in X$ which have a continuous derivative. It is worth noting that $\mathcal{D}(T)$ is not closed in X.
The operator T is unbounded, take: $x_n(t) = t^n$ with $n \in \mathbb{N}$.
Let $\{x_n\} \subset \mathcal{D}(T)$ and $\{T(x_n)\}$ be such that both sequences converge, $x_n \to x$ and $T(x_n) = x_n^{'} \to y$. The convergence in the norm of $C[0,1]$ is uniform, so from $x_n^{'} \to y$, there follows that

$$\int_0^t y(\tau)\mathrm{d}\tau = \int_0^t \lim_{n\to\infty} x_n^{'}(\tau)\mathrm{d}\tau = \lim_{n\to\infty} \int_0^t x_n^{'}(\tau)\mathrm{d}\tau = x(t) - x(0).$$

That gives that $x(t) = x(0) + \int_0^t y(\tau)\mathrm{d}\tau$, so $x \in \mathcal{D}(T)$ and $x^{'} = y$, Theorem 7.18 implies that T is closed.

Theorem 7.19

Let $T : X \to Y$ be a bijective linear operator between the normed spaces $(X, \| \cdot \|_1)$ and $(Y, \| \cdot \|_2)$. If T is closed linear operator then T^{-1} is also a closed linear operator.

Proof of Theorem 7.19

Suppose that $\{y_n\}_{n\in\mathbb{N}} \subset Y$ such that $y_n \to y$ and $T^{-1}(y_n) \to x$.
The question is, if $T^{-1}(y) = x$.
T is bijective, so there exist a sequence $\{x_n\}_{n\in\mathbb{N}} \subset X$, with $x_n \to x$, take $x_n = T^{-1}(y_n)$ for $n = 1, 2, \cdots$. So $x_n \to x$ and $T(x_n) \to y$.
Since T is closed operator, the $\mathcal{G}(T)$ is closed, so $x \in X$ and $y = T(x) \in \mathcal{D}(T^{-1})$ and $x = T^{-1}(y)$.
□

Be careful in the use of the following theorem ??, the bounded operator is defined on the whole space!

Theorem 7.20

Let $T : X \to Y$ be a linear operator between the normed spaces $(X, \|\cdot\|_1)$ and $(Y, \|\cdot\|_2)$. If T is bounded then $\mathcal{G}(T)$ is closed, so T is a closed.

Proof of Theorem 7.20

Let $x_n \to x$ in X then $x \in X$ and
$\| T(x_n) - T(x) \|_2 \leq \| T \| \| x_n - x \|_1 \to 0$.
□

Theorem 7.21

Let $(X, \|\cdot\|_0)$ and $(Y, \|\cdot\|_1)$ be Normed Spaces. Let $T : \mathcal{D}(T) \to Y$ be a bounded linear operator with $\mathcal{D}(T) \subset X$.
If $\mathcal{D}(T)$ is a closed subset of X then T is closed.

Proof of Theorem 7.21

If $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{D}(T)$ such that $\underline{x_n \to} x$ and is such that $\{T(x_n)\}$ also converges. $\mathcal{D}(T)$ is closed, so $x \in \overline{\mathcal{D}(\mathrm{T})} = \mathcal{D}(T)$ and $T(x_n) \to T(x)$, since T is bounded. Hence T is closed by **Theorem 7.18**.
□

Theorem 7.22

Let $(X, \|\cdot\|_0)$ be a Normed Space and $(Y, \|\cdot\|_1)$ a Banach Spaces. Let T be a linear operator with $\mathcal{D}(T) \subset X$ and $\mathcal{R}(T) \subset Y$. Suppose that T is closed and continuous. Then $\mathcal{D}(T)$ is closed.

Proof of Theorem 7.22

Suppose that $x \in \overline{\mathcal{D}(T)}$ then there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{D}(T)$ such that $x_n \to x$. The sequence $\{T(x_n)\}_{n\in\mathbb{N}}$ is a Cauchy sequence since $\| T(x_n) - T(x_m) \|_1 \leq \| T \| \| x_n - x_m \|_0$. So the sequence $\{T(x_n)\}_{n\in\mathbb{N}}$ has some limit $y \in Y$. T is closed, so $T(x) = y$ but then $x \in \mathcal{D}(T)$.

□

Definition 7.15

Let $(X, \| \cdot \|_0)$ and $(Y, \| \cdot \|_1)$ be two Normed Spaces and T is some linear operator, defined on its domain $\mathcal{D}(T) \subset X$.
The closure of an operator $T : X \to Y$ is the operator $\overline{\mathrm{T}}$, whose domain and action are:

- $\mathcal{D}(\overline{\mathrm{T}}) := \{x \in X \mid \exists y \in Y, \text{such that for any sequence } \{x_n\} \subset \mathcal{D}(T) \text{with} x_n \to x, T(x_n) \to y\}$
- $\overline{\mathrm{T}}(x) := y$ for any $x \in \mathcal{D}(\overline{\mathrm{T}})$.

Definition 7.15 is well-posed, because y is uniquely identified by x and $\mathcal{D}(\overline{\mathrm{T}})$ is a linear operator. Also $\mathcal{D}(T) \subset \mathcal{D}(\overline{\mathrm{T}})$ and $\overline{\mathrm{T}}(x) = T(x)$ for every $x \in \mathcal{D}(T)$.

Definition 7.16

Let (X, d_1) and (Y, d_2) be metric spaces. Then the map $T : \mathcal{D}(T) \to Y$ is called an open mapping if for every open set in $\mathcal{D}(T) \subset X$ the image is an open set in Y.

Remark 7.4

Do not confuse definition 7.16 with the property of an continuous map $T : \mathcal{D}(T) \to \mathcal{R}(T)$ of which $T^{-1}(W)$ is always open, for every open set $W \subset \mathcal{R}(T)$.
Take for instance $f : x \to \sin(x)$ at the open interval $(0, 2\pi)$, here the image is the closed interval $[-1, +1]$.

7.7.1 Baire's Category Theorem

Baire made a great contribution to the functional analysis, nowadays known as the Baire's Category Theorem. It is first given with not too much mathematical terms.

Theorem 7.23

Baire's Category Theorem

Let (X, d) be complete Metric Space and let $(F_n)_{n\geq 1}$ be a sequence of closed sets with empty interiors. Then the interior of $\cup_{n\geq 1} F_n$ is also empty.

In other words, the Euclidean plane $\mathbb{R}^2$ can not be written as the union of *countably many* straight lines.

The term (everywhere) dense is already defined in **definition 2.2**.

Definition 7.17

Let (X, d) be a Metric Space and let $M \subseteq X$ be given.

1. M is nowhere dense or rare if $X \setminus \overline{M}$ is dense in X.
2. M is meager or of first category in X, if it is the union of countable many sets each of which is nowhere dense in X.
3. M is nonmeager or second category in X, if it is not meager in X.

The next version of Baire's Category Theorem comes from the book written by (Limaye, 2008). This version gives the importance of the completeness condition.

Theorem 7.24

Baire's Category Theorem

Let (X, d) be a Metric Space.
Then the intersection of a finite number dense open subsets of X is dense in X.
If X is complete, then the intersection of a countable number of dense open subsets of X is dense in X.

Proof of Theorem 7.24

Let D_1, D_2, $\cdots$ be dense open subsets of X, so $\overline{D_i} = X$ for each $i \in \mathbb{N}$.
For $x_0 \in X$ and $r_0 > 0$, consider $U_0 = B_{r_0}(x_0, d)$.
D_1 is open and dense in X, let $x_1 \in (D_1 \cap U_0)$.
$D_1 \cap U_0$ is open in X, so there is some $r_1 > 0$ such that
$U_1 = B_{r_1}(x_1, d) \subset (D_1 \cap U_0)$.
The construction of the sets U_i can be inductively repeated.
Suppose that $U_{n-1} = B_{r_{n-1}}(x_{n-1}, d)$ and $U_n = B_{r_n}(x_n, d)$ are such that
$U_n \subset (D_n \cap U_{n-1})$.
D_{n+1} is open and dense in X, let $x_{n+1} \in (D_{n+1} \cap U_n)$.
$D_{n+1} \cap U_n$ is open in X, so there is some $r_{n+1} > 0$ such that
$U_{n+1} = B_{r_{n+1}}(x_{n+1}, d) \subset (D_{n+1} \cap U_n)$.
So there are x_1, $x_2, \cdots$ in X and positive numbers r_1, $r_2, \cdots$ such that
$U_m = B_{r_m}(x_m, d) \subset (D_m \cap U_{m-1})$ for $m = 1,\, 2 \cdots$.
So it is clear that for some given $n = 1,\, 2, \cdots$, $x_n \in ((\cap_{m=1}^{n} D_m) \cap U_0) \neq \varnothing$.
x_0 and r_0 are arbitrary chosen, and so it becomes clear that $(\cap_{m=1}^{n} D_m)$ is dense in X, exactly according to **definition 2.2**.

What in the case that X is complete? Just as above, a sequence $\{x_m\}_{m \in \mathbb{N}}$ in X and a decreasing sequence of positive numbers $\{r_m\}_{m \in \mathbb{N}}$ can be found.
There can also be assumed that $r_m \leq \dfrac{1}{m}$ as well as $\overline{U_m} \subset (D_m \cap U_{m-1})$ for $m = 1,\, 2, \cdots$.
Fix a positive number N.
If i, $j \geq N$ then follows for x_i, $x_j \in U_N = B_{r_N}(x_N, d)$ that:

$$d(x_i,\, x_j) \leq d(x_i,\, x_N) + d(x_N,\, x_j) < \frac{2}{r_N} \leq \frac{2}{N}.$$

Hence the sequence $\{x_m\}_{m\in\mathbb{N}}$ is a Cauchy sequence in X. Since X is complete the Cauchy sequence converges in X, let $x_m \to x \in X$.
But there is more to achieve. Since $x_n \in U_N$ for all $n \geq N$, it follows that $x \in \overline{U_N}$.
Since $\overline{U_N} \subset (D_N \cap U_0)$ for all $N = 1, 2, \cdots$, the result is that $x \in ((\cap_{N=1}^{\infty} D_N) \cap U_0)$, so $(\cap_{N=1}^{\infty} D_N) \cap U_0) \neq \varnothing$.
And again, since x_0 and r_0 are arbitrary chosen, it becomes clear that $(\cap_{m=1}^{\infty} D_m)$ is dense in X.
□

Actually not the intention, but is interesting to define Baire Spaces and their equivalent definitions.

Definition 7.18

Let (X, d) be a Metric Space.
X is called a Baire Space if and only if the intersection of any countable number of dense open subsets of X is dense in X.

Let $M \subseteq X$ be some set. The closure of M is denoted by $\overline{M}$ and the interior of M is denoted by M°.

Theorem 7.25

A subset M is nowhere dense in the Metric Space (X, d) if and only if $(\overline{M})^\circ = \varnothing$.

Proof of Theorem 7.25

$(\overline{M})^\circ = \varnothing \iff$ every open subset of X contains a point of $X \setminus \overline{M}$
$\iff X \setminus \overline{M}$ is dense in X, see definition ii.1.
□

Theorem 7.26

The given definition of a Baire Space in **definition 7.18** is equivalent with one of the following conditions:

1. The interior of the union of any countable number of nowhere dense closed subsets of X is empty.
2. When the union of any countable set of closed sets of X has an interior point, then one of those closed sets must have an interior point.
3. The union of any countable set of closed sets of X, whose interiors are empty, also has an interior which is empty.

Proof of Theorem 7.26

((1)$\Longleftrightarrow$(3)):

A subset M is nowhere dense in X if and only if the interior of its closure is empty, see **theorem 7.25**. So (1) and (3) are saying the same thing in different words.

((3)$\Longleftrightarrow$(2)):

Let $\{U_n\}_{n\in\mathbb{N}}$ be a countable set of closed sets in X and let $\mathcal{U} = \bigcup_{m=1}^{\infty} U_m$.

((3)$\Rightarrow$(2))

Let (3) hold.
Suppose that ${U_n}^{\circ} = \varnothing$ for $n = 1, 2, \cdots$, by (3) follows that $\mathcal{U}^{\circ} = \varnothing$. This contradicts the assumption in (2), the fact that $\mathcal{U}^{\circ} \neq \varnothing$.

((3)$\Leftarrow$(2))

Let (2) hold.
Suppose that $\mathcal{U}^{\circ} \neq \varnothing$, by (2) follows that there is some $n_0 \in \mathbb{N}$ such that $U_{n_0}^{\circ} \neq \varnothing$. This contradicts the assumption in (3), the fact that $U_n^{\circ} = \varnothing$ for $n = 1, 2, \cdots$.

((**definition 7.18**)$\Longleftrightarrow$(3)):

((**definition 7.18**)$\Rightarrow$(3))

Let (**definition 7.18**) hold.

Let $\{U_n\}_{n\in\mathbb{N}}$ be some arbitrary countable set of nowhere dense closed sets in X, so $U_m^\circ = \varnothing$ for every $m \in \mathbb{N}$. There holds that:

$$U_m^\circ = \varnothing \iff X \setminus U_m^\circ = X \iff \overline{X \setminus U_m} = X \setminus U_m^\circ = X.$$

This means that $X \setminus U_n$ is *dense* and by definition is $X \setminus U_n$ *open*, because U_n is *closed*.
Define $V_n = X \setminus U_n$ with $n = 1, 2 \cdots$. The sets V_n, with $n = 1, 2 \cdots$, are countable, dense and open. Consider $\cap_{n=1}^{\infty} V_n$, since (definition 7.18) holds

$$\overline{\bigcap_{n=1}^{\infty} V_n} = X \iff \overline{X \setminus (\bigcup_{n=1}^{\infty} U_n)} = X \iff$$
$$X \setminus ((\bigcup_{n=1}^{\infty} U_n)^\circ) = X \iff (\bigcup_{n=1}^{\infty} U_n)^\circ = \varnothing$$

So the interior of $\bigcup_{n=1}^{\infty} U_n$ is empty in X, so (3) holds.
((definition 7.18)$\Leftarrow$(3))
Let (3) hold.
Let $\{V_n\}_{n\in\mathbb{N}}$ be some arbitrary countable set of dense open sets in X, so $\overline{V_m} = X$ for every $m \in \mathbb{N}$. There holds that:

$$\overline{V_m} = X \iff X \setminus \overline{V_m} = \varnothing \iff (X \setminus V_m)^\circ = X \setminus \overline{V_m} = \varnothing.$$

This means that $X \setminus V_m$ is *nowhere dense* and by definition is $X \setminus V_m$ *closed*, because V_m is *open*.
Define $U_n = X \setminus \overline{V_n}$ with $n = 1, 2 \cdots$. The sets U_n, with $n = 1, 2 \cdots$, are countable, nowhere dense and closed. Consider $\bigcup_{n=1}^{\infty} U_n$, since (3) holds

$$(\bigcup_{n=1}^{\infty} U_n)^\circ = \varnothing \iff (X \setminus \bigcap_{n=1}^{\infty} V_n)^\circ = \varnothing \iff$$
$$X \setminus \overline{(\bigcap_{n=1}^{\infty} V_n)} = \varnothing \iff \overline{\bigcap_{n=1}^{\infty} V_n} = X$$

So $\cap_{n=1}^{\infty} V_n$ is dense in X, so definition 7.18 holds.

□

7.7.2 Closed Graph Theorem

The Closed Graph Theorem is an alternative way to check if a linear operator is bounded. If a linear operator is bounded, will be characterised by its graph. First some Lemma, which will be useful to construct an approximation with elements out of some subset of the Vector Space X.

Lemma 7.2

Let X be some Vector Space over $\mathbb{K}$.
There are subsets U, V of X and $k \in \mathbb{K}$ such that $U \subset V + k\,U$.
Then for every $x \in U$, there exists a sequence $\{v_i\}_{i \in \mathbb{N}}$ in V such that

$$x - (v_1 + k\,v_2 + \cdots + k^{n-1}\,v_n) \in k^n U,\ n = 1, 2, \cdots.$$

Proof of Lemma 7.2

Let $x \in U$. There exists some $v_1 \in U$ such that $(x - v_1) \in kU$, since there is assumed that $U \subset V + k\,U$.
Assume that there are found $v_1, \cdots, v_n \in V$ with the property that there exists some $u \in U$ such that

$$x - (v_1 + k\,v_2 + \cdots + k^{n-1}\,v_n) = k^n\,u.$$

Since $U \subset V + k\,U$, there exists some $v_{n+1} \in V$ and some $u_0 \in U$ such that $(u - v_{n+1}) = k\,u_0$, so $u = v_{n+1} + k\,u_0$ and there follows that

$$x - (v_1 + k\,v_2 + \cdots + k^{n-1}\,v_n + k^n\,v_{n+1}) = k^{n+1}\,u_0 \in k^{n+1}\,U.$$

In an inductive way there is obtained a sequence $\{v_i\}_{i \in \mathbb{N}}$ with the desired property.

□

The result of **Lemma 7.2** will be used to prove the following theorem.

Theorem 7.27

Closed Graph Theorem

Let $(X, \| \cdot \|_0)$ and $(Y, \| \cdot \|_1)$ be Banach Spaces and $T : X \rightarrow Y$ be a closed linear operator. Then T is continuous.

Proof of Theorem 7.27

Because of **Theorem 7.2**, it is enough to prove that T is bounded on X or that T is bounded on some neighbourhood of $\underline{0} \in X$, see **Theorem 7.3**.
For each $n \in \mathbb{N}$, let

$$V_n = \{x \in X \mid \| T(x) \|_1 \leq n\}.$$

The question is, if some V_n contains a neighbourhood of $\underline{0}$ in X? There holds that

$$X = \bigcup_{n=1}^{\infty} V_n = \bigcup_{n=1}^{\infty} \overline{V_n},$$

where $\overline{V_n}$ is the closure of V_n in X. This means that

$$\bigcap_{n=1}^{\infty} (X \setminus \overline{V_n}) = \varnothing.$$

Since $(X, \| \cdot \|_0)$ is a Banach Space, **Theorem 7.26** gives that there exists some $p \in \mathbb{N}$, some $x_0 \in X$ and some $\delta > 0$, such that $B_\delta(x_0, \| \cdot \|_0) \subset \overline{V_p}$.
What can be told about $B_\delta(\underline{0}, \| \cdot \|_0)$?
If $x \in X$ and $\| x \|_0 < \delta$ then $(x_0 + x) \in B_\delta(x_0, \| \cdot \|_0) \subset \overline{V_p}$. If $\{v_n\}_{n \in \mathbb{N}}$ and $\{w_n\}_{n \in \mathbb{N}}$ are sequences in V_p such that $v_n \rightarrow (x + x_0)$ and $w_n \rightarrow x_0$, then $(v_n - w_n) \rightarrow x$. Since

$$\| T(v_n - w_n) \|_1 \leq \| T(v_n) \|_1 + \| T(w_n) \|_1 \leq 2\, p,$$

there holds that $(v_n - w_n) \in \overline{V_{2p}}$, thus $x \in \overline{V_{2p}}$. So $B_\delta(\underline{0}, \| \cdot \|_0) \subset \overline{V_{2p}}$.
That means that given some $\eta > 0$ and some $x \in B_\delta(\underline{0}, \| \cdot \|_0)$, that there is

some $x_1 \in \overline{V_{2p}}$, such that $\| x - x_1 \|_0 < \eta$. Take $\eta = \alpha\,\delta$, with $0 < \alpha < 1$, for instance $\alpha = \frac{1}{3}$. Hence

$$B_\delta(\underline{0}, \| \cdot \|_0) \subset \overline{V_{2p}} + \alpha\, B_\delta(\underline{0}, \| \cdot \|_0).$$

Use **Lemma 7.2**, take some $x \in B_\delta(\underline{0}, \| \cdot \|_0)$, let $U = B_\delta(\underline{0}, \| \cdot \|_0)$ and $V = V_{2p}$ and $k = \alpha$. Then there exists a sequence $\{v_i\}_{i \in \mathbb{N}}$ in V_{2p} such that

$$x - (v_1 + \alpha\, v_2 + \cdots + \alpha^{n-1} v_n) \in \alpha^n\, B_\delta(\underline{0}, \| \cdot \|_0)$$

for $n = 1, 2 \cdots$. Let

$$w_n = v_1 + \alpha\, v_2 + \cdots + \alpha^{n-1} v_n$$

with $n = 1, 2 \cdots$. Since $\| x - w_n \|_0 < \alpha^n\, \delta$, it follows that $w_n \rightarrow x$ in X. Let $n > m$ then

$$\| T(w_n) - T(w_m) \|_1 = \| T(\sum_{i=(m+1)}^{n} \alpha^{(i-1)} v_i) \|_1 \leq \sum_{m+1}^{n} \alpha^{(i-1)} \| T(v_i) \|_1 \leq \frac{\alpha^m}{1 - \alpha} 2p.$$

Hence $\{T(v_n)\}_{n \in \mathbb{N}}$ is a Cauchy sequence in the Banach Space $(Y, \| \cdot \|_1)$, so $\{T(v_n)\}_{n \in \mathbb{N}}$ converges in $(Y, \| \cdot \|_1)$. T is a closed map, so $T(w_n) \rightarrow T(x)$ in $(Y, \| \cdot \|_1)$. Let $m = 0$ and $w_0 = \underline{0}$ then $\| T(w_n) \|_1 \leq \frac{1}{1 - \alpha} 2p$. Hence

$$\| T(x) \|_1 = \lim_{n \rightarrow \infty} \| T(w_n) \|_1 \leq \frac{1}{1 - \alpha} 2p,$$

x is an arbitrary element out of $B_\delta(\underline{0}, \| \cdot \|_0)$, $\alpha = \frac{1}{3}$ so $B_\delta(\underline{0}, \| \cdot \|_0) \subset V_{3p}$. Thus the linear map T is bounded on the neighbourhood $B_\delta(\underline{0}, \| \cdot \|_0)$ of $\underline{0}$.
□

It is of interest to mention, that in the proof of **Theorem 7.27**, are used the facts, that the spaces $(X, \| \cdot \|_0)$ and $(Y, \| \cdot \|_1)$ are Banach Spaces and that the operator T is a closed operator. The Banach Space $(X, \| \cdot \|_0)$ is of importance to use **Theorem 7.24**, the theorem of Baire. The Banach Space $(Y, \| \cdot \|_1)$ is of importance to get convergence of a constructed Cauchy sequence. The closedness of the operator is of importance to get information about the limit of the constructed Cauchy sequence in $(Y, \| \cdot \|_1)$.

7.7.3 Open Mapping Theorem

The proof of the following **Lemma 7.3** and the proof of the **Closed Graph Theorem 7.27** have much in common.

Lemma 7.3

Let $(X, \| \cdot \|_0)$ and $(Y, \| \cdot \|_1)$ be Banach Spaces. Let $T : X \rightarrow Y$ be a bounded linear operator from X onto Y. The image of the open unit ball $B^0 = B_1(\underline{0}, \| \cdot \|_0) \subset X$ contains an open ball about $\underline{0} \in Y$.

Proof of Lemma 7.3

Define $B^n = B_{2^{-n}}(\underline{0}, \| \cdot \|_0) \subset X$, with $n = 1, 2, \cdots$.
Let's try to do the proof stepwise:

a. $\overline{T(B^1)}$ contains an open ball $B_\epsilon(\underline{0}, \| \cdot \|_1)$;
b. $\overline{T(B^n)}$ contains an open ball W_n about $\underline{0} \in Y$;
c. $T(B^0)$ contains an open ball about $\underline{0} \in Y$.

Let's start with **step ii.a**:
Look at the open ball $B^1 \subset X$. Take some fixed $x \in X$ and some integer $k > 2 \| x \|_0$, then $x \in k\, B^1$, so

$$X = \sum_{k=1}^{\infty} k\, B^1.$$

The linear operator T is surjective, so

$$Y = T(X) = T(\sum_{k=1}^{\infty} k\, B^1) = \sum_{k=1}^{\infty} k\, T(B^1).$$

Since Y is complete, it is also closed, so

$$Y = \sum_{k=1}^{\infty} \overline{k\, T(B^1)}.$$

And now the use of **theorem 7.26**.
Note that in the Closed Graph Theorem, the mentioned theorem was used in $\mathcal{D}(T) = X$, the domain of the operator T, here that same theorem is used at the $\mathcal{R}(T) = Y$, the range of the operator T. Since Y is complete, there is

some $p \in \mathbb{N}$, some $y_0 \in Y$ and some $\delta > 0$, such that $B_\delta(y_0, \| \cdot \|_1) \subset \overline{p\,T(B^1)}$. This implies that $\overline{T(B^1)}$ contains an open ball, $B_\epsilon(\frac{y_0}{p}, \| \cdot \|_1) \subset \overline{T(B^1)}$ with $0 < \epsilon < \frac{\delta}{p}$. And there follows that

$$B_\epsilon(\frac{y_0}{p}, \| \cdot \|_1) - \frac{y_0}{p} = B_\epsilon(\underline{0}, \| \cdot \|_1) \subset (\overline{T(B^1)} - \frac{y_0}{p}). \tag{7.9}$$

Let's try to do **step ii.b**:
Let $y \in (\overline{T(B^1)} - \frac{y_0}{p})$ then $(\frac{y_0}{p} + y) \in \overline{T(B^1)}$. There is already known that $\frac{y_0}{p} \in \overline{T(B^1)}$. Because $\overline{T(B^1)}$ is closed, there are sequences $\{u_n\}_{n \in \mathbb{N}} \subset T(B^1)$ and $\{v_n\}_{n \in \mathbb{N}} \subset T(B^1)$ such that $u_n \to (\frac{y_0}{p} + y)$ and $v_n \to (\frac{y_0}{p})$ in Y. Since T is surjective, there are sequence $\{w_n\}_{n \in \mathbb{N}} \subset B^1$ and $\{z_n\}_{n \in \mathbb{N}} \subset B^1$ such that $u_n = T(w_n)$ and $v_n = T(z_n)$ for all $n \in \mathbb{N}$.
Since $w_n, z_n \in B^1$ there follows that

$$\| w_n - z_n \|_0 \leq \| w_n - z_n \|_0 + \| w_n - z_n \|_0 < \frac{1}{2} + \frac{1}{2} = 1,$$

such that $(w_n - z_n) \in B^0$. There is easily seen that

$$T(w_n - z_n) = T(w_n) - T(z_n) \to y \in \overline{T(B^0)},$$

such that

$$(\overline{T(B^1)} - \frac{y_0}{p}) \subset \overline{T(B^0)}. \tag{7.10}$$

The formulas 7.9 and 7.10 gives a result that

$$B_\epsilon(\frac{y_0}{p}, \| \cdot \|_1) - \frac{y_0}{p} = B_\epsilon(\underline{0}, \| \cdot \|_1) \subset \overline{T(B^0)}$$

Since the operator T is linear, so $\overline{T(B^n)} = 2^{-n}\overline{T(B^0)}$,
and that gives as result that

$$W_n = B_{(2^{-n}\,\epsilon)}(\underline{0}, \| \cdot \|_1) \subset \overline{T(B^n)}. \tag{7.11}$$

The final **step ii.c**:
The completeness of the space Y is already used, but the completeness of X not.
Let's try to prove that

$$W_1 \subset T(B^0).$$

Let $y \in W_1$, from 7.11, with $n = 1$, follows that $W_1 \subset \overline{T(B^1)}$. Since $\overline{T(B^1)}$ is closed, there exists a $w \in T(B^1)$ such that $\| y - w \|_1 < \frac{\epsilon}{4}$. The operator T is surjective, so there is some $x_1 \in B^1$ with $w = T(x_1)$, hence

$$\| y - T(x_1) \|_1 < \frac{\epsilon}{4}. \tag{7.12}$$

From 7.12, with $n = 2$, follows that $y - T(x_1) \in W_2 \subset \overline{T(B^2)}$. As before there exists some $x_2 \in B^2$, such that

$$\| (y - T(x_1)) - T(x_2) \|_1 < \frac{\epsilon}{2^3},$$

hence $y - (T(x_1) + T(x_2)) \in W_3 \subset \overline{T(B^3)}$.
In the nth step follows the existence of some $x_n \in B^n$, such that

$$\| (y - \sum_{i=1}^{n} T(x_i) \|_1 < \frac{\epsilon}{2^{(n+1)}}, \tag{7.13}$$

$n = 1, 2, \cdots$.
Look at the sequence $\{z_n\}_{n \in \mathbb{N}}$, with $z_n = x_1 + \cdots + x_n \in X$, with $x_i \in B^i$, what means that $\| x_i \|_0 < 2^{-i}$. This sequence is a Cauchy sequence, because for $n > m$,

$$\| z_n - z_m \|_0 \leq \sum_{i=(m+1)}^{n} \| x_i \|_0 < \sum_{i=(m+1)}^{n} \frac{1}{2^i} < \frac{2}{2^{(m+1)}} \rightarrow 0,$$

as $m \rightarrow \infty$. X is complete, this means that the constructed sequence $\{z_n\}_{n \in \mathbb{N}}$ converges to an element $x \in X$, so $z_n \rightarrow x$ for $n \rightarrow \infty$. It is easily seen that

$$\sum_{i=1}^{\infty} \| x_i \|_0 < \sum_{i=1}^{\infty} \frac{1}{2^i} = 1$$

and that means that $x \in B^0$. The linear operator T is continuous, because it is bounded, and that gives that $T(z_n) \rightarrow T(x)$ in Y. Out of 7.13 follows that $T(x) = y$ and that means that $y \in T(B^0)$.
Because $y \in W_1$ was arbitrary, the desired result is obtained:

$$W_1 \subset T(B^0).$$

□

Theorem 7.28

Open Mapping Theorem

Let $(X, \| \cdot \|_0)$ and $(Y, \| \cdot \|_1)$ be Banach Spaces and $T : X \rightarrow Y$ be a bounded linear operator onto Y. Then T is an open mapping.

Proof of Theorem 7.28

There has to be shown that for every open set $A \subset X$, the image $T(A)$ is open in Y. There has to be shown that for every $y = T(x) \in T(A)$, the set $T(A)$ contains an open ball about $y = T(x)$.
In **Lemma 7.3** is only proved, that the image of the open unit ball in X contains an open ball around $\underline{0} \in Y$. May be there can something be done by shifting elements to the origin and by the use of scaling?
Let A be some open subset of X and take some arbitrary $y = T(x) \in T(A)$. The existence of $x \in A$ is no problem because the operator T is surjective. Since A is open, the set A contains an open ball around x. That means that set $A - x$ contains an open set around $\underline{0} \in X$ and hence an open ball with center $\underline{0} \in X$.
Let r be the radius of that open ball, then $\frac{1}{r}(A - x)$ contains the open unit ball $B_1(\underline{0}, \| \cdot \|_0) \subset X$.
Known is that $T(\frac{1}{r}(A - x)) = \frac{1}{r} T(A - x)$, so with the use of **Lemma 7.3**, that the set $r^{-1} T(A - x)$ contains an open ball around $\underline{0} \in Y$, and so also the set $T(A - x) = T(A) - T(x)$. But this means that the set $T(A)$ contains an open ball around $T(x) = y$. y was arbitrary, so the set $T(A)$ is open.
□

7.7.4 Bounded Inverse Theorem

Theorem 7.29

Bounded Inverse Theorem

Let $(X, \| \cdot \|_0)$ and $(Y, \| \cdot \|_1)$ be Banach Spaces. If $T : X \rightarrow Y$ is a bijective bounded linear operator, then $T^{-1} : Y \rightarrow X$ is a bounded linear operator.

Proof of Theorem 7.29

The operator T^{-1} is linear, see **Theorem 7.9**, **part ii.b**. Since T is bounded, it is also a closed operator, see **Theorem 7.20**. And so the operator T^{-1} is also a closed operator, see **Theorem 7.19**. Since Y and X are Banach Spaces, the Closed Graph Theorem (**Theorem 7.27**) implies that T^{-1} is bounded.

□

7.8 Completely Continuous and Compact Linear Maps

In this section there will be tried to generalize several properties of linear transformations between finite dimensional spaces to linear transformations between infinite dimensional spaces.

Definition 7.19

Let X and Y be two Linear Spaces. A linear map $T : X \rightarrow Y$ is called of finite rank if the range of T is finite dimensional, so $\dim(R(T)) < \infty$.

Definition 7.20

Let X and Y be two Normed Spaces. A map $T : X \to Y$ is called completely continuous if the image of all weakly convergent sequences in X are convergent in norm in Y.

Definition 7.21

Let X and Y be two Normed Spaces. A map $T : X \to Y$ is called compact if the image of every bounded set in X is precompact in Y.

8 Ideas

In this chapter it is the intention to make clear why certain concepts are used.

8.1 Total and separable

First of all linear combinations, it is important to note that linear combinations are always finite. That means that if there is looked at the span of $\{1, t^1, t^2, \cdots, t^n, \cdots\}$ that a linear combination is of the form $p_n(t) = \sum_{i=0}^{n} a_i\, t^i$ with n finite.
That's also the reason that for instance $\exp t$ is written as the limit of finite sums

$$\exp(t) = \lim_{n \to \infty} \sum_{i=1}^{n} \frac{t^i}{i!},$$

see figure 8.1.

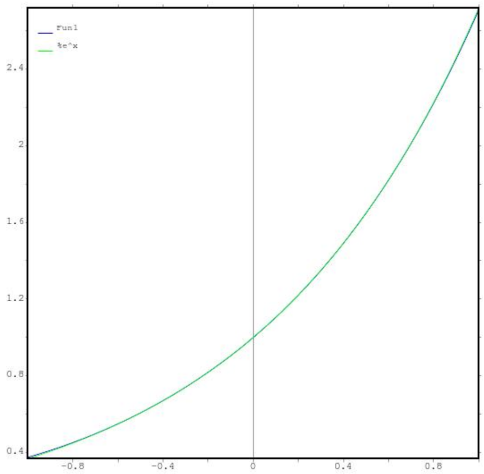


Figure 8.1 Taylor Series of $\exp(t)$ with $N = 4$.

Let's assume that $t \in [-1, +1]$ and define the inner product

$$(f, g) = \int_{-1}^{1} f(t)\, g(t)\, dt \tag{8.1}$$

with $f, g \in C[-1,+1]$, the continuous functions on the interval $[-1,+1]$.
It is of importance to note that the finite sums are polynomials. And these finite sums are elements of the space $P[-1,+1]$, equipped with the $\| \ \|_\infty$-norm, see paragraph 5.1.1. So $\exp(t)$ is not a polynomial, but can be approximated by polynomials. In certain sense, there can be said that $\exp(t) \in \overline{P[-1,+1]}$ the closure of the space of all polynomials at the interval $[-1,+1]$,

$$\lim_{n\to\infty} \| \exp(t) - \sum_{i=1}^{n} \frac{t^i}{i!} \|_\infty = 0.$$

Be careful, look for instance to the sequence $\{ \mid t^n \mid \}_{n\in\mathbb{N}}$. The pointwise limit of this sequence is

$$f : t \to \begin{cases} 1 & \text{if } t = -1 \\ 0 & \text{if } -1 < t < +1 \\ 1 & \text{if } t = +1, \end{cases}$$

so $f \notin C[-1,+1]$ and $\overline{P[-1,+1]} \neq C[-1,+1]$.
Using the sup-norm gives that

$$\lim_{n\to\infty} \| f(t) - t^n \|_\infty = 1 \neq 0.$$

Someone comes with the idea to write $\exp(t)$ not in powers of t but as a summation of cos and sin functions. But how to calculate the coefficients before the cos and sin functions and which cos and sin functions?
Just for the fun

$$\begin{aligned}
(\sin(a\,t), \sin(b\,t)) &= \frac{(b + a)\sin(b - a) - (b - a)\sin(b + a)}{(b + a)(b - a)}, \\
(\sin(a\,t), \sin(a\,t)) &= \frac{2\,a - \sin 2\,a}{2\,a}, \\
(\cos(a\,t), \cos(b\,t)) &= \frac{(b + a)\sin(b - a) + (b - a)\sin(b + a)}{(b + a)(b - a)}, \\
(\cos(a\,t), \cos(a\,t)) &= \frac{2\,a + \sin 2\,a}{2\,a},
\end{aligned}$$

with $a, b \in \mathbb{R}$. A span $\{1, \sin(a\,t), \cos(b\,t)\}_{a,b\in\mathbb{R}}$ is uncountable, a linear combination can be written in the form

$$a_0 + \sum_{\alpha\in\Lambda} (a_\alpha \sin(\alpha\,t) + b_\alpha \cos(\alpha\,t)),$$

with $\Lambda \subset \mathbb{R}$. Λ can be some interval of $\mathbb{R}$, so may be the set of α's is uncountable. It looks a system that is not nice to work with.
But with $a = n\,\pi$ and $b = m\,\pi$ with $n \neq m$ and $n, m \in \mathbb{N}$ then

$$\begin{aligned}
(\sin(a\,t), \sin(b\,t)) &= 0,\\
(\sin(a\,t), \sin(a\,t)) &= 1,\\
(\cos(a\,t), \cos(b\,t)) &= 0,\\
(\cos(a\,t), \cos(a\,t)) &= 1,
\end{aligned}$$

that looks a nice orthonormal system.
Let's examine the span of

$$\{\frac{1}{\sqrt{2}}, \sin(\pi\,t), \cos(\pi\,t), \sin(2\,\pi\,t), \cos(2\,\pi\,t), \cdots\}. \tag{8.2}$$

A linear combination out of the given span has the following form

$$\frac{a_0}{\sqrt{2}} + \sum_{n=1}^{N_0}(a_n \sin(n\,\pi\,t) + b_n \cos(n\,\pi\,t))$$

with $N_0 \in \mathbb{N}$. The linear combination can be written on such a nice way, because the elements out of the given span are countable.

Remark 8.1

Orthonormal sets versus arbitrary linear independent sets.
Assume that some given x in an Inner Product Space $(X, (\cdot, \cdot))$ has to represented by an orthonormal set $\{e_n\}$.

1. If $x \in \text{span}(\{e_1, e_2, \cdots, e_n\})$ then $x = \sum_{i=1}^{n} a_i\, e_i$. The Fourier-coefficients are relative easy to calculate by $a_i = (x, e_i)$.

2. Adding an element extra to the span for instance e_{n+1} is not a problem. The coefficients a_i remain unchanged for $1 \leq i \leq n$, since the orthogonality of e_{n+1} with respect to $\{e_1, \cdots, e_n\}$.

3. If $x \notin \text{span}(\{e_1, e_2, \cdots, e_n\})$, set $y = \sum_{i=1}^{n} a_i\, e_i$ then $(x - y) \perp y$ and $\| y \| \leq \| x \|$.

The Fourier-coefficients of the function $\exp(t)$ with respect to the given orthonormal base 8.2 are

$$a_0 = (1/\sqrt{2}, \exp(t)) = \frac{(e - (\frac{1}{e}))}{\sqrt{2}},$$

$$a_n = (\exp(t), \sin(n\,\pi\,t) = [\frac{\exp(t)(\sin(n\,\pi\,t) - \pi\,n\,\cos(n\,\pi\,t))}{((n\,\pi)^2 + 1)}]_{-1}^{1} = \frac{-\pi\,n\,(-1)^n}{((n\,\pi)^2 + 1)}\,(e - (\frac{1}{e})),$$

$$b_n = (\exp(t), \cos(n\,\pi\,t) = [\frac{\exp(t)(\cos(n\,\pi\,t) + \pi\,n\,\sin(n\,\pi\,t))}{((n\,\pi)^2 + 1)}]_{-1}^{1} = \frac{(-1)^n}{((n\,\pi)^2 + 1)}\,(e - (\frac{1}{e})),$$

for $n = 1, 2, \cdots$. See also figure 8.2, there is drawn the function

$$g_N(t) = \frac{a_0}{2} + \sum_{k=1}^{N}(a_k\,\sin(k\,\pi\,t) + b_k\,\cos(k\,\pi\,t)) \tag{8.3}$$

with $N = 40$ and the function $\exp(t)$, for $-1 \le t \le 1$.

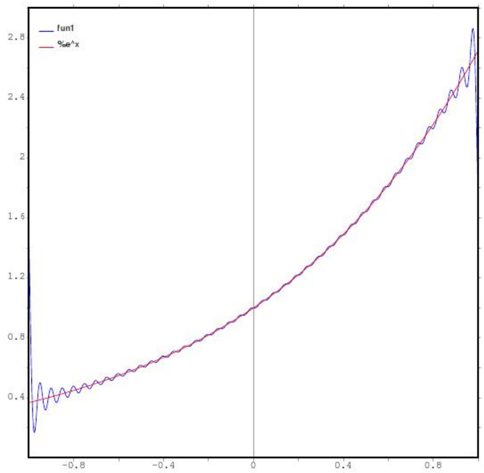


Figure 8.2 Fourier Series of $\exp(t)$ with $N = 40$.

Instead of the Fourier Series, the Legendre polynomials can also be used to approximate the function $\exp(t)$. The following Legendre polynomials are an orthonormal sequence, with respect to the same inner product as used to calculate the Fourier Series, see 8.1. The first five Legendre polynomials are given by

$$
\begin{aligned}
P_0(t) &= \frac{1}{\sqrt{2}}, \\
P_1(t) &= t\sqrt{\frac{3}{2}}, \\
P_2(t) &= \frac{(3\,t^2 - 1)}{2}\sqrt{\frac{5}{2}}, \\
P_3(t) &= \frac{(5\,t^3 \;-\; 3\,t)}{2}\sqrt{\frac{7}{2}}, \\
P_4(t) &= \frac{(35\,t^4 \;-\; 30\,t^2 \;+\; 3)}{8}\sqrt{\frac{9}{2}}.
\end{aligned}
$$

To get an idea of the approximation of $\exp(t)$, see figure 8.3.

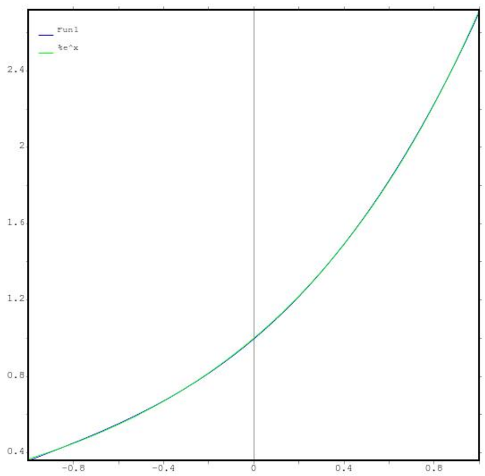


Figure 8.3 Legendre approximation of $\exp(t)$ with $N = 4$.

From these three examples the Fourier Series has a strange behaviour near $t = -1$ and $t = 1$. Using the $\| \cdot \|_\infty$-norm then the Fourier Series doesn't approximate the function $\exp(t)$ very well. But there is used an inner product and to say something about the approximation, the norm induced by that inner product is used. The inner product is defined by an integral and such an integral can hide points, which are bad approximated. Bad approximated in the sense of a pointwise limit. Define the function g, with the help of the functions g_N, see 8.3, as

$$g(t) := \lim_{N \to \infty} g_N(t)$$

for $-1 \leq t \leq +1$. Then $g(-1) = \frac{(\exp(-1) + \exp(1))}{2} = g(1)$, so $g(-1) \neq \exp(-1)$ and $g(1) \neq \exp(1)$, the functions $\exp(t)$ and $g(t)$ are pointwise not equal. For $-1 < t < +1$, the functions $g(t)$ and $\exp(t)$ are equal, but if you want to approximate function values near $t = -1$ or $t = +1$ of $\exp(t)$ with the function $g_N(t)$, N has to be taken very high to achieve a certain accuracy. The function $g(t) - \exp(t)$ can be defined by

$$g(t) - \exp(t) = \begin{cases} \frac{(-\exp(-1) + \exp(1))}{2} & \text{for } t = -1 \\ 0 & \text{for } -1 < t < +1 \\ \frac{(\exp(-1) - \exp(1))}{2} & \text{for } t = +1. \end{cases}$$

It is easily seen that $\| g(t) - \exp(t) \|_{\infty} \neq 0$ and $(g(t) - \exp(t), g(t) - \exp(t)) = \| g(t) - \exp(t) \|_2^2 = 0$. A rightful question would be, how that inner product is calculated? What to do, if there were more of such discontinuities as seen in the function $g(t) - \exp(t)$, for instance inside the interval $(-1, +1)$? Using the Lebesgue integration solves many problems, see sections **5.1.6** and **5.1.5**.
Given some subset M of a Normed Space X, the question becomes if with the span(M) every element in the space X can be descibed or can be approximated. So if for every element in X there can be found a sequence of linear combinations out of M converging to that particular element? If that is the case M is total in X, or $\overline{\text{span}(M)} = X$. In the text above are given some examples of sets, such that elements out of $L_2[-1, 1]$ can be approximated. Their span is dense in $L_2[-1, 1]$.
It is also very special that the examples, which are given, are countable. Still are written countable series, which approximate some element out of the Normed Space $L_2[-1, 1]$. If there exists a countable set, which is dense in X, then X is called separable.
Also is seen that the norm plays an important rule to describe an approximation.

8.2 Part ii.1 in the proof of Theorem 5.12, $(\mathcal{P}(\mathbb{N}) \sim \mathbb{R})$

The map

$$f : x \to \tan\left(\frac{\pi}{2}(2x - 1)\right) \tag{8.4}$$

is a one-to-one and onto map of the open interval $(0,1)$ to the real numbers $\mathbb{R}$.
If $y \in (0,1)$ then y can be written in a binary representation

$$y = \sum_{i=1}^{\infty} \frac{\eta_i}{2^i}$$

with $\eta_i = 1$ or $\eta_i = 0$.
There is a problem, because one number can have several representations. For instance, the binary representation $(0,1,0,0\cdots)$ and $(0,0,1,1,1,\cdots)$ both represent the fraction $\frac{1}{4}$. And in the decimal system, the number $0.0999999\cdots$ represents the number 0.1.
To avoid these double representation in the binary representation, there will only be looked at sequences without infinitely repeating ones.
Because of the fact that these double representations are avoided, it is possible to define a map g of the binary representation to $\mathcal{P}(\mathbb{N})$ by

$$g((z_1, z_2, z3, z_4, \cdots, z_i, \cdots)) = \{i \in \mathbb{N} \,|\, z_i = 1\}.$$

for instance $g((0,1,1,1,0,1,0,\cdots)) = \{2,3,4,6\}$ and $g((0,1,0,1,0,1,0,1,0,1,0,\cdots)) = \{2,4,6,8,10,\cdots\}$ (the even numbers).
So it is possible to define a map

$$h : \mathbb{R} \rightarrow \mathcal{P}(\mathbb{N}).$$

The map h is one-to-one but not onto, since the elimination of the infinitely repeating ones.
So there can also be defined an one-to-one map[8]

$$k : \mathcal{P}(\mathbb{N}) \rightarrow (0,1),$$

by

$$k(S) = 0.n_1 n_2 n_3 n_4 \cdots n_i \cdots \text{ where } \begin{cases} n_i = 7 & \text{if } i \in S, \\ n_i = 3 & \text{if } i \notin S. \end{cases}$$

The double representations with zeros and nines are avoided, for instance $k(\{2,3,4,7\}) = 0.377733733333333$. With the map f, see 8.4, there can be defined an one-to-one map of $\mathcal{P}(\mathbb{N})$ to $\mathbb{R}$.
With the theorem of Bernstein-Schröder, see the website **wiki-Bernstein-Schroeder**, there can be proved that there exists a bijective map between $\mathbb{R}$ and $\mathcal{P}(\mathbb{N})$,

[8] To the open interval $(0,1) \subset \mathbb{R}$.

sometimes also written as $\mathbb{R} \sim \mathcal{P}(\mathbb{N})$.
The real numbers are uncountable, but every real number can be represented by a countable sequence!

8.3 Part ii.7 in the proof of Theorem 5.12, (σ-algebra and measure)

A measure, see **Definition 8.2** is not defined on all subsets of a set Ω, but on a certain collection of subsets of Ω. That collection Σ is a subset of the power set $\mathcal{P}(\Omega)$ of Ω and is called a σ-algebra.

Definition 8.1

A σ-algebra Σ satisfies the following:

σA 1: $\Omega \in \Sigma$.

σA 2: If $M \in \Sigma$ then $M^c \in \Sigma$, with $M^c = \Omega \setminus M$, the complement of M with respect to Ω.

σA 3: If $M_i \in \Sigma$ with $i = 1, 2, 3, \cdots$, then $\bigcup_{i=1}^{\infty} M_i \in \Sigma$.

A σ-algebra is not a topology, see **Definition 3.14**. Compare for instance TS **3** with σA **ii: 3**. In TS **3** is spoken about union of an arbitrary collection of sets out of the topology and in σA **ii: 3** is spoken about a countable union of subsets out of the σ-algebra.

Remark 8.2

Some remarks on σ-algebras:

1. By σA **ii: 1**: $\Omega \in \Sigma$, so by σA **ii: 2**: $\varnothing \in \Sigma$.

2. $\cap_{i=1}^{\infty} M_i = (\bigcup_{i=1}^{\infty} M_i^c)^c$, so countable intersections are in Σ.

3. If $A, B \in \Sigma \Rightarrow A \setminus B \in \Sigma$. $(A \setminus B = A \cap B^c)$

The pair (Ω, Σ) is called a measurable space. A set $A \in \Sigma$ is called a measurable set. A measure is defined by the following definition.

Definition 8.2

A measure μ on (Ω, Σ) is a function to the extended interval $[0, \infty]$, so $\mu : \Sigma \to [0, \infty]$ and satisfies the following properties:

1. $\mu(\varnothing) = 0$.

2. μ is countable additive or σ-additive, that means that for a countable sequence $\{M_n\}_n$ of disjoint elements out of Σ

$$\mu(\bigcup_n M_n) = \sum_n \mu(M_n).$$

The triplet (Ω, Σ, μ) is called a measure space.

An outer measure need not to satisfy the condition of σ-additivity, but is σ-subadditive on $\mathcal{P}(X)$.

Definition 8.3

An outer measure μ^* on (Ω, Σ) is a function to the extended interval $[0, \infty]$, so $\mu^* : \Sigma \rightarrow [0, \infty]$ and satisfies the following properties:

1. $\mu^*(\varnothing) = 0.$
2. $\mu^*(A) \leq \mu^*(B)$ if $A \subseteq B$; μ^* is called monotone.
3. $\mu^*(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$ for every sequence $\{A_i\}$ of subsets in Ω; μ^* is σ-subadditive,

see (Aliprantis and Burkinshaw, 1998, see here).
If $\mathcal{F}$ is a collection of subsets of a set Ω containing the empty set and let $\mu : \mathcal{F} \rightarrow [0, \infty]$ be a set function such that $\mu(\varnothing) = 0$. For every subset A of Ω the outer measure generated by the set function μ is defined by

Definition 8.4

$$\mu^*(A) = \inf\{\sum_{i=1}^{\infty} \mu(A_i) \mid \{A_i\} \text{ a sequence of } \mathcal{F} \text{ with } A \subseteq \bigcup_{i=1}^{\infty} A_i\}.$$

With the outer-measure, relations can be defined which hold almost everywhere. Almost everywhere is abbreviated by a.e. and for the measurable space (Ω, Σ, μ) are here some examples of a.e. relations which can be defined:

1. $f = g$ a.e. if $\mu^*\{x \in \Omega \mid f(x) \neq g(x)\} = 0.$
2. $f \geq g$ a.e. if $\mu^*\{x \in \Omega \mid f(x) < g(x)\} = 0.$
3. $f_n \rightarrow g$ a.e. if $\mu^*\{x \in \Omega \mid f_n(x) \nrightarrow g(x)\} = 0.$
4. $f_n \uparrow g$ a.e. if $f_n \leq f_{n+1}$ a.e. for all n and $f_n \rightarrow g$ a.e.
5. $f_n \downarrow g$ a.e. if $f_n \geq f_{n+1}$ a.e. for all n and $f_n \rightarrow g$ a.e.

A σ-algebra $\mathcal{B}$ on the real numbers $\mathbb{R}$ can be generated by all kind of intervals, for instance $[a, \infty)$, $(-\infty, a)$, (a, b), or $[a, b]$ with $a \in \mathbb{R}$.
Important is to use the rules as defined in **Definition 8.1** and see also **Remark 8.2**.
Starting with $[a, \infty) \in \mathcal{B}$ then also $[a, \infty)^c = (-\infty, a) \in \mathcal{B}$. With that result it is easy to see that $[a, b) = [a, \infty) \cap (-\infty, b) \in \mathcal{B}$. Assume that $a < b$, then

evenso $[a,b] \in \mathcal{B}$ because $[a,b] = \cap_{n=1}^{\infty}[a, b+\frac{1}{n}) = ((-\infty,a) \cup (b,\infty))^c \in \mathcal{B}$, $(a,b) = ((-\infty,a] \cup [b,\infty))^c \in \mathcal{B}$ and also $\{a\} = \cap_{n=1}^{\infty}([a,\infty) \cap (-\infty, a+\frac{1}{n})) = ((-\infty,a) \cup (a,\infty))^c \in \mathcal{B}$

The same σ-algebra can also be generated by the open sets (a,b). Members of a σ-algebra generated by the open sets of a topological space are called Borel sets. The σ-algebra generated by open sets is also called a Borel σ-algebra.

The Borel σ-algebra on $\mathbb{R}$ equals the smallest family $\mathcal{B}$ that contains all open subsets of $\mathbb{R}$ and that is closed under countable intersections and countable disjoint unions. More information about Borel sets and Borel σ-algebras can be found in (Srivastava, 1998, see here).

Further the definition of a σ-measurable function .

Definition 8.5

Let the pair (Ω, Σ) be a measurable space, the function $f : \Omega \to \mathbb{R}$ is called σ-measurable, if for each Borel subset B of $\mathbb{R}$:

$$f^{-1}(B) \in \Sigma.$$

Using **Definition 8.5**, the function $f : \Omega \to \mathbb{R}$ is σ-measurable, if $f^{-1}([a,\infty)) \in \Sigma$ for each $a \in \mathbb{R}$ or if $f^{-1}((-\infty,a])) \in \Sigma$ for each $a \in \mathbb{R}$.

Theorem 8.1

If $f, g : \Omega \to \mathbb{R}$ are σ-measurable, then the set

$$\{x \in \Omega \mid f(x) \geq g(x)\}$$

is σ-measurable.

Proof of Theorem 8.1

Let r_1, r_2, $\cdots$ be an enumeration of the rational numbers of $\mathbb{R}$, then

$$
\begin{aligned}
&\{x \in \Omega \mid f(x) \geq g(x)\} \\
&= \bigcup_{i=1}^{\infty} \big(\{x \in \Omega \mid f(x) \geq r_i\} \cap \{x \in \Omega \mid g(x) \leq r_i\} \big) \\
&= \bigcup_{i=1}^{\infty} \big(f^{-1}([r_i, \infty)) \cap g^{-1}((-\infty, r_i]) \big),
\end{aligned}
$$

which is σ-measurable, because it is a countable union of σ-measurable sets.
$\square$

8.4 Discrete measure

Let Ω be a non empty set and $\mathcal{P}(\Omega)$ the family of all the subsets of Ω, the power set of Ω. Choose a finite or at most countable subset I of Ω and a sequence of strictly positive real numbers $\{\alpha_i \,|\, i \in I\}$. Consider $\mu : \mathcal{P}(\Omega) \to \{[0, \infty) \cup \infty\}$ defined by $\mu(A) = \sum_{i \in I} \alpha_i \chi_A(i)$, where

$$
\chi_A(i) = \chi_{\{i \in A\}} = \begin{cases} 1 & \text{if } i \in A, \\ 0 & \text{zero otherwise.} \end{cases} \tag{8.5}
$$

χ is called the indicator function of the set A.
By definition $\mu(\varnothing) = 0$ and μ is σ-additive , what means that if $A = \bigcup_{i=1}^{\infty} A_i$ with
$A_i \cap A_j = \varnothing$ for any $i \neq j$, then $\mu(A) = \sum_{i=1}^{\infty} \mu(A_i)$.
To define μ the values are needed of $\mu(\{i\})$ for any i in the finite or countable set I.

8.5 Development of proof of Morrison

First of all, Morrison takes some set Ω and not especially $\mathcal{P}(\mathbb{N})$, the power set of the natural numbers. A lot of information about the measure theory has been found at the webpages of **Coleman** and **Sattinger** and in the books of (Pugachev and Sinitsyn, 1999), (Rana, 2004), (Swartz, 1994) and (Yeh, 2006, see here).

Step 1:

The first step is to proof that the linear space of bounded functions $f : \Omega \to \mathbb{R}$, which are σ-measurable, denoted by $\mathcal{B}(\Omega, \Sigma)$, is a Banach Space. The norm for each $f \in \mathcal{B}(\Omega, \Sigma)$ is defined by $\| f \|_\infty = \sup\{|f(\omega)| \mid \omega \in \Omega\}$. The space $B(\Omega)$ equiped with the $\| \cdot \|_\infty$ is a Banach Space, see **Theorem 5.1.9**. In fact it is enough to prove that $\mathcal{B}(\Omega, \Sigma)$ is a closed linear subspace of $B(\Omega)$, see Theorem **3.12**.

If f, g are bounded on Ω then the functions $f + g$ and $\alpha\, f$, with $\alpha \in \mathbb{R}$, are also bounded, because $\mathcal{B}(\Omega)$ is a linear space, and $\mathcal{B}(\Omega, \Sigma) \subseteq \mathcal{B}(\Omega)$. The question becomes, if the functions $(f + g)$ and $(\alpha\, f)$ are σ-measurable?

Theorem 8.2

If f, g are σ-measurable functions and $\alpha \in \mathbb{R}$ then is

1. $f + g$ is σ-measurable and
2. $\alpha\, f$ is σ-measurable.

Proof of Theorem 8.2

Let $c \in \mathbb{R}$ be a constant, then the function $(g - c)$ is σ-measurable, because $(g - c)^{-1}([a, \infty)) = \{x \in \Omega \mid g(x) - c \geq a\} = \{x \in \Omega \mid g(x) \geq a + c\} \in \Sigma$.
If $a \in \mathbb{R}$ then

$$(f + g)^{-1}([a, \infty)) = \{x \in \Omega \mid f(x) + g(x) \geq a\} = \{x \in \Omega \mid f(x) \geq a - g(x)\}$$

is σ-measurable, with the remark just made and **Theorem 8.1**.
If $a, \alpha \in \mathbb{R}$ and $\alpha > 0$ then

$$(\alpha f)^{-1}([a, \infty)) = \{x \in \Omega \mid \alpha f(x) \geq a\} = \{x \in \Omega \mid f(x) \geq \frac{a}{\alpha}\}$$

is σ-measurable, evenso for the case that $\alpha < 0$.
If $\alpha = 0$ then $0^{-1}([a, \infty)) = \varnothing$ or $0^{-1}([a, \infty)) = \Omega$, this depends on the sign

of a, in all cases elements of Σ, so (αf) is σ-measurable. □

Use **Theorem 8.2** and there is proved that $\mathcal{B}(\Omega, \Sigma)$ is a linear subspace of $\mathcal{B}(\Omega)$. But now the question, if $\mathcal{B}(\Omega, \Sigma)$ is a closed subspace of $\mathcal{B}(\Omega)$?

Theorem 8.3

Let $\{f_n\}_{n \in \mathbb{N}}$ be a sequence of measurable functions, and $\lim_{n \to \infty} f_n = f$ a.e. then is f a measurable function.

Proof of Theorem 8.3

Since $\lim_{n \to \infty} f_n = f$ a.e., the set $A = \{x \in \Omega \mid \lim_{n \nrightarrow \infty} f_n(x) \neq f(x)\}$ has outer measure zero, so $\mu^*(A) = 0$. The set A is measurable and hence A^c is measurable set.
Important is that

$$f^{-1}((a, \infty)) = \left(A \cap f^{-1}((a, \infty))\right) \cup \left(A^c \cap f^{-1}((a, \infty))\right),$$

if both sets are measurable, then is $f^{-1}((a, \infty))$ measurable.
The set $A \cap f^{-1}((a, \infty))$ is measurable, because it is a subset of a set of measure zero. Further is

$$A^c \cap f^{-1}((a, \infty)) = A^c \cap \left(\bigcup_{n=1}^{\infty} \left(\bigcap_{i=n}^{\infty} f_i^{-1}((a + \frac{1}{n}, \infty))\right)\right)$$

since the functions f_i are measurable, the set $A^c \cap f^{-1}((a, \infty))$ is measurable.
□

The question remains if the limit of a sequence of Σ-measurable functions is also Σ-measurable? What is the relation between the outer measure and a σ-algebra? See (Melrose, 2004, page 10) or (Swartz, 1994, page 37), there is proved that the collection of μ^*-measurabe sets for any outer measure is a σ-algebra.

Hence $(\mathcal{B}(\Omega, \Sigma), \| \cdot \|_\infty)$ is a closed subspace of the Banach Space $(\mathcal{B}(\Omega), \| \cdot \|_\infty)$, so $(\mathcal{B}(\Omega, \Sigma), \| \cdot \|_\infty)$ is a Banach Space.

Step 2:

The next step is to investigate $ba(\Sigma)$, the linear space of finitely additive, bounded set functions $\mu : \Sigma \to \mathbb{R}$, see also (Dunford and Schwartz, 8 71, IV.2.15).

Linearity is meant with the usual operations. Besides finitely additive set functions, there are also countably additive set functions or σ-additive set functions.

Definition 8.6

Let $\{A_i\}_{i \in \mathbb{N}}$ be a countable set of pairwise disjoint sets in Σ, i.e. $A_i \cap A_j = \varnothing$ for $i \neq j$ with $i, j \in \mathbb{N}$.

1. A set function $\mu : \Sigma \to \mathbb{R}$ is called countably additive (or σ-additive) if

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \bigcup_{i=1}^{\infty} \mu(A_i).$$

2. A set function $\mu : \Sigma \to \mathbb{R}$ is called finitely additive if

$$\mu(\bigcup_{i=1}^{N} A_i) = \bigcup_{i=1}^{N} \mu(A_i),$$

for every finite $N \in \mathbb{N}$.

If there is spoken about bounded set functions, there is also some norm. Here is taken the so-called variational norm.

Definition 8.7

The variational norm of any $\mu \in ba(\Sigma)$ is defined by

$$\| \mu \|_n^v = \sup \left\{ \sum_{i=1}^{n} |\mu(A_i)| \,\middle|\, n \in \mathbb{N},\ A_1, \cdots, A_n \right.$$

$$\left. \text{are pairewise disjoint members of } \Sigma \right\},$$

the supremum is taken over all partitions of Ω into a finite number of disjoint measurable sets.

In the literature is also spoken about the total variation, but in that context there is some measurable space (Ω, Σ) with a measure μ. Here we have to do with a set of finitely additive, bounded set functions μ. There is

made use of the extended real numbers $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$. Sometimes is spoken about $\mathbb{R}^*$ with $\mathbb{R}^* = \mathbb{R} \cup \{\infty\}$ or $\mathbb{R}^* = \mathbb{R} \cup \{-\infty\}$, there is said to avoid problems like $(\infty + (-\infty))$. For the arithmetic operations and algebraic properties in $\overline{\mathbb{R}}$, see the website **wiki-extended-reals**.
What is the difference between a countable additive set function and a measure? A measure μ makes use of the extended real numbers $\mu : \Sigma \rightarrow [0, \infty]$, it is a countable additive set function and has the condition that $\mu(\varnothing) = 0$, see **Definition 8.2**.

Measures have positive values, a generalisation of it are signed measures, which are allowed to have negative values, (Yeh, 2006, page 202).

Definition 8.8

Given is a measurable space (Ω, Σ). A set function μ on Σ is called a signed measure on Σ if:

1. $\mu(E) \in (-\infty, \infty]$ or $\mu(E) \in [-\infty, \infty)$ for every $E \in \Sigma$,
2. $\mu(\varnothing) = 0$,
3. if finite additive: for every finite disjoint sequence $\{E_1, \cdots, E_N\}$ in Σ, $\sum_{k=1}^{N} \mu(E_k)$ exists in $\mathbb{R}^*$ and $\sum_{k=1}^{N} \mu(E_k) = \mu(\bigcup_{k=1}^{N} (E_k))$.
4. if countable additive: for every disjoint sequence $\{E_i\}_{i \in \mathbb{N}}$ in Σ, $\sum_{k \in \mathbb{N}} \mu(E_k)$ exists in $\mathbb{R}^*$ and $\sum_{k \in \mathbb{N}} \mu(E_k) = \mu(\bigcup_{k \in \mathbb{N}} (E_k))$.

If μ is a signed measure then (Ω, Σ, μ) is called a signed measure space.

Thus a measure μ on the measurable space (Ω, Σ) is a signed measure with the condition that $\mu(E) \in [0, \infty]$ for every $E \in \Sigma$.

Definition 8.9

Given a signed measure space (Ω, Σ, μ). The total variation of μ is a positive measure $|\mu|$ on Σ defined by

$$|\mu|(E) = \sup\left\{ \sum_{k=1}^{n} |\mu(E_k)| \middle| E_1, \cdots, E_n \subset \Sigma, \right.$$
$$\left. E_i \cap E_j = \varnothing (i \neq j), \bigcup_{k=1}^{n} E_k = E, n \in \mathbb{N} \right\}.$$

Important: $\| \mu \|_n^v = |\mu|(\Omega)$.
It is not difficult to prove that the expression $\| \cdot \|_n^v$, given in Defintion 8.7 is a norm. Realize that when $\| \mu \|_n^v = 0$, it means that $|\mu(A)| = 0$ for every $A \in \Sigma$, so $\mu\Big|_\Sigma = 0$. The first result is that $(ba(\Sigma), \| \cdot \|_n^v)$ is a Normed Space,
but $(ba(\Sigma), \| \cdot \|_n^v)$ is also a Banach Space.
Let $\epsilon > 0$ and $N \in \mathbb{N}$ be given. Further is given an Cauchy sequence $\{\mu_i\}_{i \in \mathbb{N}}$, so there is an $N(\epsilon) > 0$ such that for all $i, j > N(\epsilon)$, $\| \mu_i - \mu_j \|_n^v < \epsilon$. This means that for every $E \in \Sigma$:

$$\begin{aligned} |\mu_i(E) - \mu_j(E)| &\leq |\mu_i - \mu_j|(E) \\ &\leq |\mu_i - \mu_j|(X) \\ &= \| \mu_i - \mu_j \|_n^v < \epsilon. \end{aligned} \tag{8.6}$$

Hence, the sequence $\{\mu_i(E)\}_{i \in \mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. Every Cauchy sequence in $\mathbb{R}$ converges, so define

$$\lambda(E) = \lim_{n \to \infty} \mu_n(E)$$

for every $E \in \Sigma$. Remains to prove that, λ is a finitely additive, bounded set function and $\lim_{i \to \infty} \| \mu_i - \lambda \| = 0$.

Let $E = \bigcup_{k=1}^{N} E_k$, E_k are disjoint elements of Σ, then

$$|\lambda(E) - \sum_{k=1}^{N} \lambda(E_k)| \leq |\lambda(E) - \mu_i(E)| + |\mu_i(E) - \sum_{k=1}^{N} \lambda(E_k)| \quad (8.7)$$

$$\leq |\lambda(E) - \mu_i(E)| + |\sum_{k=1}^{N} \mu_i(E_k) - \sum_{k=1}^{N} \lambda(E_k)|.$$

Since $\lambda(E) = \lim_{n\to\infty} \mu_n(E)$, there is some $k_0(\epsilon)$ such that for every $i > k_0(\epsilon)$, $|\lambda(E) - \mu_i(E)| < \epsilon$. There is also some $c_k(\epsilon)$ such that for $i > c_k(\epsilon)$, $|(\mu_i(E_k) - \lambda(E_k)| < \dfrac{\epsilon}{N}$ and that for $1 \leq k \leq N$. (N is finite!)
Hence for $i > \max\{k_0(\epsilon), c_1(\epsilon), \cdots, c_N(\epsilon)\}$, $|\sum_{k=1}^{N}(\mu_i(E_k) - \lambda(E_k))| < N\dfrac{\epsilon}{N} = \epsilon$,
so λ is finitely additive, because

$$|\lambda(E) - \sum_{k=1}^{N} \lambda(E_k)| < 2\epsilon.$$

Remark 8.3

In the case of countable additivity there are more difficulties, because $E = \lim_{N\to\infty} \bigcup_{k=1}^{N} E_k$. So inequality 8.7 has to be changed to

$$|\lambda(E) - \sum_{k=1}^{M} \lambda(E_k)| \leq$$

$$|\lambda(E) - \mu_i(E)| + |\mu_i(E) - \sum_{k=1}^{M} \mu_i(E_k)| + |\sum_{k=1}^{M} \mu_i(E_k) - \sum_{k=1}^{M} \lambda(E_k)|$$

with $i \to \infty$ and $M \to \infty$.

Inequality 8.6 gives that for all $n, m > k_0(\epsilon)$ and for every $E \in \Sigma$

$$|\mu_n(E) - \mu_m(E)| < \epsilon.$$

On the same way as done to prove the uniform convergence of bounded functions, see Theorem 5.10:

$$
\begin{aligned}
& |\mu_n(E) - \lambda(E)| \\
& \leq |\mu_n(E) - \mu_m(E)| + |\mu_m(E) - \lambda(E)|
\end{aligned}
$$

There is known that

$$|\mu_n(E) - \mu_m(E)| \leq \| \mu_n - \mu_m \|_n^v < \epsilon$$

for $m, n > k_0(\epsilon)$ and for all $E \in \Sigma$, further $m > k_0(\epsilon)$ can be taken large enough for every $E \in \Sigma$ such that

$$|\mu_m(E) - \lambda(E)| < \epsilon.$$

Hence $|\mu_n(E) - \lambda(E)| < 2\,\epsilon$ for $n > k_0(\epsilon)$ and for all $E \in \Sigma$, such that $\| \mu_n - \lambda \|_n^v = |\mu_n - \lambda|(\Omega) \leq 2\,\epsilon$. The given Cauchy sequence converges in the $\| \cdot \|_n^v$-norm, so $(ba(\Sigma), \| \cdot \|_n^v)$ is a Banach Space.

Step 3:

The next step is to look to simple functions or finitely-many-valued functions. With these simple functions will be created integrals, which define bounded linear functionals on the space of simple functions. To integrate there is needed a measure, such that the linear space $ba(\Sigma)$ becomes important. Hopefully at the end of this section the connection with ℓ^∞ becomes clear, at this moment the connection is lost.

Definition 8.10

Let (Ω, Σ) be a measurable space and let $\{A_1, \cdots, A_n\}$ be a partition of disjoint subsets, out of Σ, of Ω and $\{a_1, \cdots, a_n\}$ a sequence of real numbers. A simple function $s : \Omega \rightarrow \mathbb{R}$ is of the form

$$s(\omega) \;=\; \sum_{i=1}^{n} a_i\, \chi_{A_i}(\omega) \tag{8.8}$$

with $\omega \in \Omega$ and χ_A denotes the indicator function or characteric function on A, see formula 5.16.

Theorem 8.4

The simple functions are closed under addition and scalar multiplication.

Proof of Theorem 8.4

The scalar multipication gives no problems, but the addition? Let $s = \sum_{i=1}^{M} a_i \chi_{A_i}$ and $t = \sum_{j=1}^{N} b_j \chi_{B_j}$, where $\Omega = \overset{M}{\underset{i=1}{\cup}} A_i = \overset{N}{\underset{j=1}{\cup}} B_j$. The collections $\{A_1, \cdots, A_M\}$ and $\{B_1, \cdots, B_N\}$ are subsets of Σ and in each collection, the subsets are pairwise disjoint.

Define $C_{ij} = A_i \cap B_j$. Then $A_i \subseteq X = \overset{N}{\underset{j=1}{\cup}} B_j$ and so $A_i = A_i \cap (\overset{N}{\underset{j=1}{\cup}} B_j) = \overset{N}{\underset{j=1}{\cup}} (A_i \cap B_j) = \overset{N}{\underset{j=1}{\cup}} C_{ij}$. On the same way $B_j = \overset{N}{\underset{j=1}{\cup}} C_{ij}$. The sets C_{ij} are disjoint and this means that

$$\chi_{A_i} = \sum_{j=1}^{N} \chi_{C_{ij}} \text{ and } \chi_{B_j} = \sum_{i=1}^{M} \chi_{C_{ij}}.$$

The simple functions s and t can be rewritten as

$$s = \sum_{i=1}^{M} (a_i \sum_{j=1}^{N} \chi_{C_{ij}}) = \sum_{i=1}^{M} \sum_{j=1}^{N} a_i \chi_{C_{ij}} \text{ and}$$
$$t = \sum_{j=1}^{N} (b_j \sum_{i=1}^{M} \chi_{C_{ij}}) = \sum_{i=1}^{M} \sum_{j=1}^{N} b_i \chi_{C_{ij}}.$$

Hence

$$(s + t) = \sum_{i=1}^{M} \sum_{j=1}^{N} (a_i + b_i) \chi_{C_{ij}}$$

is a simple function. □

With these simple functions s ($\in \mathcal{B}(\Omega, \Sigma)$), it is relative easy to define an integral over Ω.

Definition 8.11

Let $\mu \in ba(\Sigma)$ and let s be a simple function, see formula 8.8, then

$$\int_{\Omega} s \, d\mu = \sum_{i=1}^{n} a_i \, \mu(A_i),$$

denote that $\int_{\Omega} \cdot \, d\mu$ is a linear functional in s.

Further it is easy to see that

$$| \int_{\Omega} s \, d\mu | \leq \sum_{i=1}^{n} |a_i \mu(A_i)|$$
$$\leq \| s \|_{\infty} \sum_{i=1}^{n} |\mu(A_i)| \leq \| s \|_{\infty} \| \mu \|_{n}^{v} . \quad (8.9)$$

Thus, $\int_{\Omega} \cdot \, d\mu$ is a bounded linear functional on the linear subspace of simple functions in $\mathcal{B}(\Omega, \Sigma)$, the simple Σ-measurable functions.

Step 4:
With simple Σ-measurable functions a bounded measurable function can be approximated uniformly.

Theorem 8.5

Let $s : \Omega \rightarrow \mathbb{R}$ be a positive bounded measurable function. Then there exists a sequence of non-negative simple functions $\{\phi_n\}_{n \in \mathbb{N}}$, such that $\phi_n(\omega) \uparrow s(\omega)$ for every $\omega \in \Omega$ and the convergence is uniform on Ω.

Proof of Theorem 8.5

For $n \geq 1$ and $1 \leq k \leq n \, 2^n$, let

$$E_{nk} = s^{-1}\Big([\frac{(k-1)}{2^n}, \frac{k}{2^n}) \Big) \text{ and } F_n = s^{-1}\Big([n, \infty) \Big).$$

Then the sequence of simple functions

$$\phi_n = \sum_{k=1}^{n\,2^n} (k-1)\, 2^{-n}\, \chi_{E_{nk}} + n\, \chi_{F_n}$$

satisfy

$$\phi_n(\omega) \leq s(\omega) \text{ for all } \omega \in \Omega$$

and for all $n \in \mathbb{N}$. If $\dfrac{(k-1)}{2^n} \leq s(\omega) \leq \dfrac{k}{2^n}$ then $\phi_n(\omega) \leq s(\omega)$ for all $\omega \in E_{nk}$. Further is

$$E_{(n+1)(2k-1)} = s^{-1}\Big([\frac{2\,(k-1)}{2^{n+1}}, \frac{(2k-1)}{2^{(n+1)}})\Big) = s^{-1}\Big([\frac{(k-1)}{2^n}, \frac{k}{2^n} - \frac{1}{2}\frac{1}{2^n})\Big) \subset E_{nk}$$

and $E_{(n+1)(2k-1)} \cup E_{(n+1)(2k)} = E_{nk}$, so $\phi_{(n+1)}(\omega) \geq \phi_n(\omega)$ for all $\omega \in E_{nk}$.
Shortly written as $\phi_{(n+1)} \geq \phi_n$ for all $n \in \mathbb{N}$.
If $\omega \in \Omega$ and $n > s(\omega)$ then

$$0 \leq s(\omega) - \phi_n(\omega) < \frac{1}{2^n},$$

so $\phi_n(\omega) \uparrow s(\omega)$ and the convergence is uniform on Ω. □

Theorem 8.5 is only going about positive bounded measurable funtions. To obtain the result in Theorem 8.5 for arbitrary bounded measurable functions, there has to be made a decompostion.

Definition 8.12

If the functions s and t are given then

$$s \vee t = \max\{s, t\}, \quad s \wedge t = \min\{s, t\}$$
$$s^+ = s \vee 0, \quad s^- = (-s) \wedge 0.$$

Theorem 8.6

If s and t are measurable then are also measurable
$s \vee t$, $s \wedge t$, s^+, s^- and $|s|$.

Proof of Theorem 8.6

See Theorem 8.2, there is proved that $(s + t)$ is measurable and that if α is a scalar, that αs is measurable, hence $(s - t)$ is measurable.
Out of the fact that

$$\{s^+ > a\} = \begin{cases} \Omega & \text{if } a < 0, \\ \{x \in \Omega \,|\, s(x) > a\} & \text{if } a \geq 0, \end{cases}$$

it follows that s^+ is measurable. Using the same argument proves that s^- is measurable.
Since $|\,s\,| = s^+ + s^-$, it also follows $|\,s\,|$ is measurable.
The following two identities

$$s \vee t = \frac{(s + t) + |\,(s - t)\,|}{2}, \quad s \wedge t = \frac{(s + t) - |\,(s - t)\,|}{2}$$

show that $s \vee t$ and $s \wedge t$ are measurable. □

Theorem 8.7

Let $s : \Omega \to \mathbb{R}$ be measurable. Then there exists a sequence of simple functions $\{\phi_n\}_{n \in \mathbb{N}}$ such that $\{\phi_n\}_{n \in \mathbb{N}}$ converges pointwise on Ω with $|\phi(\omega)| \leq |s(\omega)|$ for all $\omega \in \Omega$. If s is bounded, the convergence is uniform on Ω.

Proof of Theorem 8.7

The function s can be written as $s = s^+ - s^-$. Apply Theorem 8.5 to s^+ and s^-. □

The result is that the simple Σ-measurable functions are dense in $\mathcal{B}(\Omega, \Sigma)$ with respect to the $\| \cdot \|_\infty$-norm.

Step 5:
In Step ii: 1 is proved that $(\mathcal{B}(\Omega, \Sigma), \| \cdot \|_\infty)$ is Banach Space and in Step ii: 4 is proved that the simple functions are a linear subspace of $\mathcal{B}(\Omega, \Sigma)$ and these simple functions are lying dense in $(\mathcal{B}(\Omega, \Sigma), \| \cdot \|_\infty)$. Further

is defined a bounded linear functional $\nu(\cdot) = \int_\Omega \cdot \, d\mu$, with respect to the $\| \cdot \|_\infty$-norm, on the linear subspace of simple functions in $\mathcal{B}(\Omega, \Sigma)$, see Definition 8.11.
The use of Hahn-Banach, see Theorem 4.10, gives that there exists an extension $\widetilde{\nu}(\cdot)$ of the linear functional $\nu(\cdot)$ to all elements of $\mathcal{B}(\Omega, \Sigma)$ and the norm of the linear functional $\nu(\cdot)$ is preserved, i.e. $\| \widetilde{\nu} \| = \| \nu \|$.
Hahn-Banach gives no information about the uniqueness of this extension.

Step 6:
What is the idea so far? With some element $\mu \in ba(\Sigma)$ there can be defined a bounded functional $\nu(\cdot) = \int_\Omega \cdot \, d\mu$ on $\mathcal{B}(\Omega, \Sigma)$, so $\nu \in \mathcal{B}(\Omega, \Sigma)^{'}$ and $\| \nu \| = \| \mu \|_n^v$.
The Banach Space $(\ell^\infty, \| \cdot \|_\infty)$, see Section 5.2.1, can be seen as the set of all bounded functions from $\mathbb{N}$ to $\mathcal{P}(\mathbb{N}) \sim \mathbb{R}$ (see Section 8.2), where for $x \in \ell^\infty$, $\| x \|_\infty = \sup \left\{ |x(\alpha)| \,\middle|\, \alpha \in \mathbb{N} \right\}$. So $(\ell^\infty)^{'} = \mathcal{B}(\mathbb{N}, \mathbb{R})^{'} = \mathcal{B}(\mathbb{N}, \mathcal{P}(\mathbb{N}))^{'} = ba(\mathcal{P}(\mathbb{N}))$.

The question is if $ba(\Sigma)$ and $\mathcal{B}(\Omega, \Sigma)^{'}$ are isometrically isomorph or not?

Theorem 8.8

Any bounded linear functional u on the space of bounded functions $\mathcal{B}(\Omega, \Sigma)$ is determined by the formula

$$u(s) = \int_\Omega s(\omega) \, \mu(d\omega) = \int_\Omega s \, d\mu, \tag{8.10}$$

where $\mu(\cdot)$ is a finite additive measure.

Proof of Theorem 8.8

Let u be a bounded linear functional on the space $\mathcal{B}(\Omega, \Sigma)$, so $u \in \mathcal{B}(\Omega, \Sigma)^{'}$.
Consider the values of the functional u on the characteristic functions χ_A on Ω, $A \in \Sigma$. The expression $u(\chi_A)$ defines an finite additive function $\mu(A)$.
Let $A_1, \cdots, A_n$ be a set of pairwise nonintersecting sets, $A_i \cap A_j = \varnothing$ if $i \neq j$, then

$$\mu(\bigcup_{j=1}^{n} A_j) = u(\sum_{j=1}^{n} \chi_{A_j}) = \sum_{j=1}^{n} u(\chi_{A_j}) = \sum_{j=1}^{n} \mu(A_j).$$

The additive function μ is bounded, if the values of $\mu(A_j)$ are finite, for all $j \in \{1, \cdots, n\}$. Determine now the value of the functional u on the set of simple functions

$$s(\omega) = \sum_{i=1}^{n} a_i \chi_{A_i}(\omega), \quad A_i \cap A_j = \varnothing,\, i \neq j.$$

The functional u is linear, so

$$u(s) = \sum_{i=1}^{n} a_i u(\chi_{A_i}) = \sum_{i=1}^{n} a_i \mu(\chi_{A_i}). \tag{8.11}$$

Formula 8.11, represents an integral of the simple function $s(\omega)$ with respect to the additive measure μ. Therefore

$$u(s) = \int_{\Omega} s(\omega)\mu(d\omega) = \int_{\Omega} s d\mu.$$

Thus a bounded linear functional on $\mathcal{B}(\Omega, \Sigma)$ is determined on the set of simple functions by formula 8.10.
The set of simple functions is dense in the space $\mathcal{B}(\Omega, \Sigma)$, see Theorem 8.7. This means that any function from $\mathcal{B}(\Omega, \Sigma)$ can be represented as the limit of an uniform convergent sequence of simpe functions. Out of the continuity of the functional u follows that formula 8.10 is valid for any function $s \in \mathcal{B}(\Omega, \Sigma)$. □

Theorem 8.9

The norm of the functional u determined by formula 8.10 is equal to the value of the variational norm of the additive measure μ on the whole space Ω:

$$\| u \| = \| \mu \|_n^v$$

Proof of Theorem 8.9

The norm of the functional u does not exceed the norm of the measure μ, since

$$|u(s)| \;=\; |\int_\Omega s d\mu| \leq \parallel s \parallel_\infty \parallel \mu \parallel_n^v,$$

see formula 8.9, so

$$\parallel u \parallel \leq \parallel \mu \parallel_n^v . \tag{8.12}$$

The definition of the total variation of the measure μ, see Defintion 8.9 gives that for any $\epsilon > 0$ there exists a finite collection of pairwise disjoint sets $\{A_1, \cdots, A_n\}$, $\quad (A_i \cap A_j = \varnothing,\, i \neq j)$, such that

$$\sum_{i=1}^{n} |\mu(A_i)| > |\mu|(\Omega) \;-\; \epsilon.$$

Take the following simple function

$$s(\omega) \;=\; \sum_{i=1}^{n} \frac{\mu(A_i)}{|\mu(A_i)|} \chi_{A_i}(\omega),$$

and be aware of the fact that $\parallel s \parallel_\infty = 1$, then

$$u(s) \;=\; \sum_{i=1}^{n} \frac{\mu(A_i)}{|\mu(A_i)|} \mu(A_i) \;=\; \sum_{i=1}^{n} |\mu(A_i)| \;\geq\; |\mu|(\Omega) \;-\; \epsilon.$$

Hence

$$\parallel u \parallel \geq \parallel \mu \parallel_n^v, \tag{8.13}$$

comparing the inequalities 8.12 and 8.13, the conclusion is that

$$\parallel u \parallel = \parallel \mu \parallel_n^v .$$

□

Thus there is proved that to each bounded linear functional u on $\mathcal{B}(\Omega, \Sigma)$ corresponds an unique finite additive measure μ and to each such measure corresponds the unique bounded linear functional u on $\mathcal{B}(\Omega, \Sigma)$ determined by formula 8.11. The norm of the functional u is equal to the total variation of the correspondent additive measure μ.

The spaces $\mathcal{B}(\Omega, \Sigma)^{'}$ and $ba(\Sigma)$ are isometrically isomorph.

9 Important Theorems

Most of the readers of these lecture notes have only a little knowledge of Topology. They have the idea that everything can be measured, have a distance, or that things can be separated. May be it is a good idea to read **wiki-topol-space**, to get an idea about what other kind of topological spaces their exists. A topology is needed if for instance their is spoken about convergence, connectedness, and continuity.
In first instance there will be referred to **WikipediA**, in the future their will be given references to books.

9.1 Axioms

There is assumed that there exists a nonempty set $\mathbb{R}$, the real numbers, which satisfy 10 axioms. These axioms can be divided in three groups, the *field axioms*, the *order axioms* and the *completeness axiom*. The last axiom is also known as the *least-upper-bound axiom* or the *axiom of continuity*.

9.1.1 Field Axioms

All the usual laws of arithmetic can be derived with the following axioms.

Axiom 9.1

Axiom 1: $x + y = y + x, \quad xy = yx,$
the commutative laws.

Axiom 2: $x + (y + z) = (x + y) + z, \quad x(yz) = (xy)z,$
the associative laws.

Axiom 3: $x(y + z) = xy + xz,$
the distributive law.

Axiom 4: Given any two real numbers x and y then there exists a real number z, such that $x + z = y$. This z is written by $y - x$ and $x - x$ is written by 0 and $-x$ is written for $0 - x$.

Axiom 5: There exists a real number $x \neq 0$. There exists a real number z such that $xz = y$. This real number z is written by $\frac{y}{x}$ and $\frac{x}{x}$ is written by 1. Further is $\frac{1}{x}$ written by x^{-1}, the reciprocal of x.

9.1.2 Order Axioms

The usual rules for inequalities can be done with the following axioms. The existence of a relation $<$ is assumed to exist between the real numbers.

Axiom 9.2

Axiom 6: Exactly one of the following relations holds:
$x = y$, $x < y$, $y < x$.
Note that $x > y$ means that $y < x$.

Axiom 7: If $x < y$ then for every z holds that $x + z < y + z$.

Axiom 8: If $x > 0$ and $y > 0$ then $xy > 0$.

Axiom 9: If $x > y$ and $y > z$ then $x > z$.

Note that with $x \leq y$ is meant: $x < y$ or $x = y$.

9.1.3 Completeness Axiom

Axiom 9.3

Axiom 10: Every nonempty set S of real numbers which is bounded above has a supremum. So there is a real number c such that $c = \sup S$, the least upper bound of S.

Note that a consequence of this axiom is, that if the set S is bounded below that it has a infimum, the greatest lower bound.

9.1.4 Axiom of Choice

Searching in the literature about it, it becomes more and more interesting. But let not to do too much. Keep in mind that the mathematics is based on several rules. It is nice to find out, what the minimal number of rules is to define such machinery as the mathematics, or parts of the mathematics. There can also be searched to statements, which in first instance have nothing to do with each other, but seem to be equivalent.

In set theory the axiom of choice is given by

Axiom 9.4

For every family $\{S_i\}_{i \in I}$ of nonempty sets there exists a family $\{x_i\}_{i \in I}$ of elements with $x_i \in S_i$ for every $i \in I$.

and a variant of it is given by

Axiom 9.5

Any collection of nonempty sets has a choice function.

where the definition of a choice function is given by

Definition 9.1

A choice function is a function f whose domain X is a collection of nonempty sets such that for every $S \in X$, $f(S)$ is an element of S.

There are a lot of variants, to have a nice overview of it, see the interesting book of (Herrlich, 2006).

9.2 Strange Abbreviations

Reading the literature there are sometimes used all kind of strange abbreviations. The author of such article or books thinks that everybody knows where is spoken about, but this is not always true. Here follows some list of such abbreviations, sometimes with an explanation, if not, there will be searched for it.

1. ZF: This has to do with the modern set theory. This theory based on axioms and one of these systems is named after the mathematicians Ernst Zermelo and Abraham Fraenkel.
2. ZFC: This is the Zermelo-Fraenkel set theory with the axiom of choice.
3. AC: Axiom of choice, see subsection **9.1.4**.

9.3 Background Theorems

In these lecture notes is made use of important theorems. Some theorems are proved, other theorems will only be mentioned.
In certain sections, some of these theorems are used and in other parts they play an important rule in the background. They are not always mentioned, but without these theorems, the results could not be proved or there should be done much more work to obtain the mentioned results.
But as the time passes, the mind changes, so it can happen that there is given a proof of some theorem within the written sections. See for instance Baire's category theorem, in **section 7.7.1** is given a proof of one of the variants of that theorem.

BcKTh 1: Lemma of Zorn , see Theorem **9.1** and for more information see **wiki-lemma-Zorn**.
BcKTh 2: Baire's category theorem , see **wiki-baire-category**.

9.3.1 Theorems mentioned in Section 9.3

The theorems as mentioned in the foregoing section are given. It is sometimes difficult to understand what is meant, because not every property is defined in these lecture notes.

Theorem 9.1

Lemma of Zorn:
If $X \neq \varnothing$ is a partially ordered set, in which every totally ordered subset has an upper bound in X, then X has at least one maximal element.

The Baire's category theorem seems to have several variants, which are not always equivalent. In some of these variants is also spoken about a Baire space.

Definition 9.2

A Baire space is a topological space with the property that for each countable collection of open dense sets U_n, their intersection $\cap U_n$ is dense.

9.4 Useful Theorems

A nice book, where a lot of information can be found is written by Körner, see (Körner, 2004).

10 Applications

In most books the first application, which is given by the authors, is the Banach fixed point thoerem or contraction theorem. The only thing, which is really needed, is a complete metric space. With the mentioned theorem can be proved the existence and uniqueness of the solution of some ordinary differential equations, some integral equations and linear algebraic equations. Other applications, such as partial differential equations, need soon more prior knowledge.

10.1 Banach fixed point theorem

Let X be some set and T a map of the set X into itself, so $T : X \longrightarrow X$.

Definition 10.1

A fixed point of a mapping $T : X \longrightarrow X$ is a point $x \in X$, which is mapped onto itself,

$$T(x) = x.$$

The image $T(x)$ coincides with x.

The Banach fixed point theorem gives sufficient conditions for the existence of a fixed point of certain maps T. There will be looked at contractions. A contraction can be used to calculate a fixed point.

Definition 10.2

Let (X, d) be a metric space. A mapping $T : X \longrightarrow X$ is called a contraction on X, if there exists some positive constant $\alpha < 1$, such that for all $x, y, \in X$

$$d(T(x), T(y)) \leq \alpha\, d(x, y).$$

With such a contraction it is sometimes possible to construct an approximation of a solution of some problem. The used procedure is called an iteration. There is started with some arbitrary x_0 in a given set and there is recusively calculated a sequence x_0, x_1, ... with the following relation

$$x_{n+1} = T(x_n) \quad \text{for} \quad n = 0, 1, 2, \ldots.$$

If the constructed sequence converges then it will converge to a fixed point of the map T. The Banach fixed point theorem gives sufficient conditions such that the constructed sequence converges and that the fixed point is unique.

Theorem 10.1

Banach fixed point theorem

Let (X, d) be a complete metric space, where $X \neq \varnothing$ and let the mapping $T : X \longrightarrow X$ be a contraction on X. Then T has an unique fixed point.

Proof of Theorem 10.1

The idea of the proof is to construct a Cauchy sequence $\{x_n\}_{n \in \{\mathbb{N} \cup 0\}}$ in the complete space X. The limit x of the constructed Cauchy sequence is a fixed point of T and it is the only fixed point of T.
Let's start with an arbitrary $x_0 \in X$ and construct the iterative sequence $\{x_n\}_{n \in \{\mathbb{N} \cup 0\}}$ by

$$x_{n+1} = T(x_n) \quad \text{for} \quad n = 0, 1, 2, \ldots.$$

So is obtained a sequence of images of x_0 under repeated application of the mapping T. Is that constructed sequence $\{x_n\}_{n \in \{\mathbb{N} \cup 0\}}$ a Cauchy sequence?
Let $n > m$ and let's look to the distance between x_n and x_m.
First is looked at the difference between two consecutive terms out of the constructed sequence. Since T is a contraction

$$\begin{aligned} d(x_{m+1}, x_m) &= d(T(x_m), T(x_{m-1})) \\ &\leq \alpha\, d(x_m, x_{m-1}) \\ &\leq \alpha\, d(T(x_{m-1}), T(x_{m-2})) \\ &\leq \alpha^2\, d(x_{m-1}, x_{m-2}) \leq \ldots \\ &\leq \alpha^m\, d(x_1, x_0) \end{aligned}$$

By the triangle inequality is obtained that

$$d(x_m, x_n) \leq d(x_m, x_{m+1}) + \ldots + d(x_{n-1}, x_n)$$
$$\leq (\alpha^m + \ldots + \alpha^{n-1})\, d(x_0, x_1)$$
$$\alpha^m\, (\frac{1 - \alpha^{n-m}}{1 - \alpha})\, d(x_0, x_1) \quad \text{for} \quad n > m.$$

From α is known that $0 < \alpha < 1$, so

$$d(x_m, x_n) \leq \frac{\alpha^m}{1 - \alpha}\, d(x_0, x_1) \quad (n > m). \tag{10.1}$$

The right-hand side of (10.1) can be made as small as wanted, because $d(x_0, x_1)$ is fixed and $0 < \alpha < 1$. So the constructed sequence $\{x_n\}_{n \in \{\mathbb{N} \cup 0\}}$ is a Cauchy sequence and since the space X is complete, that Cauchy sequence converges, to the limit is given some name $\lim_{n\to\infty} x_m = x$.
The next problem is to prove that x is a fixed point of the mapping T.
Again is used the triangle inequality and the fact that T is a contraction

$$d(x, T(x)) \leq d(x, x_m) + d(x_m, T(x)) \leq d(x, x_m) + d(T(x_{m-1}), T(x))$$
$$\leq d(x, x_m) + \alpha\, d(x_{m-1}, x).$$

Since $\lim_{n\to\infty} x_m = x$, it is easily seen that

$$\lim_{n\to\infty} d(x, T(x)) \leq \lim_{n\to\infty} (d(x, x_m) + \alpha\, d(x_{m-1}, x)) = 0.$$

So $d(x, T(x)) = 0$ and out that follows that $T(x) = x$ and it is shown that x is a fixed point of the mapping T.
The next question is, if x is the only fixed point of the mapping T? This is relative easy to prove by contradiction. Assume that there is some other fixed point $\widehat{x}$ of the mapping T, so $T(\widehat{x}) = \widehat{x}$, with $\widehat{x} \neq x$. Since x and $\widehat{x}$ are fixed points

$$d(x, \widehat{x}) = d(T(x), T(\widehat{x})) \leq \alpha\, d(x, \widehat{x}),$$

since $0 < \alpha < 1$, there follows that $d(x, \widehat{x}) = 0$. So $x = \widehat{x}$ and that is in comparison with the assumption. This means that the fixed point x of the mapping T is unique. The theorem is completely proved. $\square$

May be that later on there will be given some other theorems about fixed points of operators. Important in these theorems is the continuity of the operator.

Theorem 10.2

A contraction T on a metric space (X, d) is a continuous mapping.

Proof of Theorem 10.2

If $\epsilon > 0$ is given, there is easily find some $\delta(\epsilon) > 0$, etc.. □

10.2 Fixed Point Theorem and the Ordinary Differential Equations

Can the given fixed point theorem of Banach, see Theorem 10.1, be used to prove the existence and uniqueness of a solution of some ordinary differential equation?
Let's consider an explicit ordinary differential equation of order one, with some initial condition. For instance the following problem

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} x = f(t, x), & (10.2) \\ x(t_0) = x_0, & (10.3) \end{cases} \tag{10.4}$$

where x is an unknown function of the variable t, f is a given function and t_0 and x_0 are known values.
There are several theorems to prove the existence and uniqueness of a solution x of the given problem. Here Picard's theorem will be proved and there will also be given a method to obtain an approximation of the solution x of (10.4).
The function f is assumed to be continuous at a rectangle

$$R = \{(t, x) \mid \mid t - t_0 \mid \leq a, \mid x - x_0 \mid \leq b\},$$

where $a(> 0)$ and $b(> 0)$ are known positive constants. Sometimes the continuity of f is enough to prove the existence of a solution x of (10.4). But in the theorem of Picard there will be assumed more than that. With this extra assumption, about f, the uniqueness of the solution x is obtained.
The rectangle R is compact, so the function f is bounded on R. This means that there exists some positive constant $c(> 0)$ such that

$$\mid f(t, x) \mid \leq c \quad \text{for all} \quad (x, t) \in R.$$

Theorem 10.3

Picard's existence and uniqueness theorem ODE

Suppose that the function f, besides the continuity on the rectangle R, also satisfies the following Lipschitz condition. There exists a constant k such that

$$| f(t, x) - f(t, v) | \leq k | x - v |,$$

for all (t, x), $(t, v) \in R$.
Then the initial value problem (10.4) has an unique solution. This solution exists on the interval $t_0 - \beta \leq t \leq t_0 - \beta$ with $\beta < \min\{a, \frac{b}{c}, \frac{1}{k}\}$.

Proof of Theorem 10.3

The idea is to use the Banach fixed point theorem, so there is needed a metric space. Let $C(I)$ be that metric space, the space of real-valued continous functions on the interval $I = [t_0 - \beta, t_0 + \beta]$ with the metric

$$d(x, y) = \sup_{t \in I} | x(t) - y(t) |.$$

The space $(C(I), d)$ is complete, see Theorem 5.4.
There is some problem, that is there will be searched for a solution in some subspace $\widehat{C}$ of $C(I)$. $\widehat{C}$ are those functions x out of $C(I)$, which satisfy the condition

$$| x(t) - x_0 | \leq c\beta.$$

If $\{y_i\}_{i \in \mathbb{N}}$ is a convergent sequence in $\widehat{C}$, it is also a Cauchy sequence in $C(I)$. The space $(C(I), d)$ is complete, so the sequence $\{y_i\}_{i \in \mathbb{N}}$ converges in $C(I)$, define

$$\lim_{i \to \infty} y_i = y,$$

with $y \in C(I)$. because of the fact that

$$| y_i(t) - x_0 | \leq c\beta \quad \forall i,$$

there follows that

$$| y(t) - x_0 | \leq c\beta,$$

so $f \in \widehat{C}$. This means that the subspace $\widehat{C}$ is closed in $C(I)$ and also complete, see Theorem **3.7**.

$\square$

11 History

It is always a conflict where to place something about the history, at the beginning of the lectures notes or at the end? The end has been chosen because otherwise such a chapter can not be written. Every strange mathematical expression has to be defined, before it can be used.
So this chapter is written with the idea that everyone has read the chapters before. If not, the hope is that every strange mathematical word can be found in the Index, or can be found on the Internet.
To write this chapter was because of the question, where the word "functional" comes from? It is easy to point to the linear functionals, but if that is really the case? No idea, but after reading this chapter may be something becomes clear.

11.1 History

An very helpful article, to read something about the history, was written by (Carothers, 1993). The "Examensarbete" from (Lindstrom, 2008) is also a nice piece of work with much more mathematical details, as given in (Carothers, 1993). Further are given nice short biografies of important people, which have contributed to the functional analysis, in the book of (Saxe, 2002). At the very least may not be forgotten the book of (Dieudonné, 1981). It is not so readable, may be because of the use of a typewriter, but that was common in the time that book was written. That has also to do with history!
Names of people very much mentioned in the Functional Analysis are Fredholm, Lebesgue, Hilbert, Fréchet, Riesz, Helly, Banach and Hahn. These people lived around 1900, the time that the functional analysis started to become a discipline. There are more people who have contributed to the functional analysis, such as Fourier, d'Alembert, Poisson and Poincaré and go so on. They lived in the $18^{\text{th}} - 19^{\text{th}}$ century and had may be no idea that there work would become of great importance, or better would become a great source of inspiration, for the functional analysis in the 20^{th} century and later.
A nice book about the history of Functional Analysis is written by (Pietsch, 2007). It is focused on Banach Spaces and linear operators. It is not easy to read, but a lot of information is given, about proofs and all kind of other things. It is an interesting book. The writer of the book sees it's book as a historical supplement to the two books of (Johnson and Lindenstrauss, 2001) and (Johnson and Lindenstrauss, 2003).

12 Spectral Theory

Be careful in reading the definitions of the different spectra, for more details see textbooks as (Müller, 2007), (Bonsall, 1973) and (Kreyszig, 1978).

12.1 Complexified Operator

If there is done something with spectra, most of the time there are used Vector Spaces over the field of the complex numbers, $\mathbb{C}$. The real operators have to be adjusted, they have to be complexified.

Let $(X, \|\cdot\|)$ be a complex Normed Space. Let T be a linear operator with domain $\mathcal{D}(T) \subset X$ and range $\mathcal{R}(T) \subset X$. The scalar field may be either real or complex.
If the operator T is defined on a real Normed Space X, such an operator is to adjust for the complex case, but it is not as easy as it looks. A problem is to get an isometric isomorphism between certain Normed Vector Spaces, see **SubSection 12.6.1**.

Remark 12.1

Let X is a real Normed Space and let be $T \in L(X, X)$ a real bounded operator. The operator T can be complexified to the operator $T^{`}$ at the complex Normed Space $X_{\mathbb{C}} := X \times X = X \oplus \mathrm{i}\, X$.

If $x + \mathrm{i}\, y \in X_{\mathbb{C}}$ then $T^{`}(x + \mathrm{i}y) := T(x) + \mathrm{i}\, T(y)$, see **Theorem 12.4**.

12.2 Definition of the Spectrum

Very often people are interested in finding invariant subspaces of a linear operator.

Definition 12.1

Given a linear vector space X over a complex field $\mathbb{C}$ and a linear operator $T : X \longrightarrow X$, a subspaces M of X is called an invariant subspace of the operator T, if for every $x \in M$ holds that $T(x) \in M$, so $T(M) \subseteq M$.

The operator can be a matrix transformation, a linear integral operator, a linear differential operator and any other kind of a linear transformation.

The equation

$$T(x) - \lambda x = y \tag{12.1}$$

and the respective homogeneous equation

$$T(x) = \lambda x \tag{12.2}$$

play an important rule in the theory of linear operators; λ is a complex parameter, y is a given element of the space X and x is the unknown element of X. The equation **12.2** has a trivial solution $x = 0$ for every λ, but it may have also a solution different from zero at certain values of the parameter λ. These values play an exceptional role in the linear operator theory, the eigenvalues of T and the corresponding eigenvectors of the operator T, see **Definition 7.4**.

Definition 12.2

Given a linear vector space X over a complex field $\mathbb{C}$ and a linear operator $T : X \longrightarrow X$. Let the set $\{x_\alpha\}$ be the set of eigenvectors of the operator T, corresponding to the eigenvalue λ. The span, see **Definition 3.10**, of these eigenvectors is called the eigensubspace of the operator T, corresponding to the eigenvalue λ. This eigenspace, corresponding to the eigenvalue λ of the operator T, is notated by $E(\lambda)$, or $E(T)(\lambda)$, if there is spoken about more operators then T alone.

An eigensubspace is an invariant subspace of X, but an invariant subspace may be not an eigensubspace.

Example 12.1

A litte example to show that an invariant subspace has not to be an eigensubspace. Let M_1 and M_2 be invariant subspaces of T, then
$M_1 + M_2 = \{x_1 + x_2 \mid x_1 \in M_1, x_2 \in M_2\}$.
Let λ_1 and λ_2 be two different eigenvalues of T and let $M := E(\lambda_1) + E(\lambda_2)$. It is clear that M is an invariant subspace of T, but T restricted to M is not a multiple of the identity operator on M. If $x_1 \in E(\lambda_1)$ and $x_2 \in E(\lambda_2)$ then $M \ni T(x_1 + x_2) = \lambda_1 x_1 + \lambda_2 x_2 \neq \mu (x_1 + x_2)$.

Busy with Functional Analysis, the attention will go to the infinite dimensional spaces. There will be often searched to the largest invariant subspace, but be careful. Certainly with the dimension of these spaces, if the set of eigenvectors $\{x_\alpha\}$ is infinite, then the eigensubspace is an infinite dimensional subspace of X.
But also in thinking about, what is meant with the "largest" invariant subspace. Given an eigenvalue λ of T, then $E(\lambda)$ is the largest subspace M of X such that T restricted to M is λ times the identity operator on M. A nice question to be answered is: "Is $E(\lambda)$ the largest subspace M of X that is invariant under T and such that T restricted to M has λ as the only eigenvalue?". The following example will give the answer to that question.

Example 12.2

Let $M(\lambda)$ be a two dimensional Vector Space X. Let $B := \{\phi_1, \phi_2\}$ be a basis of X. And let $T : X \to X$ be a linear operator defined by the following equations: $T(\phi_1) = \lambda\phi$ and $T(\phi_2) = \phi_1 + \lambda\phi_2$, with λ some scalar. It is obvious that ϕ_1 is an eigenvector of T and ϕ_2 is not an eigenvector of T. The space $M(\lambda) := \text{span}(\phi_1, \phi_2)$ is invariant under T. It is easy to see that $\phi_1 = (T - \lambda I)\phi_2$ and since ϕ_1 is an eigenvector of T, associated with λ, $(T - \lambda I)^2\phi_2 = 0$. Hence

$$E(T)(\lambda) = \mathcal{N}(T - \lambda I) \subsetneq \mathcal{N}((T - \lambda I)^2) = M(\lambda).$$

Let $S : X \to X$ be a linear operator defined by the following equations: $S(\phi_1) = \lambda\phi$ and $S(\phi_2) = \lambda\phi_2$, with λ some scalar. It is obvious that ϕ_1 and ϕ_2 are eigenvectors of S and on the other hand

$$E(S)(\lambda) = \mathcal{N}(S - \lambda I) = \mathcal{N}((S - \lambda I)^2).$$

Theorem 12.1

If a parameter λ is an eigenvalue of the operator T, then the solution of the equation, given in 12.1, can not be unique.

Proof of Theorem 12.1

Assume that the equation, given in 12.1, has a solution x_0. Let ϕ be an eigenvector corresponding to the eigenvalue λ. Then $x_1 = x_0 + c\phi$ is also a solution of the equation given in 12.1, with c some arbitrary constant.
So equation 12.1 has no solutions, or it has infinitely many solutions. The operator $T - \lambda I$, with I the identity operator, has no inverse, if λ is an eigenvalue of the operator T.

□

Here just some more examples of invariant subspaces.

Example 12.3

Let's define the operator $T : X \longrightarrow X$, with $X = C^{\infty}(\mathbb{R})$, by

$$T : x \longrightarrow \frac{d}{dt} x, \quad \text{for} \quad x \in X.$$

1. If no other conditions are imposed on $x \in C^{\infty}(\mathbb{R})$, then every $\lambda \in \mathbb{C}$ is called an eigenvalue of T. The function $\phi : t \longrightarrow \exp(\lambda t)$ is called an eigenvector of T corresponding to the eigenvalue λ.

2. If there is looked at the linear space of all bounded functions in $C^{\infty}(\mathbb{R})$. Since $\phi : t \longrightarrow c \exp(\lambda t)$ defines a bounded function if and only if $\mathrm{Re}(\lambda) = 0$, the set of eigenvalues of T is defined by $\{\lambda \in \mathbb{C} \mid \mathrm{Re}(\lambda) = 0\}$.

Example 12.4

Let's define the operator $T : X \longrightarrow X$, with $X = C^{\infty}(\mathbb{R})$, by

$$T : x \longrightarrow \frac{d}{dt} x, \quad \text{for} \quad x \in X.$$

If there is looked at the linear space of functions $x \in C^{\infty}(\mathbb{R})$, such that: $x(t) = 0$ if $\mid t \mid \geq 1$, then T has no eigenvalues at all.
If $\lambda \in \mathbb{C}$ and $\phi \in C^{\infty}(\mathbb{R})$ then $\phi : t \longrightarrow c \exp(\lambda t)$ for some constant $c \neq 0$.
But the condition: $x(t) = 0$ if $\mid t \mid \geq 1$, implies that $c = 0$.
The described subspace is not empty, the linear subspace contains:

$$x(t) = \begin{cases} \exp\left(\frac{-1}{1 - t^2}\right) & \text{if } \mid t \mid < 1, \\ 0 & \text{otherwise.} \end{cases}$$

In the finite dimensional case there is an equivalence between injectivity and surjectivity. If there is looked at a linear map of a finite dimensional linear space to a space of the same dimension, there holds that the linear map is injective if and only if it is surjective. This equivalence doesn't hold in infinite dimensional linear spaces.

Example 12.5

An injective or a surjective operator has not to be bijective. To illustrate this fact, here a simple example. Let $X = \ell^2$ and let $\{e_k\}_{k\in\mathbb{B}}$ be the standard basis for X. For $x = \sum_{k=1}^{\infty} x_k\, e_k$, define

$$T_1(x) = \sum_{k=1}^{\infty} x_k\, e_{k+1} \quad \text{and}$$

$$T_2(x) = \sum_{k=1}^{\infty} x_{k+1}\, e_k$$

The operator T_1 is injective but not surjective, while the operator T_2 is surjective but not injective.

And busy with shift operators T_1 and T_2 in Example **12.5**, it is also illustrative to do a similar thing in the $\mathbb{L}_2(\mathbb{R})$.

Example 12.6

Define for some fixed value $h \in \mathbb{R}$, the operator T_h on $\mathbb{L}_2(\mathbb{R}$ by

$$T_h f(x) \;=\; f(x - h), \;\; \text{for all } \; x \in \mathbb{R}.$$

The linearity of the operator T is obvious and

$$\| \, T_h(f) \, \|_2^2 = \int_{-\infty}^{\infty} |\, f(x-h)\,|^2 \, dx \;=\; \int_{-\infty}^{\infty} |\, f(x)\,|^2 \, dx \;= \| \, f \, \|_2^2 \, .$$

So the operator T is bounded and $\| \, T_h \, \| = 1$.
The operator T_h is also regular for all $h \in \mathbb{R}$, since

$$(T_h)^{-1} f(x) \;=\; f(x+h) \;=\; T_{-h} f(x), \;\; \text{for all } \; x \in \mathbb{R}.$$

and also $\| \, (T_h)^{-1} \, \| = 1$.

Let's define some frequently used expressions for operators. With I is meant the identity operator on X, or $\mathcal{D}(T)$. With the operator T_λ is associated the operator

$$T_\lambda \;=\; (T \;-\; \lambda\, I), \tag{12.3}$$

where $\lambda \in \mathbb{C}$. If T_λ has an inverse, that is denoted by $R_\lambda(T)$, so

$$R_\lambda(T) = T_\lambda^{-1} = (T - \lambda I)^{-1}, \tag{12.4}$$

this operator is called the the resolvent operator of T. Sometimes only R_λ is written for the resolvent operator, if it is clear to what operator T is referred.

Remark 12.2

The definition of T_λ and $R_\lambda(T)$ is not unique. Other ways of writing are $T_\lambda = (\lambda I - T)$ or $T_\lambda = (I - \lambda T)$, and $R_\lambda(T) = (\lambda I - T)^{-1}$ or $R_\lambda(T) = (I - \lambda T)^{-1}$. Before reading publications, about spectral theory, is important to check what kind of definitions an author has used.
In these lecture notes are used the definitions **12.3** and **12.4**.

The term regular is already used in **Example 12.6**, but a definition was not given.

Definition 12.3

Let $(X, \|\cdot\|)$ be a complex Normed Space, with $X \neq \{0\}$ and let $T : \mathcal{D}(T) \rightarrow X$ be a linear operator with domain $\mathcal{D}(T) \subset X$. A regular value λ of T is a complex number such that

(R1) $R_\lambda(T)$ exists,
(R2) $R_\lambda(T)$ is bounded,
(R3) $R_\lambda(T)$ is defined on a set which is dense in X.

With **Definition 12.3**, the definition of the spectrum of T can be given.

Definition 12.4

The set of all the regular values of T is the resolvent set $\rho(T)$ and its complement $\rho(T)^c = \mathbb{C} \setminus \rho(T) =$ σ(T) is the spectrum of T.

An element of $\sigma(T)$ is called a spectral value of T.

Example 12.7

Look at the operator of multiplication by the independent variable in the space $C[a,b]$

$$T(x)(t) = tx(t).$$

This operator has no eigenvalues, because there is not a function $x \neq 0$ that satisfies the equation

$$tx(t) = \lambda x(t), \text{ for all } t \in [a,b],$$

at some λ.
On the other hand, if $\lambda \in [a,b]$ then the equation

$$tx(t) - \lambda x(t) = y(t)$$

has the solution

$$x(t) = \frac{y(t)}{(t - \lambda)}$$

for all those functions $y(t)$ which are representable in the form

$$y(t) = (t - \lambda)z(t), \text{ with } z \in C[a,b].$$

Important to mention is that the set of functions, with a zero at $t = \lambda$ are not dense in $C[a,b]$.
All the values of $\lambda \notin [a,b]$ are regular values. The resolvent operator is represented by a multiplication,

$$R_\lambda(T)x(t) = \frac{1}{(t - \lambda)}x(t) \quad (\in C[a,b]).$$

Example 12.7 makes clear that the spectrum $\sigma(T)$ can also exist out of other values than only eigenvalues.

It is possible to divide the spectrum $\sigma(T)$ into three mutually exclusive parts.

12.3 Spectrum (with state of an operator)

In this section the decomposition of the spectrum is done with the method used in (Taylor, 1958).

12.3.1 The states of an operator

In this section is considered a linear operator $T : X \rightarrow Y$, whose domain $\mathcal{D}(T)$ is a dense subspace of a normed linear space X and whose range $\mathcal{R}(T)$ is in a normed linear space Y. There will be made a ninefold classification of what is called the state of an operator.
First is listed three possibilities for $\mathcal{R}(T)$:

I. $\mathcal{R}(T) = Y$.

II. $\overline{\mathcal{R}(T)} = Y$, but $\mathcal{R}(T) \neq Y$.

III. $\overline{\mathcal{R}(T)} \neq Y$.

As regards T^{-1}, there are also listed three possibilities:

1. T^{-1} exists and is continuous.
2. T^{-1} exists but is not continuous.
3. T^{-1} does not exist.

If these possibilities are combined there are nine different situations. State II_2 for T means that $\overline{\mathcal{R}(T)} = Y$, but $\mathcal{R}(T) \neq Y$ and T^{-1} exists but is not continuous, also can be said that T is in state II_2.

In defining the different parts of the spectrum of an operator T, the state of the operator T_λ can be used.

12.3.2 Decomposition of Spectrum

With the use of the state of the operator T_λ, see **Section 12.3.1**, it is possible to divide the spectrum of the operator T into three mutually exclusive parts.

Definition 12.5

Let X be some Banach Space and let $T \in BL(X, X)$.

The resolvent set, denoted by $\rho(T)$:
$\lambda \in \rho(T)$ if and only if T_λ is in class I_1 or II_1.

The spectrum, denoted by $\sigma(T)$:
$\sigma(T) = \rho(T)^c = \mathbb{C} \setminus \rho(T)$

The continuous spectrum, denoted by $C_\sigma(T)$:
T_λ is in class I_2 or II_2.

The residual spectrum, denoted by $R_\sigma(T)$:
T_λ is in class III_1 or III_2.

The point spectrum, denoted by $P_\sigma(T)$:
T_λ is in class I_3, II_3 or III_3.

The mentioned subsets are disjoint and their union is the whole complex plane:

$$\mathbb{C} = \sigma(T) \cup \rho(T) = P_\sigma(T) \cup R_\sigma(T) \cup C_\sigma(T) \cup \rho(T).$$

12.4 Decomposition of Spectrum

In the literature is made use of all kind of decompositions of the spectrum. A little overview will be given, but not every part can be spoken into detail.

First is given the most common decomposition, as for instance given in (Kreyszig, 1978), with the help of **Definition 12.3**.

Definition 12.6

Let X be a complex Normed Space and $T \in L(X, X)$.
The resolvent set of T is the set

$$\rho(T) = \{\lambda \in \mathbb{C} \,|\, R_\lambda(T) \text{ exists and satisfies (R2) and (R3) } \}$$

and the spectrum of T is the set

$$\sigma(T) = \mathbb{C} \setminus \rho(T) = P_\sigma(T) \cup C_\sigma(T) \cup R_\sigma(T),$$

with the point spectrum $P_\sigma(T)$:

Definition 12.7

$P_\sigma(T) = \{\lambda \in \mathbb{C} \,|\, R_\lambda(T) \text{ does not exist } \},$

the residual spectrum $R_\sigma(T)$:

Definition 12.8

$R_\sigma(T) = \{\lambda \in \mathbb{C} \,|\, R_\lambda(T) \text{ exists, but does not satisfy (R3) } \}.$

and the continuous spectrum $C_\sigma(T)$:

Definition 12.9

$C_\sigma(T) = \{\lambda \in \mathbb{C} \,|\, R_\lambda(T) \text{ exists and satisfies (R3), but not (R2) } \},$

The mentioned subsets are disjoint and their union is the whole complex plane:

$$\mathbb{C} = \rho(T) \cup \sigma(T) = \rho(T) \cup P_\sigma(T) \cup R_\sigma(T) \cup C_\sigma(T).$$

The conditions of the different spectra are summarized in the following table. In short: ((R1): R_λ exists, (R2) R_λ bounded, (R3): R_λ defined on dense set in X).

Satisfied			**Not satisfied**			λ belongs to:
(R1),	(R2),	(R3)				$\rho(T)$
			(R1)			$P_\sigma(T)$
(R1)					(R3)	$R_\sigma(T)$
(R1),		(R3)		(R2)		$C_\sigma(T)$

Table 12.1 Conditions different spectra.

See also the flowchart at page 314.

12.4.1 Differences between classifications

There are a lot of differences in the way the authors classify the points of the spectrum. But most of the time the given classifications are equivalent. Here follows a list of differences and also given some alternative conditions.

a. Busy with linear Functional Analysis the operators are linear, but often is given the extra assumption, that these operators have to be continuous. Or there is spoken in the definition about bounded linear operators. Continuity and boundedness are equivalent for linear operators at a Normed Space. So speaking about $L(X, X)$ with the assumption that the operators have to be continuous or speaking about $BL(X, X)$ makes no difference.

b. A great difference is, if there is taken an operator at a Normed Space or a Banach Space. Possible consequence? A Normed Space is not necessarily complete, but the continuous dual space of a normed space over a complete field is necessarily complete. In the case of a Banach Space both are complete.

c. The condition that an operator is bijective. Be careful what is meant, bijective at the whole space or at the range of the corresponding operator? The range of an operator has not to be the whole space. And is the operator defined at its domain or at the whole space?

d. Injectivity of T_λ, also called one-to-one, that can be done at several ways:

i. $\mathcal{N}(T_\lambda) = \{0\}$, so *only(!)* the 0-element of X than T_λ is injective. But be careful, is meant injectivity at $\mathcal{D}(T_\lambda)$ or at the whole space X?

ii. R_λ exist and is bounded if and only if T_λ is bounded from below, see **Theorem 7.10**.

But if this is given is this way, again the difficult question, what is meant: $T_\lambda : \mathcal{D}(T_\lambda) \rightarrow \mathcal{R}(T_\lambda)$ or $T_\lambda : X \rightarrow X$? Is there looked at the whole space or only at subsets, subspaces of it?

Most of the time $T : X \rightarrow \mathcal{R}(T)$, so $T_\lambda : X \rightarrow \mathcal{R}(T_\lambda)$ and $R_\lambda : \mathcal{R}(T_\lambda) \rightarrow X$. R_λ is not always defined at the whole space X.

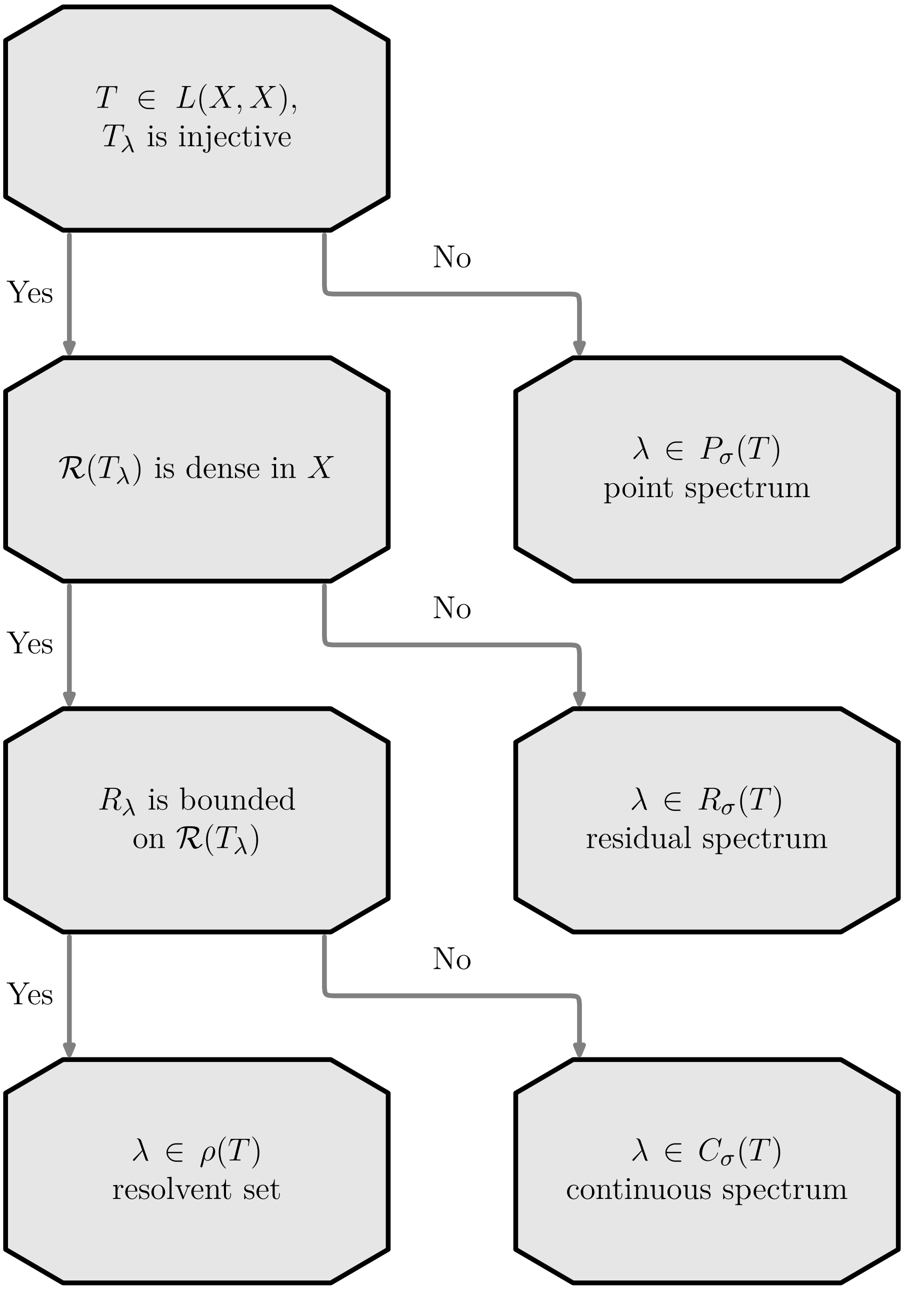


Figure 12.1 A flowchart of the spectrum.

12.5 Spectral Properties of Bounded Linear Operators

In this chapter the power series of **Chapter 7.4.1** will be used. **Theorem 12.2** has been already been proved, in certain sense, in **Theorem 7.13**.

Lemma 12.1

Let X be a complex Banach Space, $T : X \to X$ a linear operator and $\lambda \in \rho(T)$. Assume that

a. T is closed or

b. T is bounded.

Then $R_\lambda(T)$ is defined on the whole space X and is bounded.

Proof of Theorem 12.1

a. If T is closed, so is T_λ by **Theorem 7.18**. Hence $R_\lambda(T)$ is closed and $R_\lambda(T)$ is also bounded, see (R2) in **Definition 12.3**. Hence $\mathcal{D}(R_\lambda(T))$ is closed, use **Theorem 7.22**, and so follows with (R3) of **Definition 12.3** that
$\mathcal{D}(R_\lambda(T)) = \overline{\mathcal{D}(R_\lambda(T))} = X$.

b. Since $\mathcal{D}(T) = X$ is closed, there follows that T is closed, see **Theorem 7.21**, and with **part ii.a** of this Lemma, the statement follows.

□

Theorem 12.2

The resolvent set $\rho(T)$ of a bounded linear operator T on a complex Banach Space X is open; hence the spectrum $\sigma(T)$ is closed.

Proof of Theorem 12.2

Let's start with the last part of the proposition.
If $\rho(T)$ is open, then its complement $\rho^c(T) = \mathbb{C} \setminus \rho(T) = \sigma(T)$ is closed.
If $\rho(T) = \emptyset$, it is open. (There will be proved that $\rho(T) \neq \emptyset$, see ??)
So let $\rho(T) \neq \emptyset$ and let $\lambda_0 \in \rho(T)$. For any $\lambda \in \mathbb{C}$, there can be written

$$T_\lambda = T_{\lambda_0} - (\lambda - \lambda_0)I = T_{\lambda_0}(I - (\lambda - \lambda_0)T_{\lambda_0}^{-1}), \tag{12.5}$$

so **formula 12.5** can be written in the form

$$T_\lambda = T_{\lambda_0} V \text{ where } V = I - (\lambda - \lambda_0)R_{\lambda_0}(T). \tag{12.6}$$

Since $\lambda_0 \in \rho(T)$ and T is bounded, follows with **Lemma 12.1**,**part ii.b** that $R_{\lambda_0}(T)$ is bounded.
With **Theorem 7.12**, the Neumann series, the inverse of V is given by

$$V^{-1} = \sum_{n=0}^{\infty}((\lambda - \lambda_0)R_{\lambda_0}(T))^n = \sum_{n=0}^{\infty}(\lambda - \lambda_0)^n R_{\lambda_0}(T)^n.$$

Furthermore V^{-1} is bounded, for all λ with $\| (\lambda - \lambda_0)R_{\lambda_0}(T) \| < 1$, so for all λ with

$$|(\lambda - \lambda_0)| < \frac{1}{\| R_{\lambda_0}(T) \|}. \tag{12.7}$$

Since $R_{\lambda_0}(T)$ is bounded, there follows that for every λ, which satisfies **inequality 12.7**, T_λ has a bounded inverse

$$R_\lambda(T) = T_\lambda^{-1} = (T_{\lambda_0} V)^{-1} = V^{-1} R_{\lambda_0}(T).$$

So **inequality 12.7** gives a neighbourhood of λ_0, consisting of regular values λ of T. $\lambda_0 \in \rho(T)$ was arbitrary chosen, so $\rho(T)$ is open.
□

Theorem 12.3

The spectrum $\sigma(T)$ of a bounded linear operator T on a complex Banach Space X is compact.

Proof of Theorem 12.3

If $|\lambda| > \| T \|$ then $\frac{\| T \|}{|\lambda|} < 1$ and the operator $T_\lambda =$ âĹŠÎż$(I$âĹŠ$\frac{T}{\lambda})$ has the inverse

$$R_\lambda(T) = T_\lambda^{-1} = -\sum_{n=0}^{\infty} \lambda^{(-n-1)} T^n,$$

use Theorem 7.12 and therefore

$$\sigma(T) \subset \{z \in \mathbb{C} \mid |z| \leq \| T \|\}.$$

So $\sigma(T)$ is bounded and in Theorem 12.2 is proved that $\sigma(T)$ is closed, so $\sigma(T)$ is compact. □

12.6 Banach Algebras

Reading about Spectral Theory and very fast there will be the confrontation with the Banach Algebras. Speaking about a spectrum of an operator T, that operator will be an linear operator from some space X to the same space X. Otherwise there can not be spoken about eigenvectors, eigenspaces, invariant subspaces and so on.
The space $L(X, X)$ is the Vector Space of all linear operators of X into itself and that means that if $S, T \in L(X, X)$, there can also be spoken about a product between the operators S and T, see Definition ??.
Combining the linearity of the space $L(X, X)$ and the possibility to define products between linear operators gives the possibility to speak about an algebra.

Definition 12.10

An algebra over $\mathbb{K}$ is a linear space A over $\mathbb{K}$ together with a mapping $(x, y) \to xy$ of $A \times A$ into A, that satisfies for every $x, y, z \in A$ and for all $\alpha \in \mathbb{K}$:

Alg 1: $\mathrm{x(yz) = (xy)z}$,

Alg 2: $\mathrm{x(y+z) = xy+yz}$, $\mathrm{(x+y)z = xz+yz}$,

Alg 3: $(\alpha \mathrm{x})\mathrm{y} = \alpha(\mathrm{xy}) = \mathrm{x}(\alpha \mathrm{y})$.

Example 12.8

Let X be a linear space over $\mathbb{K}$ and for $L(X, X)$, see **Definition 7.8** and take $Y = X$.
$L(X, X)$ with the product, defined by the composition

$$(ST)(x) \;=\; S(T(x)), \text{ for every } x \in X,$$

is an algebra, also denoted by $L(X)$.

Be careful in the next definition when reading 0. There can be meant the scalar 0 of the field $\mathbb{K}$, or there is meant the 0-element of the algebra A. From the text should be clear what is meant.

Definition 12.11

Let A be an algebra over $\mathbb{K}$. An algebra seminorm is a mapping $p : A \to \mathbb{R}$ such that for all $x, y \in A$ and $\alpha \in \mathbb{K}$:

ASN 1: $p(x) \;\geq\; 0$,

ASN 2: $p(\alpha x) \;= \mid \alpha \mid p(x)$,

ASN 3: $p(x \,+\, y) \;\leq\; p(x) \,+\, p(y)$,

ASN 4: $p(xy) \;\leq\; p(x)p(y)$.

Definition 12.12

Let A be an algebra over $\mathbb{K}$. By an algebra norm is meant a mapping $\|\cdot\|: A \to \mathbb{R}$ such that:

AN 1: $(A, \|\cdot\|)$ is a Normed Space over $\mathbb{K}$,

AN 2: $\| xy \| \leq \| x \| \| y \|$ for all $x, y \in A$.

Definition 12.13

A Normed Algebra is a pair $(A, \|\cdot\|)$, where A is non-zero algebra and $\|\cdot\|$ is a given algebra-norm on A.

A Banach Algebra is Normed Algebra that is complete in its norm (i.e. it is Banach Space).

By an Unital Normed (or Banach) Algebra is meant a Normed (respectively Banach) Algebra with an identity element I_A such that

UA 1: $\| I_A \| = 1$.

Example 12.9

An important class of Banach Algebras is made up $C[a,b]$, the continuous functions on the compact interval $[a,b]$.
The algebraic operations are the usual pointwise addition and the multiplication of functions

1. $(f + g)(t) = f(t) + g(t)$,

2. $(f{\cdot}g)(t) = f(t) \cdot g(t)$.

The norm is defined by $\| f \|_\infty = \sup_{x \in [a.b]} \mid f(t) \mid$, note that:

$$\| f \cdot g \|_\infty \leq \| f \|_\infty \cdot \| g \|_\infty .$$

Further is $C[a,b]$, in the given norm, a Banach Space, so $A = (C[a,b], \| \cdot \|_\infty)$ is a Banach Algebra.
The constant function 1 is the identity of A, so it is also a Unital Banach Algebra.

12.6.1 Complexification of real Normed Algebras

The theory of Banach Algebras is most of the time concerned with algebras over the complex numbers. The study of real algebras can be reduced to a study of complex algebras. The real algebra A will be embedded isometrically and isomorphically in a certain complex algebra $A_{\mathbb{C}}$.

Definition 12.14

The complex algebra $A_{\mathbb{C}}$ is the cartesian product $A \times A$ with the algebraic operations similar as in the field $\mathbb{C}$, the element (x, y) behaves like $x + i\,y$, so

a. $(x, y) + (u, v) = (x + u, y + v)$,
b. $(\alpha + i\,\beta)(x, y) = (\alpha\,x - \beta\,y, \alpha\,y + \beta\,x)$,
c. $(x, y) \cdot (u, v) = (x\,u - y\,v, x\,v + y\,u)$,

for all $x, y, u, v \in A$ and $\alpha, \beta \in \mathbb{R}$.
The mapping $x \rightarrow (x,\, 0)$ is an isomorphism of A into $A_{\mathbb{C}}$.
The algebra $A_{\mathbb{C}}$ is the complexification of A.

There is a little problem with the requirement of an isometric embedding, but that will be solved. That problem is may illustrative for the fact that sometimes is required that the norm of the identity has to be one, see in the definition of a normed algebra ii: 1.
Let's define

$$|\,(x,\, y)\,| = \|\, x \,\| \,+\, \|\, x \,\|, \tag{12.8}$$

and define

$$\|\,(x,\, y)\,\| = \sup_{\theta}(|\,\exp{(i\theta)}\,(x, y)\,|), \tag{12.9}$$

Now the $A_{\mathbb{C}}$ becomes a complex Normed Algebra, with $\|\,(x,\, y)\,\|$ as norm. If the algebra A has the identity 1 with norm 1, then $(1, 0)$ is an identity for $A_{\mathbb{C}}$, but $\|\,(1,\, 0)\,\| = \sup_{\theta}(|\,\cos{(\theta)}\,| \,+\, |\,\sin{(\theta)}\,|) \,=\, \sqrt{2} \,\neq\, 1$.
Now first a theorem about the comlexification of an arbitrary real Normed Vector Space X and after that the complexification of an real Normed Algebra A.

Theorem 12.4

Let $X_{\mathbb{C}}$ be the complexification of an arbitrary real Normed Vector Space X. Then $X_{\mathbb{C}}$ can be given a norm $\| (x,\, y) \|$ so that is a complex Normed Vector Space wit the following properties:

i. The isomorphism $x \rightarrow (x,\, 0)$ of X into $X_{\mathbb{C}}$ is an isometry.

ii. $X_{\mathbb{C}}$ is a Banach Space if and only if X is a Banach Space.

iii. Let $T \in BL(X, X)$ and define $T^{‘}(x,\, y) = (T(x),\, T(y))$ for $(x,\, y) \in X_{\mathbb{C}}$, then the mapping $T \rightarrow T^{‘}$ is an isometric isomorhism of the algebra $BL(X, X)$ into $BL(X_{\mathbb{C}}, X_{\mathbb{C}})$.

Proof of Theorem 12.4

The properties (i) and (ii):
$X_{\mathbb{C}}$ is the Cartesian product $X \times X$, with the same operations as defined in ii.a and ii.b. With the use of 12.8, $X_{\mathbb{C}}$ becomes a real Normed Vector Space with $| (x,\, y) |$ as norm. It is easy to observe that $X_{\mathbb{C}}$ is complete if and only if X is complete in its norm.
Now define

$$\| (x,\, y) \| = \frac{1}{\sqrt{2}} \sup_{\theta} (| \exp{(i\theta)}\, (x, y) |),$$

so $X_{\mathbb{C}}$ becomes a complex Normed Vector Space, with as norm $\| (x,\, y) \|$. (The same in the case of Banach Spaces.) Since $\| x \| = | (x,\, 0) | = \| (x,\, 0) \|$ the embedding $x \rightarrow (x,\, 0)$ of X into $X_{\mathbb{C}}$ becomes an isometry. The norms $| (x,\, y) |$ and $\| (x,\, y) \|$ are equivalent since

$$\frac{1}{\sqrt{2}} \,|\, (x,\, y) \,| \leq \| (x,\, y) \| \leq | (x,\, y) | \,.$$

Propertie (iii):
That $T^{‘}$ is complex linear on $X_{\mathbb{C}}$ is easy to check, evenas the fact that $T \rightarrow T^{‘}$ is a real isomorphism of $BL(X, X)$ into the algebra of $BL(X_{\mathbb{C}}, X_{\mathbb{C}})$.
In the following inequality is only spoken about T

$$\| T \| = \sup_{x} \frac{\| T(x) \|}{\| x \|} \leq \sup_{x,y} \frac{\| T(x) \| + \| T(y) \|}{\| x \| + \| y \|} \leq \| T \|,$$

since

$$| T^{‘} | = \sup_{(x,y)} \frac{| T^{‘}(x,y) |}{| (x,\, y) |} = \sup_{x,y} \frac{\| T(x) \| + \| T(y) \|}{\| x \| + \| y \|},$$

there is obtained that $\| T \| = | T^{‘} |$. Furthermore

$$\| T^{‘}(x,y) \| = \frac{1}{\sqrt{2}} \sup_{\theta}(| \exp{(i\theta)}T^{‘}(x,y) |) = \frac{1}{\sqrt{2}} \sup_{\theta}(| T^{‘}(\exp{(i\theta)})(x,y)) |)$$

$$\leq \frac{| T^{‘} |}{\sqrt{2}} \sup_{\theta}(| \exp{(i\theta)}(x,y) |) = \| T \| \| (x,y) \| .$$

Hence $\| T^{‘} \| \leq \| T \|$, but also $\| T(x) \| = \| T^{‘}(x,\, 0) \| \leq \| T^{‘} \| \| x \|$, so that $\| T \| \leq \| T^{‘} \|$. Therefore $\| T \| = \| T^{‘} \|$. ☐

12.7 Examples of Spectra

In this chapter there will be given examples of spectra of all kind of linear operators.

12.7.1 Right-, Left-Shift Operators

Let's define the right-shift operator $RS : \ell^p \rightarrow \ell^p$, with $1 \leq p < \infty$, by

$$RS(x_1, x_2, x_3, \cdots) = (0, x_1, x_2, x_3, \cdots),$$

sometimes also called the forward-shift operator. Since $(\ell^p)' = \ell^q$, with $\frac{1}{p} + \frac{1}{q} = 1$, see **Theorem 5.15**, the duality is defined by

$$(x, y) = \sum_{i=1}^{\infty} x_i\, y_i$$

for all $x \in \ell^p$ and $y \in \ell^q$. It is easy to verify that

$$(RS(x), y) \;=\; \sum_{1}^{\infty} x_i \, y_{i+1} \;=\; (x, RS^*(y)).$$

That means that the adjoint operator RS^* is defined by

$$RS^*(y_1, y_2, y_3, \cdots) \;=\; (y_2, y_3, \cdots) \;=\; LS(y_1, y_2, y_3, \cdots).$$

The adjoint operator RS^* is also known as the left-shift operator, or the backward-shift operator.

Let $1 < p < \infty$:

i. Point spectrum of RS:
Since $RS(x) = \lambda \, x \;\Leftrightarrow\; (0, x_1, x_2, \cdots) = \lambda \, (x_1, x_2, \cdots) \;\Leftrightarrow$
$x = (0, 0, \cdots)$, so the point spectrum of RS is empty,
$P_\sigma(RS) = \emptyset$.

ii. Point spectrum of LS:
Since $LS(x) = \lambda \, x \;\Leftrightarrow\; (x_2, x_3, \cdots) = \lambda \, (x_1, x_2, \cdots) \;\Leftrightarrow$
$x_{i+1} = \lambda x_i$ for all $i \in \mathbb{N} \;\Leftrightarrow\; x = (x_1, x_2, x_3, \cdots) = (1, \lambda, \lambda^2, \cdots)$
with $\lambda \in \mathbb{C}$, $x \in \ell^q$ if $\mid \lambda \mid < 1$.
If $\mid \lambda \mid = 1$, it can not be an eigenvalue of $LS = RS^*$.

iii. Spectral radius of RS:
Since $\parallel (RS)^n(z) \parallel_p = \parallel z \parallel_p$,

$$r(RS) \;=\; \lim_{n \to \infty} \parallel (RS)^n \parallel^{(\frac{1}{n})} = \parallel RS \parallel = \; 1.$$

That means that $\sigma(RS) \;=\; \sigma(LS) \;=\; \{\lambda \in \mathbb{C} \mid \mid \lambda \mid \leq 1\}$

13 Exercises-All

It has to become a chapter with exercises of all kind, most of the time also with a solution. Problems can be solved on different ways. So sometimes there are given several ways to solve.
First there has to be thought about how to solve an exercise. If the solution of an exercise is read, before solving such an exercise, the step about how an exercise should be resolved is beaten. In practice the solution of a problem will never be earlier than the problem is asked. But in a lot of lecture notes and most eductional books, they exist unfortunately at the same time.
The most important thing of solving problems is to understand the problem and to know what can be done to solve such a problem. In an exercise is usually given the information needed to solve such an exercise. But for problems in practice that is most of the time not the case.

13.1 Lecture Exercises

Ex-1:

If $f(x) = f(y)$ for every bounded linear functional f on a Normed Space X.
Show that $x = y$.

Ex-2:

Define the metric space $B[a, b]$, $a < b$, under the metric

$$d(f, g) = \sup_{x \in [a,b]} \{| f(x) - g(x) |\},$$

by all the bounded functions on the compact interval $[a, b]$.
If $f \in B[a, b]$ then there exists some constant $M < \infty$ such that $| f(x) | \leq M$ for every $x \in [a, b]$.
Show that $(B[a, b], d)$ is not separable.

Ex-3:

Let (X, d) be a Metric Space and A a subset of X. Show that $x_0 \in \overline{A} \Leftrightarrow d(x_0, A) = 0$.

Ex-4:

Let X be a Normed Space and X is reflexive and separable. Show that $X^{''}$ is separable.

Ex-5:

Given is some sequence $\{u_n\}_{n \in \mathbb{N}}$.
Prove the following theorems:

a. First Limit-Theorem of Cauchy:
If the

$$\lim_{n \to \infty} (u_{n+1} - u_n)$$

exists, then the limit

$$\lim_{n \to \infty} \frac{u_n}{n}$$

exists and

$$\lim_{n \to \infty} \frac{u_n}{n} = \lim_{n \to \infty} (u_{n+1} - u_n).$$

b. Second Limit-Theorem of Cauchy:
If $u_n > 0$ and the limit

$$\lim_{n \to \infty} \frac{u_{n+1}}{u_n}$$

exists, then the limit

$$\lim_{n \to \infty} \sqrt[n]{u_n}$$

exists and

$$\lim_{n \to \infty} \sqrt[n]{u_n} = \lim_{n \to \infty} \frac{u_{n+1}}{u_n}.$$

Ex-6:

Given is some sequence $\{u_n\}_{n \in \mathbb{N}}$ and $\lim_{n \to \infty} u_n = L$ exists then also

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} u_i = L.$$

Ex-7:

X and Y are Normed Spaces.
Let $T : X \to Y$ be a linear operator.
Then the following are equivalent:

i. For every bounded set S of X, $T(S)$ is bounded in Y.

ii. The set $\{T(x) \mid \parallel x \parallel = 1\}$ is bounded in Y.

iii. There exists a $c > 0$ such that $\parallel T(x) \parallel \leq c \parallel x \parallel$ for all $x \in X$.

iv. T is uniformly continuous.

v. T is continuous at 0.

Solution, see **Sol- ii: 7**.

13.2 Revision Exercises

Ex. 1:

What is a "norm"?

For solution, see **Sol. ii:1**.

Ex. 2:

What does it mean if a Metric Space is "complete"?

For solution, see **Sol. ii:2**.

Ex. 3:

Give the definition of a "Banach Space" and give an example.

For solution, see **Sol. ii:3**.

Ex. 4:

What is the connection between bounded and continuous linear maps?

For solution, see **Sol. ii:4**.

Ex. 5:

What is the Dual Space of a Normed Space?

For solution, see **Solution ii:5**.

Ex. 6:

What means "Hilbert space"? Give an example.

For solution, see **Sol. ii:6**.

13.3 Exam Exercises

Ex-1: Consider the normed linear space $(c, \| \cdot \|_\infty)$ of all convergent sequences, i.e., the space of all sequences $x = \{\lambda_1, \lambda_2, \lambda_3, \ldots\}$ for which there exists a scalar L_x such that $\lambda_n \to L_x$ as $n \to \infty$. Define the functional f on c by

$$f(x) = L_x.$$

a. Show that $| L_x | \leq \| x \|_\infty$ for all $x \in c$.

b. Prove that f is a continous linear functional on $(c, \| \cdot \|_\infty)$.

Solution, see **Sol- ii: 1**.

Ex-2: Consider the Hilbert space $L_2[0, \infty)$ of square integrable real-valued functions, with the standard inner product

$$\langle f, g \rangle = \int_0^\infty f(x) g(x) dx = \lim_{R \to \infty} \int_0^R f(x) g(x) dx.$$

Define the linear operator $T : L_2[0, \infty) \to L_2[0, \infty)$ by

$$(Tf)(x) = f(\frac{x}{5}) \text{ where } f \in L_2[0, \infty) \text{ and } x \in [0, \infty).$$

a. Calculate the Hilbert-adjoint operator T^*.
Recall that $\langle Tf, g \rangle = \langle f, T^*(g) \rangle$ for all $f, g \in L_2[0, \infty)$.

b. Calculate the norm of $\| T^*(g) \|$ for all $g \in L_2[0, \infty)$ with $\| g \| = 1$.

c. Calculate the norm of the operator T.

Solution, see **Sol- ii: 2**.

Ex-3: Let $A : [a, b] \to \mathbb{R}$ be a continuous function on $[a, b]$. Define the operator $T : L_2[a, b] \to L_2[a, b]$ by

$$(Tf)(t) = A(t) f(t).$$

a. Prove that T is a linear operator on $L_2[a, b]$.

b. Prove that T is a bounded linear operator on $L_2[a, b]$.

Solution, see Sol- ii: 3.

Ex-4: Show that there exist unique real numbers a_0 and b_0 such that for every $a, b \in \mathbb{R}$ holds

$$\int_0^1 |t^3 - a_0 t - b_0|^2 dt \leq \int_0^1 |t^3 - at - b|^2 dt.$$

Moreover, calculate the numbers a_0 and b_0.
Solution, see Sol- ii: 4.

Ex-5: Consider the inner product space $C[0,1]$ with the inner product

$$(f, g) = \int_0^1 f(t) g(t) dt.$$

The sequence of functions $\{f_n\}_{n \in \mathbb{N}}$ is defined by

$$f_n(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq \frac{1}{2} \\ 1 - n(t - \frac{1}{2}) & \text{if } \frac{1}{2} < t < \frac{1}{2} + \frac{1}{n} \\ 0 & \text{if } \frac{1}{2} + \frac{1}{n} \leq t \leq 1 \end{cases}$$

a. Sketch the graph of f_n.

b. Calculate the pointwise limit of the sequence $\{f_n\}$ and show that this limit function is not an element of $C[0,1]$.

c. Show that the sequence $\{f_n\}$ is a Cauchy sequence.

d. Show that the the sequence $\{f_n\}$ is not convergent.

Solution, see Sol- ii: 5.

Ex-6: Define the operator $A : \ell^2 \to \ell^2$ by

$$(A\,\mathbf{b})_n = (\frac{3}{5})^n\, b_n$$

for all $n \in \mathbb{N}$ and $b_n \in \mathbb{R}$ and $\mathbf{b} = (b_1, b_2, \cdots) \in \ell^2$.

a. Show that A is a linear operator on ℓ^2.

b. Show that A is a bounded linear operator on ℓ^2 and determine $\| A \|$. (The operator norm of A.)

c. Is the operator A invertible?

Solution, see Sol- ii: 6.

Ex-7: Given is the following function $f : \mathbb{R}^3 \to \mathbb{R}$

$$f(a,b,c) = \int_{-\pi}^{\pi} |\sin(\frac{t}{2}) - a - b\cos(t) - c\sin(t)|^2 dt,$$

which depends on the real variables a, b and c.

a. Show that the functions $f_1(t) = 1$, $f_2(t) = \cos(t)$ and $f_3(t) = \sin(t)$ are linear independent on the interval $[-\pi, \pi]$.

b. Prove the existence of unique real numbers a_0, b_0 and c_0 such that

$$f(a_0, b_0, c_0) \leq f(a,b,c)$$

for every $a, b, c \in \mathbb{R}$. (Don't calculate them!)

c. Explain a method, to calculate these coefficients a_0, b_0 and c_0. Make clear, how to calculate these coefficients. Give the expressions you need to solve, if you want to calculate the coefficients a_0, b_0 and c_0.

Solution, see Sol- ii: 7.

Ex-8: Consider the space $C[0,1]$, with the sup-norm $\| \, . \, \|_\infty$,

$$\| g \|_\infty = \sup_{x \in [0,1]} |g(x)| \quad (g \in C[0,1]).$$

The sequence of functions $\{f_n\}_{n \in \mathbb{N}}$ is defined by

$$f_n(x) = \arctan(nx), x \in [0,1].$$

a. Sketch the graph of f_n.
For $n \to \infty$, the sequence $\{f_n\}$ converges pointwise to a function f.

b. Calculate f and prove that f is not an element of $C[0,1]$.
Let's now consider the normed space $L_1[0,1]$ with the L_1-norm

$$\| g \|_1 = \int_0^1 |g(x)| dx \quad (g \in L_1[0,1]).$$

c. Calculate

$$\lim_{n \to \infty} \int_0^1 |f(t) - f_n(t)| dt$$

(Hint: $\int \arctan(a\,x)\,dx = x \arctan(a\,x) - \frac{1}{2\,a} \ln(1 + (a\,x)^2) + C$, with $C \in \mathbb{R}$ (obtained with partial integration))

d. Is the sequence $\{f_n\}_{n \in \mathbb{N}}$ a Cauchy sequence in the space $L_1[0,1]$?

Solution, see Sol- ii: 8.

Ex-9: Consider the normed linear space ℓ^2. Define the functional f on ℓ^2 by

$$f(\mathbf{x}) = \sum_{n=1}^{\infty} (\frac{3}{5})^{(n-1)} x_n,$$

for every $\mathbf{x} = (x_1, x_2, \cdots) \in \ell^2$.

a. Show that f is a linear functional on ℓ^2.

b. Show that f is a continous linear functional on ℓ^2.

Solution, see Sol- ii: 9.

Ex-10: Consider $A : L_2[-1,1] \to L_2[-1,1]$ given by

$$(Af)(x) = x\,f(x).$$

a. Show that $(Af) \in L_2[-1,1]$ for all $f \in L_2[-1,1]$.

b. Calculate the Hilbert-adjoint operator A^*. Is the operator A self-adjoint?

Solution, see Sol- ii: 10.

Ex-11: Define the operator $T : C[-1,1] \longrightarrow C[0,1]$ by

$$T(f)(t) = \int_{-t}^{t} (1 + \tau^2)\,f(\tau)\,d\tau$$

for all $f \in C[-1,1]$.

a. Take $f_0(t) = \sin(t)$ and calculate $T(f_0)(t)$.

b. Show that T is a linear operator on $C[-1,1]$.

c. Show that T is a bounded linear operator on $C[-1,1]$.

d. Is the operator T invertible?

Solution, see Sol- ii: 11.

Ex-12: Define the following functional

$$F(x) = \int_0^1 \tau\, x(\tau)\, d\tau,$$

on $(C[0,1], \|\cdot\|_\infty)$.

a. Show that F is a linear functional on $(C[0,1], \|\cdot\|_\infty)$.

b. Show that F is bounded on $(C[0,1], \|\cdot\|_\infty)$.

c. Take $x(t) = 1$ for every $t \in [0,1]$ and calculate $F(x)$.

d. Calculate the norm of F.

Solution, see Sol- ii: 12.

Ex-13: Let $x_1(t) = t^2$, $x_2(t) = t$ and $x_3(t) = 1$.

a. Show that $\{x_1, x_2, x_3\}$ is a linear independent set in the space $C[-1,1]$.

b. Orthonormalize x_1, x_2, x_3, in this order, on the interval $[-1,1]$ with respect to the following inner product:

$$< x, y > = \int_{-1}^1 x(t)\, y(t)\, dt.$$

So $e_1 = \alpha\, x_1$, etc.

Solution, see Sol- ii: 13.

Ex-14: Let H be a Hilbert space, $M \subset H$ a closed convex subset, and (x_n) a sequence in M, such that $\| x_n \| \rightarrow d$, where $d = \inf_{x \in M} \| x \|$, this means that $\| x \| \geq d$ for every $x \in M$.

a. Show that (x_n) converges in M.
(Hint: $(x_n + x_m) = 2\,(\frac{1}{2}x_n + \frac{1}{2}x_m)$)

b. Give an illustrative example in $\mathbb{R}^2$.

Solution, see **Sol- ii: 14**.

Ex-15: Some questions about ℓ^2 and ℓ^1.

a. Give a sequence $a \in \ell^2$, but $a \notin \ell^1$.

b. Show that $\ell^1 \subset \ell^2$.

Solution, see **Sol- ii: 15**.

Ex-16: Define the operator $A : \ell^2 \rightarrow \ell^2$ by

$$(A\,a)_n = \frac{1}{n^2}\,a_n \text{ for all } n \in \mathbb{N},\ a_n \in \mathbb{C} \text{ and } a = (a_1, a_2, \cdots) \in \ell^2.$$

a. Show that A is linear.

b. Show that A is bounded; find $\| A \|$.

c. Is the operator A invertible? Explain your answer.

Ex-17: Given are the functions $f_n : [-1, +1] \rightarrow \mathbb{R}$, $n \in \mathbb{N}$,

$$f_n(x) = \sqrt{(\frac{1}{n} + x^2)}.$$

a. Show that $f_n : [-1, +1] \rightarrow \mathbb{R}$ is differentiable and calculate the derivative $\frac{\partial f_n}{\partial x}$.

b. Calculate the pointwise limit $g : [-1, +1] \rightarrow \mathbb{R}$, i.e.

$$g(x) = \lim_{n \rightarrow \infty} f_n(x),$$

for every $x \in [-1, +1]$.

c. Show that

$$\lim_{n \to \infty} \| f_n - g \|_\infty = 0,$$

with $\| \cdot \|_\infty$, the sup-norm on $C[-1, +1]$.

d. Converges the sequence $\{\frac{\partial f_n}{\partial x}\}_{n \in \mathbb{N}}$ in the normed space $(C[-1, +1], \| \cdot \|_\infty)$?

Ex-18: Let $C[-1, 1]$ be the space of continuous functions at the interval $[-1, 1]$, provided with the inner product

$$\langle f, g \rangle = \int_{-1}^{+1} f(\tau)\, g(\tau)\, d\tau$$

and $\| f \| = \sqrt{\langle f, f \rangle}$ for every $f, g \in C[-1, 1]$.
Define the functional $h_n : C[-1, 1] \to \mathbb{R}$, $n \in \mathbb{N}$ by

$$h_n(f) = \int_{-1}^{+1} (\tau^n)\, f(\tau)\, d\tau.$$

a. Show that the functional h_n, $n \in \mathbb{N}$ is linear.

b. Show that the functional h_n, $n \in \mathbb{N}$ is bounded.

c. Show that

$$\lim_{n \to \infty} \| h_n \| = 0.$$

Solution, see **Sol- ii: 17**.

Ex-19: Let (e_j) be an orthonormal sequence in a Hilbert space H, with inner product $\langle \cdot, \cdot \rangle$.

a. Show that if

$$x = \sum_{j=1}^{\infty} \alpha_j\, e_j \text{ and } y = \sum_{j=1}^{\infty} \beta_j\, e_j$$

then

$$\langle x, y \rangle = \sum_{j=1}^{\infty} \alpha_j \overline{\beta_j},$$

with $x, y \in H$.

b. Show that $\sum_{j=1}^{\infty} | \alpha_j \overline{\beta_j}|$ converges.

Ex-20: In $L_2[0,1]$, with the usual inner product $\langle \cdot, \cdot \rangle$, is defined the linear operator $T : f \rightarrow T(f)$ with

$$T(f)(x) = \frac{1}{\sqrt[4]{(4\,x)}} f(\sqrt{x}\,).$$

a. Show that T is a bounded operator and calculate $\| T \|$.

b. Calculate the adjoint operator T^* of T.

c. Calculate $\| T^* \|$.

d. Is $T^* T = I$, with I the identity operator?

Solution, see **Sol- ii: 16**.

Ex-21: For $n \in \mathbb{N}$, define the following functions $g_n,\, h_n,\, k_n : \mathbb{R} \rightarrow \mathbb{R}$

$$\begin{cases} g_n(x) = \sqrt{n} & \text{if } 0 < x < \frac{1}{n} \text{ and 0 otherwise,} \\ h_n(x) = n & \text{if } 0 < x < \frac{1}{n} \text{ and 0 otherwise,} \\ k_n(x) = 1 & \text{if } n < x < (n+1) \text{ and 0 otherwise.} \end{cases}$$

a. Calculate the pointwise limits of the sequences $(g_n)_{(n \in \mathbb{N})}$, $(h_n)_{(n \in \mathbb{N})}$ and $(k_n)_{(n \in \mathbb{N})}$.

b. Show that none of these sequences converge in $L_2(\mathbb{R})$. The norm on $L_2(\mathbb{R})$ is defined by the inner product $\langle\, f,\, g \rangle = \int_{-\infty}^{\infty} f(x)\, g(x)\, dx$.

Ex-22: Consider the space $\mathbb{R}^{\infty}$ of all sequences, with addition and (scalar) multiplication defined termwise.
Let $S : \mathbb{R}^{\infty} \rightarrow \mathbb{R}^{\infty}$ denote a *shift* operator, defined by $S((a_n)_{(n \in \mathbb{N})} = (a_{n+1})_{(n \in \mathbb{N})}$ for all $(a_n)_{(n \in \mathbb{N})} \in \mathbb{R}^{\infty}$. The operator S working on the sequence $(a_1,\, a_2,\, a_3,\, \ldots)$ has as image the sequence $(a_2,\, a_3,\, a_4,\, \ldots)$.

a. Prove that S^2 is a linear transformation.

b. What is the kernel of S^2?

c. What is the range of S^2?

Ex-23: Let $L_2[-1,1]$ be the Hilbert space of square integrable real-valued functions, on the interval $[-1,+1]$, with the standard inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(x)\, g(x)\, dx$$

Let $a,\, b \in L_2[-1,1]$, with $a \neq 0,\, b \neq 0$ and let the operator $T : L_2[-1,1] \longrightarrow L_2[-1,1]$ be given by

$$(T\, f)(t) = \langle f,\, a\rangle\, b(t)$$

for all $f \in L_2[-1,1]$.

a. Prove that T is a linear operator on $L_2[-1,1]$.

b. Prove that T is a continuous linear operator on $L_2[-1,1]$.

c. Compute the operator norm $\| T \|$.

d. Derive the null space of T, $\mathcal{N}(T) = \{g \in L_2[-1,1] \,|\, T(g) = 0\}$ and the range of T, $\mathcal{R}(T) = \{T(g) \,|\, g \in L_2[-1,1]\}$.

e. What condition the function $a \in L_2[-1,1]$ ($a \neq 0$) has to satisfy, such that the operator T becomes idempotent, that is $T^2 = T$.

f. Derive the operator $S : L_2[-1,1] \longrightarrow L_2[-1,1]$ such that

$$\langle T(f),\, g\rangle = \langle f,\, S(g)\rangle$$

for all $f,\, g \in L_2[-1,1]$.

g. The operator T is called self-adjoint, if $T = S$. What has to be taken for the function a, such that T is a self-adjoint operator on $L_2[-1,1]$.

h. What has to be taken for the function a ($a \neq 0$), such that the operator T becomes an orthogonal projection?

Ex-24: a. Let V be a vectorspace and let $\{V_n | n \in \mathbb{N}\}$ be a set of linear subspaces of V. Show that $\bigcap_{n=1}^{\infty} V_n$ is a linear subspace of V.

b. Show that c_{00} is not complete in ℓ^1.

Ex-25: In $L_2[0,1]$, with the usual inner product $(\cdot,\cdot)$, is defined the linear operator $S : u \rightarrow S(u)$ with

$$S(u)(x) \;=\; u(1-x).$$

Just for simplicity, the functions are assumed to be real-valued.
The identity operator is notated by I. ($I(u) = u$ for every $u \in L_2[0,1]$.)
An operator P is called idempotent, if $P^2 = P$.

a. Compute S^2 and compute the inverse operator S^{-1} of S.

b. Derive the operator $S^* : L_2[0,1] \longrightarrow L_2[0,1]$ such that

$$(S(u),\, v) \;=\; (u,\, S^*(v))$$

for all $u,\, v \in L_2[0,1]$. The operator S^* is called the adjoint operator of S. Is S selfadjoint? (selfadjoint means that: $S^* = S$.)

c. Are the operators $\frac{1}{2}(I+S)$ and $\frac{1}{2}(I-S)$ idempotent?

d. Given are fixed numbers $\alpha, \beta \in \mathbb{R}$ with $\alpha^2 \neq \beta^2$. Find the function $u : [0,1] \rightarrow \mathbb{R}$ such that

$$\alpha\, u(x) + \beta\, u(1-x) = \sin(x).$$

(Suggestion(s): Let $v \in L_2[0,1]$. What is $\frac{1}{2}(I+S)v$? What is $\frac{1}{2}(I-S)v$? What is $\frac{1}{2}(I+S)v + \frac{1}{2}(I-S)v$? What do you get, if you take $v(x) = \sin(x)$?)

Solution, see **Sol- ii: 20**.

Ex-26: The functional f on $(C[-1,1], \|\cdot\|_\infty)$ is defined by

$$f(x) \; = \; \int_{-1}^{0} x(t)\, dt \; - \; \int_{0}^{1} x(t)\, dt$$

for every $x \in C[-1, 1]$.

a. Show that f is linear.

b. Show that f is continuous.

c. Show that $\| f \| = 2$.

d. What is $\mathcal{N}(f)$?
$\mathcal{N}(f) = \{x \in C[-1, 1] \mid f(x) = 0\}$ is the null space of f.

Solution, see Sol- ii: 19.

Ex-27: Some separate exercises, they have no relation with each other.

a. Show that the vector space $C[-1, 1]$ of all continuous functions on $[-1, 1]$, with respect to the $\| \cdot \|_\infty$-norm, is the direct sum of the set of all even continuous functions and the set of all odd continuous functions on $[-1, 1]$.

b. Given are the functions $f_n : [-1, +1] \to \mathbb{R}$, $n \in \mathbb{N}$,

$$f_n(t) \; = \; \begin{cases} 1 & \text{for } -1 \leq t \leq -\frac{1}{n} \\ -n\, x & \text{for } -\frac{1}{n} < t < \frac{1}{n} \\ -1 & \text{for } \frac{1}{n} \leq t \leq 1. \end{cases}$$

Is the sequence $\{f_n\}_{n \in \mathbb{N}}$ a Cauchy sequence in the Banach space $(C[-1, 1], \| \cdot \|_\infty)$?

Solution, see Sol- ii: 18.

Ex-28: Just some questions.

a. What is the difference between a Normed Space and a Banach Space?

b. For two elements f and g in an Inner Product Space holds that $\| f + g \|^2 = \| f \|^2 + \| g \|^2$. What can be said about f and g? What can be said about f and g, if $\| f + g \| = \| f \| + \| g \|$?

c. What is the difference between a Banach Space and a Hilbert Space?

Ex-29: The sequence $x = (x_n)_{n\in\mathbb{N}}$ and the sequence $y = (y_n)_{n\in\mathbb{N}}$ are elements of c, with c the space of all convergent sequences, with respect to the $\|\cdot\|_\infty$-norm. Assume that

$$\lim_{n\to\infty} x_n = \alpha \text{ and } \lim_{n\to\infty} y_n = \beta.$$

Show that

$$(\alpha\, x + \beta\, y) \in c.$$

Solution, see Sol- ii: 21.

Ex-30: Let $\mathbb{F}(\mathbb{R})$ be the linear space of all the functions f with $f : \mathbb{R} \to \mathbb{R}$. Consider $f_1,\ f_2,\ f_3$ in $\mathbb{F}(\mathbb{R})$ given by

$$f_1(x) = 1,\ f_2(x) = \cos^2(x),\ f_3(x) = \cos(2x).$$

a. Prove that $f_1,\ f_2$ and f_3 are linear dependent.

b. Prove that f_2 and f_3 are linear independent.

Solution, see Sol- ii: 22.

Ex-31: Consider the operator $A : \ell^2 \to \ell^2$ defined by

$$A(a_1,\ a_2,\ a_3,\ \cdots) =$$
$$(a_1 + a_3, a_2 + a_4, a_3 + a_5,\ \cdots,\ a_{2k-1} + a_{2k+1}\,,\ a_{2k} + a_{2k+2},\ \cdots),$$

with $\ell^2 = \{(a_1,\ a_2,\ a_3,\ \cdots) \mid a_i \in \mathbb{R} \text{ and } \sum_{i=1}^{\infty} a_i^2 < \infty\}$.

a. Prove that A is linear.

b. Prove that A is bounded.

c. Find $N(A)$.

Solution, see Sol- ii: 23.

Ex-32: The linear space P_2 consists of all polynomials of degree ≤ 2. For $p, q \in P_2$ is defined

$$(p, q) = p(-1)q(-1) + p(0)q(0) + p(1)q(1).$$

a. Prove that (p, q) is an inner product on P_2.

b. Prove that q_1, q_2 and q_3, given by

$$q_1(x) = x^2 - 1, \quad q_2(x) = x^2 - x, \quad q_3(x) = x^2 + x$$

are mutually orthogonal.

c. Determine $\| q_1 \|, \| q_2 \|, \| q_3 \|$.

Solution, see **Sol- ii: 24**.

13.4 Solutions Lecture Exercises

Sol-1: $f(x) - f(y) = f(x - y) = 0$ for every $f \in X^{'}$, then

$$\| x - y \| = \sup_{\{f \in X', f \neq 0\}} \{\frac{|f(x-y)|}{\| f \|}\} = 0,$$

see **theorem 4.13**. Hence, $x = y$. □

Sol-2: For each $c \in [a, b]$ define the function f_c as follows

$$f_c(t) = \begin{cases} 1 & \text{if } t = c \\ 0 & \text{if } t \neq c \end{cases}.$$

Then $f_c \in B[a, b]$ for all $c \in [a, b]$. Let M be the set containing all these elements, $M \subset B[a, b]$. If $f_c, f_d \in M$ with $c \neq d$ then $d(f_c, f_d) = 1$.

Suppose that $B[a, b]$ has a dense subset D. Consider the collection of balls $B_{\frac{1}{3}}(m)$ with $m \in M$. These balls are disjoint. Since D is dense in $B[a, b]$, each ball contains an element of D and D is also countable, so the set of balls is countable.

The interval $[a, b]$ is uncountable, so the set M is uncountable and that is in contradiction with the fact that the set of disjoint balls is countable.
So the conclusion is that $B[a, b]$ is not separable.

Sol-3: Has to be done.

Sol-4: The Normed Space X is separable. So X has a countable dense subset S.
If $f \in X$ there is a countable sequence $\{f_n\}_{n \in \mathbb{N}}$, with $f_n \in S$, such that

$$\lim_{n \to \infty} \| f_n - f \| = 0.$$

X is reflexive, so the canonical map $C : X \to X^{''}$ is injective and onto. Let $z \in X^{''}$, then there is some $y \in X$, such that $z = C(y)$. X is separable, so there is some sequence $\{y_i\}_{i \in \mathbb{N}} \subset S$ such that $\lim_{i \in \mathbb{N}} \| y_i - y \| = 0$. This means that

$$0 = \lim_{i \to \infty} \| y_i - y \| = \lim_{i \to \infty} \| C(y_i - y) \| = \lim_{i \to \infty} \| C(y_i) - z \| .$$

S is countable, that means that $C(S)$ is countable. There is found a sequence $\{C(y_i)\}_{i \in \mathbb{N}} \subset C(S)$ in $X^{''}$, which converges to $z \in X^{''}$. So $C(S)$ lies dense in $X^{''}$, since $z \in X^{''}$ was arbitrary chosen, so $X^{''}$ is separable.

Sol-5: Every proof wil be done in several steps.
Let $\epsilon > 0$ be given.

ii.5.a

1. The limit $\lim_{n \to \infty}(u_{n+1} - u_n)$ exist, so there is some L such that $\lim_{n \to \infty}(u_{n+1} - u_n) = L$. This means that there is some $N(\epsilon)$ such that for every $n > N(\epsilon)$:

$$L - \epsilon < u_{n+1} - u_n < L + \epsilon.$$

2. Let M be the first natural natural number greater then $N(\epsilon)$ such that

$$L - \epsilon < u_{M+1} - u_M < L + \epsilon.$$

then

$$L - \epsilon < u_{(M+1+i)} - u_{(M+i)} < L + \epsilon,$$

for $i = 0, 1, 2, \cdots, n-(M+1)$, with $n > (M+1)$.
Summation of these inequalities gives that:

$$(n - M)(L - \epsilon) < u_n - u_M < (n - M)(L + \epsilon),$$

so

$$(L-\epsilon) + \frac{u_M - M(L - \epsilon)}{n} < \frac{u_n}{n} < (L+\epsilon) + \frac{u_M - M(L + \epsilon)}{n}.$$

3. u_m, M and ϵ are fixed numbers, so

$$\lim_{n\to\infty} \frac{u_M - M(L - \epsilon)}{n} = 0$$

and

$$\lim_{n\to\infty} \frac{u_M - M(L + \epsilon)}{n} = 0.$$

That means that there are numbers $N_1(\epsilon)$ and $N_2(\epsilon)$ such that

$$|\frac{u_M - M(L - \epsilon)}{n}| < \epsilon$$

and

$$|\frac{u_M - M(L + \epsilon)}{n}| < \epsilon.$$

Take $N_3(\epsilon) > \max(N(\epsilon), N_1(\epsilon), N_2(\epsilon))$ then

$$(L - 2\epsilon) < \frac{u_n}{n} < (L + 2\epsilon),$$

for every $n > N_3(\epsilon)$, so

$$\lim_{n\to\infty} \frac{u_n}{n} = L.$$

ii.5.b

It can be proven with ϵ and $N_i(\epsilon)$'s, but it gives much work.
Another way is, may be, to use the result of part ii.5.a?
Since $u_n > 0$, for every $n \in \mathbb{N}$, $\ln(u_n)$ exists.
Let $v_n = \ln(u_n)$ then $\lim_{n\to\infty}(v_{n+1} - v_n) = \lim_{n\to\infty} \ln(\frac{u_{n+1}}{u_n})$ exists, because $\lim_{n\to\infty} \frac{u_{n+1}}{u_n}$ exists. The result of part ii.5.a can be used.
First:

$$\lim_{n\to\infty} \frac{v_n}{n} = \lim_{n\to\infty} \frac{\ln(u_n)}{n}$$
$$= \lim_{n\to\infty} \ln \sqrt[n]{u_n},$$

and second:

$$\lim_{n\to\infty} (v_{n+1} - v_n) = \lim_{n\to\infty} (\ln(u_{n+1}) - \ln(u_n))$$
$$= \lim_{n\to\infty} \ln(\frac{u_{n+1}}{u_n}),$$

and with the result of **ii.5.a**:

$$\lim_{n\to\infty} \ln \sqrt[n]{u_n} = \lim_{n\to\infty} \ln(\frac{u_{n+1}}{u_n}),$$

or,

$$\lim_{n\to\infty} \sqrt[n]{u_n} = \lim_{n\to\infty} \frac{u_{n+1}}{u_n}.$$

Sol-6: Define

$$s_n = \sum_{i=1}^{n} u_i,$$

then is

$$\lim_{n\to\infty} (s_{n+1} - s_n) = \lim_{n\to\infty} (\sum_{i=1}^{n+1} u_i - \sum_{i=1}^{n} u_i) = \lim_{n\to\infty} u_{n+1} = L.$$

Using the result of exercise **ii.5.a** gives that

$$\lim_{n\to\infty} \frac{s_n}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} u_i = L.$$

Sol-7: a. (i)⇒(ii):

Let $S = \{x \in X \mid \; \| x \| = 1\}$ then S is bounded and so $T(S) = \{T(x) \mid \; \| x \| = 1\}$ is bounded.

b. (ii)⇒(iii):

Let $x \in X$ then $\| \frac{x}{\| x \|} \| = 1$. So there is some $c > 0$, independent of x, such that $\| T(\frac{x}{\| x \|}) \| \leq c$. Since T is linear operator, there follows that $\| T(x) \| \leq c \| x \|$.

c. (iii)⇒(iv):

Let $\epsilon > 0$ and take $\delta = \dfrac{\epsilon}{2\,c}$,
then for every $x, y \in X$ with $\| x - y \| < \delta$
$\| T(x) - T(y) \| = \| T(x - y) \| \leq c \| x - y \| < \epsilon$.

d. (iv)⇒(v):

T is uniform continuous, so T is continuous in $x = 0$.

e. (v)⇒(i):

Let S be a bounded set in X, then there is some $c > 0$ such that $\| x \| \leq c$ for all $x \in S$.
T is continuous in $x = 0$.
Take $\epsilon = 1$, then there exists some $\delta(\epsilon) > 0$ such that for all $x \in X$ with $\| x - 0 \| < \delta(\epsilon)$, $\| T(x) - T(0) \| < \epsilon$, because of the continuity of T in $x = 0$.
Let $x \in S$ then $\| \dfrac{x}{c} \dfrac{\delta(\epsilon)}{2} \| < \delta(\epsilon)$ and $\| T(\dfrac{x}{c} \dfrac{\delta(\epsilon)}{2} \| < \epsilon$. This means that $T(S)$ is bounded, because $\| T(x) \| < 2 \dfrac{\epsilon}{\delta(\epsilon)} c$ for all $x \in S$.

Go back to exercise **Ex. ii: 7**.

13.5 Solutions Revision Exercises

Sol. 1:

See **definition 3.23**.
Go back to exercise **Ex. ii:1**.

Sol. 2:

A Metric Space is complete if every Cauchy sequence converges in that Metric Space.
Go back to exercise **Ex. ii:2**.

Sol. 3:

A Banach Space is a complete Normed Space, for instance $C[a, b]$ with the $\| \cdot \|_\infty$ norm.
Go back to exercise **Ex. ii:3**.

Sol. 4:

Bounded linear maps at Normed Spaces are continuous and continuous maps at Normed Spaces are bounded, see **theorem 7.2**. Be careful, use the mentioned theorem in a good way, be aware of the Normed Spaces!
Go back to exercise **Ex. ii:4**.

Sol. 5:

See the **section 4.5**.
Go back to exercise **Ex. ii:5**.

Sol. 6:

For the definition, see **3.34**. An example of a Hilbert Space is the ℓ^2, see **5.2.4**.
Go back to exercise **Ex. ii:6**.

13.6 Solutions Exam Exercises

Some of the exercises are worked out into detail. Of other exercises the outline is given about what has to be done.

Sol-1:

a. Let $x = \{\lambda_1, \lambda_2, \lambda_3, \ldots\} \in c$, then $|\lambda_i| \leq \| x \|_\infty$ for all $i \in \mathbb{N}$, so $|L_x| \leq \| x \|_\infty$.

b. $|f(x)| = |L_x| \leq \| x \|_\infty$, so

$$\frac{|f(x)|}{\| x \|_\infty} \leq 1,$$

the linear functional is bounded, so continuous on $\| x \|_\infty$.

Sol-2:

a. $\langle T f, g \rangle = \lim_{R \to \infty} \int_0^R f(\frac{x}{5}) g(x) dx = \lim_{R \to \infty} \int_0^{\frac{R}{5}} f(y) g(5\,y)\, 5 dy = \langle f, T^* g \rangle$, so $T^* g(x) = 5\, g(5\, x)$.

b. $\| T^*(g) \|^2 = \lim_{R \to \infty} \int_0^R |5\, g(5\, x)|^2 dx$, so $\| T^*(g) \|^2 = 25 \lim_{R \to \infty} \int_0^{5\,R} \frac{1}{5} |g(y)|^2 d$ $5 \, \| g \|^2$ and this gives that $\| T^* \| = \sqrt{5}$.

c. $\| T \| = \| T^* \|$.

Sol-3:

a. Let $f, g \in L_2[a, b]$ and $\alpha \in \mathbb{R}$ then $T(f + g)(t) = A(t)(f + g)(t) = A(t) f(t) + A(t) g(t) = T(f)(t) + T(g)(t)$ and $T((\alpha\, f))(t) = A(t)(\alpha\, f)(t) = \alpha\, A(t)(f)(t) = \alpha\, T(f)(t)$.

b. $\| (T f) \| \leq \| A \|_\infty \| f \|$, with $\| \cdot \|_\infty$ the sup-norm, A is continuous and because $[a, b]$ is bounded and closed, then $\| A \|_\infty = \max_{t \in [a,b]} |A(t)|$.

Sol-4: Idea of the exercise. The span of the system $\{1, t\}$ are the polynomials of degree less or equal 1. The polynomial t^3 can be projected on the subpace span$(1, t)$. Used is the normal inner product $\langle f, g \rangle = \int_0^1 f(t)\, g(t)\, dt$. The Hilbert Space theory gives that the minimal distance of t^3 to the span$(1, t)$ is given by the length of the difference of t^3 minus its projection at the span$(1, t)$. This latter gives the existence of the numbers a_0 and b_0 as asked in the exercise.
The easiest way to calculate the constants a_0 and b_0 is done by $\langle t^3 - a_0\, t - b_0, 1 \rangle = 0$ and $\langle t^3 - a_0\, t - b_0, t \rangle = 0$, because the

difference $(t^3 - a_0\, t\, -\, b_0)$ has to be perpendicular to $\text{span}(1, t)$.

Sol-5: a. See figure 13.1.

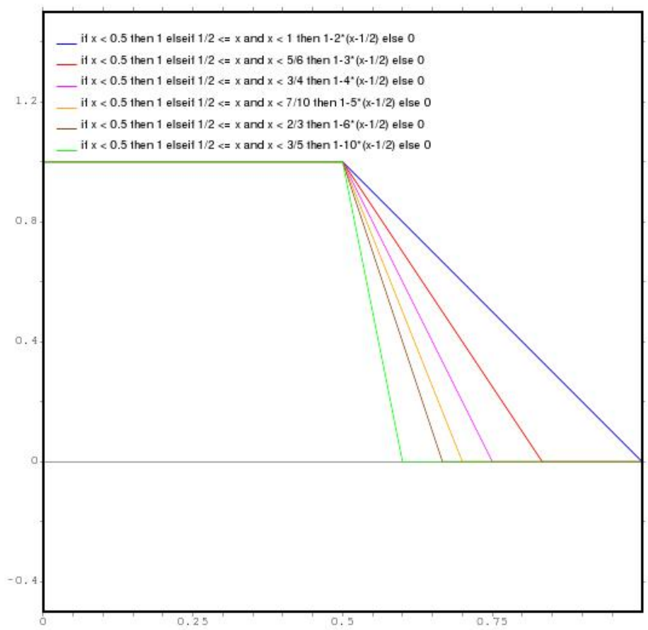


Figure 13.1 f_n certain values of n

b. Take x fixed and let $n \rightarrow \infty$, then

$$f(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq \frac{1}{2} \\ 0 & \text{if } 0\frac{1}{2} < t \leq 1. \end{cases}$$

It is clear that the function f makes a jump near $t = \frac{1}{2}$, so the function is not continuous.

c. There has to be looked to $\| f_n - f_m \|$ for great values of n and m. Exactly calculated this gives $\frac{|m - n|}{\sqrt{(3\, m^2\, n)}}$. Remark: it is not the intention to calculate the norm of $\| f_n - f_m \|$ exactly!
Because of the fact that $|f_n(t) - f_m(t)| \leq 1$ it is easily seen that

$$\int_0^1 |f_n(t) - f_m(t)|^2 dt \leq \int_0^1 |f_n(t) - f_m(t)| dt \leq \frac{1}{2}(\frac{1}{n} - \frac{1}{m})$$

for all $m > n$. The bound is the difference between the areas beneath the graphic of the functions f_n and f_m.
Hence, $\| f_n - f_m \| \rightarrow 0$, if n and m are great.

d. The functions f_n are continuous and the limit function f is not continous. This means that the sequence $\{f_n\}_{n \in \mathbb{N}}$ does not converge in the Normed Space $(C[0,1], \| \cdot \|)$, with $\| g \| = \sqrt{\langle g, g \rangle}$.

Sol-6: a. Take two arbitrary elements $\mathbf{c}, \mathbf{d} \in \ell^2$, let $\alpha \in \mathbb{R}$, show that

$$\begin{cases} A(\mathbf{c} + \mathbf{d}) = A(\mathbf{c}) + A(\mathbf{d}) \\ A(\alpha \mathbf{c}) = \alpha A(\mathbf{c}). \end{cases}$$

by writing out these rules, there are no particular problems.

b. Use the norme of the space ℓ^2 and

$$\| A(\mathbf{b}) \|^2 = (\frac{3}{5})^2 (b_1)^2 + (\frac{3}{5})^4 (b_2)^2 + (\frac{3}{5})^6 (b_3)^2 + \cdots \leq (\frac{3}{5})^2 \; \| \mathbf{b} \|^2,$$

so $\| A(\mathbf{b}) \| \leq \frac{3}{5} \; \| \mathbf{b} \|$.
Take $\mathbf{p} = (1, 0, 0, \cdots)$, then $\| A(\mathbf{p}) \| = \frac{3}{5} \; \| \mathbf{p} \|$, so $\| A \| = \frac{3}{5}$ (the operator norm).

c. If A^{-1} exists then $(A^{-1}(\mathbf{b}))_n = (\frac{5}{3})^n \, (\mathbf{b})_{\mathbf{n}}$. Take $b = (1, \frac{1}{2}, \frac{1}{3}, \cdots) \in \ell^2$ and calculate $\| A^{-1}(\mathbf{b}) \|$, this norm is not bounded, so $A^{-1}(\mathbf{b}) \notin \ell^2$. This means that A^{-1} does not exist for every element out of the ℓ^2, so A^{-1} does not exist.

Sol-7: a. Solve $\lambda_1 f_1(t) + \lambda_2 f_2(t) + \lambda_3 f_3(t) = 0$ for every $t \in [-\pi, \pi]$. If it has to be zero for every t then certainly for some particular t's, for instance $t = 0, t = \frac{\pi}{2}, t = \pi$ and solve the linear equations.

b. Same idea as the solution of exercise **Ex- ii: 4**. Working in the Inner Product Space $L_2[-\pi, \pi]$. Project $\sin(\frac{t}{2})$ on the $\text{span}(f_1, f_2, f_3)$. The length of the difference of $\sin(\frac{t}{2})$ with the projection gives the minimum distance. This minimizing vector exists and is unique, so a_0, b_0, c_0 exist and are unique.

c. $(\sin(\frac{t}{2}) - a_0 - b_0 \cos(t) - c_0 \sin(t))$ is perpendicular to f_1, f_2, f_3,so the inner products have to be zero. This gives three linear equations which have to be solved to get the values of a_0, b_0 and c_0.
The solution is rather simple $a_0 = 0, b_0 = 0$ and $c_0 = \frac{8}{3\pi}$. Keep in mind the behaviour of the functions, if they are even or odd at the interval $[-\pi, \pi]$.

Sol-8: a. See figure **13.2**.

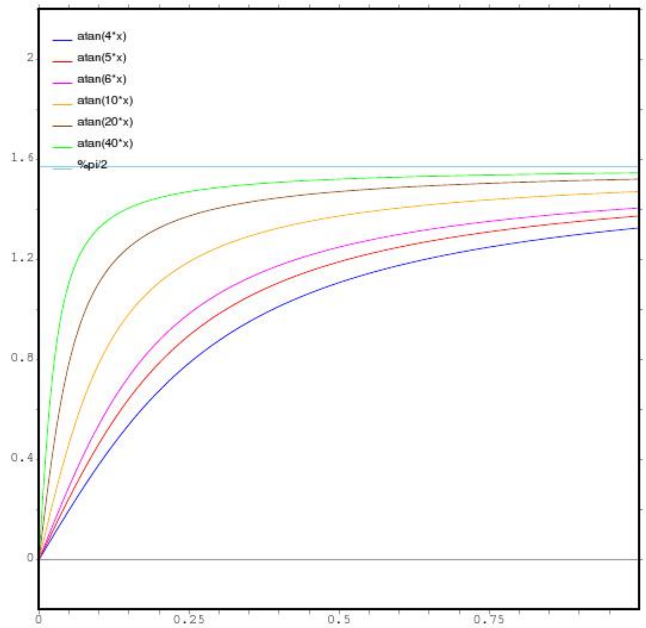


Figure 13.2 f_n certain values of n

b. Take $x = 0$ then $f_n(0) = 0$ for every $n \in \mathbb{N}$. Take $x > 0$ and fixed then $\lim_{n\to\infty} f_n(x) = \frac{\pi}{2}$, the pointwise limit f is defined by

$$f(x) = \begin{cases} 0 & \text{if } x = 0 \\ \dfrac{\pi}{2} & \text{if } 0 < x \leq 1, \end{cases}$$

it is clear that the function f makes a jump in $x = 0$, so f is not continuous at the interval $[0, 1]$.

c.

$$\lim_{n\to\infty}\left(\int_0^1 |\frac{\pi}{2} - \arctan{(n\,x)}| dx\right) \;=\; \lim_{n\to\infty}\left(\frac{\log{(1+n^2)} \;-\; 2\,n\,\arctan{(n)} \;+\; \pi\,n}{2\,n}\right) \;=\; 0$$

d. The sequence $\{f_n\}_{n\in\mathbb{N}}$ converges in the space $L_1[0,1]$ and every convergent sequence is a Cauchy sequence.

Sol-9: a. Take $\mathbf{x},\mathbf{y}\in\ell^2$ and $\alpha\in\mathbb{R}$ and check if

$$\begin{cases} f(\mathbf{x}+\mathbf{y}) \;=\; f(\mathbf{x}) \;+\; f(\mathbf{y}),\\ f(\alpha\mathbf{x}) \;=\; \alpha\, f(\mathbf{x}). \end{cases}$$

There are no particular problems.

b. The functional can be read as an inner product and the inequality of Cauchy-Schwarz is useful to show that the linear functional f is bounded.

$$|f(\mathbf{x})| \leq \sqrt{(\sum_{n=1}^{\infty}(\frac{3}{5})^{2(n-1)})}\;\parallel x\parallel = \sqrt{(\frac{1}{1-\frac{9}{25}})}\;\parallel x\parallel$$

A bounded linear functional is continuous.

Sol-10: a. Since $|x|\leq 1$, it follows that $\parallel (A\,f)\parallel^2 = \int_{-1}^{1}(x\,f(x))^2\,dx \;\leq\; \int_{-1}^{1}(f(x))^2\,dx \;=\parallel f\parallel^2$, so $(A\,f)\in L_2[-1,1]$.

b.

$$\langle Af, g\rangle \;=\; \int_{-1}^{1} x\,f(x)\,g(x)\,dx \;=\; \int_{-1}^{1} f(x)\,x\,g(x)\,dx \;=\; \langle f, A^{*}g\rangle,$$

so $(A^{*}g)(x) \;=\; x\,g(x) \;=\; (A\,g)(x)$, so A is self-adjoint.

Sol-11: a. $(T\,f_0)(t) \;=\; 0$ because $(1\;+\;t^2)\,f_0(t)$ is an odd function.

b. Take $f,g\in C[-1,1]$ and $\alpha\in\mathbb{R}$ and check if

$$\begin{cases} T(f+g) \;=\; T(f) \;+\; T(g),\\ T(\alpha f) \;=\; \alpha\,T(f). \end{cases}$$

There are no particular problems.

c. The Normed Space $C[0,1]$ is equiped with the sup-norm $\| \cdot \|_\infty$, so

$$|(T\,f)(t)| \leq 2\ \|\,(1\,+\,t^2)\,\|_\infty \|\,f\,\|_\infty = \,4\ \|\,f\,\|_\infty,$$

the length of the integration interval is ≤ 2. Hence, $\|\,(T\,f) \leq 4\ \|\,f\,\|_\infty$ and the linear operator T is bounded.

d. Solve the equation $(Tf) = 0$. If $f = 0$ is the only solution of the given equation then the operator T is invertible. But there is a solution $\neq 0$, see **part ii.11.a**, so T is not invertible.

Sol-12: a. Take $f, g \in C[0,1]$ and $\alpha \in \mathbb{R}$ and check if

$$\begin{cases} F(f+g) = F(f) + F(g), \\ F(\alpha f) = \alpha\, F(f). \end{cases}$$

There are no particular problems.

b. $|F(x)| \leq 1\ \|\,x\,\|_\infty$, may be too coarse. Also is valid $|F(x)| \leq \int_0^1 \tau\, d\tau\ \|\,x\,\|_\infty = \frac{1}{2}\ \|\,x\,\|_\infty$.

c. $F(1) = \frac{1}{2}$.

d. With **part ii.12.b** and **part ii.12.c** it follows that $\|\,F\,\| = \frac{1}{2}$.

Sol-13: a. Solve $\lambda_1\, x_1(t) + \lambda_2\, x_2(t) + \lambda_3\, x_3(t) = 0$ for every $t \in [-1,1]$. $\lambda_i = 0,\ i = 1,2,3$ is the only solution.

b. Use the method of Gramm-Schmidt: $e_1(t) = \sqrt{\frac{5}{2}}\, t^2$, $e_2(t) = \sqrt{\frac{3}{2}}\, t$ and $e_3(t) = \sqrt{\frac{9}{8}}\,(1 - \frac{2\sqrt{5}}{3\sqrt{2}}\, e_1(t))$. Make use of the fact that functions are even or odd.

Sol-14: a. A Hilbert Space and convergence. Let's try to show that the sequence (x_n) is a Cauchy sequence. Parallelogram identity: $\|\,x_n - x_m\,\|^2 + \|\,x_n + x_m\,\|^2 = 2\,(\|\,x_n\,\|^2 + \|\,x_m\,\|^2)$ and $(x_n + x_m) = 2\,(\frac{1}{2}x_n + \frac{1}{2}x_m)$. So $\|\,x_n - x_m\,\|^2 = 2\,(\frac{1}{2}x_n + \frac{1}{2}x_m) - 4\ \|\,\frac{1}{2}\,x_n + \frac{1}{2}\,x_m\,\|^2$.$M$ is convex so $\frac{1}{2}\,x_n + \frac{1}{2}\,x_m \in M$ and $4\ \|\,\frac{1}{2}\,x_n + \frac{1}{2}\,x_m\,\|^2 \geq 4\,d^2$.

Hence, $\| x_n - x_m \|^2 \leq 2\,(d^2 - \| x_n \|^2) + 2\,(d^2 - \| x_m \|^2) \rightarrow 0$ if $n, m \rightarrow \infty$.
The sequence (x_n) is a Cauchy sequence in a Hilbert Space H, so the sequence converge in H, M is closed. Every convergent sequence in M has it's limit in M, so the given sequence converges in M.

b. See **figure 3.6**, let $x = 0$, $\delta = d$ and draw some x_i converging to the closest point of M to the origin 0, the point y_0.

Sol-15: a. $a = (1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{n}, \cdots)$. $\int_1^\infty \frac{1}{t} dt$ does not exist, $\int_1^\infty \frac{1}{t^2} dt$ exists, so $a \in \ell^2$, but $a \notin \ell^1$.

b. Take an arbitrary $x \in \ell^1$, since $\| x \|_1 = \sum_{i=1}^\infty |x_i| < \infty$ there is some $K \in \mathbb{N}$ such that $|x_i| < 1$ for every $i > K$. If $|x_i| < 1$ then $|x_i|^2 < |x_i|$ and $\sum_{i=(K+1)}^\infty |x_i|^2 \leq \sum_{i=(K+1)}^\infty |x_i| < \infty$ since $x \in \ell^1$, so $x \in \ell^2$.

Sol-16: a. Use the good norm!

$$\| Tf \|^2 = \int_0^1 |(Tf)(x)|^2 dx = \int_0^1 \frac{1}{\sqrt{(4x)}}\, f^2(x)\, dx,$$

take $y = \sqrt{(x)}$ then $dy = \frac{1}{2\sqrt{x}}\, dx$ and

$$\| Tf \|^2 = \int_0^1 f^2(y) dy = \| f \|^2,$$

so $\| T \| = 1$.

b. The adjoint operator T^*, see the substitution used in **Sol- ii.16.a**,

$$\langle TF, g \rangle = \int_0^1 \frac{1}{\sqrt[4]{(4\,x)}}\, f(\sqrt{x}\,)\, g(x)\, dx = \int_0^1 f(y)\, \sqrt{2}\, \sqrt{y}\, g(y^2)\, dy = \langle f, T^* g \rangle,$$

so $T^* g(x) = \sqrt{2}\, \sqrt{x}\, g(x^2)$.

c. $\| T \| = \| T^* \|$.

d. $T^*((Tf)(x)) = T^*(\frac{1}{\sqrt[4]{(4\,x)}}\, f(\sqrt{x}\,)) = \sqrt{2}\, \sqrt{x}\, (\frac{1}{\sqrt{2}\,\sqrt{x}}\, f(\sqrt{x^2})) = f(x) = (If)(x)$.

Sol-17: a. Take $f, g \in C[-1,1]$ and $\alpha \in \mathbb{R}$ and check if

$$\begin{cases} h_n(f+g) \;=\; h_n(f) \,+\, h_n(g), \\ h_n(\alpha f) \;=\; \alpha\, h_n(f). \end{cases}$$

There are no particular problems.

b. It is a linear functional and not a function, use Cauchy-Schwarz

$$|h_n(f)| \;=\; |\int_{-1}^{+1} (\tau)^n\, f(\tau)\, d\tau| \;\leq\; (\int_{-1}^{+1} (\tau)^{2\,n} d\tau)^{\frac{1}{2}}\, (\int_{-1}^{+1} f^2(\tau)\, d\tau)^{\frac{1}{2}}$$

$$= (\frac{2}{2\,n\,+\,1})^{\frac{1}{2}} \;\parallel f \parallel .$$

c.

$$\lim_{n \to \infty} \parallel h_n \parallel \leq \frac{\sqrt{2}}{\sqrt{(n\,+\,1)}} \to 0$$

if $n \to \infty$, so $\lim_{n \to \infty} \parallel h_n \parallel =\; 0.$

Sol-18: a. $f(t) = \dfrac{1}{2}(f(t) - f(-t)) + \dfrac{1}{2}(f(t) + f(-t))$, the first part is odd ($g(-t) = -g(t)$) and the second part is even ($g(-t) = g(t)$). Can there be a function h which is even and odd? $h(t) = -h(-t) = -h(t) \Rightarrow h(t) = 0!$

b. If the given sequence is a Cauchy sequence, then it converges in the Banach space $(C[-1,1], \| \cdot \|_\infty)$. The limit should be a continuous function, but $\lim_{n \to \infty} f_n$ is not continuous, so the given sequence is not a Cauchy sequence.

Sol-19: a. Take $x, y \in C[-1,1])$ and $\alpha \in \mathbb{R}$ and let see that $f(x+y) = f(x) + f(y)$ and $f(\alpha x) = \alpha x$, not difficult.

b.

$$|f(x)| \;=\; |\int_{-1}^{0} x(t)\, dt \;-\; \int_{0}^{1} x(t)\, dt| \;\leq\; |\int_{-1}^{0} x(t)\, dt\,| + |\int_{0}^{1} x(t)\, dt|$$

$$\leq \parallel x \parallel_\infty \,+\, \parallel x \parallel_\infty =\; 2\; \parallel x \parallel_\infty,$$

so f is bounded and so continuous.

c. Take $x_n : [-1, +1] \to \mathbb{R}$, $n \in \mathbb{N}$,

$$x_n(t) = \begin{cases} 1 & \text{for } -1 \leq t \leq -\dfrac{1}{n} \\ -n\,x & \text{for } -\dfrac{1}{n} < t < \dfrac{1}{n} \\ -1 & \text{for } \dfrac{1}{n} \leq t \leq 1. \end{cases}$$

then $f(x_n) = 2 - \dfrac{1}{n}$. Therefore the number 2 can be approximated as close as possible, so

$$\| f \| = 2.$$

d. Even functions are a subset of $\mathcal{N}(f)$, but there are more functions belonging to $\mathcal{N}(f)$. It is difficult to describe $\mathcal{N}(f)$ otherwise then for all functions $x \in C[-1, 1]$ such that $\int_{-1}^{0} x(t)dt = \int_{0}^{1} x(t)dt$.

Sol-20: a. $S^2(u)(x) = S(u(1-x)) = u(1-(1-x)) = u(x) = I(u)(x)$, so $s^{-1} = S$.

b. $(S(u), v) = \int_0 1 u(1-x)v(x)\,dx = -\int_1^0 u(y)v(1-y)\,dy = (u, Sv)$, so $S^* = S$.

c. $\frac{1}{2}(I - S)\frac{1}{2}(I - S) = \frac{1}{4}(I - IS - SI + S^2) = \frac{1}{2}(I - S)$, so idempotent, evenso $\frac{1}{2}(I + S)$.
Extra information: the operators are idempotent and self-adjoint, so the operators are (orthogonal) projections and $\frac{1}{2}(I - S)\frac{1}{2}(I + S) = 0$!

d. Compute $\frac{1}{2}(I - S)(sin(x))$ and compute $\frac{1}{2}(I - S)(\alpha\, u(x) + \beta\, u(1-x))$. The last one gives $\frac{1}{2}(\alpha\, u(x) + \beta\, u(1-x) - (\alpha\, u(1-x) + \beta\, u(x))) = \frac{1}{2}((\alpha-\beta)u(x) - (\alpha-\beta)u(1-x))$. Do the same with the operator $\frac{1}{2}(I+S)$. The result is two linear equations, with the unknowns $u(x)$ and $u(1-x)$, compute $u(x)$ out of it. The linear equations become:

$$\begin{aligned} \sin(x) - \sin(1-x) &= (\alpha - \beta)(u(x) - u(1-x)) \\ \sin(x) + \sin(1-x) &= (\alpha + \beta)(u(x) + u(1-x)). \end{aligned}$$

(Divide the equations by $(\alpha - \beta)$ and $(\alpha + \beta)$!)

Sol-21: The question is if $x_n\alpha + y_n\beta$ converges in the $\| \cdot \|_\infty$-norm for $n \rightarrow \infty$?
And it is easily seen that

$$\| (x_n\alpha + y_n\beta) - (\alpha^2 + \beta^2) \|_\infty \leq \| (x_n - \alpha) \|_\infty | \alpha | + \| (y_n - \beta) \|_\infty | \beta | \rightarrow 0$$

for $n \rightarrow \infty$. It should be nice to write a proof which begins with: Given is some $\epsilon > 0 \cdots$.
Because $\lim_{n\rightarrow\infty} x_n = \alpha$, there exists a $N_1(\epsilon)$ such that for all $n > N_1(\epsilon)$, $| x_n - \alpha | < \frac{\epsilon}{2 | \alpha |}$. That gives that $\| (x_n - \alpha) \|_\infty | \alpha | < \frac{\epsilon}{2}$ for all $n > N_1(\epsilon)$.
Be careful with $\frac{\epsilon}{2 | \alpha |}$, if $\alpha = 0$ (or $\beta = 0$).
The sequence $(y_n)_{n\in\mathbb{N}}$ gives a $N_2(\epsilon)$. Take $N(\epsilon) = \max(N_1(\epsilon), N_2(\epsilon))$ and make clear that $| (x_n\alpha + y_n\beta) - (\alpha^2 + \beta^2) | < \epsilon$ for all $n > N(\epsilon)$.
So $\lim_{n\rightarrow\infty}(x_n\alpha + y_n\beta)$ exists and $(x_n\alpha + y_n\beta)_{n\in\mathbb{N}} \in c$.

Sol-22:

a. The easiest way is $\cos(2x) = 2\cos^2(x) - 1$. Another way is to formulate the problem $\alpha\, 1 + \beta \cos^2(x) + \gamma \cos(2x) = 0$ for every x. Fill in some nice values of x, for instance $x = 0$, $x = \frac{\pi}{2}$ and $x = \pi$, and let see that $\alpha = 0$, $\beta = 0$ and $\gamma = 0$ is not the only solution, so the given functions are linear dependent.

b. To solve the problem: $\beta \cos^2(x) + \gamma \cos(2x) = 0$ for every x. Take $x = \frac{\pi}{2}$ and there follows that $\gamma = 0$ and with $x = 0$ follows that $\beta = 0$. So $\beta = 0$ and $\gamma = 0$ is the only solution of the formulated problem, so the functions f_2 and f_3 are linear independent.

Sol-23:

a. Linearity is no problem.

b. Boundednes is also easy, if the triangle-inequality is used

$$\begin{aligned} &\| A(a_1, a_2, a_3, \cdots) \| \leq \\ &\| (a_1, a_2, a_3, \cdots) \| + \| (a_3, a_4, a_5, \cdots) \| \leq \\ 2 &\| (a_1, a_2, a_3, \cdots) \| \end{aligned}$$

c. The null space of A is, in first instance, given by the span S, with $S = \mathrm{span}((1,0,-1,0,1,0,-1,\cdots), (0,1,0,-1,0,1,0,\cdots))$.

Solve: $A(a_1,\, a_2,\, a_3,\, \cdots) = (0,0,0,0,\cdots)$.
But be careful: $S \not\subseteq \ell^2$, so $N(A) = \{0\}$ with respect to the domain of the operator A and that is ℓ^2.

Sol-24: a. Just control the conditions given in **Definition 3.29**. The most difficult one is may be condition **3.29**(IP **1**). If $(p,p) = 0$ then $p(-1)p(-1)+p(0)p(0)+p(1)p(1) = 0$ and this means that $p(-1) = 0$, $p(0) = 0$ and $p(1) = 0$. If $p(x) = \alpha\, 1 + \beta\, x + \gamma\, x^2$, p has at most degree 2, then with $x = 0 \rightarrow \alpha = 0$ and with $x = 1$, $x = -1$ there follows that $\beta = 0$ and $\gamma = 0$, so $p(x) = 0$ for every x.

b. Just calculate (q_1, q_2), (q_1, q_3) and (q_2, q_3) and control if there comes 0 out of it.

c. $\| q_1 \| = 1$, $\| q_2 \| = 2$, $\| q_3 \| = 2$.

14 List of figures

15 Index

16 References

1 Aliprantis, C. (2006). *Infinite Dimensional Analysis (A Hitchhiker's Guide), pdf-file*. Springer.

2 Aliprantis, C. and Burkinshaw, O. (1998). *Principles of Real Analysis , pdf-file*. Academic Press.

3 Bonsall (1973). *Complete Normed Algebras.*.

4 Bouziad, A. and Clabrix, J. (1993). *Théorie de la Mesure et de l'Integration, djvu-file*. j'Universite the Rouen.

5 Carothers, N. (1993). *Théorie de la Mesure et de l'Integration, pdf-file.*.

6 Dieudonné, J. (1981). *History of Functional Analysis, djvu-file*. North-Holland.

7 Dunford, N. and Schwartz, T. (1958-71). *Linear Operators (3 parts), djvu-file*. Wiley.

8 Fabian, M. (2001). *Functional Analysis and Infinite-Dimensional Geometry*. Springer.

9 Griffel, D. (1981). *Applied Functional Analysis*. Ellis Horwood.

10 Herrlich, H. (2006). *Axiom of Choice, pdf-file*. Springer.

11 Jain, P., Ahuja, O. and K.Ahmad (1996). *Functional Analysis*. New Age Publishers.

12 Johnson, W. and Lindenstrauss, J. (2003). *Handbook of The Geometry of Banach Spaces, Volume 2*. North-Holland.

13 Johnson, W. and Lindenstrauss, J. (2001). *Handbook of The Geometry of Banach Spaces, Volume 1*. North-Holland.

14 Körner, T. (2004). *A Companion to Analysis, pdf-file*. American Mathematical Society.

15 Kolmogorv, A. and Fomin, S. (1961). *Measure, Lebesgue Integrals and Hilbert Space, djvu-file*. Academic Press.

16 Kreyszig, E. (1978). *Introductory Functional Analysis with Applications, djvu-file*. Wiley.

17 Kuttler (2009). *Topics in Analysis.*.

18 Limaye, B. (2008). *Functional Analysis (sec.ed.)*. New Age Publishers.

19 Lindstrom, J. (2008). *On the origin and early history of functional analysis, pdf-file.*.

20 Müller, V. (2007). *Spectral Theor of Linear Operators (sec. ed.)*. Birkhäuser.

21 Megginson, R. (1998). *An Introduction to Banac Space Theory*. Springer.

22 Melrose, R. (2004). *Lecture Notes for 18.155, fall 2004, pdf-file*. MIT.

23 Moller, J. (-). *General Topology, pdf-file*. M.I., K∅benhaven.

24 Morrison, T. (2001). *Functional Analysis, Introduction to Banach Space Theory, pdf-file*. Wiley.

25 Pietsch, A. (2007). *History of Banach Spaces and Linear Operators*. Birkhäuser.

26 Pugachev, V. and Sinitsyn, I. (1999). *Lectures on Functional Analysis and Applications*. World Scientific.

27 Rana, I. (2004). *An Introduction to Measure And Integration*. Alpha Science International.

28 Rao, K. P. S. B. and Rao, M. B. (1983). *Theory of charges: a study of finitely additive measures, pdf-file*. Academic Press.

29 Royden, H. (1988). *Real Analysis, third edition*. Macmillan Publishing Company.

30 Saxe, K. (2002). *Beginning Functional Analysis, djvu-file*. Springer.

31 Searcoid, M. (2007). *Metric Spaces*. Springer.

32 Srivastava, S. (1998). *A Course on Borel Sets, pdf-file*. Springer.

33 Sutherland, W. (1975). *Introduction to metric and topological spaces*. Oxford University Press.

34 Swartz, C. (1994). *Measure, Integration and Function Spaces*. World Scientific.

35 Taylor, A. (1958). *Introduction to Functional Analysis, djvu-file*. Wiley.

36 Yeh, J. (2006). *Real Analysis, Theory of Measure and Integration, djvu-file*. World Scientific.

17 Busy With

Here is written, where I'm busy with the writing of these lecture notes:

a. Write about: compact operator $\Rightarrow$ completely continuous operator. In general completely continuous operator $\not\Rightarrow$ compact operator, see 7.20, (start: 140117, oke: ??).

b. Writing about spectral properties of bounded linear operators, see Chapter 12.5,
(start: 131230, oke: ??).

c. Writing weak and weak* convergence, see 4.9.

d. Writing completely continous and compact operators, see 7.8.

e. Writing example of spectra, shift-operators, see 12.7.1.

f. Writing relation between residual spectrum T and point spectrum T^*.

What to do?

a. Have forgotten, see 5.2.4, 5.2.5, 5.2.6?? May be more forgotten, have to search for: todo,
(start: 14017, oke: ??).

b. Uniform Boundedness Theorem.

c. Totally bounded $\leftrightarrow$ precompact or relatively compact, are there differences or not?

What has been done?

a. Dual space of ℓ^2 and ℓ^p, see 5.15, comment: $1 < p < \infty$ has been done, so also $p = 2$,
$p = 1$ has already been done,
(start: 140117, oke: 140120).

b. Writing about Schur's property, see 4.9, see 4.19, see 4.18,
(start: 140117, oke: 140118).

c. Complete Bounded Inverse Theorem, see SubSection 7.7.4,
(start 140101, oke: 140101).

d. Proved is $\sigma(T)$ is closed, or $\rho(T)$ is open, see **Theorem 12.2**,
(start:131230, oke: 140101).

e. Closedness does not imply boundedness of a linear operator.
Boundedness does not imply closededness of a linear operator.
Have no idea what I have written in **Section 7.7.2**,
(start: 131230, oke: 131231).

f. Had no idea what I had written in **Section 7.7.2**, but now some things rewritten,
(start: 131230, oke: 131231).

g. Worked at closed operators, see Theorem ??,
(start: 131230, oke: 131230).

h. Writing the theorem and proof of INV is continuous, see **7.13**,
(start: 131227, oke: 131227).

i. Busy with proof of **7.12**, last step has to be done,
(oke: 131226).

j. Solution **Sol- ii: 7**, oke.

k. Writing relations between nullspace and range operator at Hilbert space, see **7.16**. $R^{\perp\perp} = \overline{R}$?? Oke.

In doubt about?

a. What have I done at the beginning of **Chapter 7.7** and in **Section 7.7.2**?
(start:131230, oke: ??).

Ideas?

a. Hilbert-Schmidt operators? (start: 140118, oke: ??).