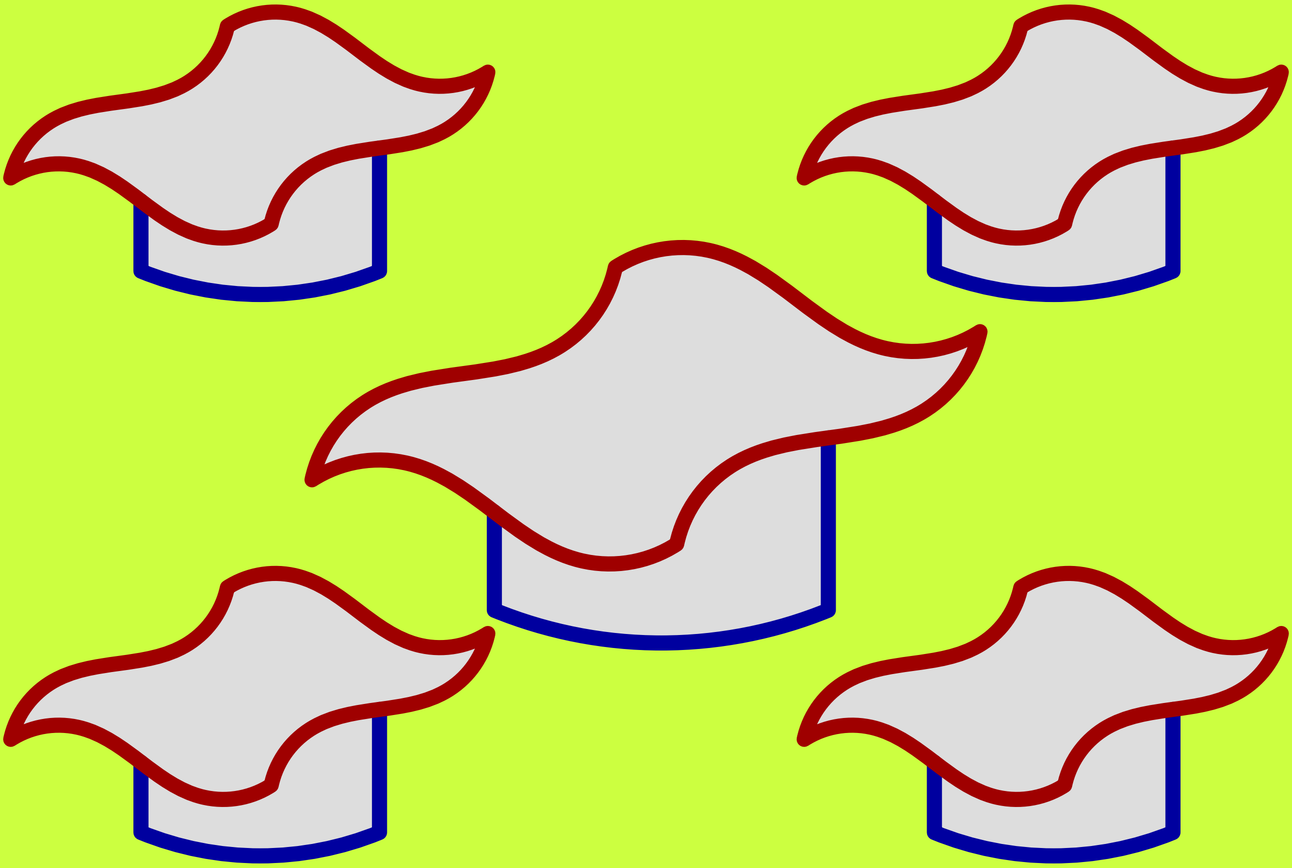
Own Lecture Notes Functional Analysis



R.R. van Hassel
Helmond
date: 6 November 2009
email: r.hassel1@chello.nl

Contents

1 Preface

You are reading some lecture notes of an introduction to Functional Analysis.
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!
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:

   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

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

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

There will be very much spoken about spaces. Most of the time, there is something to measure the distance between two elements in these spaces, a so-called metric. A good definition of spaces with a metric will be given in section 3.5.

2.1 Mappings

If X and Y are sets and $A \subseteq X$ any subset. 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 sets X and Y.
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 $D(T)$.
The set of all images of T is called the range of T, denoted by $R(T)$,

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

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

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

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

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

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.
If $T : D(T) \to Y$, the mapping is said to be onto if $R(T) = Y$, or

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

note that $T : D(T) \to R(T)$ is always onto.

If T is onto, it is also called surjective , or T is an surjection.
If $T : D(T) \to Y$ 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. So $T^{-1} : Y \to D(T)$, for every $y \in Y$ there exists an unique $x \in D(T)$ such that $T(x) = y$.

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 $\mathbb{R}^n$. This mapping is called a norm and written by $\| \,.\, \|$.
A subset $A \subset \mathbb{R}^n$ is bounded, if there exists an $a \in A$ and a $K \in \mathbb{R}$ such that

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

for all $x \in A$.
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

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

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.

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 $n > N(\epsilon)$, $\| \lambda_n - \lambda \| < \epsilon$.
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 $m, n > N(\epsilon)$, $\| \lambda_m - \lambda_n \| < \epsilon$. 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.

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 $\mathbb{Q} = \{\frac{p}{q} | p, q \in \mathbb{Z}\}$, the real numbers $\mathbb{R}$ and the complex numbers $\mathbb{C} = \{a + \mathrm{i}\, b | a, b \in \mathbb{R} \text{ and } \mathrm{i}^2 = -1\}$.
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 $\overline{M}$, is the union of M with all its accumulation points.

Theorem 2.5.1 $x \in \overline{M}$ *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 The proof exists out of two parts.

($\Rightarrow$) If $x \in \overline{M}$ 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.

□

Theorem 2.5.2 M is closed *if and only if* the limit of every convergent sequence in M is an element of M.

Proof 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 2.5.3 *M is closed if and only if $M = \overline{\mathrm{M}}$.*

Proof 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 $\overline{\mathrm{M}} \setminus M = \emptyset$.

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

2.6 Dense subset

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 $\lim_{n\to\infty} y_n = x$. There can also be said that Y is a dense subset of X if $\overline{\mathrm{Y}} = X$.
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 $\overline{\text{span}}\{y_n \in M \,|\, n \in \mathbb{N}\} = X$.
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 compactnes 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, carefuleness 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

$$M \subset \cup_{\alpha \in IS} S_\alpha \tag{2.9}$$

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

Explanation of Example 2.8.1
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!

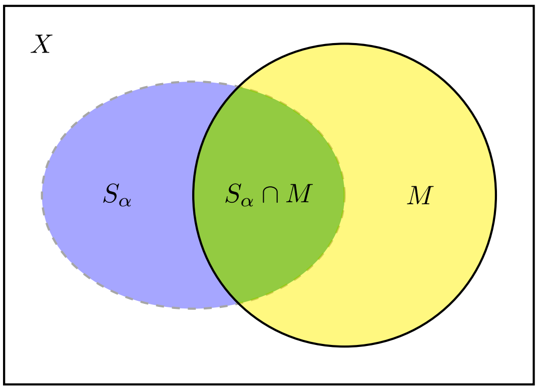


Figure 2.1 Compactness and open sets

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

Proof 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 2.5.1. The subset M is compact in the metric space (X, d). 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. □

2.9 Supremum and infimum

Axiom: 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 2.9.1 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 and uniformly 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 $x \in B_\delta(x_0) = \{y \in X | \parallel y - x_0 \parallel_1 < \delta\}$, there is satisfied that $T(x) \in B_\epsilon(T(x_0))$, this means that $\parallel T(x) - T(x_0) \parallel_2 < \epsilon$, 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 $\parallel x - y \parallel_1 < \delta(\epsilon)$, there is satisfied that $\parallel T(x) - T(y) \parallel_2 < \epsilon$.
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.

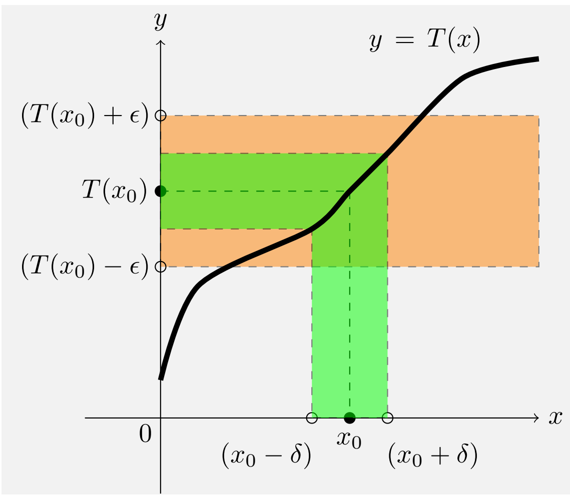


Figure 2.2 Continuous map

Theorem 2.10.1 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 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 $T(x_n) \rightarrow T(x_0)$.

($\Leftarrow$) 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 $T(x_n) \rightarrow T(x_0)$. □

2.11 Continuity and compactness

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

Theorem 2.11.1 If $T : X \rightarrow Y$ is a continuous map and $V \subset X$ is compact *then* is $T(V) \subset Y$ compact.

Proof

($\Rightarrow$) 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 2.11.2 Let (X, d_1) and (Y, d_2) be metric spaces and $T : X \rightarrow Y$ a continuous mapping then is the image $T(V)$, of a compact subset $V \subset X$, closed and bounded.

Proof The image $T(V)$ is compact, see theorem 2.11.1 and a compact subset of a metric space is closed and bounded, see theorem 2.8.1. □

Definition 2.11.1 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 2.11.1 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 2.11.3 Let (X, d_1) and (Y, d_2) be two Compact Metric Spaces, then every continuous function $f : X \rightarrow Y$ is uniformly continuous.

Proof 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 $\{w_{n_k}\}$, ($v_{n_k} \rightarrow v_0$ and $w_{n_k} \rightarrow w_0$), so

$$d_1(v_{n_k}, w_{n_k}) < \frac{1}{n_k} \text{ and } d_2(f(v_{n_k}), f(w_{n_k})) \geq \epsilon 0. \qquad (2.10)$$

The sequences $\{v_{n_k}\}$ and $\{w_{n_k}\}$ converge to the same point and since f is continuous, statement 2.10 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 $\sup_{t \in [a,b]} \mid f(t) \mid$

$$\| f \|_\infty = \sup_{t \in [a,b]} \mid f(t) \mid \qquad (2.11)$$

with $f \in C[a,b]$. The fact that $[a,b]$ is a compact set of $\mathbb{R}$, means that the $\sup_{t\in[a,b]} \mid f(t) \mid = \max_{t\in[a,b]} \mid f(t) \mid$.
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 $\mid f(x) - f_n(x) \mid < \epsilon$ 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 2.12.1 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 $\{f_n\}_{n\in\mathbb{N}}$[3] converges uniform to the function f, if for every $\epsilon > 0$, there exist a $N(\epsilon)$ such that for every $n > N(\epsilon)$ $\parallel f - f_n \parallel_\infty < \epsilon$. Note that $N(\epsilon)$ does not depend of x anymore. So $\mid f(x) - f_n(x) \mid < \epsilon$ for all $n > N(\epsilon)$ and for all $x \in [a,b]$.

Theorem 2.12.1 If the sequence of functions $\{f_n\}_{n\in\mathbb{N}}$[3] converges uniform to the function f on the interval $[a,b]$, then the function f is continuous on $[a,b]$.

Proof Let $\epsilon > 0$ be given, and there is proven 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 $n > N(\epsilon)$,
$\mid f(s) - f_n(s) \mid < \frac{\epsilon}{3}$ and $\mid f_n(x) - f(x) \mid < \frac{\epsilon}{3}$ ($N(\epsilon)$ 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 $\mid s - x \mid < \delta(\epsilon)$, $\mid f_n(s) - f_n(x) \mid < \frac{\epsilon}{3}$. So the function f is continous in x, because

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

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

[3] $f_n \in C[a,b]$, $n \in \mathbb{N}$.

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 2.13.1 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 2.13.2 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 2.13.3 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 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 2.14.1 Let $\{x_n\}$ be real sequence. The limit superior of $\{x_n\}$ is the extended real number

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

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

Definition 2.14.2 Let $\{x_n\}$ be real sequence. The limit inferior of $\{x_n\}$ is the extended real number

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

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 $f(x) = (1 + 4\exp(-x/10))\sin(5x)$ in figure 2.3.

The definitions of lim sup and lim inf, given the Definitions 2.14.1 and 2.14.2, 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.

2.15 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 $E_k \supset E_{k+1}$ 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.

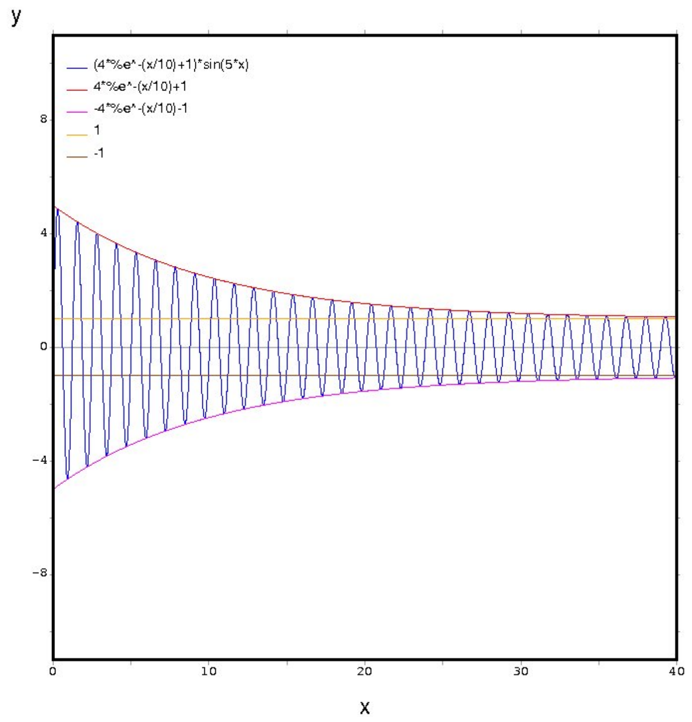


Figure 2.3 Illustration of lim sup and lim inf.

Definition 2.15.1 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 $\emptyset$. □

If $E_k \uparrow$ then $\lim_{k\to\infty} E_k = \emptyset \Leftrightarrow E_k = \emptyset$ for every $k \in \mathbb{N}$.
If $E_k \downarrow$ then $\lim_{k\to\infty} E_k = \emptyset$ can be the case, even if $E_k \neq \emptyset$ 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 2.15.2 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 $\emptyset$. □

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

1. $\limsup_{k\to\infty} E_k = \{s \in S \mid s \in E_k$ 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 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 2.15.1: 2 and 1. □

Example 2.15.1 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 2.15.3 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 2.15.2 It is clear that the sequence of subsets defined in Example 2.15.1 has no limit, because $\limsup_{k \to \infty} E_k \neq \liminf_{k \to \infty} E_k$.
But the subsequence $(E_{2k} \mid k \in \mathbb{N})$ is an increasing sequence with

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

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

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

□

2.16 Essential supremum and essential infimum

Busy with limit superior and limit inferior, see the Sections 2.14 and 2.15, 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 147.
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 147.
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 2.16.1 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

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

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

Example 2.16.2 This example is based on Example 2.16.1. 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 Vector 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 24.

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.

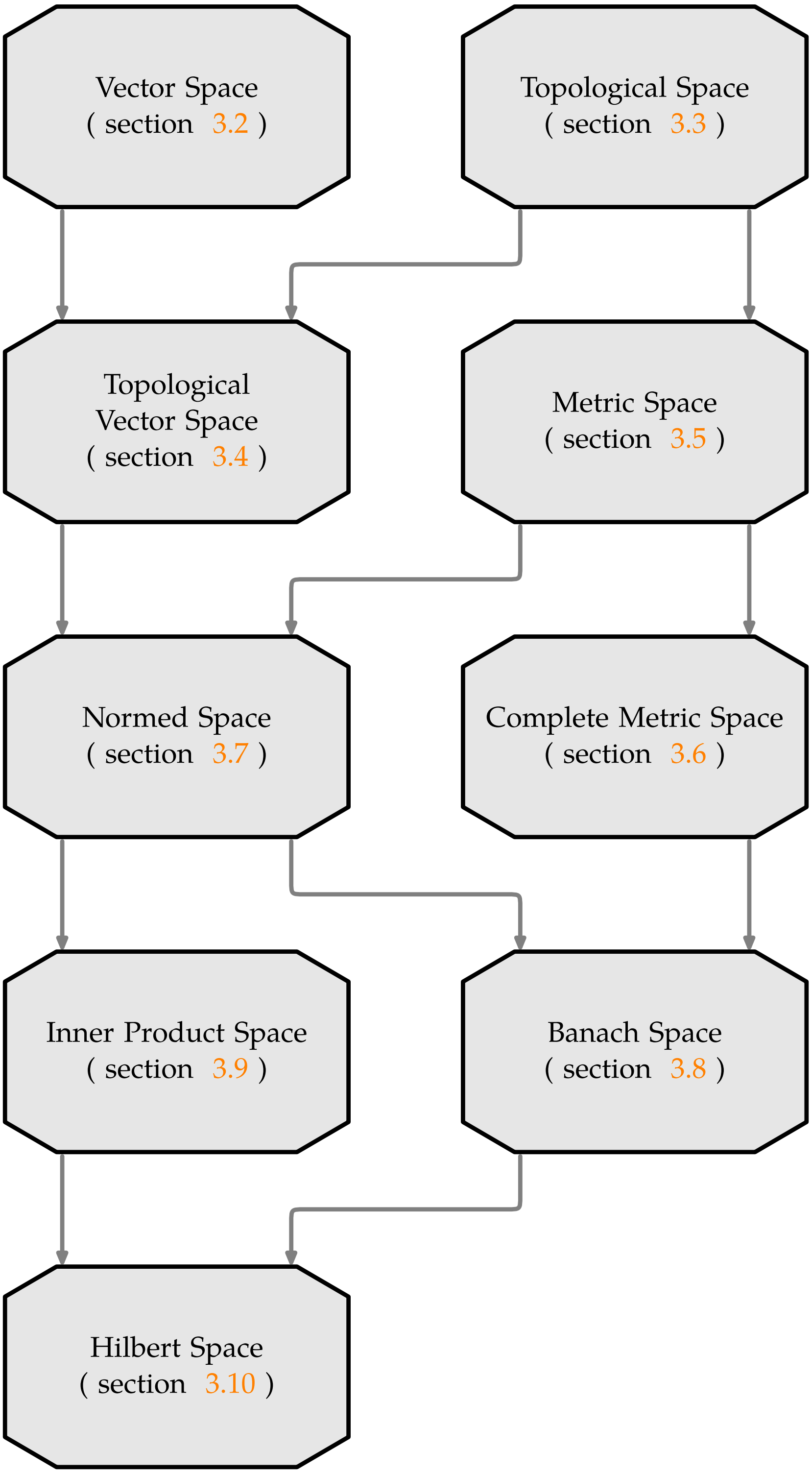


Figure 3.1 A flowchart of spaces.

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 24.

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.

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 3.2.1 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.

3.2.1 Linear Subspaces

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

Definition 3.2.2 Let $\emptyset \neq Y \subseteq X$, 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 $\emptyset \neq Y \subseteq X$ could be a linear subspace of the Vector Space X, the following theorem is very useful.

Theorem 3.2.1 If $\emptyset \neq Y \subseteq X$ is a linear subspace of the Vector Space X then $0 \in Y$.

Proof 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 3.2.3 Let U and V be two linear subspaces of a Vector Space X. The sum $U + V$ is defined by

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

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 $U \cap V = \emptyset$ 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 3.2.4 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 $U \cap V = \emptyset$. Every $x \in X$ has an unique representation

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

If $X = U \oplus V$ then V is callled the algebraic complement of U and vice versa.
There will be very much worked with products of Vector Spaces .

Definition 3.2.5 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. □

3.2.2 Quotient Spaces

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

Definition 3.2.6 The coset of an element $x \in V$ with respect to W is defined by the set

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

□

Any two cosets are either disjoint (distinct) or identical and the distinct cosets form a partition of V. The quotient space is written by

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

Definition 3.2.7 The linear operations on $V \,/\, W$ are defined by

QS 1: $(x + W) + (y + W) = (x + y) + W$,

QS 2: $\alpha (x + W) = \alpha x + W$. □

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

3.2.3 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 3.2.8 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 to 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 3.2.9 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 3.2.10 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 3.2.11 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 $X = \infty$. □

3.2.4 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

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

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 algebraïc basis, or Hamel basis.

3.2.5 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 3.2.9 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 3.2.1 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 3.2.3. 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 3.2.2 Define the set of functions $\mathrm{Id}_r : \mathbb{R} \to \{0, 1\}$ by

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

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 3.2.12 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 3.2.2 Every Vector Space $X \neq 0$ has a Hamel basis H.

Proof 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 3.2.1 and 3.2.2.

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. $\square$

In the section about Normed Spaces, the definition of an infinite sequence is given, see definition 3.7.4. An infinite sequence will be seen as the limit of finite sequences, if possible.

3.3 Topological Spaces

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 3.3.1 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: $A, \emptyset \in \Psi$,

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 collecion 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.

3.4 Topological Vector Spaces

Definition 3.4.1 A Topological Vector Space space $TVS = \{VS, \Psi\}$ consists of a non-empty vectorspace VS together with a topology Ψ. □

3.5 Metric Spaces

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

Definition 3.5.1 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: $d(x, y) = 0 \Longleftrightarrow x = y$,

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

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

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

Definition 3.5.2 The set $\{x \mid x \in X, d(x, x_0) < r\}$ is called an open ball of radius r around the point x_0 and denoted by $B_r(x_0)$.
A closed ball of radius r around the point x_0 is defined and denoted by $\overline{B_r(x_0)} = \{x \mid x \in X, d(x, x_0) \leq r\}$. □

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

Definition 3.5.3 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 $B_r(x_0) \subset G$.
The set of all interior points of G is called the interior of G and is denoted by G°. □

Theorem 3.5.1 The distance function $d(\cdot, \cdot) : X \times X \rightarrow \mathbb{R}$ is continuous.

Proof 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

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

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 3.5.4
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

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

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

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

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

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

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

Remark 3.5.1 The distance function $\text{dist}(\cdot, A)$ is most of the time denoted by $d(\cdot, A)$. □

Theorem 3.5.2 The distance function $d(\cdot, A) : X \to \mathbb{R}$, defined in 3.5.4 is continuous.

Proof 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

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

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

3.6 Complete Metric Spaces

Definition 3.6.1 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 3.6.1 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

($\Rightarrow$) Take some $x \in \overline{M}$. Then there exists a convergent sequence $\{x_n\}$ to x, see theorem 2.5.1. 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 3.6.2 For $1 \leq p \leq \infty$, the metric space ℓ^p is complete.

Proof

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

$n \to \infty$.
For all n, $m > N(\epsilon)$

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

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

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

for $k = 1, 2, \cdots$. Now letting $k \to \infty$ 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 5.2.8 b, there follows that $x = x_n + (x - x_n) \in \ell^p$.
Inequality 3.2 implies that $x_n \to x$ for $n \to \infty$.
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 3.7.1 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 3.7.2 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 3.7.1. □

Remark 3.7.1 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:

$$\begin{aligned} p(0) &= 0, \\ |p(x) - p(y)| &\leq p(x - y). \end{aligned}$$

□

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 3.7.1 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 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 3.7.3 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 $x \in X$

$$m \parallel x \parallel_0 \leq \parallel x \parallel_1 \leq M \parallel x \parallel_0 .$$

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

Proof

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 5.31, 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 3.2.5. In a normed space, the norm can be used to overcome some problems.
Every Vector Space has a Hamel basis, see Theorem 3.2.2, 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 3.7.4 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 3.7.3 Every Normed Space $(X, \|\cdot\|)$ has a completion $(\hat{X}, \|\cdot\|_1)$.

Proof 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 to 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[4].

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

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.

[4] For isometric isomorphisms, see page 109

Example 3.7.1 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 3.7.1
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 3.7.2 Look at the space c_{00}, see example 3.7.1. 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 3.7.2
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

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

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 3.2.3.
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 3.7.5 A total set (or fundamental set) in a Normed Space X is a subset $M \subset X$ whose span 3.2.9 is dense in X. □

Remark 3.7.2 According the definition:

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

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

3.8 Banach Spaces

Definition 3.8.1 If every Cauchy row in a Normed Space NS_1 converges to an element of that same space NS_1 then that Normed Space NS_1 is called complete in the metric induced by the norm.
A complete normed space is called a Banach Space. □

Theorem 3.8.1 Let Y be a subspace of a Banach Space BS. Then,
Y is closed *if and only if* Y is complete.

Proof

($\Rightarrow$) 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 $x_n \to x$. 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 3.9.1 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[5] 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.

[5] If $z = (a + ib) \in \mathbb{C}$, with $a, b \in \mathbb{R}$ and $i^2 = -1$, then $\bar{z} = a - ib$. Sometimes j is used instead of i.

Theorem 3.9.1 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 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 3.9.2 If $(X, (.,.))$ is an Inner Product Space, then is the inner product $(.,.) : X \times X \rightarrow \mathbb{K}$ continuous. This means that if

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

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

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

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

So the norm and the inner product are continuous, see theorem 3.7.1 and theorem 3.9.2.

3.9.2 Parallelogram Identity and Polarization Identity

An important equality is the parallelogram equality, see figure 3.2.

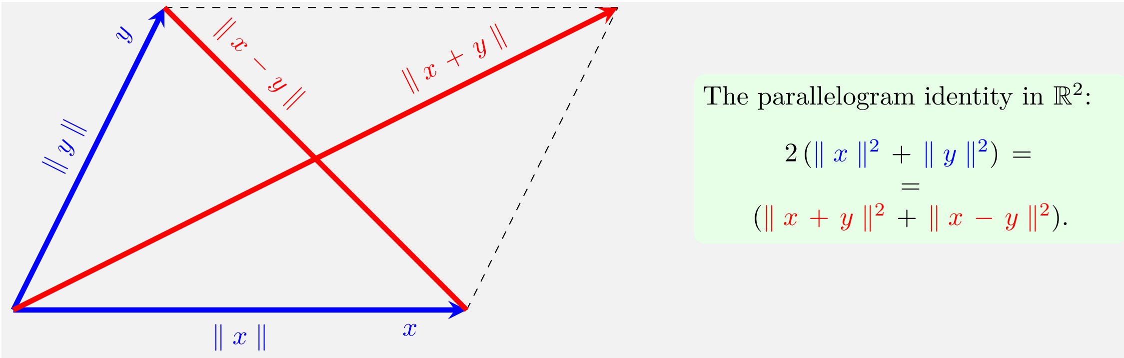


Figure 3.2 Parallelogram Identity

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 3.9.3 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

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

$$\| x + y \|^2 = (x + y, x + y)$$
$$= \| x \|^2 + (x, y) + (y, x) + \| y \|^2 = \| x \|^2 + (x, y) + \overline{(x, y)} + \| y \|^2 .$$ [5]

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.

(⇐) 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

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

The polarization identity and the continuity of the norm ensures that $|\,((q_n \;-\; \alpha)\, x,\; y)\,| \rightarrow 0$ 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

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

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

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

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 3.9.2 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 3.9.3 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 3.2.3, 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 3.9.4 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 3.9.4

a. If A be an non-empty 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 an non-empty subset of X then $A \subset A^{\perp\perp}$.
d. If A, B are non-empty subsets of X and $A \subset B$, then $A^{\perp} \supset B^{\perp}$.

Proof

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 $x \in A \cap A^{\perp} \neq \emptyset$ 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 $B^{\perp} \subset A^{\perp}$. □

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 3.2.3.

Definition 3.9.5 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 3.9.1 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 3.9.5 Orthogonal systems are linear independent systems.

Proof 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

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

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

Theorem 3.9.6 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

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

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

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

for every $x \neq y$ in S.

Proof

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 3.9.7 (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 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 $(X, (\cdot, \cdot))$, 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 3.9.1.

Theorem 3.9.8 Let the elements of the set $S = \{x_i \mid i \in \mathbb{N}\}$ be a linear independent set of the Inner Product Spaces $(X, (\cdot, \cdot))$. Then there exists an orthonormal set $ONS = \{e_i \mid i \in \mathbb{N}\}$ of the Inner Product Spaces $(X, (\cdot, \cdot))$, such that

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

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

Proof 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 $\text{span}(x_1) = \text{span}(y_1)$. 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

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

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 $\text{span}(y_1, \cdots, y_{(n-1)}) = \text{span}(x_1, \cdots, x_{(n-1)})$. 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,

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

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 $\text{span}(y_1, \cdots, y_n) = \text{span}(y_1, \cdots, y_{(n-1)}, x_n) = \text{span}(x_1, \cdots, x_{(n-1)}, x_n)$.
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 3.10.1 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 46. The next theorem makes clear why the word *pre-* is written before Hilbert. For the definition of an isometric isomorphism see page 109.

Theorem 3.10.1 If X is an Inner Product Space, then there exists a Hilbert Space H and an isometric isomorphism $T : X \to W$, where W is a dense subspace of H. The Hilbert Space H is unique except for isometric isomorphisms.

Proof The Inner Product Space with its inner product is a Normed Space. So there exists a Banach Space H and an isometry $T : X \to W$ onto a subspace of H, which is

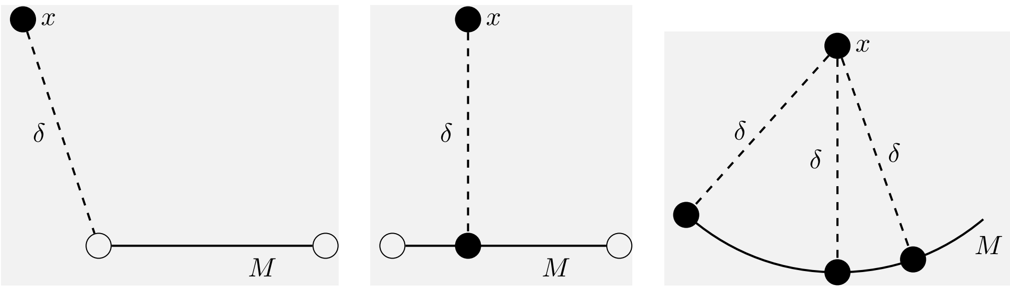
dense in H, see theorem 3.7.3 and the proof of the mentioned theorem.
The problem is the inner product. But with the help of the continuity of the inner product, see theorem 3.9.2, there can be defined an inner product on H by

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

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 3.7.3. 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 3.7.3 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 3.5.4.
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 $\mathrm{dist}(x, M) = \| x - y \|$. And if there exists such a $y \in M$, the question becomes if this y is unique? See the figures 3.3 for several complications which can occur.



No solution. An unique solution. Infinitely many solutions.

Figure 3.3 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 3.10.2 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 3.10.2 Let X be an Inner Product Space and $M \neq \emptyset$ 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 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 3.10.3 See theorem 3.10.2, 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

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

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

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 $\lambda \rightarrow 0$ and the desired result is obtained. □

Theorem 3.10.3 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 3.4. Several possibilities of such a hyperplane are drawn.

If there is only an unique hyperplane than the direction of $(x - y_0)$ is perpendicular to that plane, see figure 3.5.

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

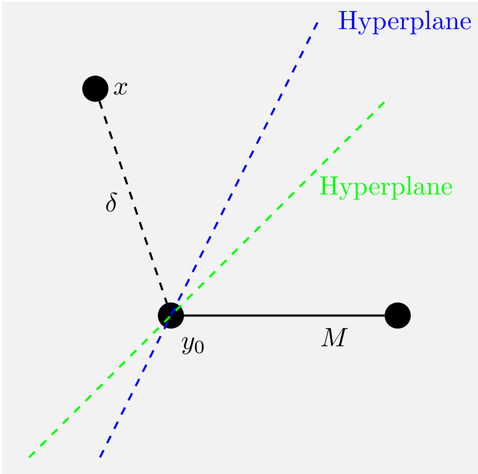


Figure 3.4 Some hyperplanes through y_0.

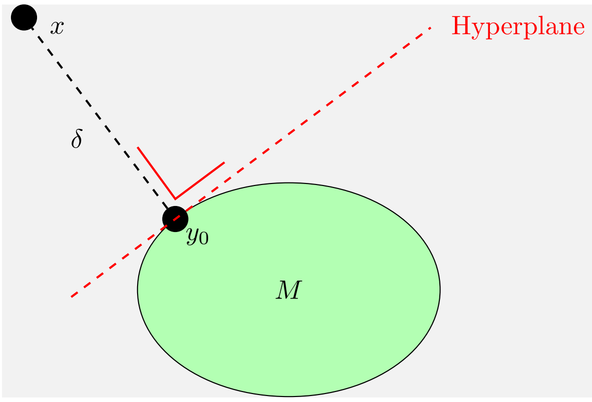


Figure 3.5 Unique hyperplane through y_0.

be calculated. The next theorem generalizes the above mentioned fact. Read theorem 3.10.2 very well, there is spoken about a non-empty convex subset, in the next theorem is spoken about a linear subspace.

Theorem 3.10.4 See theorem 3.10.2, but now with M a complete subspace of X, then $z = x - y_0$ is orthogonal to M.

Proof

A subspace is convex, that is easy to verify. So theorem 3.10.2 gives the existence of an element $y_0 \in M$, such that $\text{dist}(x, M) = \| x - y_0 \| = \delta$.
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

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

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 3.10.4, 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 3.10.5 If M is closed subspace of a Hilbert Space H. Then

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

Proof Since M is a closed subspace of H, M is also a complete subspace of H, see theorem 3.6.1. Let $x \in H$, theorem 3.10.4 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 3.10.6 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

($\Rightarrow$) 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 $M^{\perp} = \{0\}$.

(⇐) 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 $\overline{\text{span}(M)}$ is a subspace of H. With theorem 3.10.5 is obtained that $\overline{\text{span}(M)} = H$, so M is total in H. □

Remark 3.10.1 In Inner Product Spaces theorem 3.10.6 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! □

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 3.10.3 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 3.10.2 An orthogonal base is sometimes also called a complete orthogonal system . Be careful, the word "complete" has nothing to do with the topological concept: completeness. □

Theorem 3.10.7 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 There exists a $x \neq 0$ in H. The set, which contains only $\dfrac{x}{\| x \|}$ 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 3.10.8 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 The proof is splitted up into several steps.

1. First will be proved the inequality of Bessel. In the proof given in theorem 3.9.7 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

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

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

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

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

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

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

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

□

3.10.3 Representation of bounded linear functionals

In the chapter about Dual Spaces, see chapter 6, 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 6.4.1 and for the vectorspace ℓ^1 see 6.6.1. 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 3.10.3 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 3.10.9 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 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 $x = (x - \frac{f(x)}{f(z_1)} z_1) + \frac{f(x)}{f(z_1)} z_1 = x_0 + x_1$. 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, $(x, z_1) = \frac{f(x)}{f(z_1)} (z_1, z_1) = f(x) \frac{\parallel z_1 \parallel^2}{f(z_1)}$. Take $a = \frac{\overline{f(z_1)}}{\parallel z_1 \parallel^2} z_1$ 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 $\parallel b - a \parallel^2 = 0$ 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 3.10.4 Let X and Y be two Vector Spaces over the same field $\mathbb{K}$. A mapping

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

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 3.10.5 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 3.10.10 Let H_1 and H_2 be Hibert Spaces over the field $\mathbb{K}$ and

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

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

$$\parallel S \parallel = \parallel h \parallel .$$

Proof 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 3.10.9 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 $S : H_1 \rightarrow H_2$ given by

$$z = S(x).$$

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

$$(S(x_1 + x_2), y)_{H_2} = h((x_1 + x_2), y) = h(x_1, y) + h(x_2, y)$$
$$= (S(x_1), y)_{H_2} + (S(x_2), y)_{H_2} = ((S(x_1) + S(x_2)), y)_{H_2}$$

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 $T : H_1 \rightarrow H_2$ 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 $S = T$. □

4 Linear Maps

4.1 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 $T : X \rightarrow Y$, 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 called a linear transformation indexlinear 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 4.1.1 Let X and Y be two Vector Spaces. A map $T : X \rightarrow Y$ 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 mentoined then the fields of the Vector Spaces X and Y are asumed 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 6.2.1.

Now follow several notation, which are of importance, see section 2.1 and figure 4.1:

Domain: $\mathcal{D}(T) \subset X$ 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 .

Further: T is an operator *from* $\mathcal{D}(T)$ *onto* $\mathcal{R}(T)$, $T : \mathcal{D}(T) \rightarrow \mathcal{R}(T)$; T is an operator *from* $\mathcal{D}(T)$ *into* Y, $T : \mathcal{D}(T) \rightarrow Y$; if $\mathcal{D}(T) = X$ then $T : X \rightarrow Y$.
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.

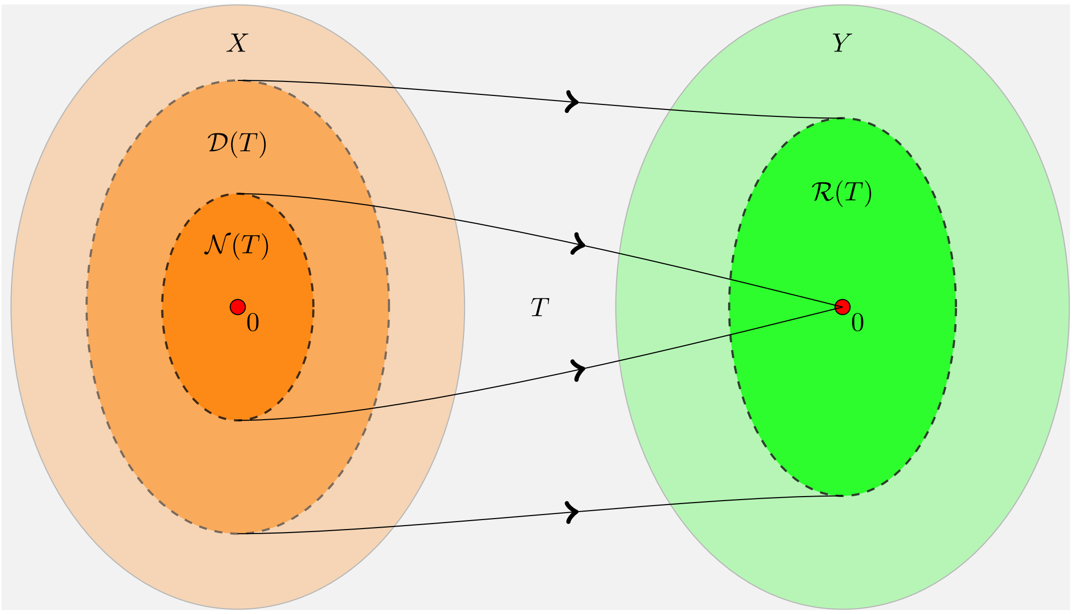


Figure 4.1 Domain, Range, Nullspace

Theorem 4.1.1 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.

Proof

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 $(x_1 + \alpha\, x_2) \in \mathcal{N}(T)$. □

Linear operators can be added together and multiplied by a scalar, the obvious way to do that is as follows.

Definition 4.1.2 If T, $S : X \to Y$ 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 $X \to Y$ is a Vector Space, the zero-operator $\tilde{0} : X \to Y$ in that Vector Space maps every element of X to the zero element of Y.

Definition 4.1.3 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 4.1.4 Let X, Y and Z be Vector Spaces, if $T : X \to Y$ and $S : Y \to Z$ are linear operators then the product $S\,T : X \to Z$ of these linear operators is defined by

$$(S\,T)x = S(T\,x)$$

for every $x \in X$. □

The product operator $ST : X \to Z$ 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 $\alpha \in \mathbb{K}$.

4.2 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 4.2.1 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

$$\| T(x) \| \leq M \| x \|, \tag{4.1}$$

for every $x \in \mathcal{D}(T)$. □

Read formula 4.1 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 4.2.2 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 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

$$\frac{\| T(x) \|}{\| x \|} = \| \frac{1}{\| x \|} T(x) \| = \| T(\frac{x}{\| x \|}) \| \text{ and } \| \frac{x}{\| x \|} \| = 1$$

for all $x \in \mathcal{D}(T) \setminus \{0\}$.

A very nice property of linear operators is that boundedness and continuity are equivalent.

Theorem 4.2.1 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 Let $\epsilon > 0$ be given.

a. ($\Rightarrow$) 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 arbitrarily chosen.

($\Leftarrow$) 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 $\mathcal{D}(T)$. $\square$

b. ($\Rightarrow$) If T is continuous in $x_0 \in \mathcal{D}(T)$ then is T bounded, see part a (($\Rightarrow$)), so T is continuous, see Theorem 4.2.1 a. $\square$

Theorem 4.2.2 Let $T : \mathcal{D}(T) \to Y$ 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 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 $\| x - x_n \| \to 0$ if $n \to \infty$.
Using the linearity and the boundedness of the operator T gives that

$$\| T(x_n) - T(x) \| = \| T(x_n - x) \| \leq \| T \| \| x_n - x \| \to 0 \, (n \to \infty). \tag{4.2}$$

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 4.2 follows that $T(x) = 0$, this means that $x \in \mathcal{N}(T)$, so $\mathcal{N}(T)$ is closed, see Theorem 2.5.2. □

4.3 Space of bounded linear operators

Let X and Y be in first instance arbitrary Vector Spaces. Later on there can also be looked to 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 4.3.1 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

$$\begin{cases} (S + T)(x) = S(x) + T(x), \\ (\alpha\, S)(x) = \alpha\, (S(x)) \end{cases} \tag{4.3}$$

for all $x \in X$ and for all $\alpha \in \mathbb{K}$. □

Theorem 4.3.1 The set $L(X, Y)$ is a Vector Space under the linear operations given in 4.3.

Proof It is easy to check the conditons given in definition 3.2.1 of a Vector Space. □

There will be looked to a special subset of $L(X, Y)$, but then it is of importance that X and Y are Normed Spaces. There will be looked to the bounded linear operators of the Normed Space X into the Normed Space Y, denoted by $BL(X, Y)$.

Theorem 4.3.2 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 The set $BL(X, Y) \subset L(X, Y)$ and $BL(X, Y) \neq \emptyset$, for instance $0 \in BL(X, Y)$, the zero operator. For a linear subspace two conditions have to be checked, see definition 3.2.2.
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 4.3.3 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 The norm of an operator is already defined in definition 4.2.2. It is not difficult to verify that the defined expression satisfies the conditions given in definition 3.7.1. □

Remark 4.3.1 $\| T \|$ 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 4.3.4 If Y is a Banach Space, then $BL(X, Y)$ is a Banach Space.

Proof 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 $T_n \rightarrow T$ for $n \rightarrow \infty$? 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 $T_n(x) \rightarrow T(x)$ for $n \rightarrow \infty$. Hence, there is constructed an operator $T : X \rightarrow Y$, 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 $n \rightarrow \infty$ then always

$$\| T(x) \|_Y \leq K \| x \|_X,$$

since $T_n(x) \rightarrow T(x)$, 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 $T_n \rightarrow T$ 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 $n, m > N(\epsilon) :\ \| T_n - T_m \| < \frac{\epsilon}{2}$. Hence,

$$\| T_n(x) - T_m(x) \|_Y < \frac{\epsilon}{2} \| x \|_X$$

for every $x \in X$. Let $m \to \infty$ 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, $T_n \to T$, for $n \to \infty$ in $(BL(X, Y), \| \cdot \|)$.

The last step completes the proof of the theorem. □

4.4 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 4.4.1 Let $T : X \to Y$ be a linear operator and X and Y Vector Spaces. T is (algebraic) invertible, if there exists an operator $S : Y \to x$ 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 4.2. 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.

Theorem 4.4.1 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.

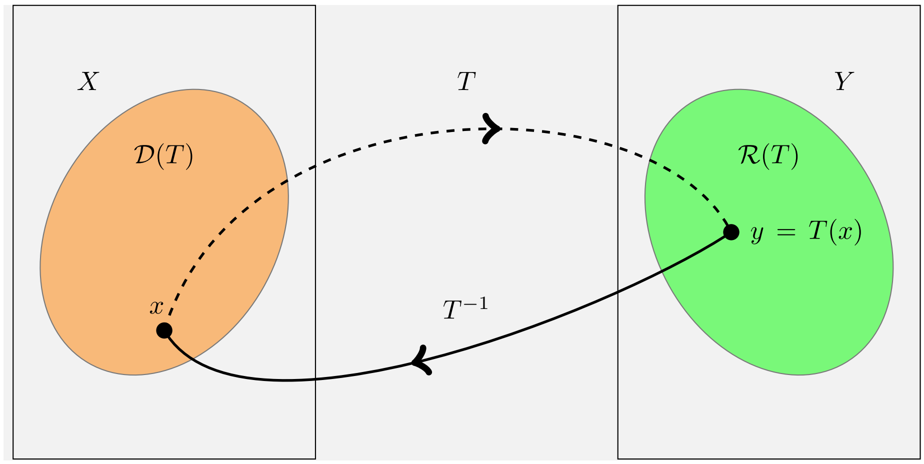


Figure 4.2 The inverse operator: T^{-1}

Proof

a. (⇒) If T^{-1} exists, then is T injective and is obtained out of $T(x) = T(0) = 0$ that $x = 0$. □

(⇐) 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 4.1.1 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 condtions of linearity, see Definition 4.1.1. (α is some scalar.) □

In this paragraph is so far only looked to 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 4.4.1 Look to the operator $T : \ell^\infty \to \ell^\infty$ 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 $T^{-1} : \mathcal{R}(T) \to \ell^\infty$ exists and is unbounded. □

Explanation of Example 4.4.1
The linearity of the operator T is no problem.
The operator is bounded because

$$\| T(x) \|_\infty = \sup_{i \in \mathbb{N}} | \frac{\alpha_i}{i} | \leq \sup_{i \in \mathbb{N}} | \frac{1}{i} | \sup_{i \in \mathbb{N}} | \alpha_i | = \| x \|_\infty . \tag{4.4}$$

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 4.4 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 $T : \ell^\infty \rightarrow \mathcal{R}(T)$. 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 $T^{-1} : \mathcal{R}(T) \rightarrow \ell^\infty$ 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

$$\lim_{n \rightarrow \infty} \| y - y_n \|_\infty = \lim_{n \rightarrow \infty} \frac{1}{\sqrt{n + 1}} = 0.$$

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 $\lim_{n \rightarrow \infty} \| x_n \|_\infty = \lim_{n \rightarrow \infty} \sqrt{n}$ 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 $n \rightarrow \infty$ of

$$\frac{\| T^{-1}(y_n) \|_\infty}{\| y_n \|_\infty} = \frac{\sqrt{n}}{1}$$

does not exist. The inverse operator $T^{-1} : \mathcal{R}(T) \rightarrow \ell^\infty$ is not bounded. □

In example 4.4.1, 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 4.4.2 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 4.4.1 makes clear that the inverse of a bounded operator need not to be bounded. The inverse operator is sometimes bounded.

Theorem 4.4.2 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

$$\| T(x) \|_2 \geq K \ \| x \|_1 \tag{4.5}$$

for every $x \in X$.

Proof

($\Rightarrow$) 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

$$\| x \|_1 = \| T^{-1}(T(x)) \|_1 \leq C_1 \ \| T(x) \|_2 \Rightarrow \| T(x) \|_2 \geq \frac{1}{C_1} \ \| x \|_1 \tag{4.6}$$

Take $K = \dfrac{1}{C_1}$.

($\Leftarrow$) If $T(x) = 0$ then $\| T(x) \|_2 = 0$, using equality 4.5 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 4.4.1 b is proved that T^{-1} is linear.
Almost on the same way as in 4.6 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 4.3.

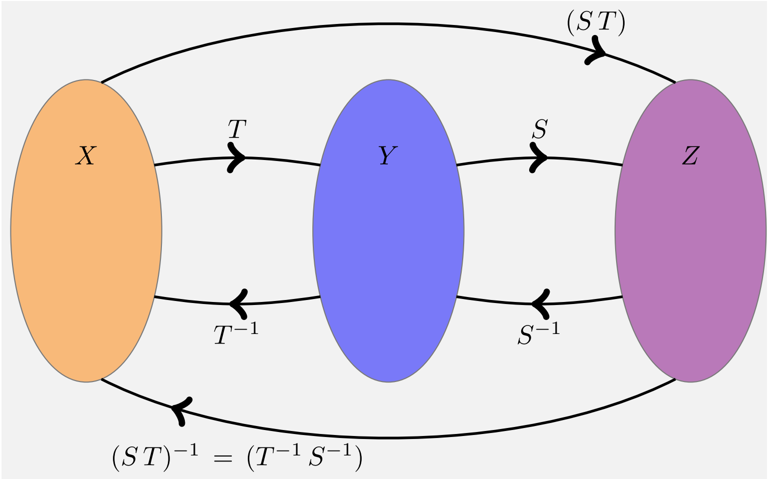


Figure 4.3 Inverse Composite Operator

Theorem 4.4.3 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 The operator $(S\,T) : X \to Z$ 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. □

4.5 Projection operators

For the concept of a projection operator, see section 3.10.1.

Definition 4.5.1 See theorem 3.10.2, y_0 is called the projection of x on M, denoted by

$$P_M : x \to y_0, \text{ or } y_0 = P_M(x),$$

P_M is called the projection operator on M, $P_M : X \to M$. □

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 4.5.1, is not linear. To get a linear projection operator M has to be a closed linear subspace of a Hilbert Space H.
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 $M^{\perp}$.
The projection operator P_M maps

a. X onto M and

b. M onto itself

c. $M^{\perp}$ onto $\{0\}$.

and is idempotent .

Definition 4.5.2 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 4.5.1 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 $(P_M P_M)(x) = P_M(x)$.

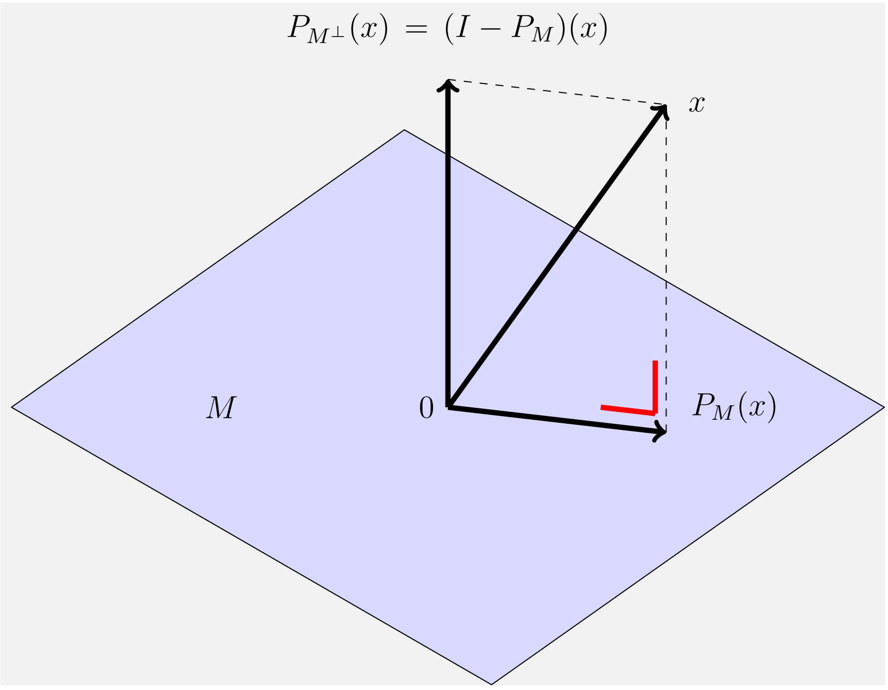


Figure 4.4 Orthogonal projection on a subspace M.

The projection operator P_M is called an orthogonal projection on M, see figure 4.4, 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

$$x = y + z = P_M(x) + P_{M^{\perp}}(x) = P_M(x) + (I - P_M)(x).$$

4.6 Adjoint operators

In first instance, it is the easiest way to introduce adjoint operators in the setting of Hilbert Spaces, see page 56. But the concept of the adjoint operator can also be defined in Normed Spaces.

Theorem 4.6.1 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 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 3.10.9, gives that there exists an unique element $u \in H$ such that

$$(Tx, y) = (x, u) \quad \text{for all} \quad x \in H. \tag{4.7}$$

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 4.7.

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

$$\| T^* \| \leq \| T \| . \tag{4.8}$$

Formula 4.8 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 $\| T^* \| = \| T \|$. □

Definition 4.6.1 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 4.6.2 If $T : H \to H$ 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 4.1.3.

Proof

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.

□

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 to the set of all functions $S = \{f \mid f : I \to \mathbb{K}\}$.

Definition 5.1.1 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 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 to subsets of the Vector Space $(S, \mathbb{K}, (+), (\cdot))$. For instance there is looked to 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

$$\| p_n \| = \max_{i=0,\cdots,n} | a_i | . \tag{5.1}$$

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

$$\| p_n \|_\infty = \sup_{a \leq t \leq b} | p_n(t) |,$$

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 5.1.1 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 First is defined the function $G(s)$ by

$$G(s) = (st + (1-t))^n, \tag{5.3}$$

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

$$G'(s) = n\,t\,(st + (1-t))^{n-1} = \sum_{k=0}^{n} k \binom{n}{k} t^k (1-t)^{(n-k)} s^{k-1}$$

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 $t = \dfrac{x - a}{b - a}$, $0 \leq t \leq 1$ if $x \in [a, b]$. Therefore will now be looked to 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 5.1.2 The Normed Space $(C([0,1]), \|\cdot\|_\infty)$ is the completion of the Normed Space $(\mathbb{P}([0,1]), \|\cdot\|_\infty)$.

Proof 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 2.11.3. Further f is bounded on the compact interval $[0,1]$, see theorem 2.11.2, 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 $t \in [0,1]$

$$|f(t) - p_n(f)(t)| = |\sum_{k=0}^{n} (f(t) - f(\frac{k}{n})) \binom{n}{k} t^k (1-t)^{(n-k)}|$$

$$\leq \sum_{k=0}^{n} |(f(t) - f(\frac{k}{n}))| \binom{n}{k} t^k (1-t)^{(n-k)}$$

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

$$|f(t) - p_n(f)(t)| \leq \sum_{|t-\frac{k}{n}|<\delta} |(f(t) - f(\frac{k}{n}))| \binom{n}{k} t^k (1-t)^{(n-k)}$$

$$+ \sum_{|t-\frac{k}{n}|\geq\delta} |(f(t) - f(\frac{k}{n}))| \binom{n}{k} t^k (1-t)^{(n-k)}$$

$$\leq \sum_{k=0}^{n} \epsilon \binom{n}{k} t^k (1-t)^{(n-k)} + \sum_{|t-\frac{k}{n}|\geq\delta} 2M \binom{n}{k} t^k (1-t)^{(n-k)}.$$

The fact that $|t - \frac{k}{n}| \geq \delta$ 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 5.1.1 results in

$$\begin{aligned} |f(t) - p_n(f)(t)| &\leq \epsilon + \frac{2M}{\delta^2} \sum_{k=0}^{n} (t - \frac{k}{n})^2 \binom{n}{k} t^k (1-t)^{(n-k)} \\ &= \epsilon + \frac{2M}{\delta^2} \frac{1}{n} t(1-t) \\ &\leq \epsilon + \frac{2M}{\delta^2} \frac{1}{n} \frac{1}{4}, \end{aligned}$$

for all $t \in [0,1]$. The upper bound $(\epsilon + \frac{2M}{\delta^2} \frac{1}{n} \frac{1}{4})$ does not depend on t and for $n > \frac{M}{2\delta^2 \epsilon}$, this implies that

$$\| f(t) - p_n(f)(t) \|_\infty < 2\epsilon.$$

The consequence is that

$$p_n(f) \rightarrow f \text{ for } n \rightarrow \infty \text{ in } (C([0,1]), \|\cdot\|_\infty).$$

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 5.1.3 The Normed Space $(C([a,b]), \|\cdot\|_\infty)$ is separable.

Proof According the Weierstrass Approximation Theorem, theorem 5.1.2, every continous function f on the bounded en 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 5.1.4 The Normed Space $(C([a,b]), \|\cdot\|_\infty)$ is a Banach Space.

Proof 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

$$\| f \|_p = (\int_a^b | f(x) |^p \, dx)^{(\frac{1}{p})}. \tag{5.7}$$

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 to 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 5.1.1 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 147. 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 5.1.1 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 147. 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 to 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 3.7.2. 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 3.2.2) 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.

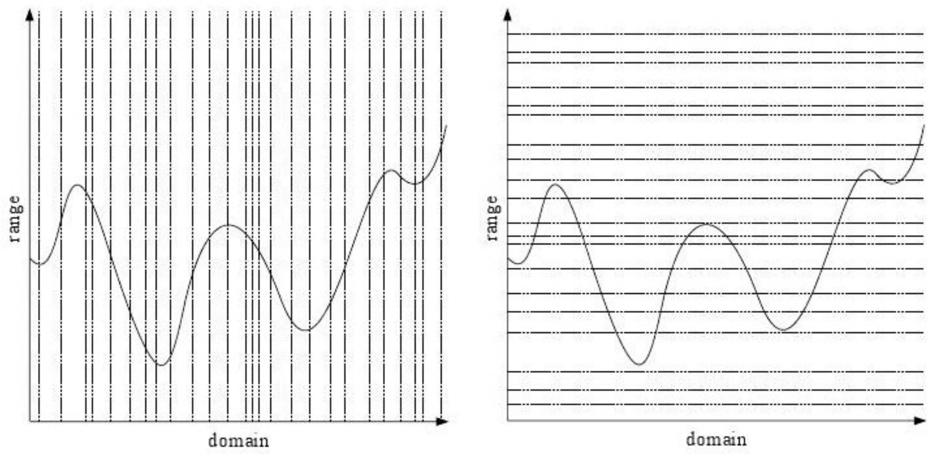


Figure 5.1 Left: Riemann-integral, right: Lebesgue-integral.

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

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases} \tag{5.15}$$

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,

and if $\alpha = 1 + 2^{2\,n}$

$$B_{n, 1 + 2^{2\,n}} = \{x \in [a,b] \,|\, f(x) \geq 2^n\}.$$

Define the sequence $\{f_n\}$ of functions by

$$f_n = \sum_{\alpha=1}^{1+2^{2\,n}} \frac{(\alpha - 1)}{2^n} \chi_{B_{n,\alpha}}.$$

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 $m(B_{n,\alpha})$.

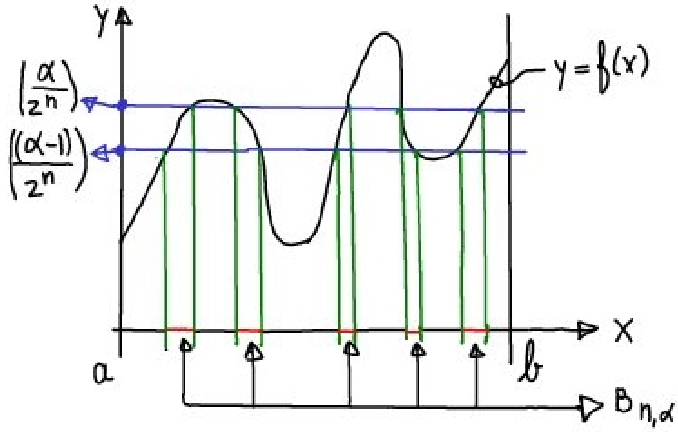


Figure 5.2 A subset $B_{n,\alpha}$.

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,

$$\lim_{n\to\infty} \sum_{\alpha=1}^{1+2^{2n}} \frac{(\alpha - 1)}{2^n} m(\chi_{B_{n,\alpha}}) = \int_a^b f(x)\,dx,$$

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 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 5.1.5 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

$$|\,(f, g)\,| \leq \parallel f \parallel_2 \parallel g \parallel_2, \tag{5.16}$$

with $\parallel \cdot \parallel_2$ defined as in 5.9.

Proof 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 $\parallel g \parallel_2 = 0$. □

5.1.8 $B(\Omega)$ with $\parallel \cdot \parallel_\infty$-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

$$\parallel f \parallel_\infty = \sup_{x \in \Omega} |\, f(x)\,|\,. \tag{5.17}$$

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 5.1.6 The Normed Space $(B(\Omega), \parallel \cdot \parallel_\infty)$ is a Banach Space.

Proof 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 $\parallel g - f_n \parallel_\infty < 2\epsilon$ for $n > N(\epsilon)$ and this means that the convergence is uniform.
The inequality $\parallel g \parallel \leq \parallel g - f_n \parallel + \parallel f_n \parallel$ gives that, for an appropriate choice of n. The constructed function g is bounded, so $g \in B(\Omega)$. □

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 $\mathbb{K}^{\mathbb{N}} = \{f \mid f : \mathbb{N} \to \mathbb{K}\}$[6]. 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.)

Definition 5.2.1 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 to subsets of the Vector Space $(\mathbb{K}^{\mathbb{N}}, \mathbb{K}, (+), (\cdot))$, there is looked to 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 to 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.

[6] Important: The sequence spaces are also function spaces, only their domain is most of the time $\mathbb{N}$ or $\mathbb{Z}$.

5.2.1 ℓ^∞ with $\| \, . \, \|_\infty$-norm

The norm used in this space is the $\| \, . \, \|_\infty$-norm, which is defined by

$$\| \underline{\xi} \|_\infty = \sup_{i \in \mathbb{N}} | \xi_i | \tag{5.18}$$

and $\underline{\xi} \in \ell^\infty$, if $\| \underline{\xi} \|_\infty < \infty$.
The Normed Space $(\ell^\infty, \| \cdot \|_\infty)$ is complete.

Theorem 5.2.1 The space ℓ^∞ is not separable.

Proof Let $S = \{x \in \ell^\infty \,|\, x(j) = 0 \text{ or } 1, \text{ for } j = 1, 2, \cdots\}$ 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 5.2.2 The dual space $(\ell^\infty)' = ba(\mathcal{P}(\mathbb{N}))$.

Proof This will become a difficult proof[7].

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 2 of Section 8.5 for more information.
   -This part is finished.

[7] At the moment of writing, no idea if it will become a succesful proof.

3. An additive function f preserves the addition operation:

$$f(x + y) = f(x) + f(y),$$

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 5.2.3.
-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[8].
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 8.5.1, but these charges?

7. Reading the literature, there is much spoken about σ-algebras and measures, for more information about these subjects, see section 8.3.

[8] At the moment of writing, no idea what this means!

-This part gives some information.

8. In the literature, see (Morrison, 2001, page 50), can be read a way to prove theorem 5.2.2. For more information, see section 8.5.
   -This part gives a way to a proof of Theorem 5.2.2, it uses a lot of information out of the steps made above.

Theorem 5.2.2 is proved yet, see 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 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 5.2.3 are called Banach Limits.

Theorem 5.2.3 There is a bounded linear functional $L : \ell^\infty \rightarrow \mathbb{R}$ such that

a. $\| L \| = 1$.
b. If $x \in c$ then $L(x) = \lim_{n\rightarrow\infty} x_n$.
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 $\sigma : \ell^\infty \rightarrow \ell^\infty$ is the shift-operator, defined by $(\sigma(x))_n = x_{n+1}$.

Proof The proof is splitted up in several parts and steps.
First the parts a and d. Here Hahn-Banach, theorem 6.7.3, 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 $\lim_{n\rightarrow\infty} x_n$ exists. Thus $\lim_{n\rightarrow\infty}(x - \sigma(x))_n = 0$ and $\| e - (x - \sigma(x)) \|_\infty \geq 1$.
   This proves that $\mathrm{dist}(e, M) = 1$.
3. By theorem 6.7.3 there is linear functional $L : \ell^\infty \rightarrow \mathbb{R}$ such that $\| L \| = 1$, $L(e) = 1$ and $L(M) = 0$. The bounded functional L satisfies a and d of the theorem. $(L(x - \sigma(x)) = 0$, L is linear, so $L(x) = L(\sigma(x))$.) □

Part 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 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

$$|L(x)| = | L(\sigma^{(n)}(x)) | < \epsilon. \tag{5.19}$$

Since $\epsilon > 0$ is arbitrary chosen, inequality 5.19 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 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 a, so $L(v) \geq 0$. □

Theorem 5.2.3, about the Banach Limits, is proved. □

Example 5.2.1 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

$$\| \underline{\xi} \|_1 = \sum_{i=1}^{\infty} | \xi_i | \tag{5.20}$$

and $\underset{\sim}{\xi} \in \ell^1$, if $\| \underset{\sim}{\xi} \|_1 < \infty$.
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

$$\| \underset{\sim}{\xi} \|_p = (\sum_{i=1}^{\infty} | \xi_i |^p)^{\frac{1}{p}} \tag{5.21}$$

and $\underset{\sim}{\xi} \in \ell^p$, if $\| \underset{\sim}{\xi} \|_p < \infty$.
The Normed Space $(\ell^p, \| \cdot \|_p)$ is complete.

Theorem 5.2.4 The space ℓ^p is separable.

Proof The set $S = \{\underset{\sim}{y} = (\eta_1, \eta_2, \cdots, \eta_n, 0, 0, \cdots) \,|\, \eta_i \in \mathbb{Q}, 1 \leq i \leq n, \mathbb{N}\}$ is a countable subset of ℓ^p.
Given $\epsilon > 0$ and $\underset{\sim}{x} = (\xi_1, \xi_2, \xi_3, \cdots) \in \ell^p$ 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 $\underset{\sim}{y} \in S$ such that

$$\sum_{j=1}^{N(\epsilon)} | \eta_j - \xi_j |^p < \frac{\epsilon^p}{2}.$$

Hence, $\| \underset{\sim}{x} - \underset{\sim}{y} \|_p = (\sum_{j=1}^{N(\epsilon)} | \eta_j - \xi_j |^p + \sum_{j=N(\epsilon)+1}^{\infty} | \xi_j |^p)^{\frac{1}{p}} < \epsilon$,
so $\overline{S} = \ell^p$. □

5.2.4 ℓ^2 with $\| \, . \, \|_2$-norm

The norm used in this space is the $\| \, . \, \|_2$-norm, which is defined by

$$\| \underset{\sim}{\xi} \|_2 = \sqrt{\sum_{i=1}^{\infty} | \xi_i |^2} \tag{5.22}$$

and $\underset{\sim}{\xi} \in \ell^2$, if $\| \underset{\sim}{\xi} \|_2 < \infty$.
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 5.2.5 If $x \in \ell^2$ and $y \in \ell^2$ then $(x + y) \in \ell^2$.

Proof 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 3.9.1, 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

$$\lim_{N\to\infty} (\sum_{i=1}^{N} | x_i + y_i |^2)^{\frac{1}{2}} \leq \lim_{N\to\infty} ((\sum_{i=1}^{N} | x_i |^2)^{\frac{1}{2}} + (\sum_{i=1}^{N} | y_i |^2)^{\frac{1}{2}})$$

$$= (\sum_{i=1}^{\infty} | x_i |^2)^{\frac{1}{2}} + (\sum_{i=1}^{\infty} | y_i |^2)^{\frac{1}{2}} < \infty.$$

□

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 5.2.6 The mapping $T : c \to c_0$ 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 $x_\infty = \lim_{i\to\infty} x_i$.
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 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

$$|x_i - x_\infty| \leq |x_i| - |x_\infty| \leq 2 \parallel x \parallel_\infty \tag{5.23}$$

and

$$|x_i| \leq |x_i| - |x_\infty| + |x_\infty| \leq 2 \parallel T(x) \parallel_\infty . \tag{5.24}$$

With the inequalities 5.23 and 5.24, it follows that

$$\frac{1}{2} \parallel x \parallel_\infty \leq \parallel T(x) \parallel_\infty \leq 2 \parallel x \parallel_\infty . \tag{5.25}$$

T is continuous, because T is linear and bounded. Further is T bijective and bounded from below. With theorem 4.4.2, it follows that T^{-1} is continuous.
The bounds given in 5.25 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

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

$$\mathbb{R}^N = \{\underline{x} \mid \underline{x} = (x_1, \cdots, x_N), x_i \in \mathbb{R}, 1 \leq i \leq N\}, \tag{5.26}$$

replace $\mathbb{R}$ by $\mathbb{C}$ and we have the definition of $\mathbb{C}^N$.
An inner product is given by

$$(\underline{x}, \underline{y}) = \sum_{i=1}^{N} x_i \overline{y_i}, \tag{5.27}$$

with $\overline{y_i}$ the complex conjugate[5] 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

$$(\underline{x}, \underline{y}) = \underline{x} \bullet \underline{y} = < \underline{x}, \underline{y} > \tag{5.28}$$

Often the elements out of $\mathbb{R}^N$ or $\mathbb{C}^N$ are presented by columns, i.e.

$$\underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \tag{5.29}$$

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

$$(\underline{x}, \underline{y}) = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}^T \begin{bmatrix} \overline{y_1} \\ \vdots \\ \overline{y_N} \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_N \end{bmatrix} \begin{bmatrix} \overline{y_1} \\ \vdots \\ \overline{y_N} \end{bmatrix} \tag{5.30}$$

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 5.2.7 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

$$|\,(\underline{x}, \underline{y})\,| \leq \|\,\underline{x}\,\|_2 \,\|\,\underline{y}\,\|_2\,. \tag{5.31}$$ [9]

Proof 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

$$\begin{aligned} 0 &\leq (\underline{x}, \underline{x}) - (\underline{x}, \alpha\,\underline{y}) - (\alpha\,\underline{y}, \underline{x}) + (\alpha\,\underline{y}, \alpha\,\underline{y}) \\ &= (\underline{x}, \underline{x}) - \overline{\alpha}\,(\underline{x}, \underline{y}) - \alpha(\underline{y}, \underline{x}) + \overline{\alpha}\,\alpha\,(\underline{y}, \underline{y}) \end{aligned} \tag{5.32}$$

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 5.32 and use that

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

see definition 3.9.1. Writing out, and some calculations, gives the inequality of Cauchy-Schwarz. □

[9] With $\|\cdot\|_2$ is meant expression 5.22, but not the length of a vector. Nothing is known about how the coördinates are chosen.

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 5.2.8 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

a. If $a \geq 0$ and $b \geq 0$ then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \tag{5.33}$$

If $b = 0$, the inequality 5.33 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

$$t^{\frac{1}{p}} \leq \frac{1}{q} + \frac{t}{p}, \; t \geq 0. \tag{5.34}$$

Take $t = \frac{a^p}{b^q}$ and fill in formula 5.34, multiply the inequality by b^q and inequality 5.33 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 $a = \frac{a_j}{\alpha}$ and $b = \frac{b_j}{\alpha}$ and fill in in formula 5.33, $j = 1, \cdots, n$.
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 3.9.1.

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 $\frac{x}{\| x \|_p}$ then it follows that $\| x \|_r \leq \| x \|_p$ for $1 \leq p < r < \infty$. □

Remark 5.2.2 Jensen's inequality 5.2.8 c implies that $\ell^p \subset \ell^r$ and if $x_n \rightarrow x$ in ℓ^p then $x_n \rightarrow x$ in ℓ^r. □

6 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 $\dim X = n$. In 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 $\mathbb{C}$.

6.1 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

$$\begin{cases} T(x + y) = T(x) + T(y) \\ T(\beta x) = \beta T(x) \end{cases} \tag{6.1}$$

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

$$\tilde{d}(T(x), T(y)) = d(x, y) \tag{6.2}$$

for all $x, y \in X$, also called an distance-preserving isomorphism.

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

$$\| T(x) \| = \| x \| \tag{6.3}$$

for all $x \in X$, also called an isometric isomorphism.

6.2 Linear functional and sublinear functional

Definition 6.2.1 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 6.2.2 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 6.2.1 The *norm* on a Normed Space is an example of a sublinear functional. □

Example 6.2.2 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

$$f(\underline{x}) = \begin{bmatrix} a_1 & \cdots & a_N \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \tag{6.4}$$

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 (6.4), 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 $\underline{x} \in \mathbb{R}^N$

$$g(\underline{x}) = \underline{b} \bullet \underline{x},$$

for the notation, see (5.28).

6.3 Algebraïc dual space of X, denoted by X^*

Let X be a Vector Space and take the set of <u>all</u> linear functionals $f : X \to \mathbb{K}$. This set of <u>all</u> 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

$$\begin{cases} (f_1 + f_2)(x) = f_1(x) + f_2(x) \\ f_1(\beta\, x) = \beta\, f_1(x) \end{cases} \tag{6.5}$$

for all $x \in X$ and for all $\beta \in \mathbb{K}$.
The set of <u>all</u> linear functionals on X, together with the above defined addition and scalar multiplication, see (6.5), 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 (6.5).

6.4 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

$$x = \alpha_1 e_1 + ... + \alpha_n e_n \tag{6.6}$$

and the coefficients α_i, with $1 = i \leq n$, are unique.

6.4.1 Unique representation of linear functionals

let f be a linear functional on X, the image of x is

$$f(x) \in \mathbb{K}.$$

Theorem 6.4.1 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 Choose a basis $\{e_1, \cdots, e_n\}$ then every $x \in X$ has an unique representation

$$x = \sum_{i=1}^{n} \alpha_i e_i. \tag{6.7}$$

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 6.7 is unique, the result is obtained. □

6.4.2 Unique representation of linear operators between finite dimensional spaces

let T be a linear operator between the finite dimensional Vector Spaces X and Y

$$T : X \to Y.$$

Theorem 6.4.2 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 Take the basis $\{e_1, \cdots, e_n\}$ of X then x has an unique representation

$$x = \sum_{i=1}^{n} \alpha_i e_i. \tag{6.8}$$

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 6.8 is unique, the result is obtained. □

Let $\{b_1, \cdots, b_k\}$ be a basis of Y.

Theorem 6.4.3 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 6.10 for τ_{ij}.)

Proof Since $y = T(x)$ and $y_k = T(e_k)$ are elements of Y they have an unique representation with respect tot the basis $\{b_1, \cdots, b_k\}$,

$$y = \sum_{i=1}^{k} \beta_i b_i, \tag{6.9}$$

$$T(e_j) = \sum_{i=1}^{k} \tau_{jk} b_i, \tag{6.10}$$

Substituting the formulas of 6.9 and 6.10 together gives

$$T(x) = \sum_{j=1}^{k} \beta_j b_j = \sum_{i=1}^{n} \alpha_i T(e_i) = \sum_{i=1}^{n} \alpha_i (\sum_{j=1}^{k} \tau_{ij} b_j) = \sum_{j=1}^{k} (\sum_{i=1}^{n} \alpha_i \tau_{ij}) b_j. \tag{6.11}$$

Since $\{b_1, \cdots, b_k\}$ is basis of Y, the coefficients

$$\beta_j = \sum_{i=1}^{n} \alpha_i \tau_{ij}$$

for $1 \leq j \leq k$. □

6.4.3 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 6.4.1, and see also example 6.2.2.
Look at the following n-tuples:

$$\begin{gathered}(1, 0, \ldots\ldots\ldots\ldots, 0), \\ (0, 1, 0, \ldots\ldots\ldots, 0), \\ \ldots\ldots, \\ (0, \ldots\ldots 0, 1, 0, .., 0), \\ \ldots\ldots, \\ (0, \ldots\ldots\ldots\ldots, 0, 1),\end{gathered}$$

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 $f = \gamma_1 f_1 + \gamma_2 f_2 + \ldots\ldots + \gamma_n f_n$ and realize that when $x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots\ldots + \alpha_n e_n$ that $f_j(x) = f_j(\alpha_1 e_1 + \alpha_2 e_2 + \ldots\ldots + \alpha_n e_n) = \alpha_j$, so $f(x) = f(\alpha_1 e_1 + \alpha_2 e_2 + \ldots\ldots + \alpha_n e_n) = \alpha_1 \gamma_1 + \ldots\ldots + \alpha_n \gamma_n$.
It is interesting to note that: $\dim X^* = \dim X = n$.

Theorem 6.4.4 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 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 $1 \leq j \leq n$. □

6.4.4 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 6.4.5 The canonical mapping C is injective.

Proof 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 6.4.4 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 6.4.6 The canonical mapping C is surjective.

Proof 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 to 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 6.4.1 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.

6.5 The dual space X' of a Normed Space X

In section 6.4 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 to 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 $\| \, . \, \|$. This norm is needed to speak about a norm of a linear functional on X.

Definition 6.5.1 The norm of a linear functional f is defined by

$$\| f \| = \sup_{\left\{ \begin{array}{c} x \in X \\ x \neq 0 \end{array} \right\}} \frac{| f(x) |}{\| x \|} = \sup_{\left\{ \begin{array}{c} x \in X \\ \| x \| = 1 \end{array} \right\}} | f(x) | \tag{6.12}$$

□

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 6.5.2 A functional f is bounded if there exists a number A such that

$$| f(x) | \leq A \parallel x \parallel \tag{6.13}$$

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

$$\frac{| f(x) |}{\parallel x \parallel} = | f(\frac{x}{\parallel x \parallel}) |$$

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 4.3.4 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 $f : X \to \mathbb{K}$. Results derived for linear operators are also applicable to linear functionals.

Theorem 6.5.1 A linear functional, on a Normed Space, is bounded *if and only if* it is continuous.

Proof The proof exists out of two parts.

($\Rightarrow$) Suppose f is linear and bounded, then there is a positive constant A such that $| f(x) | \leq A \parallel x \parallel$ for all x.
If $\epsilon > 0$ is given, take $\delta = \frac{\epsilon}{A}$ and for all y with $\parallel x - y \parallel \leq \delta$

$$| f(x) - f(y) | = | f(x - y) | \leq A \parallel x - y \parallel \leq A \, \delta = A \frac{\epsilon}{A} = \epsilon.$$

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

$$| f(x) | < 1 \text{ for } \parallel x \parallel < \delta.$$

For some arbirary y, in the Normed Space, it follows that

$$| f(y) | = \frac{2 \parallel y \parallel}{\delta} f(\frac{\delta}{2 \parallel y \parallel} y) < \frac{2}{\delta} \parallel y \parallel,$$

since $\parallel \frac{\delta}{2 \parallel y \parallel} y \parallel = \frac{\delta}{2} < \delta$. Take $A = \frac{2}{\delta}$ in formula 6.13, this positive constant A is independent of y, the functional f is bounded. □

6.6 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 6.4.4. Be careful if X is an infinite dimensional Normed Space.

Theorem 6.6.1 $(\ell^1)' = \ell^\infty$ and $(c_0)' = \ell^1$

Proof See the sections 6.6.1 and 6.6.2. □

Theorem 6.6.1 gives that $((c_0)')' = (\ell^1)' = \ell^\infty$. One thing can always be said and that is that $X \subseteq X''$, see theorem 6.8.1. 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.

6.6.1 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

$$\parallel x \parallel_1 = \sum_{k=1}^{\infty} | \alpha_k | \; (< \infty)$$

and the norm of $x \in \ell^\infty$ is

$$\parallel x \parallel_\infty = \sup_{k \in \mathbb{N}} | \alpha_k | (< \infty).$$

A bounded linear functional f of ℓ^1, $(f : \ell^1 \rightarrow \mathbb{R})$ 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^∞.

6.6.2 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 6.5.1), 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

$$\mid f(x_0^n) \mid = \mid \sum_{k=1}^{n} \eta_k^0 \gamma_k \mid = \sum_{k=1}^{n} \mid \gamma_k \mid = \sum_{k=1}^{n} \mid f(e_k) \mid \leq \| f \|_\infty \| x_0^n \|_\infty \leq \| f \|_\infty, \tag{6.14}$$

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 6.14, follows that

$$\sum_{k=1}^{n} \mid \gamma_k \mid = \sum_{k=1}^{\infty} \mid f(e_k) \mid \leq \| f \|_\infty . \tag{6.15}$$

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

$$\mid f(x) \mid \leq \sum_{k=1}^{\infty} \mid \lambda_k \| \gamma_k \mid \leq \| x \|_\infty \sum_{k=1}^{\infty} \mid \lambda_k \mid \leq \| x \|_\infty \sum_{k=1}^{\infty} \mid f(e_k) \mid,$$

and this means that

$$\frac{\mid f(x) \mid}{\| x \|_\infty} \leq \sum_{k=1}^{\infty} \mid f(e_k) \mid,$$

together inequalty 6.15, 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 arbitrary $x = (\lambda_k)_{k \in \mathbb{N}} \in c_0$ 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

$$\frac{\mid f(x) \mid}{\| x \|_\infty} \leq \sum_{k=1}^{\infty} \mid \alpha_k \mid \leq \| (\alpha_k)_{k \in \mathbb{N}} \|_1 .$$

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.

6.7 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 3.10.3 could be read as that it is possible to construct hyperplanes through y_0, which separate x form the subset M, see figures 3.4 and 3.5. 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 3.5.4. The question is if there exists some bounded linear functional $g \in X'$ such that

$$g(x_0) = 1,\ g|_M = 0, \text{ and may be } \| g \| = \frac{1}{d}? \tag{6.16}$$

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 6.7.1 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 10.1.1. 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 6.16 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

$$\frac{|g(m + \alpha x_0)|}{\| m + \alpha x_0 \|} \leq \frac{|\alpha|}{|\alpha| \, d} = \frac{1}{d}, \tag{6.17}$$

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

$$\| g \| \geq \frac{1}{d} \tag{6.18}$$

on $\widehat{M}$. With the inequalities 6.18 and 6.17 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 6.16. 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 6.2.2. If $f \in X'$ then is an example of a sublinear functional p given by

$$p(x) = \| f \| \, \| x \|, \tag{6.19}$$

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 6.19 of the sublinear functional p. The conditions SLF 1 and SLF 2 are easy to check. First will be proved the Lemma of Hahn-Banach .

Theorem 6.7.1 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 The proof is splitted up in several steps.

1. First will be looked to the set of all possible extensions of (M, f) and the question will be if there exists some maximal extension? See Step 1.
2. If there exists some maximal extension, the question will be if that is equal to $(X, f_{\mathcal{E}})$? See Step 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 6.7 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 10.1. 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 2.13.1. 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 to 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

$$h(z + \alpha\, y_1) \leq p(z + \alpha\, y_1) \tag{6.20}$$

for all elements out of $\widehat{M_\epsilon}$. The existence of such a C is proved in Step 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

$$h(z + \alpha\, y_1) = f_\epsilon(z) + \alpha\, C \leq p(z + \alpha\, y_1) \tag{6.21}$$

for all $z \in M_\epsilon$ and $\alpha \in \mathbb{R} \setminus \{0\}$. Replace z by $\alpha\, z$ and divide both sides of formula 6.21 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 $v, w \in M_\epsilon$

$$\begin{aligned} h(w) - h(v) &= h(w - v) \leq p(w - v) \\ &= p(w + y_1 - v - y_1) \leq p(w + y_1) + p(-\,v - y_1), \end{aligned}$$

and therefore

$$-\,h(v) - p(-\,v - y_1) \leq -\,h(w) + p(w + y_1).$$

Hence, there exists a real constant C such that

$$\sup_{v \in M_\epsilon} (-h(v) - p(-v - y_1)) \leq C \leq \inf_{w \in M_\epsilon} (-h(w) + p(w + y_1)). \tag{6.22}$$

With the choice of a real constant C, which satisfies inequality 6.22, the extended functional h can be constructed, as used in Step 2. □

In the Lemma of Hahn-Banach, see theorem 6.7.1, 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 6.7.2 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 The proof is splitted up in two cases.

1. The real case $\mathbb{K} = \mathbb{R}$, see Case 1.
2. The complex case $\mathbb{K} = \mathbb{C}$, see Case 2.

Case 1: Set $p(x) = \| f \| \| x \|$, p is a sublinear functional on X and by the Lemma of Hahn-Banach, see theorem 6.7.1, 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

$$\| g \| \leq \| f \| . \tag{6.23}$$

Take some $y \in M$ then

$$\| g \| \geq \frac{|g(y)|}{\| y \|} = \frac{|f(y)|}{\| y \|}.$$

Hence,

$$\| g \| \geq \| f \| . \tag{6.24}$$

The inequalities 6.23 and 6.24 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 6.7.1, 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

$$g(x) = \widehat{u}(x) - \imath \widehat{u}(\imath x) \tag{6.25}$$

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 6.25 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

$$\| g \| \geq \| f \| . \tag{6.26}$$

Let $x \in X$ then there is some real number ϕ such that

$$g(x) = | g(x) | \exp(\imath \phi).$$

Then

$$
\begin{aligned}
|\, g(x)\,| &= \exp\left(-\imath\,\phi\right) g(x) \\
&= \mathrm{Re}(\exp\left(-\imath\,\phi\right) g(x)) \;=\; \mathrm{Re}(g(\exp\left(-\imath\,\phi\right) x)) \\
&= \widehat{u}(\exp\left(-\imath\,\phi\right) x) \;\leq \| \, f \,\| \| \exp\left(-\imath\,\phi\right) x \,\| \\
&= \| \, f \,\| \| \, x \,\| \,.
\end{aligned}
$$

This shows that g is bounded and $\| \, g \,\| \leq \| \, f \,\|$, together with inequality 6.26 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

$$g(x_0) = 1, \; g|_M \;=\; 0, \text{ and may be } \| \, g \,\| = \frac{1}{d}, \tag{6.27}$$

with $x_0 \in X$ such that $d(x_0, M) = d > 0$.
Before the Lemma of Hahn-Banach, theorem 6.7.1, there was constructed a bounded linear functional g on $\widehat{M}$, the span of M and x_0, which satisfied the condition given in 6.27. 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 6.7.2, 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 6.27, is a fact.

The result of the question in 6.16 can be summarized into the following theorem:

Theorem 6.7.3 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 Read this section 6.7. □

Remark 6.7.2 With the result of theorem 6.7.3 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 6.7.1. □

6.7.1 Useful results with Hahn-Banach

There are enough bounded linear functionals on a Normed Space X to distinguish between the points of X.

Theorem 6.7.4 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 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

$$f(x) = f(\alpha\, x_0) = \alpha \| x_0 \|,$$

with $\alpha \in \mathbb{K}$. f is a linear functional on M and

$$|f(x)| = |f(\alpha\, x_0)| = |\alpha| \| x_0 \| = \| x \|$$

for every $x \in M$. Hence, f is bounded and $\| f \| = 1$.
By the theorem of Hahn-Banach, theorem 6.7.2, there exists a functional $g \in X^{'}$, such that $g|M = f$ and $\| g \| = \| f \|$. Hence, $g(x_0) = f(x_0) = \| x_0 \|$, and $\| g \| = 1$. □

Theorem 6.7.5 Let X be a Normed Space over the field $\mathbb{K}$ and $x \in X$, then

$$\| x \| = \sup\{\frac{| f(x) |}{\| f \|} | f \in X^{'} \text{ and } f \neq 0\}.$$

Proof The case that $x = 0$ is trivial.
Let $0 \neq x \in X$. With theorem 6.7.4 there exists a $g \in X'$, such that $g(x) = \| x \|$, and $\| g \| = 1$. Hence,

$$\sup\{\frac{|f(x)|}{\| f \|} \,|\, f \in X' \text{ and } f \neq 0\} \geq \frac{|g(x)|}{\| g \|} = \| x \| . \tag{6.28}$$

Further,

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

for every $f \in X'$, therefore

$$\sup\{\frac{|f(x)|}{\| f \|} \,|\, f \in X' \text{ and } f \neq 0\} \leq \| x \| . \tag{6.29}$$

The inequalities 6.28 and 6.29 complete the proof. □

6.8 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 proven 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 4.3.4. 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 6.7.2, 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 6.8.1 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 The proof is splitted up in several steps.

1. Several steps are already done in section 6.4.4. The linearity of $g_x : X' \to X''$ and $C : X \to X''$ 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 6.7.5. Be careful the normed space X is may be not finite dimensional anymore, so theorem 6.4.4 cannot be used. Hence, the mapping C is injective.
3. The mapping C preserves the norm, because

$$\| C(x) \| = \| g_x \| = \sup\{\frac{|g_x(f)|}{\| f \|} \mid f \in X' \text{ and } f \neq 0\} = \| x \|,$$

   see theorem 6.7.5.

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 6.8.2 Let $(X, \| \cdot \|)$ be a Normed Space. If X' is separable *then* X is separable.

Proof 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 1.
2. Secondly there is proved, by a contradiction, that $\overline{S} = X$, see Step 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 6.5.1, $\| f_n \| = \sup_{\|x\|=1} |f_n(x)|$, 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 6.27 is that there exists a functional $g \in X^{'}$ such that

$$g(w) \neq 0,$$
$$g(\overline{S}) = 0,$$
$$\| g \| = 1. \quad (6.30)$$

In particular $g(v_n) = 0$ for all $n \in \mathbb{N}$ and

$$\frac{1}{2} \| f_n \| \leq |f_n(v_n)| = |f_n(v_n) - g(v_n) + g(v_n)|$$
$$\leq |f_n(v_n) - g(v_n)| + |g(v_n)|.$$

Since $\| v_n \| = 1$ and $g(v_n) = 0$ for all $n \in \mathbb{N}$, it follows that

$$\frac{1}{2} \| f_n \| \leq \| f_n - g \| . \tag{6.31}$$

Since M is dense in $X^{'}$, choose f_n such that

$$\lim_{n \to \infty} \| f_n - g \| = 0. \tag{6.32}$$

Using the formulas 6.30, 6.30 and 6.32, the result becomes that

$$1 = \| g \| = \| g - f_n + f_n \|$$
$$\leq \| g - f_n \| + \| f_n \|$$
$$\leq \| g - f_n \| + 2 \| g - f_n \|,$$

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 6.8.1 and $X^{''}$ is a Banach Space.

Theorem 6.8.3 A Normed Space X is isometrically isomorphic to a dense subset of a Banach Space.

Proof 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 3.8.1. Hence, X is isometrically isomorphic with to the dense subspace $C(X)$ of the Banach Space $\overline{C(X)}$. □

An nice use of theorem 6.8.2 is the following theorem.

Theorem 6.8.4 Let $(X, \| \cdot \|)$ be a separable Normed Space. If X' is non-separable *then* X is non-reflexive.

Proof The proof will be done by a contradiction.
Assume that X is reflexive. Then is X'' isometrically isomorphic to X under the canonical mapping $C : X \rightarrow X''$. X is separable, so X'' is separable and with the use of theorem 6.8.2, the space X' is separable. But that contradicts the hypothesis that X' is non-separable. □

7 Examples

7.1 Block-Wave and Fourier Series

An approximation of a block-wave, with a period of 2π:

$$f(x) = \begin{cases} -1 & \text{if } -\pi < x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } 0 < x < \pi \end{cases}$$

is for instance given by:

$$\lim_{M \to \infty} f_M(x)$$

where the function f_M is defined by

$$f_M(x) = \sum_{j=0}^{M} \frac{1}{2j+1} \sin((2m + 1)x),$$

a Fourier series of the function f. See figure 7.1 and figure 7.2 for the graphs of the approximations of f by different values of M. More information about Fourier series can be found at wiki-four-series.

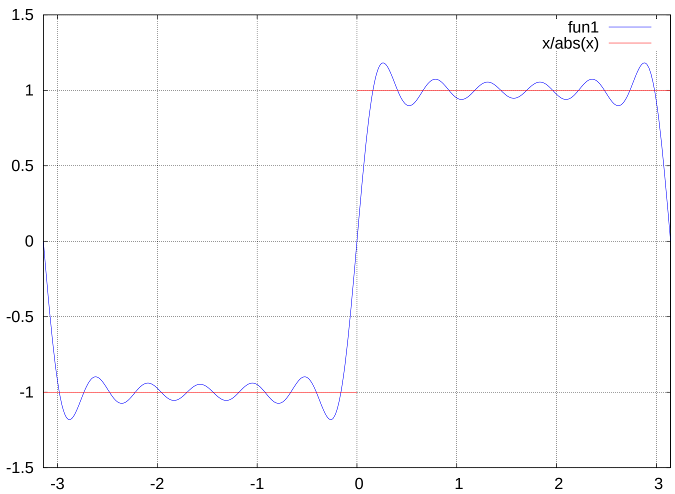


Figure 7.1 Fourier Series $M = 5$

But what kind of an approximation?

$\| \,.\, \|_\infty$-norm: Gives problems! Take for instance $y = 0.25$ and there can be constructed a sequence x_M, such that $f_M(x_M) = 0.25$. This means that $\| f - f_M \|_\infty \geq 0.75 \nrightarrow 0$, if $M \to \infty$.

$\mathbb{L}_2$-norm: It will be proven that $\| f - f_M \|_2 \to 0$, if $M \to \infty$. Numerically, it goes very slow, see table 7.1.

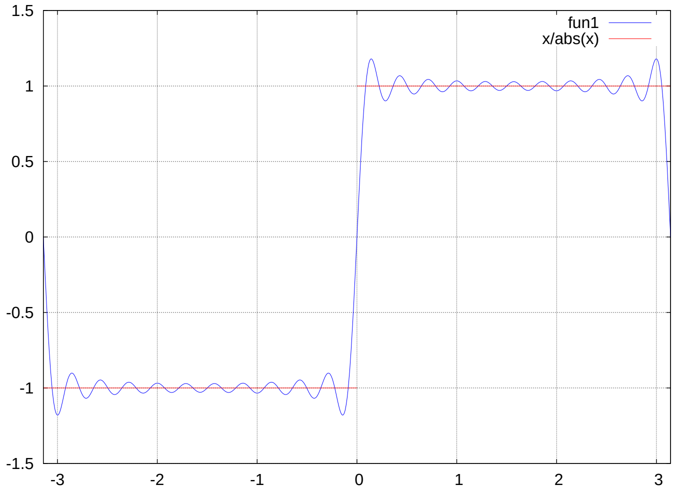


Figure 7.2 Fourier Series $M = 10$

$$\text{To get some idea: } \begin{cases} \| f - f_{50} \|_2 \approx 0.158\cdots, \\ \| f - f_{100} \|_2 \approx 0.112\cdots, \\ \| f - f_{150} \|_2 \approx 0.092\cdots. \end{cases} \tag{7.1}$$

It is interesting to look what happens in the neighbourhood of $x = 0$, the so-called Gibbs-phenomenon , see figure 7.3. More interesting things are to find in wiki-gibbs and for those, who want to have a complete lecture in Fourier-analysis, see lect-fourier.

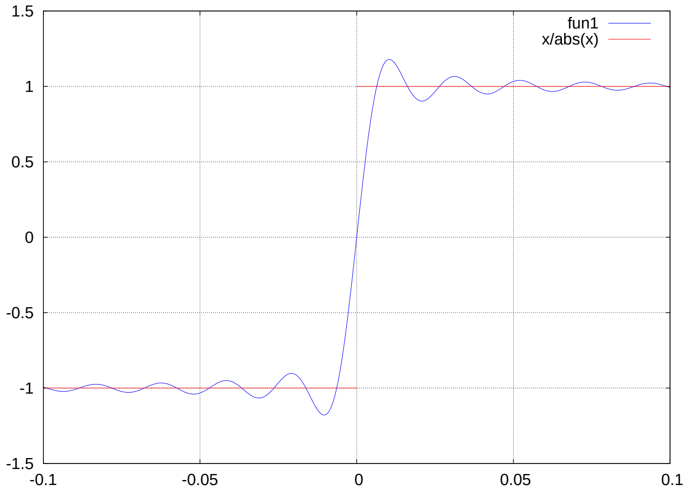


Figure 7.3 Fourier Series $M = 150$

If the coëfficients, before the sin and/or cos functions are calculated, the so-called Fourier-coefficients, there is a sequence of numbers. The functional analysis tries to find some relation between functions and such a sequences of numbers. Is such a sequence of numbers in certain sense the same as a function? Can the norm of the function be calculated by the numbers of that given sequence of numbers?
Another question could be, if the inner product between two functions f, g can be calculated by an inner product between the sequences of the Fourier-coefficients of f

and g? Is there a relation between the inner product space $\mathbb{L}_2(-\pi, \pi)$, see page 91, and the space ℓ^2, see page 102?

7.2 Sturm-Liouville Problem BUSY

In very much books an application of the functional analysis is given by the Sturm-Liouville Problem.
Look to a wave which vibrates along the direction of travel, a longitudinal wave. It is the vibration of an elastic rod, with stiffness $p(x)$ and density $r(x)$. The stiffness and the density are positive functions, so $p(x) > 0$ and $r(x) > 0$. The displacement of the rod from its equilibrium position is given by $u(x, t)$ and satisfies the wave equation

$$r(x)\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}(p(x)\,\frac{\partial u}{\partial x}).$$

If there exists a solution $u(x, t)$ and this solution can be calculated by separation of variables, for instance $u(x, t) = v(x)\,w(t)$, then

$$\frac{(\frac{\partial^2 w}{\partial t^2})}{w(t)} = \frac{\frac{\partial}{\partial x}(p(x)\,\frac{\partial v}{\partial x})}{r(x)\,v(x)} = \lambda \in \mathbb{R}.$$

Out of the physical boundary conditions

$$u(0, t) = 0 \quad \text{and} \quad u(L, t) = 0,$$

follows that

$$v(0) = v(L) = 0.$$

In this particular case, we get the Sturm-Liouville eigenvalue problem:

$$\begin{cases} -\frac{\partial}{\partial x}(p(x)\,\frac{\partial v}{\partial x}) + \lambda\, r(x)\, v = 0 \\ v(0) = v(L) = 0. \end{cases} \tag{7.2}$$

It is easily seen that the zero function, $v \equiv 0$, is a solution of 7.2. The Sturm-Liouville problem consists in finding values of λ for which there exist a nonzero solution v of 7.2. Such λ's are called eigenvalues of the Sturm-Liouville problem and the corresponding v is called an eigenfunction, or eigenvector.

7.3 General Sturm-Liouville Problem

A more general form of the Sturm-Liouville problem is

$$\begin{cases} -\frac{\partial}{\partial x}(p(x)\,\frac{\partial y}{\partial x}) + q(x)\,y = \lambda\, w(x)\, y \\ y(0)\,\cos(\alpha) - p(0)\,y'(0)\,\sin(\alpha) = 0 \\ y(L)\,\cos(\beta) - p(L)\,y'(L)\,\sin(\beta) = 0. \end{cases} \tag{7.3}$$

where $\alpha, \beta \in [0, \pi)$ and with $p(x) > 0$ and $w(x) > 0$. The used functions are assumed to be regular enough, such that there exists a solution y of the Sturm-Liouville problem, defined in 7.3.
A possible space to search for a solution is the space $\mathbb{L}_2(0, L)$, see 5.1.5. On this space can be defined a so-called weighted inner product by

$$(f, g) = \int_0^L \overline{f(x)}\, w(x) g(x)\, dx. \tag{7.4}$$

The function w is called the weight function , it has to be positive $w : [0.L] \to \mathbb{R}^+$ and the integral $\int_0^L w(x)\, dx$ has to exist.
Further the functions f, g have to be such that $\int_0^L \overline{f(x)}\, w(x) g(x)\, dx$ exists. Most of the time, there is required that

$$\int_0^L |\, f(x)\, |\; w(x)\, dx < \infty$$

and the same for g, the integrals have to be bounded. The functions f and g have to be absolute integrable , with respect to the weight w.
The linear space of functions can be made smaller to take only those functions f and g, which satisfy certain boundary conditions.

There can also be defined a linear operator

$$L(y) = \frac{1}{w(x)}(-\frac{\partial}{\partial x}(p(x)\,\frac{\partial y}{\partial x}) + q(x)\, y) \tag{7.5}$$

and problem 7.3 can be written as $L(y) = \lambda\, y$, with y in the inner product space $\mathbb{L}_2(0, L)$, with the weighted inner product defined in 7.4.

If $y \in C^2[0, L], p \in C^1[0, L]$ and $q \in C[0.L]$ then $L(y) \in C[0.L]$. And if $f, g \in C^2[0, L]$, both they satisfy he boundary conditions given in 7.3 then

$$\begin{aligned}
&(L(f), g) - (f, L(g)) = \\
&\int_0^L \overline{(-\frac{\partial}{\partial x}(p(x)\,\frac{\partial f(x)}{\partial x}) + q(x)\, f(x))}\, g(x) - \overline{f(x)}(-\frac{\partial}{\partial x}(p(x)\,\frac{\partial g(x)}{\partial x}) + q(x)\, g(x))\, dx \\
&= \int_0^L (-p(x)\frac{\partial^2 \overline{f(x)}}{\partial x^2} g(x) - \frac{\partial p(x)}{\partial x}\frac{\partial \overline{f(x)}}{\partial x} g(x) + \frac{\partial p(x)}{\partial x}\overline{f(x)}\frac{\partial g(x)}{\partial x} + p(x)\overline{f(x)}\frac{\partial^2 g(x)}{\partial x^2} = \\
&\int_0^L -\frac{\partial}{\partial x}(p(x)\,(\frac{\partial \overline{f(x)}}{\partial x}\, g(x) - \overline{f(x)}\frac{\partial g(x)}{\partial x}))dx = \\
&p(0)\,(\overline{f'(0)}\, g(0) - \overline{f(0)} g'(0)) - p(L)\,(\overline{f'(L)}\, g(L) - \overline{f(L)} g'(L)) = 0.
\end{aligned}$$

Used are the boundary conditions given in 7.3 and the fact that both functions are in $C^2[0, L]$.

This means that the linear operator L, defined in 7.5, is self-adjoint, see theorem 4.6.1. And a self-adjoint operator has real eigenvalues, see theorem 4.6.2 and the eigenvectors corresponding to two different eigenvalues are orthogonal, see also theorem 4.6.2. The operator L is selfadjoint on a subspace of $C^2[0, L]$, those functions that are in $C^2[0, L]$ and satisfy the boundary conditions given in 7.3.

7.4 Transform differential equations

7.4.1 Transformation of not Self-Adjoint to Self-Adjoint

Let L be the following differential operator

$$L(y) = a(x)\,\frac{\partial^2 y}{\partial x^2} + b(x)\,\frac{\partial y}{\partial x} + c(x)\,y, \tag{7.6}$$

with $a(x) > 0$ on $[0, L]$ and for security $a \in C^2[0, L]$, $b \in C^1[0, L]$ and $c \in C[0, L]$. This operator is formally self-adjoint if

$$\frac{\partial a}{\partial x}(x) = b(x), \tag{7.7}$$

see section 7.3.
If

$$\frac{\partial a}{\partial x}(x) \neq b(x), \tag{7.8}$$

then the the operator L can be transformed to a formally self-adjoint operator when it is multiplied by a suitable function ρ. This new operator is defined by

$$\widehat{L}(y) = \rho(x)\,L(y) = (\rho(x)\,a(x))\,\frac{\partial^2 y}{\partial x^2} + (\rho(x)\,b(x))\,\frac{\partial y}{\partial x} + (\rho(x)\,c(x))\,y \tag{7.9}$$

with

$$\frac{\partial(\rho(x)\,a(x))}{\partial x} = (\rho(x)\,b(x)).$$

The function ρ satisfies the following first-order differential equation

$$\frac{\partial \rho(x)}{\partial x}\,a(x) + \frac{\partial a(x)}{\partial x}\,\rho(x) = \rho(x)\,b(x)$$

and a solution is given by

$$\rho(x) = \frac{\gamma}{a(x)}\,\exp\Big(\int_0^x \frac{b(t)}{a(t)}dt\Big)$$

with γ a constant. If $\gamma > 0$ then is ρ a positive function and $\rho \in C^2[0.L]$. If $b(x) = \frac{\partial a}{\partial x}(x)$ then ρ reduces to a constant.
The result is that the eigenvalue equation

$$L(y) = \lambda\, y$$

can be multiplied by the positive function ρ to the equivalent equation

$$(\rho\, L)(y) = \lambda\, \rho\, y,$$

where the operator $\rho\, L$ is formally self-adjoint. See also the Sturm-Liouville eigenvalue problem given in 7.3. But be careful, ρ is function of x and $(\rho\, L)(y) \neq L(\rho\, y)$.

7.4.2 Get rid of term with $\frac{\partial}{\partial x}$

If there is looked to the zeros of a solution $v(x)$ of the differential equation

$$\frac{\partial^2 v}{\partial x^2} + b(x)\,\frac{\partial v}{\partial x} + c(x)\, v = 0, \tag{7.10}$$

it is convenient to get rid of the term with $\frac{\partial v}{\partial x}$.
Assume that $v(x) = u(x)\, w(x)$ such that

$$\frac{\partial v}{\partial x} = \frac{\partial u}{\partial x}\, w(x) + \frac{\partial w}{\partial x}\, u(x) \tag{7.11a}$$

$$\frac{\partial^2 v}{\partial x^2} = \frac{\partial^2 u}{\partial x^2}\, w(x) + 2\,\frac{\partial w}{\partial x}\,\frac{\partial u}{\partial x} + \frac{\partial^2 w}{\partial x^2}\, u(x) \tag{7.11b}$$

Substitute the results of 7.11a and 7.11b into 7.10 and rearrange the terms, such that the differential equation becomes

$$w\,\frac{\partial^2 u}{\partial x^2} + (2\,\frac{\partial w}{\partial x} + b(x)\, w)\,\frac{\partial u}{\partial x} + (\frac{\partial^2 w}{\partial x^2} + b(x)\,\frac{\partial w}{\partial x} + c(x)\, w)\, u = 0. \tag{7.12}$$

Choose the function w such that

$$2\,\frac{\partial w}{\partial x} + b(x)\, w = 0$$

this can be done by

$$w(x) = \exp\,(-\frac{1}{2}\int_0^x b(t)dt). \tag{7.13}$$

With the positive function w, out of 7.13, the original differential equation can be transformed to

$$\frac{\partial^2 u}{\partial x^2} + \rho(x)\, u = 0,$$

with

$$\rho(x) = (c(x) - \frac{1}{4}\, b^2(x) - \frac{1}{2}\,\frac{\partial b(x)}{\partial x}).$$

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 to 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.

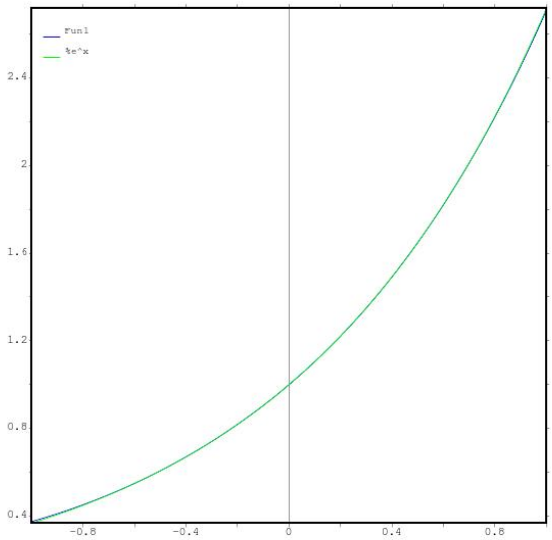


Figure 8.1 Taylor Series of $\exp(t)$ with $N = 4$.

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 $\{| \, t^n \, |\}_{n \in \mathbb{N}}$. 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

$$\begin{aligned}
(\sin(a\,t), \sin(b\,t)) &= \frac{(b+a)\sin(b-a) - (b-a)\sin(b+a)}{(b+a)(b-a)}, \\
(\sin(a\,t), \sin(a\,t)) &= \frac{2a - \sin 2a}{2a}, \\
(\cos(a\,t), \cos(b\,t)) &= \frac{(b+a)\sin(b-a) + (b-a)\sin(b+a)}{(b+a)(b-a)}, \\
(\cos(a\,t), \cos(a\,t)) &= \frac{2a + \sin 2a}{2a},
\end{aligned}$$

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 8.1.1 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 $x \in \mathrm{span}(\{e_1, e_2, \cdots, e_n\})$ 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 $x \notin \mathrm{span}(\{e_1, e_2, \cdots, e_n\})$, 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 $-1 \leq t \leq 1$.

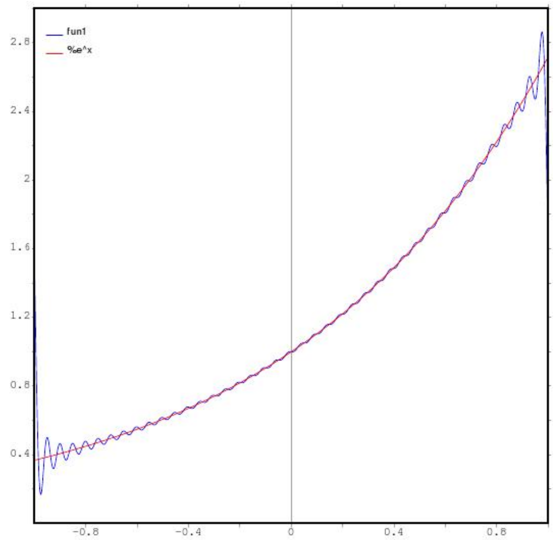


Figure 8.2 Fourier Series of $\exp(t)$ with $N = 40$.

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.

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

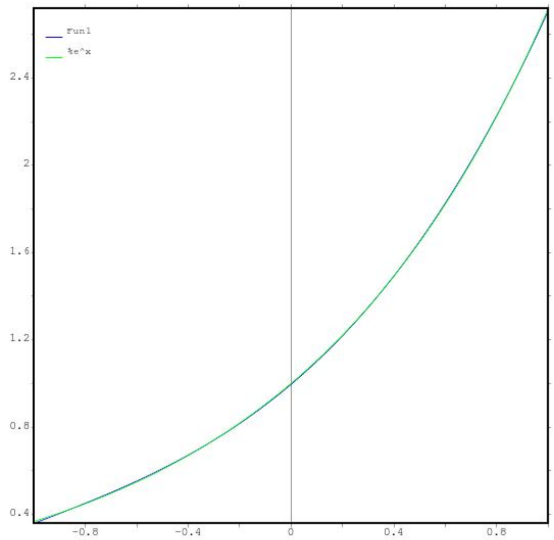


Figure 8.3 Legendre approximation of $\exp(t)$ with $N = 4$.

$$g(t) := \lim_{N \to \infty} g_N(t)$$

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 1 in the proof of Theorem 5.2.2, ($\mathcal{P}(\mathbb{N}) \sim \mathbb{R}$)

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 to 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

$$h : \mathbb{R} \to \mathcal{P}(\mathbb{N}).$$

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[10]

$$k : \mathcal{P}(\mathbb{N}) \to (0, 1),$$

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
$k(\{2, 3, 4, 7\}) = 0.37773373333333$. With the map f, see 8.4, there can be defined an

[10] To the open interval $(0, 1) \subset \mathbb{R}$.

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 $\mathcal{P}(\mathbb{N})$, 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 7 in the proof of Theorem 5.2.2, (σ-algebra and measure)

A measure, see Definition 8.3.2 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 8.3.1 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 3.3.1. Compare for instance TS 3 with σA 3. In TS 3 is spoken about union of an arbitrary collection of sets out of the topology and in σA 3 is spoken about a countable union of subsets out of the σ-algebra.

Remark 8.3.1 Some remarks on σ-algebras:

1. By σA 1: $\Omega \in \Sigma$, so by σA 2: $\emptyset \in \Sigma$.
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 8.3.2 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. $\mu(\emptyset) = 0$.
2. μ is countable additive or σ-additive, that means that for a countable sequence $\{M_n\}_n$ of disjoint elements out of Σ

$$\mu(\bigcup_n M_n) = \sum_n \mu(M_n). \square$$

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 8.3.3 An outer 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. $\mu^*(\emptyset) = 0$.
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 [extras=, see here][Aliprantis-2].
If $\mathcal{F}$ is a collection of subsets of a set Ω containing the empty set and let $\mu : \mathcal{F} \to [0, \infty]$ be a set function such that $\mu(\emptyset) = 0$. For every subset A of Ω the outer measure generated by the set function μ is defined by

Definition 8.3.4

$$\mu^*(A) = \inf\{\sum_{i=1}^{\infty} \mu(A_i) \mid \{A_i\} \text{ a sequence of } \mathcal{F} \text{ with } A \subseteq \bigcup_{i=1}^{\infty} A_i\}.$$

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. $f_n \to g$ 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 $f_n \to g$ a.e.
5. $f_n \downarrow g$ a.e. if $f_n \geq f_{n+1}$ a.e. for all n and $f_n \to g$ 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 8.3.1 and see also Remark 8.3.1. 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 8.3.5 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 8.3.5, 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 8.3.1 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 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 $\{\alpha_i \mid i \in I\}$. 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 $\mu(\emptyset) = 0$ and μ is σ-additive , what means that if $A = \bigcup_{i=1}^{\infty} A_i$ with $A_i \cap A_j = \emptyset$ 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 $f : \Omega \rightarrow \mathbb{R}$, 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 5.1.8. In fact it is enough to prove that $\mathcal{B}(\Omega, \Sigma)$ is a closed linear subspace of $B(\Omega)$, see Theorem 3.8.1.

If f, g are bounded on Ω then the functions $f + g$ and αf, 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 (αf) are σ-measurable?

Theorem 8.5.1 If f, g are σ-measurable functions and $\alpha \in \mathbb{R}$ then is

1. $f + g$ is σ-measurable and
2. αf is σ-measurable.

Proof 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 8.3.1.
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 $0^{-1}([a, \infty)) = \emptyset$ 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 8.5.1 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 8.5.2 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 Since $\lim_{n \to \infty} f_n = f$ a.e., the set $A = \{x \in \Omega \mid \lim_{n \to \infty} f_n(x) \neq f(x)\}$ 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 $\mu : \Sigma \rightarrow \mathbb{R}$, 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 8.5.1 Let $\{A_i\}_{i \in \mathbb{N}}$ be a countable set of pairwise disjoint sets in Σ, i.e. $A_i \cap A_j = \emptyset$ for $i \neq j$ with $i, j \in \mathbb{N}$.

1. A set function $\mu : \Sigma \rightarrow \mathbb{R}$ 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 $\mu : \Sigma \rightarrow \mathbb{R}$ 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 8.5.2 The variational norm of any $\mu \in ba(\Sigma)$ is defined by

$$\| \mu \|_n^v = \sup \left\{ \sum_{i=1}^{n} |\mu(A_i)| \,\middle|\, n \in \mathbb{N}, A_1, \cdots, A_n \text{ are pairewise disjoint members of } \Sigma \right\},$$

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 $\mu(\emptyset) = 0$, see Definition 8.3.2.
Measures have positive values, a generalisation of it are signed measures, which are allowed to have negative values, (Yeh, 2006, page 202).

Definition 8.5.3 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. $\mu(\emptyset) = 0$,
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 $\sum_{k=1}^{N} \mu(E_k) = \mu(\overset{N}{\underset{k=1}{\cup}}(E_k))$.
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 $\sum_{k \in \mathbb{N}} \mu(E_k) = \mu(\underset{k \in \mathbb{N}}{\cup}(E_k))$.

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 8.5.4 Given a signed measure space (Ω, Σ, μ). The total variation of μ is a positive measure $|\mu|$ on Σ defined by

$$|\mu|(E) = \sup\left\{ \sum_{k=1}^{n} |\mu(E_k)| \Big| E_1, \cdots, E_n \subset \Sigma, E_i \cap E_j = \emptyset (i \neq j), \overset{n}{\underset{k=1}{\cup}} E_k = E, n \in \mathbb{N} \right\}. \ \square$$

Important: $\| \mu \|_n^v = |\mu|(\Omega)$.

It is not difficult to prove that the expression $\| \cdot \|_n^v$, given in Defintion 8.5.2 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 $E \in \Sigma$:

$$\begin{aligned} |\mu_i(E) - \mu_j(E)| &\leq |\mu_i - \mu_j|(E) \\ &\leq |\mu_i - \mu_j|(X) \\ &= \| \mu_i - \mu_j \|_n^v < \epsilon. \end{aligned} \qquad (8.6)$$

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 $i > c_k(\epsilon)$, $|(\mu_i(E_k) - \lambda(E_k)| < \frac{\epsilon}{N}$ and that for $1 \leq k \leq N$. (N is finite!)
Hence for $i > \max\{k_0(\epsilon), c_1(\epsilon), \cdots, c_N(\epsilon)\}$, $|\sum_{k=1}^{N}(\mu_i(E_k) - \lambda(E_k))| < N\frac{\epsilon}{N} = \epsilon$, so λ is finitely additive, because

$$|\lambda(E) - \sum_{k=1}^{N} \lambda(E_k)| < 2\epsilon.$$

Remark 8.5.1 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 5.1.6:

$$|\mu_n(E) - \lambda(E)|$$
$$\leq |\mu_n(E) - \mu_m(E)| + |\mu_m(E) - \lambda(E)|$$

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 8.5.5 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 5.15. □

Theorem 8.5.3 The simple functions are closed under addition and scalar multiplication.

Proof 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 $\Omega = \bigcup_{i=1}^{M} A_i = \bigcup_{j=1}^{N} B_j$. 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 $A_i \subseteq X = \bigcup_{j=1}^{N} B_j$ and so $A_i = A_i \cap (\bigcup_{j=1}^{N} B_j) = \bigcup_{j=1}^{N} (A_i \cap B_j) = \bigcup_{j=1}^{N} C_{ij}$. On the same way $B_j = \bigcup_{j=1}^{N} C_{ij}$. 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 8.5.6 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 8.5.4 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 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 $\frac{(k-1)}{2^n} \leq s(\omega) \leq \frac{k}{2^n}$ 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 8.5.4 is only going about positive bounded measurable funtions. To obtain the result in Theorem 8.5.4 for arbitrary bounded measurable functions, there has to be made a decompostion.

Definition 8.5.7 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 8.5.5 If s and t are measurable then are also measurable $s \vee t$, $s \wedge t$, s^+, s^- and $|s|$.

Proof See Theorem 8.5.1, 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 8.5.6 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 The function s can be written as $s = s^+ - s^-$. Apply Theorem 8.5.4 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 1 is proved that $(\mathcal{B}(\Omega, \Sigma), \|\cdot\|_\infty)$ is Banach Space and in Step 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 8.5.6.
The use of Hahn-Banach, see Theorem 6.7.2, 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 8.5.7 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 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, $A_i \cap A_j = \emptyset$ 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

$$s(\omega) = \sum_{i=1}^{n} a_i \chi_{A_i}(\omega), \quad A_i \cap A_j = \emptyset,\ i \neq j.$$

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

$$u(s) = \int_{\Omega} s(\omega)\mu(d\omega) = \int_{\Omega} sd\mu.$$

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 8.5.6. 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 8.5.8 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 The norm of the functional u does not exceed the norm of the measure μ, since

$$|u(s)| = |\int_{\Omega} sd\mu| \leq \| s \|_{\infty} \| \mu \|_n^v,$$

see formula 8.9, so

$$\| u \| \leq \| \mu \|_n^v . \tag{8.12}$$

The definition of the total variation of the measure μ, see Defintion 8.5.4 gives that for any $\epsilon > 0$ there exists a finite collection of pairwise disjoint sets $\{A_1, \cdots, A_n\}$, $(A_i \cap A_j = \emptyset,\ i \neq j)$, 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 $\| s \|_{\infty} = 1$, 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

$$\| u \| \geq \| \mu \|_n^v, \tag{8.13}$$

comparing the inequalities 8.12 and 8.13, the conclusion is that

$$\| u \| = \| \mu \|_n^v .$$

□

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 Questions

In this chapter, questions are tried to be answered, if possible at several manners.

9.1 Is $\mathbb{L}_2(-\pi, \pi)$ separable? BUSY

Several things are already said about the space $\mathbb{L}_2(a, b)$, see in 5.1.5. See also in 8.1, where the answer is given to the question, if $\mathbb{L}_2(-\pi, \pi)$ is separable, but not really proved.

In this section will be looked to functions at the interval $[-\pi, \pi]$, but that is no problem. If a and b are finite, the interval $[-\pi, \pi]$ can always be rescaled to the interval $[a, b]$, by $t = a + (\frac{x + \pi}{2\pi})(b - a)$, if $-\pi \leq x \leq b$ then $a \leq t \leq b$.
Reading books or articles about quadratic integrable functions, sometimes is written $\mathbb{L}_2(-\pi, \pi)$ or $\mathcal{L}_2(-\pi, \pi)$, see section 5.1.5 for the difference between these spaces. It means, that if you take some $f \in \mathbb{L}_2(-\pi, \pi)$, that the result which is obtained for that f, is also valid to a function $g \in \mathbb{L}_2(-\pi, \pi)$ with $\| f - g \|_2 = 0$. The functions f and g are called equal *almost everywhere*.
If f should be taken out of $\mathcal{L}_2(-\pi, \pi)$. It should have meant, that an obtained result is only valid for that only function f and not for an equivalence class of functions.
But most of the time, that difference between $\mathcal{L}_2$ and $\mathbb{L}_2$ is not really made. There will be worked with some $f \in \mathbb{L}_2$, without to be aware of the fact, that f is a representant of some class of functions.

9.1.1 The space $\mathbb{L}_2(-\pi, \pi)$ is a Hilbert Space

Monkey todo!

9.1.2 Dirichlet and Fejer kernel

Theorem 9.1.1 For any real number $\alpha \neq 2 m \pi$, with $m \in \mathbb{Z}$ and $n \in \mathbb{N}$,

$$\frac{1}{2} + \sum_{k=1}^{n} \cos(k\,\alpha) = \frac{\sin((n + \frac{1}{2})\alpha)}{2 \sin(\frac{\alpha}{2})} \tag{9.1}$$

Proof The best way to start is to write the cosine functions into the complex exponential expression, the left hand side of 9.1 becomes

$$\frac{1}{2} + \sum_{k=1}^{n} \cos(k\,\alpha) = \frac{1}{2} \sum_{k=-n}^{n} \exp(k\,\alpha\,\imath) =$$

$$\frac{1}{2} \exp(-n\,\alpha\,\imath) \sum_{k=0}^{2n} \exp(k\,\alpha\,\imath) = \frac{1}{2} \exp(-n\,\alpha\,\imath) \sum_{k=0}^{2n} (\exp(\alpha\,\imath))^k$$

There holds that

$$\exp(\alpha\,\imath) = 1 \quad \text{if and only if} \quad \alpha = n\,2\,\pi \text{ with } n \in \mathbb{Z}.$$

If $\alpha \neq n\,2\,\pi$ with $n \in \mathbb{Z}$ then

$$\frac{1}{2} \exp(-n\,\alpha\,\imath) \sum_{k=0}^{2n} (\exp(\alpha\,\imath))^k = \frac{1}{2} \exp(-n\,\alpha\,\imath) \frac{1 - (\exp(\alpha\,\imath))^{2n+1}}{1 - \exp(\alpha\,\imath)} =$$

$$\left(\frac{\exp((-\alpha/2)\imath)}{\exp((-\alpha/2)\imath)}\right)\left(\frac{\exp(-n\,\alpha\,\imath) - \exp((n+1)\,\alpha\,\imath)}{1 - \exp(\alpha\,\imath)}\right)$$

With these manipulations the result out of the right-hand side of 9.1 is obtained. □

The expression

$$\mathrm{D}_n(\alpha) = \frac{1}{2\,\pi} \sum_{k=-n}^{n} \exp(k\,\alpha\,\imath) \tag{9.2}$$

is called the Dirichlet kernel . The Dirichlet kernel is an even function of α and periodic with a period of $2\,\pi$. Look to Theorem 9.1.1 and it is easily seen that the Dirichlet kernel is represented by

$$\mathrm{D}_n(\alpha) = \begin{cases} \dfrac{\sin((n + \frac{1}{2})\,\alpha)}{2\,\pi\,\sin(\frac{1}{2}\,\alpha)} & \text{if } \alpha = n\,2\,\pi \text{ with } n \notin \mathbb{N} \cup 0, \\ \frac{2\,n+1}{2\,\pi} & \text{if } \alpha = n\,2\,\pi \text{ with } n \in \mathbb{N} \cup 0, \end{cases} \tag{9.3}$$

see also Figure 9.1.

Of interest is the fact that

$$\int_{-\pi}^{\pi} \mathrm{D}_n(\alpha)\,d\alpha = \frac{1}{\pi} \int_{-\pi}^{\pi} \left(\frac{1}{2} + \sum_{k=1}^{n} \cos(k\,\alpha)\right) d\alpha = 1 \text{ for every } n \in \mathbb{N}. \tag{9.4}$$

To calculate the mean value of $\mathrm{D}_0(t), \cdots, \mathrm{D}_{(n-1)}$, the value of the summation of certain sin-functions is needed. The result looks very similar to the result of Theorem 9.1.1.

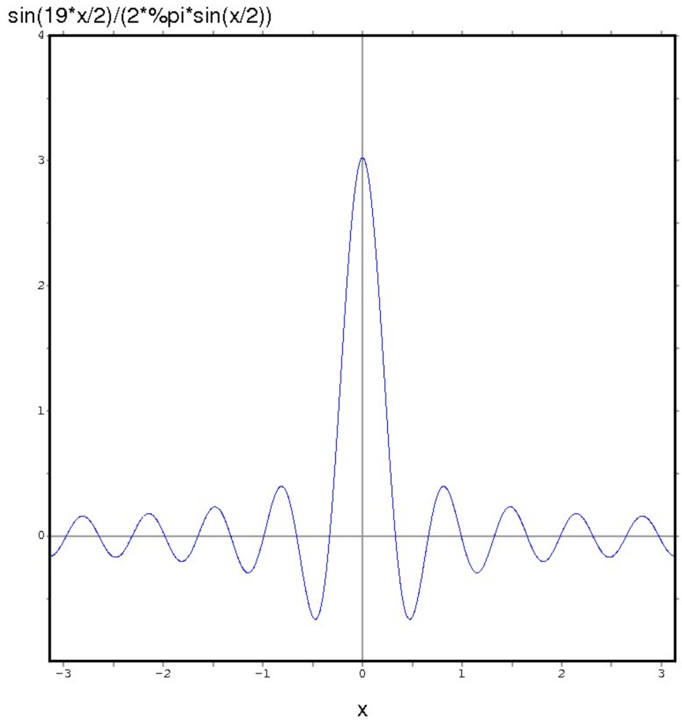


Figure 9.1 Dirichlet kernel D_9, see 9.3.

Theorem 9.1.2 For any $n \in \mathbb{N}$ and $\alpha \neq n\,2\,\pi$ with $n \in \mathbb{N}$

$$\sum_{k=0}^{(n-1)} \sin\left((k + \frac{1}{2})\,\alpha\right) = \frac{(\sin(\frac{n}{2}\,\alpha))^2}{\sin(\frac{1}{2}\,\alpha)}. \tag{9.5}$$

Proof The proof is almost similar as the proof of Theorem 9.1.1. The start is to write the sine functions into the complex exponential expression,

$$\frac{\exp(\frac{\alpha}{2}\imath)}{(2\,\imath)} \sum_{k=0}^{(n-1)} (\exp(\alpha\,\imath))^k - \frac{\exp(-\frac{\alpha}{2}\imath)}{(2\,\imath)} \sum_{k=0}^{(n-1)} (\exp(-\alpha\,\imath))^k =$$

$$\frac{1}{(\exp(-\frac{\alpha}{2}\imath))\,(2\,\imath)} \frac{1-(\exp(\alpha\,\imath))^n}{1-\exp(\alpha\,\imath)} - \frac{1}{(\exp(\frac{\alpha}{2}\imath))\,(2\,\imath)} \frac{1-(\exp(-\alpha\,\imath))^n}{1-\exp(-\alpha\,\imath)} =$$

$$\frac{(1-(\exp(\alpha\,\imath))^n)}{(2\imath)(-2\imath)(\sin(\alpha/2))} - \frac{(1-(\exp(-\alpha\,\imath))^n)}{(2\imath)(2\imath)(\sin(\alpha/2))} =$$

$$\frac{1}{(4\,\sin(\alpha/2))}\,(2-\exp(n\,\alpha\,\imath) - \exp(-n\,\alpha\,\imath)) = \frac{(\sin(\frac{n}{2}\,\alpha))^2}{\sin(\alpha/2)}.$$

□

The mean value of the first n Dirichlet kernels is called the Fejer kernel

$$F_n(\alpha) = \frac{1}{n} \sum_{k=0}^{(n-1)} D_k(\alpha) = \frac{1}{2\,\pi\,n} \left(\frac{\sin\left(\frac{n}{2}\,\alpha\right)}{\sin\left(\alpha/2\right)} \right)^2, \tag{9.6}$$

see also Figure 9.2.

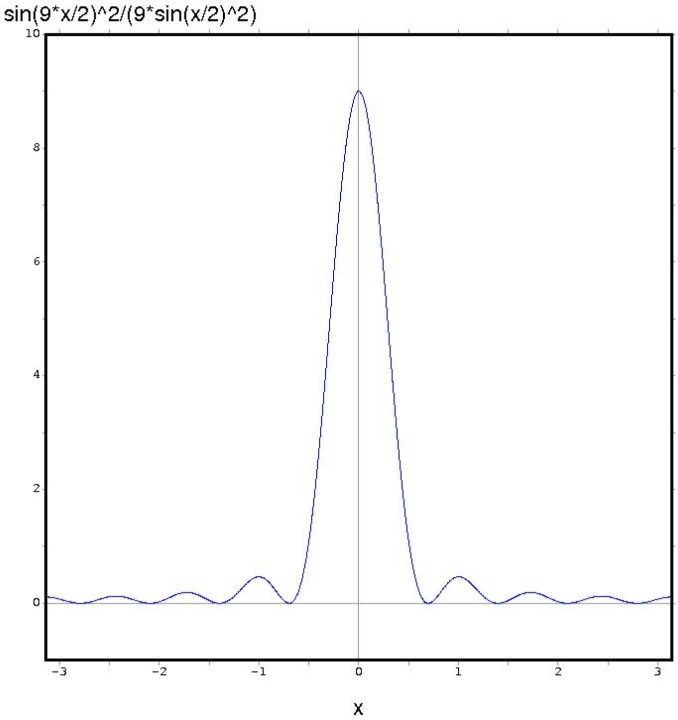


Figure 9.2 Fejer kernel F_9, see 9.6.

The Fejer kernel is positive unlike the Dirichlet kernel. Both kernels are even functions and of interest is also to note that

$$\int_{-\pi}^{\pi} F_n(\alpha)\,d\alpha = \frac{1}{n} \sum_{k=0}^{(n-1)} \int_{-\pi}^{\pi} D_n(\alpha)\,d\alpha = 1.$$

9.1.3 The functions $e_k(x) = \frac{1}{\sqrt{2\,\pi}} \exp\left(\imath\,k\,x\right)$, with $k \in \mathbb{Z}$, are complete

In first instance there will be looked to continuous functions u, on the interval $[-\pi, \pi]$ with the condition that $u(-\pi) = u(\pi)$. There will be made use of theorem 3.10.6, if of the set of functions e_k, $k \in \mathbb{Z}$ is said, to describe a whole space, the orthoplement (see definition 3.9.4) of the span of that set has to be $\{0\}$.

Theorem 9.1.3 Let $W = \{u \in C[-\pi, \pi] \mid u(-\pi) = u(\pi)\}$ and $e_k(x) = \frac{1}{\sqrt{2\pi}} \exp(\imath k x)$ with $k \in \mathbb{Z}$.

a. The functions out of the set $\{e_k\}_{k\in\mathbb{Z}}$ are orthonormal in $\mathbb{L}_2(-\pi, \pi)$.

b. There holds for $u \in W$ that

$$\int_{-\pi}^{\pi} u(x)\, e_k(x)\, dx = 0 \quad \forall k \in \mathbb{Z} \Longrightarrow u = 0.$$

Proof

a. The elements e_k are elements of $\mathbb{L}_2(-\pi, \pi)$, that is not difficult to control. Calculate the inner products between the elements e_k and the result is that

$$\int_{-\pi}^{\pi} e_k(x)\, \overline{e_j}(x) dx = \int_{-\pi}^{\pi} \frac{1}{2\pi} \exp \imath(k - j) x dx = \begin{cases} 0 & \text{if } k \neq j, \\ 1 & \text{if } k = j. \end{cases}$$

So the functions e_k with $k \in \mathbb{Z}$ are orthonormal in $\mathbb{L}_2(-\pi, \pi)$.

b. □

There is obtained some result but not for the whole $\mathbb{L}_2(-\pi, \pi)$! There is only looked to the set of functions $W \subset \mathbb{L}_2(-\pi, \pi)$. How to extend the obtained result to the space $\mathbb{L}_2(-\pi, \pi)$?

Theorem 9.1.4 The set $\{e_k\}_{k\in\mathbb{Z}}\}$ form an orthonormal base of $\mathbb{L}_2(-\pi, \pi)$.

Proof □

9.2 How to generalise the derivative?

In an explicit example is tried to give an idea about how the differential operator $\frac{d}{dt}$ can be generalised. To get more information, see wiki-Distribution-math.
The differential operator works well on the space $C^1[0, 1]$, the space of the continuously differentiable functions at the interval $[0, 1]$,

$$\frac{d}{dt} : C^1[0, 1] \to C[0, 1],$$

if $f \in C^1[0, 1]$ then

$$\frac{d}{dt}(f)(t) = f'(t) \in C[0, 1], \tag{9.7}$$

sometimes also called the strong derivative of the function f. What can be done if $f \notin C^1[0,1]$? For instance, if f is only continuous or worse?
Let $\phi \in C^\infty[0,1]$ and $\phi(0) = 0 = \phi(1)$, so ϕ can be continuously differentiated as many times as wanted. Use partial integration and

$$\int_0^1 \frac{d}{ds}(f)(s)\,\phi(s)\,ds = -\int_0^1 f(s)\frac{d}{ds}(\phi)(s)ds + [f(s)\,\phi(s)]_0^1 =$$
$$-\int_0^1 f(s)\frac{d}{ds}(\phi)(s)ds + f(1)\,\phi(1) - f(0)\,\phi(0) =$$
$$-\int_0^1 f(s)\frac{d}{ds}(\phi)(s)ds,$$

the result becomes

$$\int_0^1 \frac{d}{ds}(f)(s)\,\phi(s)\,ds = -\int_0^1 f(s)\frac{d}{ds}(\phi)(s)ds. \tag{9.8}$$

Take f such that the last integral is proper defined. The Riemann-integral gives much problems, but the use of Lebesgue integration prevents much problems, see section 5.1.6.
The idea is, that if you take all kind of functions ϕ and calculate the integral on the right hand side of formula 9.8, that you get an idea of the shape or the behaviour of $\frac{d}{ds}(f)$. May be $\frac{d}{ds}(f)$ is not a function anymore, in the sence still used.
These functions ϕ are, most of the time, chosen such that ϕ is not equal to zero at some compact interval. Because ϕ is very smooth, the derivatives of ϕ at the boundaries of such a compact interval are also zero. These functions ϕ are often called test functions. The expression $\frac{d}{ds}(f)$ as defined in 9.8 is sometimes also called the weak derivative of f.

Example 9.2.1 Define the following function

$$f(t) = \begin{cases} 0 & \text{if } 0 < t < \frac{1}{2}, \\ \frac{1}{2} & \text{if } t = 0, \\ 1 & \text{if } \frac{1}{2} < t < 1. \end{cases} \tag{9.9}$$

Calculate

$$\int_0^1 \frac{d}{ds}(f)(s)\,\phi(s)\,ds = -\int_0^1 f(s)\frac{d}{ds}(\phi)(s)ds =$$

$$-\int_{\frac{1}{2}}^1 \frac{d}{ds}(\phi)(s)ds = -[\phi(s)]_{\frac{1}{2}}^1 =$$

$$\phi(\frac{1}{2}) = \delta_{\frac{1}{2}}(\phi) \Rightarrow "f' = \delta_{\frac{1}{2}}",$$

for $\delta_{\frac{1}{2}}$, see wiki-Dirac-delta-function for the Dira delta function. The internet site, which is referred to, gives also a nice animation of the Dira delta function. The Dira delta function is not really a function, but it is a linear functional, see definition 6.2, at the space of testfunctions. A linear functional at the space of testfunctions is called a distribution.
The Dira delta function looks like a pulse at some point of the domain, in this case $t = \frac{1}{2}$. The function f comes from the left get a pulse at $t = \frac{1}{2}$ and jumps from 0 to 1, So the derivative of the discontinuous function f is given some sense. □

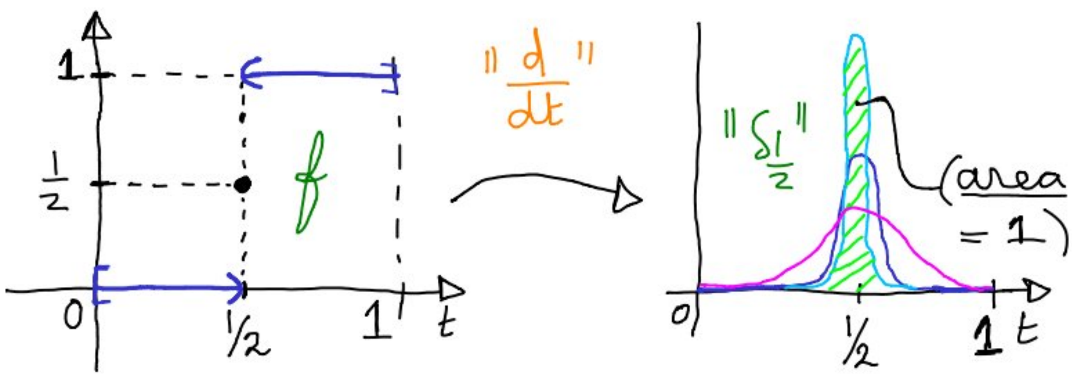


Figure 9.3 Derivative of function f, see 9.9.

Here another easier example of a function, which is not differentiable in the usual way.

Example 9.2.2 Define the following function

$$f(x) = \begin{cases} -x & \text{if } -1 \leq x < 0, \\ 0 & \text{if } x = 0, \\ x & \text{if } 0 < x \leq 1. \end{cases} \tag{9.10}$$

Calculate

$$\int_{-1}^{1} \frac{d}{ds}(f)(s)\,\phi(s)\,ds = -\int_{-1}^{1} f(s)\frac{d}{ds}(\phi)(s)ds =$$
$$-\,(\int_{-1}^{0} (-s)\,\frac{d}{ds}(\phi)(s)ds + \int_{0}^{1} (s)\,\frac{d}{ds}(\phi)(s)ds) =$$
$$-\,(-\int_{-1}^{0} (-1)\,\phi(s)\,ds - [(-s)\,\phi(s)]_{-1}^{0} + -\int_{0}^{1} (1)\,\phi(s)\,ds - [(s)\,\phi(s)]_{0}^{1}) =$$
$$(\int_{-1}^{0} (-1)\,\phi(s)\,ds + \int_{0}^{1} (1)\,\phi(s)\,ds = \int_{-1}^{1} sgn(s)\,\phi(s)\,ds,$$

with

$$sgn(s) = \begin{cases} -1 & \text{if } s < 0, \\ 0 & \text{if } s = 0, \\ 1 & \text{if } s > 0, \end{cases}$$

the Sign-function, see wiki-Sign-function. The result is that

$$\frac{d}{dx}(f)(x) = sgn(x).$$

□

10 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.

10.1 Background Theorems

In these lecture notes is made use of important theorems. Some theorems are proved, other theorems will only be mentioned.
In some sections, some of these theorems are very much used and in other parts they play an important rule in the background. They are not mentioned, but without these theorems, the results could not be proved or there should be done much more work to obtain the mentioned results.

BcKTh 1:Lemma of Zorn , see wiki-lemma-Zorn.

Theorem 10.1.1 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. □

BcKTh 2:Baire's category theorem , see wiki-baire-category.

11 Exercises-All

11.1 Lecture Exercises

Ex-1: If $f(x) = f(y)$ for every bounded linear functional f on a Normed Space X. Show that $x = y$.
Sol- 1

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.
Sol- 2

Ex-3: Let (X,d) be a Metric Space and A a subset of X. Show that $x_0 \in \overline{A} \Leftrightarrow d(x_0,A) = 0$.
Sol- 3

Ex-4: Let X be a Normed Space and X is reflexive and separable. Show that X'' is separable. Sol- 4

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

$$\lim_{n\to\infty} \sqrt[n]{u_n} = \lim_{n\to\infty} \frac{u_{n+1}}{u_n}.$$

Sol- 5

Ex-6: Given is some sequence $\{u_n\}_{n\in\mathbb{N}}$ and $\lim_{n\to\infty} u_n = L$ exists then also

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} u_i = L.$$

Sol- 6

11.2 Revision Exercises

Ex-1: What is a "norm"? Sol- 1.

Ex-2: What does it mean if a Metric Space is "complete"? Sol- 2.

Ex-3: Give the definition of a "Banach Space" and give an example. Sol- 3.

Ex-4: What is the connection between bounded and continuous linear maps? Sol- 4.

Ex-5: What is the Dual Space of a Normed Space? Sol- 5.

Ex-6: What means "Hilbert space"? Give an example. Sol- 6.

11.3 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- 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- 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- 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- 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- 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- 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- 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: $\int \arctan(a\,x)\,dx = x\arctan(a\,x) - \frac{1}{2a}\ln(1 + (a\,x)^2) + C$, 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- 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- 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- 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- 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- 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- 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- 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- 15.

Ex-16: Define the operator $A : \ell^2 \to \ell^2$ 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 $f_n : [-1, +1] \to \mathbb{R}$, $n \in \mathbb{N}$,

$$f_n(x) = \sqrt{(\frac{1}{n} + x^2)}.$$

a. Show that $f_n : [-1, +1] \to \mathbb{R}$ is differentiable and calculate the derivative $\frac{\partial f_n}{\partial x}$.

b. Calculate the pointwise limit $g : [-1, +1] \to \mathbb{R}$, i.e.

$$g(x) = \lim_{n\to\infty} f_n(x),$$

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- 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- 16.

Ex-21: For $n \in \mathbb{N}$, define the following functions $g_n, h_n, k_n : \mathbb{R} \rightarrow \mathbb{R}$

$$\begin{cases} g_n(x) = \sqrt{n} & \text{if } 0 < x < \frac{1}{n} \text{ and } 0 \text{ otherwise,} \\ h_n(x) = n & \text{if } 0 < x < \frac{1}{n} \text{ and } 0 \text{ otherwise,} \\ k_n(x) = 1 & \text{if } n < x < (n+1) \text{ and } 0 \text{ otherwise.} \end{cases}$$

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 $S : \mathbb{R}^\infty \to \mathbb{R}^\infty$ 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, $N(T) = \{g \in L_2[-1,1] \,|\, T(g) = 0\}$ and the range of T, $R(T) = \{T(g) \,|\, g \in L_2[-1,1]\}$.
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- 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- 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 $f_n : [-1,+1] \to \mathbb{R}$, $n \in \mathbb{N}$,

$$f_n(t) = \begin{cases} 1 & \text{for } -1 \le t \le -\frac{1}{n} \\ -n\,x & \text{for } -\frac{1}{n} < t < \frac{1}{n} \\ -1 & \text{for } \frac{1}{n} \le t \le 1. \end{cases}$$

Is the sequence $\{f_n\}_{n\in\mathbb{N}}$ a Cauchy sequence in the Banach space $(C[-1,1], \| \cdot \|_\infty)$?

Solution, see Sol- 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- 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- 22.

Ex-31: Consider the operator $A : \ell^2 \rightarrow \ell^2$ 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- 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- 24.

11.4 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 6.7.5. 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.

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

$$\lim_{n\to\infty} \frac{u_n}{n} = L.$$

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 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 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 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 5.a gives that

$$\lim_{n\to\infty} \frac{s_n}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} u_i = L.$$

11.5 Solutions Revision Exercises

Sol-1: See definition 3.7.1.

Sol-2: A Metric Space is complete if every Cauchy sequence converges in that Metric Space.

Sol-3: A Banach Space is a complete Normed Space, for instance $C[a,b]$ with the $\|\cdot\|_\infty$ norm.

Sol-4: Bounded linear maps are continuous and continuous maps are bounded, see theorem 4.2.1.

Sol-5: See the section 6.5.

Sol-6: Fot the definition, see 3.10.1. An example of a Hilbert Space is the ℓ^2, see 5.2.4.

11.6 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. $\langle Tf, g \rangle = \lim_{R\to\infty} \int_0^R f(\frac{x}{5}) g(x) dx = \lim_{R\to\infty} \int_0^{\frac{R}{5}} f(y) g(5\,y)\, 5 dy = \langle f, T^* g \rangle$, 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 $\| T^*(g) \|^2 = 25 \lim_{R\to\infty} \int_0^{5\,R} \frac{1}{5} |g(y)|^2 dy = 5 \| g \|^2$ 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. $\| (Tf) \| \leq \| A \|_\infty \| f \|$, 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 span$(1, t)$.

Sol-5: a. See figure 11.1.

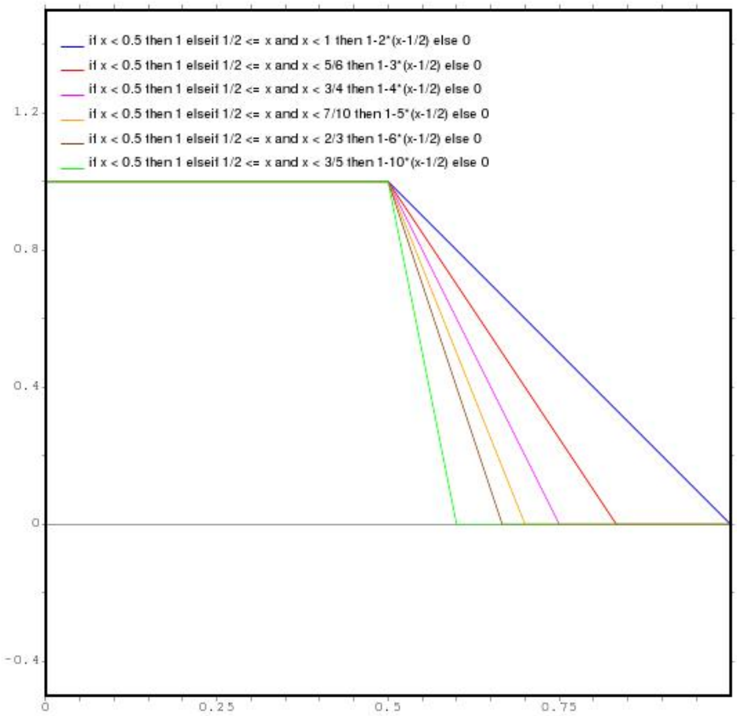


Figure 11.1 f_n certain values of n

b. Take x fixed and let $n \rightarrow \infty$, then

$$f(t) = \begin{cases} 1 & \text{if } 0 \le t \le \frac{1}{2} \\ 0 & \text{if } 0\frac{1}{2} < t \le 1. \end{cases}$$

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 $|f_n(t) - f_m(t)| \le 1$ it is easily seen that

$$\int_0^1 |f_n(t) - f_m(t)|^2 dt \le \int_0^1 |f_n(t) - f_m(t)| dt \le \frac{1}{2}(\frac{1}{n} - \frac{1}{m})$$

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- 4. Working in the Inner Product Space $L_2[-\pi, \pi]$. Project $\sin(\frac{t}{2})$ on the 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 11.2.

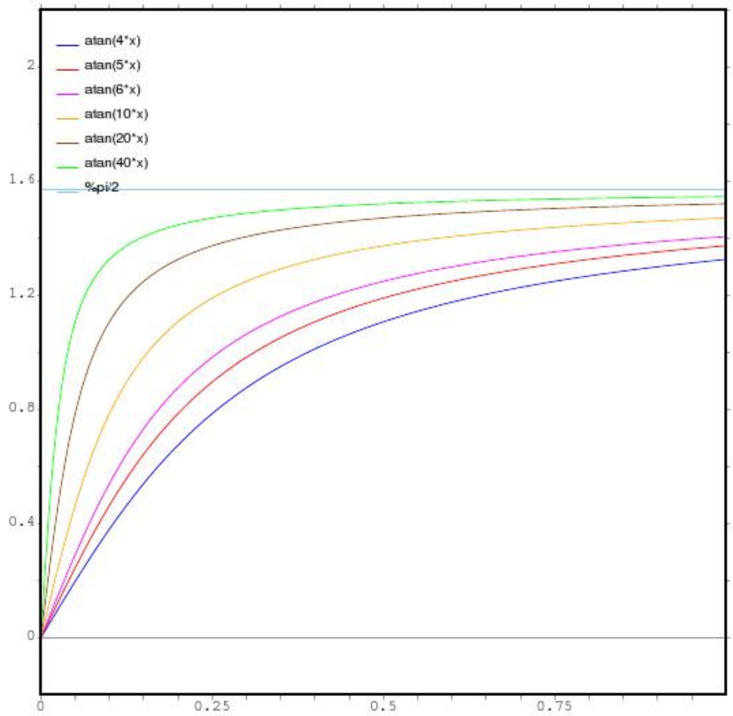


Figure 11.2 f_n certain values of n

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.

$$\lim_{n\to\infty} (\int_0^1 |\frac{\pi}{2} - \arctan(n\,x)|dx) = \lim_{n\to\infty} (\frac{\log(1+n^2) - 2\,n\,\arctan(n) + \pi\,n}{2\,n}) = 0.$$

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

$$|(T f)(t)| \leq 2 \parallel (1 + t^2) \parallel_\infty \parallel f \parallel_\infty = 4 \parallel f \parallel_\infty,$$

the length of the integration interval is ≤ 2. Hence, $\parallel (T f) \leq 4 \parallel f \parallel_\infty$ 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 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. $|F(x)| \leq 1 \parallel x \parallel_\infty$, may be too coarse. Also is valid $|F(x)| \leq \int_0^1 \tau \, d\tau \parallel x \parallel_\infty = \frac{1}{2} \parallel x \parallel_\infty$.

c. $F(1) = \frac{1}{2}$.

d. With part 12.b and part 12.c it follows that $\parallel F \parallel = \frac{1}{2}$.

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: $\parallel x_n - x_m \parallel^2 + \parallel x_n + x_m \parallel^2 = 2 (\parallel x_n \parallel^2 + \parallel x_m \parallel^2)$ and $(x_n + x_m) = 2 (\frac{1}{2} x_n + \frac{1}{2} x_m)$. So $\parallel x_n - x_m \parallel^2 = 2 (\frac{1}{2} x_n + \frac{1}{2} x_m) - 4 \parallel \frac{1}{2} x_n + \frac{1}{2} x_m \parallel^2$. M is convex so $\frac{1}{2} x_n + \frac{1}{2} x_m \in M$ and $4 \parallel \frac{1}{2} x_n + \frac{1}{2} x_m \parallel^2 \geq 4 d^2$.

Hence, $\parallel x_n - x_m \parallel^2 \leq 2 (d^2 - \parallel x_n \parallel^2) + 2 (d^2 - \parallel x_m \parallel^2) \rightarrow 0$ 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 3.5, 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 $\parallel x \parallel_1 = \sum_{i=1}^\infty |x_i| < \infty$ 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!

$$\parallel Tf \parallel^2 = \int_0^1 |(Tf)(x)|^2 dx = \int_0^1 \frac{1}{\sqrt{(4x)}}\, f^2(x)\, dx,$$

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- 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

$$|h_n(f)| = |\int_{-1}^{+1} (\tau)^n\, f(\tau)\,d\tau| \leq (\int_{-1}^{+1} (\tau)^{2\,n} d\tau)^{\frac{1}{2}}\,(\int_{-1}^{+1} f^2(\tau)\,d\tau)^{\frac{1}{2}}$$

$$= (\frac{2}{2\,n + 1})^{\frac{1}{2}}\ \| f \|\,.$$

c.

$$\lim_{n\to\infty} \| h_n \| \leq \frac{\sqrt{2}}{\sqrt{(n + 1)}} \to 0$$

if $n \to \infty$, so $\lim_{n\to\infty} \| h_n \| = 0$.

Sol-18: a. $f(t) = \frac{1}{2}(f(t) - f(-t)) + \frac{1}{2}(f(t) + f(-t))$, 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.

$$|f(x)| = |\int_{-1}^{0} x(t)\,dt - \int_{0}^{1} x(t)\,dt| \leq |\int_{-1}^{0} x(t)\,dt| + |\int_{0}^{1} x(t)\,dt|$$
$$\leq \| x \|_\infty + \| x \|_\infty = 2 \| x \|_\infty,$$

so f is bounded and so continuous.

c. Take $x_n : [-1,+1] \rightarrow \mathbb{R}$, $n \in \mathbb{N}$,

$$x_n(t) = \begin{cases} 1 & \text{for } -1 \leq t \leq -\frac{1}{n} \\ -n\,x & \text{for } -\frac{1}{n} < t < \frac{1}{n} \\ -1 & \text{for } \frac{1}{n} \leq t \leq 1. \end{cases}$$

then $f(x_n) = 2 - \frac{1}{n}$. 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 selfadjoint, 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:

$$\sin(x) - \sin(1-x) = (\alpha-\beta)(u(x) - u(1-x))$$
$$\sin(x) + \sin(1-x) = (\alpha+\beta)(u(x) + u(1-x)).$$

(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 $S = \text{span}((1, 0, -1, 0, 1, 0, -1, \cdots), (0, 1, 0, -1, 0, 1, 0, \cdots))$.
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 3.9.1. The most difficult one is may be condition 3.9.1(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. $\| q_1 \| = 1$, $\| q_2 \| = 2$, $\| q_3 \| = 2$.

Figures

References

1 Aliprantis, C. (2006). *Infinite Dimensional Analysis (A Hitchhiker's Guide)*, *pdf-file*. Springer.

2 Bouziad, A. and Clabrix, J. (1993). *Théorie de la Mesure et de l'Integration*, *djvu-file*. j'Universite the Rouen.

3 Dunford, N. and Schwartz, T. (1958-71). *Linear Operators (3 parts)*, *djvu-file*. Wiley.

4 Griffel, D. (1981). *Applied Functional Analysis*. Ellis Horwood.

5 Jain, P., Ahuja, O. and K.Ahmad (1996). *Functional Analysis*. New Age Publishers.

6 Kreyszig, E. (1978). *Introductory Functional Analysis with Applications*, *djvu-file*. Wiley.

7 Melrose, R. (2004). *Lecture Notes for 18.155, fall 2004*, *pdf-file*..

8 Morrison, T. (2001). *Functional Analysis, Introduction to Banach Space Theory*, *pdf-file*. Wiley.

9 Pugachev, V. and Sinitsyn, I. (1999). *Lectures on Functional Analysis and Applications*. World Scientific.

10 Rana, I. (2004). *An Introduction to Measure And Integration*. Alpha Science International.

11 Rao, K. P. S. B. and Rao, M. B. (1983). *Theory of charges: a study of finitely additive measures*, *pdf-file*. Academic Press.

12 Srivastava, S. (1998). *A Course on Borel Sets*, *pdf-file*. Springer.

13 Sutherland, W. (1975). *Introduction to metric and topological spaces*. Oxford University Press.

14 Swartz, C. (1994). *Measure, Integration and Function Spaces*. World Scientific.

15 Yeh, J. (2006). *Real Analysis, Theory of Measure and Integration*, *djvu-file*. World Scientific.

Index