TMA4145 – Linear Methods

Franz Luef

2010 *Mathematics Subject Classification.* Primary

ABSTRACT. These notes are for the course TMA4145 – Linear Methods at NTNU and cover the following topics: Linear and metric spaces. Completeness, Banach spaces and Banach's fixed point theorem. Picard's theorem. Linear transformations. Inner product spaces, projections, and Hilbert spaces. Orthogonal sequences and approximations. Linear functionals, dual space, and Riesz' representation theorem. Spectral theorem, Jordan canonical form, and matrix decompositions.

CHAPTER 1

Real numbers and its topology

1.1. Real Numbers

1.1.1. Notation. We introduce some notation:

(1) $\mathbb{N} = \{1, 2, 3, ...\}$ the set of natural numbers,
(2) $\mathbb{Q} = \{p/q : p, q \in \mathbb{Z}\}$ the set of rational numbers,
(3) $\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$ the set of integers.
(4) For real numbers a, b with $a < b$ we denote by $[a, b]$ the closed bounded interval, and by (a, b) the open bounded interval. The length of these bounded intervals is $b - a$.

1.1.2. Real numbers. The set $\mathbb{Q}$ of rational numbers does not contain all the numbers one encounters in geometry or analysis, e.g. $x^2 - 5 = 0$ has no ratinonal solution or Euler's number e is an irrational number.

For the moment we do not introduce the set of real number $\mathbb{R}$ in an informal manner. In the chapter on metric spaces $\mathbb{R}$ will be constructed as the completion of $\mathbb{Q}$, as was originally done by A. L. Cauchy.

Real numbers may be realized as points on a line, the real line, where the irrational numbers correspond to the points that are not given by rational numbers $\mathbb{R}\backslash\mathbb{Q}$.

The real numbers have the Archimedean property:

LEMMA 1.1 (Archimedean property). *For any $x, y \in \mathbb{R}$ there exists a natural number n such that $nx > y$.*

As a consequence we deduce a close relation between $\mathbb{Q}$ and $\mathbb{R}$.

PROPOSITION 1.1.1. *For $x, y \in \mathbb{R}$ with $x < y$ there exists a $r \in \mathbb{Q}$ such that $x < r < y$.*

PROOF. Goal: Find $m, n \in \mathbb{Z}$ such that

$$x < \frac{m}{n} < y. \tag{1.1}$$

First step: Choose the denominator of n large such that there exists an $m \in \mathbb{Z}$ such that $x \in (\frac{m-1}{n}, \frac{m}{n})$ are separating x and y. The Archimedean property of $\mathbb{R}$ allows us to a $n \in \mathbb{N}$ with this property. More concretely, we pick $n \in \mathbb{N}$ large enough such that $1/n < y - x$ or equivalently

$$x < y - \frac{1}{n} \tag{1.2}$$

Second step: Inequality (1.1) is equivalent to $nx < m < ny$. From the first step we have n already chosen. Now we choose $m \in \mathbb{Z}$ to be the smallest integer greater than nx. In other words, we pick $m \in \mathbb{Z}$ such that $m-1 \leq nx < m$. Thus we have $m-1 \leq nx$, i.e. $m \leq nx+1$. By inequality (1.2)

$$m \leq nx + 1 < n(y - \frac{1}{n}) + 1 = ny,$$

hence we have $m < ny$, i.e. $m/n < y$. Once more by (1.2) we have $x \leq m/n$. These two inequalites yield the desired assertion: $x < m/n < y$. □

In an similiar manner one may deduce the statement for irrational numbers.

Proposition 1.1.2. *For $x, y \in \mathbb{R}$ with $x < y$ there exists a $r \in \mathbb{R}\backslash\mathbb{Q}$ such that $x < r < y$.*

Proof. Pick your favorite irrational number, a popular choice is $\sqrt{2}$. Then by the density of the rational numbers there exists a rational number $r \in (x/\sqrt{2}, y/\sqrt{2})$. Hence $r\sqrt{2} \in (x,y)$. Note that $r\sqrt{2}$ is an irrational number in (x,y) that completes our argument. □

The absolute value of $x \in \mathbb{R}$, denoted by $|x|$, is defined by

$$|x| = \begin{cases} -x & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ x & \text{if } x > 0. \end{cases}$$

Note that $|x| = \max\{x, -x\}$. We define the positive, x^+ and negative part, x^- of $x \in \mathbb{R}$:

$$x^+ = \max\{x, 0\}, \quad \text{and} \quad x^- = \max -x, 0,$$

so we have $x = x^+ - x^-$ and $|x| = x^+ + x^-$.
For $x, y \in \mathbb{R}$ we measure the distance between x and y in $\mathbb{R}$ by

$$d(x,y) = |x-y|, \tag{1.3}$$

the standard distance. By definition of d we have $d(x,y) = d(y,x)$.

Lemma 1.2 (Triangle inequality). *For x, y in $\mathbb{R}$ we have $|x+y| \leq |x| + |y|$.*

Proof. For all $x \in \mathbb{R}$ we have $x \leq |x|$ and thus for $x, y \in \mathbb{R}$ we obtain $x + y \leq |x+y|$. By definition of $|.|$ we also get that $-x-y \leq |x| + |y|$. Thus we have proved the desired assertion. □

The triangle inequality has numerous consequences, such as

$$||x| - |y|| \leq |x-y|. \tag{1.4}$$

The triangle inequality for $x = y + x - y$ yields $|x| - |y| \leq |x-y|$, and the interchange of x and y, i.e. $y = x + y - x$ gives $-(|x| - |y|) \leq |x-y|$. Hence we have the desired assertion.
We introduce two crucial notions: the infimum and supremum of a set. First we provide some preliminaries.

Definition 1.1.3. Let A be a subset of $\mathbb{R}$

- If there exists $M \in \mathbb{R}$ such that $a \leq M$ for all $a \in A$, then M is an *upper bound* of A. We call A *bounded above.*
- If there exists $m \in \mathbb{R}$ such that $m \leq a$ for all $a \in A$, then m is a *lower bound* of A.

- If there exist lower and upper bounds, then we say that A is *bounded*. We call A *bounded below*.

DEFINITION 1.1.4 (Infimum and Supremum). Let A be a subset of $\mathbb{R}$.

- If m is a lower bound of A such that $m \geq m'$ for every lower bound m', then m is called the *infimum* of A, denoted by $m = \inf A$. Furthermore, if $\inf A \in A$, then we call it the minimum of A, $\min A$.
- If M is an upper bound of A such that $m' \geq M$ for every upper bound M', then M is called the *supremum* of A, denoted by $M = \sup A$.Furthermore, if $\sup A \in A$, then we call it the maximum of A, $\max A$.

Note that the infimum of a set A, as well as the supremum, are unique. The elementary argument is left as an exercise.
If $A \subset \mathbb{R}$ is not bounded above, then we define $\sup A = \infty$. Suppose that a subset A of $\mathbb{R}$ is not bounded below, then we assign $-\infty$ as its infimum.
We state a different formulation of the notions $\inf A$ and $\sup A$ that is just a reformulation of the definition.

LEMMA 1.3. *Let A be a subset of $\mathbb{R}$.*

- *Suppose A is bounded above. Then $M \in \mathbb{R}$ is the supremum of A if and only if the following two conditions are satisfied:*
  (1) *For every $a \in A$ we have $a \leq M$.*
  (2) *Given $\varepsilon > 0$, there exists $a \in A$ such that $M - \varepsilon < a$.*
- *Suppose A is bounded below. Then $m \in \mathbb{R}$ is the infimum of A if and only if the following two conditions are satisfied:*
  (1) *For every $a \in A$ we have $m \leq a$.*
  (2) *Given $\varepsilon > 0$, there exists $a \in A$ such that $a < m + \varepsilon$.*

LEMMA 1.4. *Suppose A is a bounded subset of A. Then* $\inf A \leq \sup A$

For $c \in \mathbb{R}$ we define the *dilate* of a set A by $cA := \{b \in \mathbb{R} : \; b = ca \quad \text{for } a \in A\}$.

LEMMA 1.5 (Properties). *Suppose A is a subset of $\mathbb{R}$.*

(1) *For $c > 0$ we have $\sup cA = c \sup A$ and $\inf cA = c \inf A$.*
(2) *For $c < 0$ we have $\sup cA = c \inf A$ and $\inf cA = c \sup A$.*
(3) *Suppose A is contained in a subset B. If $\sup A$ and $\sup B$ exist, then $\sup A \leq \sup B$. In words, making a set larger, increases its supremum.*
(4) *Suppose A is contained in a subset B. If $\inf A$ and $\inf B$ exist, then $\inf A \geq \inf B$. In words, making a set smaller increases its infimum.*
(5) *Suppose $A \subset B$ are non-empty subsets of $\mathbb{R}$ such that $x \leq y$ for all $x \in A$ and $y \in B$. Then $\sup A \leq inf B$.*
(6) *If A and B are non-empty subsets of $\mathbb{R}$, then $\sup(A + B) = \sup A + \sup B$ and $\inf(A + B) = \inf A + \inf B$*

PROOF. (1) We prove that $\sup cA = c \sup A$ for positive c. Suppose $c > 0$. Then $cx \leq M \Leftrightarrow x \leq M/c$. Hence M is an upper bound of cA if and only if M/c is an upper bound of A. Consequently, we have the desired result.
(2) Analogously to (i).
(3) Since $\sup B$ os an upper bound of B, it is also an upper bound of A, i.e. $\sup A \leq \sup B$.
(4) Analogously to (iii).

(5) Since $x \leq y$ for all $x \in A$ and $y \in B$, y is an upper bound of A. Hence $\sup A$ is a lower bound of B and we have $\sup A \leq \inf B$.
(6) By definition $A + B = \{c : c = a + b \,\text{for some}\, a \in A, b \in B\}$ and thus $A+B$ is bounded above if and only if A and B are bounded above. Hence $\sup(A + B) < \infty$ if and only if $\sup A$ and $\sup B$ are finite. Take $a \in A$ and $b \in B$, then $a + b \leq \sup A + \sup B$. Thus $\sup A + \sup B$ is an upper bound of $A + B$:
$$\sup(A + B) \leq \sup A + \sup B.$$
The reverse direction is a little bit more involved. Let $\varepsilon > 0$. Then there exists $a \in A$ and $b \in B$ such that
$$a > \sup A - \varepsilon/2, \quad b > \sup B - \varepsilon/2.$$
Thus we have $a+b > \sup A+\sup B-\varepsilon$ for every $\varepsilon > 0$, i.e. $\sup(A+B) \geq \sup A + \sup B$.

The other statements are assigned as exercises. □

A property of utmost importance is the *completeness* of the real numbers.

THEOREM 1.6. *Let A be a non-empty subset of $\mathbb{R}$ that is bounded above. Then there exists a supremum of A. Equivalently, if A is a non-empty subset of $\mathbb{R}$ that is bounded below, then A has an infimum.*

We have noted above that the supremum of a bounded above set is unique. A different form to express the completeness property of $\mathbb{R}$ is to consider the set of all upper bounds of a bounded above set A and the Theorem asserts that this set of upper bounds has a least element.

One reason for the relevance of the notions of supremum and infimum is in the formulation of properties of functions.

DEFINITION 1.1.5. Let f be a function with domain X and range $Y \subseteq \mathbb{R}$. Then
$$\sup_X f = \sup\{f(x) : x \in X\}, \quad \inf_X f = \inf\{f(x) : x \in X\}.$$
If $\sup_X f$ is finite, then f is bounded from above on A, and if $\inf_X f$ is finite we call f bounded from below. A function is bounded if both the supremum and infimum are finite.

LEMMA 1.7. *Suppose that $f, g : X \to \mathbb{R}$ and $f \leq g$, i.e. $f(x) \leq g(x)$ for all $x \in X$. If g is bounded from above, then $\sup_X f \leq \sup_A g$. Assume that f is bounded from below. Then $\inf_X f \leq \inf_X g$.*

PROOF. Follows from the definitions. □

The supremum and infimum of functions do not preserve strict inequalities. Define $f, g : [0, 1] \to \mathbb{R}$ by $f(x) = x$ and $g(x) = x + 1$. Then we have $f < g$ and
$$\sup_{[0,1]} f = 1, \quad \inf_{[0,1]} f = 0, \quad \sup_{[0,1]} g = 2, \quad \inf_{[0,1]} g = 1.$$
Hence we have $\sup_{[0,1]} f > \inf_{[0,1]} g$.

LEMMA 1.8. *Suppose f, g are bounded functions from X to $\mathbb{R}$ and c a positive constant. Then*
$$\sup_X(f + cg) \leq \sup_X f + c \sup_X g \quad \inf_X(f + cg) \geq \inf_X f + c \inf_X g.$$

The proof is left as an exercise. Try to convice yourself that the inequalities are in general strict, since the functions f and g may take values close to their suprema/infima at different points in X.

LEMMA 1.9. *Suppose f, g are bounded functions from X to $\mathbb{R}$. Then*

$$|\sup_X f - \sup_X g| \leq \sup_X |f-g|, \quad |\inf_X f - \inf_X g| \leq \sup_X |f-g|$$

LEMMA 1.10. *Suppose f, g are bounded functions from X to $\mathbb{R}$ such that*

$$|f(x) - f(y)| \leq |g(x) - g(y)| \quad \textit{for all } x, y \in X.$$

Then

$$\sup_X f - \inf_X f \leq \sup_X g - \inf_X g.$$

Recall that a sequence (x_n) of real numbers is an ordered list of numbers x_n, indexed by the natural numbers. In other words, (x_n) is a function f from $\mathbb{N}$ to $\mathbb{R}$ with $f(n) = x_n$. Hence we may define the if a sequence (x_n) is *bounded from above*, *bounded from below* and *bounded* as a special case of the above definitions, i.e. if there eixts $M \in \mathbb{R}$ such that $x_n \leq M$ for all $n \in \mathbb{N}$, if there exists $m \in \mathbb{R}$ such that $x_n \geq m$ for all $n \in \mathbb{N}$ and if there exist m, M such that $m \leq x_n \leq M$.

We define the *lim sup* and *lim inf* of a sequence (x_n). These notions reduce quebtions about the convergence of a sequence to ones about monotone sequences. We introduce two sequences associated to (x_n) by taking the supremum and infimum, respectively of the tails of $((x_k)_{k\geq n})_k$:

$$y_n = \sup\{x_k : k \geq n\}, \quad z_n = \inf\{x_k : k \geq n\}.$$

The sequences (y_n) and (z_n) are monotone sequences, because the supremum and infimum are taken over smaller sets for increasing n. Moreover, (y_n) is monotone decreasing and (z_n) is monotone decreasing. Hence the limits of these sequences exist:

$$\limsup_{n\to\infty} x_n := \lim_{n\to\infty} y_n = \inf_{n\in\mathbb{N}} (\sup_{k\geq n} x_k),$$
$$\liminf_{n\to\infty} x_n := \lim_{n\to\infty} z_n = \sup_{n\in\mathbb{N}} (\inf_{k\geq n} x_k).$$

We allow $\limsup$ and $\liminf$ to be $+\infty$ and $-\infty$. Note that we have $z_n \leq y_n$ and so by taking the limit as $n \to \infty$

$$\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n$$

. We illustrate these notions with some examples.

EXAMPLES 1.1.6. Consider the sequences.

(1) $(x_n) = \left((-1)^{n+1}\right)$ has $\limsup x_n = 1$ and $\liminf x_n = -1$.
(2) $(x_n) = (n^2)$ has $\limsup x_n = \infty$ and $\liminf x_n = \infty$.
(3) $(x_n) = (2 - 1/n)$ has $\limsup x_n = 2$ and $\liminf x_n = 2$.

EXERCISE 1.1.7. *Let (x_n) and (y_n) be sequences in $\mathbb{R}$.*

(1) $\liminf(x_n + y_n) \geq \liminf x_n + \liminf y_n$,
(2) $\limsup(x_n + y_n) \leq \limsup x_n + \limsup y_n$,
(3) $\limsup(-x_n) = -\liminf x_n$ *and* $\liminf(-x_n) = -\limsup x_n$.

Note that for convergent sequences $\limsup$ and $\liminf$ are finte and equal. We recommend to prove this property.

Proposition 1.1.8. *Let (x_n) be a sequence in $\mathbb{R}$. Then (x_n) converges if and only if* $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$.

Note that a sequence diverges to ∞ if and only if $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = \infty$ and that it diverges to $-\infty$ if and only if $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = -\infty$.

These considerations suggests that for non-convergent seqences the difference $\liminf_{n\to\infty} x_n - \limsup_{n\to\infty} x_n$ measures the size of the oscillations in the sequene.

A central notion in analysis is the notion of a Cauchy sequence of objects, here we define it for real numbers.

Definition 1.1.9. A sequence (x_n) in $\mathbb{R}$ is a *Cauchy sequence* if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $|x_m - x_n|\varepsilon$ for all $m, n \geq N$.

A theorem of utmost importance is that every Cauchy sequence converges to a real number.

Theorem 1.11. *A sequence (x_n) converges in $\mathbb{R}$ if and only if it is a Cauchy sequence.*

Proof. **One direction:** Suppose (x_n) converges to a real number x. Then for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $|x_n - x| < \varepsilon/2$ for all $n > N$. Hence by the triangle inequality we have

$$|x_n - x_m| \leq |x_n - x| + |x - x_m| \quad for\ m, n > N,$$

i.e (x_n) is a Cauchy sequence.

Other direction: Suppose that (x_n) is a Cauchy sequence. Then there exists $N_1 \in \mathbb{N}$ such that $|x_m - x_n| < 1$ for all $m, n > N_1$, and that for $n > N_1$ we have

$$|x_n| \leq |x_n - x_{N_1}| + |x_{N_1+1}| \leq 1 + |x_{N_1+1}|.$$

Hence a Cauchy sequence is bounded with $|x_n| \leq \max\{|x_1|, ..., |x_{N_1}|, 1 + |x_{N_1+1}|\}$ and $\limsup, \liminf$ exist.
The aim is to show that $\limsup x_n = \liminf x_n$.
By the Cauchy property of (x_n) we have for a given $\varepsilon > 0$ a $N \in \mathbb{N}$ such that

$$x_n - \varepsilon < x_m < x_n + \varepsilon \quad \text{for all } m \geq n > N.$$

Consequently, we have for all $n > N$

$$x_n - \varepsilon \leq \inf\{x_m : m \geq n\} \quad and \quad \sup\{x_m : m \geq n\} \leq x_n + \varepsilon.$$

Thus we have

$$\sup\{x_m : m \geq n\} - \varepsilon \leq \inf\{x_m : m \geq n\} + \varepsilon$$

and for $n \to \infty$ we get that

$$\limsup x_n - \varepsilon \leq \liminf x_n + \varepsilon$$

for arbitray $\varepsilon > 0$ and so

$$\limsup x_n \leq \liminf x_n.$$

□

In the proof we established that Cauchy sequences are bounded. Let us record this for later use.

LEMMA 1.12. *A Cauchy sequence (x_n) in $\mathbb{R}$ is bounded.*

We define the notion of a subsequence of a sequence (x_n).

DEFINITION 1.1.10. Suppose (x_n) is a sequence in $\mathbb{R}$. Then a *subsequence* is a sequence of the form (x_{n_k}) where $n_1 < n_2 < \cdots < x_{n_k} < \cdots$.

An elementary observation is

LEMMA 1.13. *Every subsequence of a convergent sequence converges to the limit of the sequence.*

PROOF. Suppose that (x_n) is a convergent sequence with $\lim x_n = x$ and (x_{n_k}) is a subsequence. Given $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that $|x_n - x| < \varepsilon$ for all $n > N$. Since $n_k \to \infty$ as $k \to \infty$, there exists a $K \in \mathbb{N}$ such that $n_k > N$ for $k > K$, but then we have $|x_{n_k} - x| < \varepsilon$. Hence $\lim_{k\to\infty} x_{n_k} = x$. □

COROLLARY 1.1.11. *If a sequence has subsequences that converge to different limits, then the sequence diverges.*

A well-known theorem due to Buolzano and Weierstraßdeduces the convergence of a subsequence from its boundedness.

THEOREM 1.14 (Bolzano-Weierstraß). *Every bounded sequence (x_n) in $\mathbb{R}$ has a convergent subsequence.*

PROOF. Suppose that (x_n) is a bounded sequence in $\mathbb{R}$. Hence there are m and M such that

$$m = \inf_n x_n \quad M = \sup_n x_n.$$

We define the closed interval $I_0 = [m, M]$ and divide it into two closed intervals L_0, R_0:

$$L_0 = [m, (m+M)/2], \quad R_0 = [(m+M)/2, M].$$

Now, at least one of the intervals L_0, R_0 contains infinitely many terms of (x_n). Choose I_1 to be the interval that contains infinitely many terms and pick $n_1 \in \mathbb{N}$ such that $x_{n_1} \in I_1$. Divide $I_1 = L_1 \cup R_1$, again one of these intervals contains infinitely many terms of (x_n). Choose I_2 to be one of these intervals that contains infinitely many terms. We continue by dividing I_2 into two closed intervals, pick $n_2 > n_1$ such that $x_{n_2} \in I_2$. Continue in this manner we get a sequence of nested intervals (I_k) with $|I_k| = (M - m)/2^k$, and a sequence (x_{n_k}) such that $x_{n_k} \in I_{n_k}$. Given $\varepsilon > 0$. Since $|I_k| \to 0$ as $k \to \infty$, there exists a $K \in \mathbb{N}$ such that $|I_k| < \varepsilon$ for all $k > K$. Furthermore we have $|x_{n_j} - x_{n_k}|\varepsilon$ for $j, k > K$, i.e. (x_{n_k}) is a Cauchy sequence and thus converges by Theorem 1.11. □

The Bolzano-WeierstraßTheorem does not claim that the subsequence is unique, i.e. there might be convergent subsequences with different limits depending on the choice of L_k or R_k.

THEOREM 1.15. *If (x_n) is a bounded sequence in $\mathbb{R}$ such that every convergent subsequence has the same limit x, then (x_n) converges to x.*

Proof. We will show the contrapositive statement: Suppose a bounded sequence does not converge to x. Then (x_n) has a convergent subsequence with limit different from x.

If (x_n) does not converge to x, then there exists $\varepsilon_0 > 0$ such that $|x_n - x| \geq \varepsilon_0$ for infinitely many $n \in \mathbb{N}$. Hence there exists a subsequence (x_{n_k}) such that $|x_{n_k} - x| \geq \varepsilon_0$ for every $k \in \mathbb{N}$. Note that (x_{n_k}) is a bounded sequence and so by Bolzano-Weierstraßthere exists a convergent subsequence $(x_{n_{k_j}})$. If $\lim_j x_{n_{k_j}} = y$, then $|x - y| \geq \varepsilon_0$. In other words, x is not equal to y. □

1.1.3. Topology of $\mathbb{R}$. In this section we treat some basic notions of topology for the real line. Generalizations of these notions and its manifestations in normed spaces and general metric spaces are going to be the pillars of this course.

We generalize the notion of open intervals (a, b) and closed intervals $[a, b]$.

Definition 1.1.12 (Open sets). A subset O of $\mathbb{R}$ is called *open* if for every $x \in S$ there exists an open interval I contained in O with $x \in I$.

Definition 1.1.13 (Closed sets). A subset C of $\mathbb{R}$ is called *closed* if the complement $C^c = \mathbb{R}\backslash C = \{x \in \mathbb{R} : x \notin C\}$ is open.

Note that the interval (a, b) is an open set and $[a, b]$ is closed. Observe further that by definition the empty set $\emptyset$ and $\mathbb{R}$ are open and closed.

Proposition 1.1.14. *Suppose $\{I_j\}_{j\in J}$ is a collection of open intervals in $\mathbb{R}$ with non-empty intersection $\cap_{j\in J} I_j \neq \emptyset$.*

(1) *If J has finitely many elements, then $\cap_{j\in J} I_j$ is an open interval.*
(2) *$\cup_{j\in J} I_j$ is an open interval for an arbitrary index set J.*

Proof. We define open intervals $I_j = (a_j, b_j)$ for real numbers $a_j < b_j$, the interval bounds are also allowed to be $\pm\infty$, and set $I := \cup_{j\in J} I_j$.

(1) We pick a point x in $\cup_{j=1}^n I_j$ and set $a := \max\{a_j : j = 1, ..., n\}$ and $b = \min\{b_j : j = 1, ..., n\}$. If all the $a_j' s$ are $-\infty$, then $a = -\infty$, and if all the b_j's are ∞, then we have $b = \infty$.
Since $a_j < x < b_j$ for $j = 1, ..., n$ we get that $x \in (a, b)$. Furthermore, we have that $\cap_{j\in J}(a_j, b_j) = (a, b)$.
(2) We choose $x \in \cap_{j\in J} I_j$. Suppose $y \in \cup_{j\in J} I_j$. Then $y \in I_j$ for some $j \in J$. Since $x \in I_j$, the interval $(x, y) \subset I_j$ and thus in I. Hence I is the interval (a, b), where $a = \inf\{a_j : j \in J\}$ or $-\infty$ and $b = \sup\{b_j : j \in J\}$ or ∞.

□

The assumption in (i) cannot be weakend, e.g. $\cap_{n=1}^\infty(-1.n, 1/n) = \{0\}$. Hence an infinite intersection of open intervals is not necessarily an open interval. We show that the preceding statement is true for a more general class of sets, the open sets.

Proposition 1.1.15. *Let $\{O_j : j \in J\}$ be a family of open sets of $\mathbb{R}$.*

(1) *$\cap_{j=1}^n O_j$ is an open set for any $n \in \mathbb{N}$.*
(2) *$\cup_{j\in J} O_j$ is open for a general index set J.*

PROOF. (1) We set $O = \cap_{j=1}^{n} O_j$. If $x \in O$, then $x \in O_j$ for $j = 1, ..., n$. Since O_j's are open, there are open intervals $I_j \subset O_j$ containing x. Hence, we have that $\cap_{j=1}^{n} I_j \subset \cap_{j=1}^{n} O_j$, the desired assertion.

(2) Let x be in $\cup_{j\in J} O_j$. Then there exists some j such that $x \in U_j$ and thus an open interval I_j contained in U_j with $x \in I_j$ and consequently $I_j \subset O$. Hence O is an open set.

□

We are in the positon to introduce a notion of closedness between points, known as neighborhoods.

DEFINITION 1.1.16. Given $x \in \mathbb{R}$. Then a subset U of $\mathbb{R}$ is called a *neighborhood of* x if there exists an open subset O of $\mathbb{R}$ such that $x \in O \subset U$.

Due to the structure of $\mathbb{R}$ we have that U is a neighborhood of x if and only if there exists a $\delta > 0$ such that $(x - \delta, x + \delta) \subset U$.

DEFINITION 1.1.17. For a subset A we introduce some notions.

(1) The *closure* of a subset A of $\mathbb{R}$, denoted by $\overline{A}$, is the intersection of all closed sets containg A.
(2) The *interior* of a subset of A of $\mathbb{R}$, denoted by $\mathrm{int}A$, is the union of all open subsets of $\mathbb{R}$ contained in A.
(3) The boundary of a subset A of $\mathbb{R}$, denoted by $\mathrm{bd}A$, is the set $\overline{A}\backslash\mathrm{int}A$.

Note that $\mathrm{bd}A$ is a closed set and that the closure of a bounded subset of $\mathbb{R}$ is bounded, too.

Here are some useful facts.

LEMMA 1.16. *Suppose A is a subset of* $\mathbb{R}$.

(1) $\overline{A} = (\mathrm{Int}(A^c))^c$ *and* $\mathrm{int}(A) = (\overline{A^c})^c$
(2) $\mathrm{bd}A = \mathrm{bd}(A^c) = \overline{A} \cap \overline{A^c}$
(3) $\overline{A} = A \cup \mathrm{br}A = \mathrm{int}A \cup \mathrm{bd}A$

PROOF. (1) These identities are a consequence of the following general fact: B is a closed containing A if and only if B^c is open and $B^c \subset A^c$. The statement about the interior of A is the first statement for A^c instead of A.

(2) $\mathrm{bd}A = \overline{A}\backslash \int A = \overline{A} \cap (\mathrm{int}A)^c = \overline{A} \cap \overline{A^c}$, where we used (i) in the last step. Let us compute $\mathrm{bd}A^c$: $\mathrm{bd}A^c = \overline{A^c}\backslash \int A^c = \overline{A^c} \cap (\int A^c)^c = \overline{A^c} \cap \overline{A}$. Hence we have the desired assertions.

(3) First note that $\mathrm{int}A \cup \mathrm{bd}A \subset A \cup \mathrm{bd}A \subset \overline{A}$. Furthermore we have $\mathrm{int}A \cup \mathrm{bd}A = \int A \cup (\overline{A}\backslash A) = \mathrm{int}A \cup (\overline{A} \cap (\mathrm{int}A)^c) = ((\mathrm{int}A) \cup \overline{A}) \cap (\mathrm{int}A \cup (\mathrm{int}A)^c) = \overline{A}$.

□

LEMMA 1.17. *Suppose A is a subset of* $\mathbb{R}$.

(1) $\overline{A} = \{x \in \mathbb{R} :$ *every neighborhood of x intersects A*$\}$
(2) $\mathrm{int}(A) = \{x \in \mathbb{R} :$ *some neighborhood of x is contained in A*$\}$
(3) $\mathrm{bd}(A) = \{x \in \mathbb{R} :$ *every neighborhood of x intersects A and its complement*$\}$

Proof. (1) We choose an open neighborhood U of $x \in \mathbb{R}$ that does not intersect A, i.e. $A \subset U^c$. Since U^c is closed, we have that $\overline{A} \subset U^c$ and from $x \neq U^c$ we also have that $x \neq \overline{A}$. On the other hand, if $x \neq \overline{A}$, then $(\overline{A})^c$ is an open set containing x that is disjoint from A.

(2) Follows from (i) and the preceding proposition.

(3) Follows from (i), (ii) and the preceding proposition.

□

Definition 1.1.18. Let A be a subset of $\mathbb{R}$.

(1) A point $x \in A$ is *isolated* in A if there exists a neighborhood U of x such that $U \cap A = \{x\}$.

(2) A point $x \in \mathbb{R}$ is said to be an *accumulation point* of A if every neighborhood of x contains points in $A \backslash \{x\}$.

Note: Accumulation points of a set are not necessarily elements of the set. A well-known example is $A = \{1/n : n \in \mathbb{N}\}$ with 0 as accumulation point, which is clearly not in A.

The definition of an accumulation point makes only sense for sets with infinitely many elements.

Finally, an infinite closed set may not have accumulation points, e.g. $\mathbb{N} \subset \mathbb{R}$ has no accumulation points in $\mathbb{R}$.

Lemma 1.18. *A point $x \in \mathbb{R}$ is an accumulation point of A if and only if every neighborhood of x contains infinitely many points of A.*

Proof. One direction: Suppose every neighborhood of x contains infinitely many points of A, then x is an accumulation point of A.
Other direction: Suppose x is an accumulation point of A. For a neighborhood U of x, we choose $n_1 \in \mathbb{N}$ such that $(x - 1/n_1, x + 1/n_1) \subset U$. Take a point x_1 different from x in $A \cap (x - 1/n_1, x + 1/n_1)$. Now we repeat the procedure: Take $n_2 \geq n_1$ such that $x_1 \neq\in (x - 1/n_2, x + 1/n_2)$ and pick $x_2 \in A \cap (x - 1/n_2, x + 1/n_2)$ with $x_2 \neq x$. We continue in this way and get a sequence of points $(x_n) \subset A \cap U$. □

Proposition 1.1.19. *Let A be a subset of $\mathbb{R}$. Then $\overline{A} = \{isolated\ points\ of\ A\} \cup \{accumulation points of A\}$.*

Proof. Suppose $x \in \overline{A}$. Then if $x \in A$, then either x is isolated in A or every neighborhood of x contains points in A different from x. In the later case x is an accumulation point of A. Now assume $x \in \overline{A}$ and $x \neq A$. Then every neighborhood of x has a non-trivial intersection with A, and thus x is an accumulation point of A. In summary, we have that the closure of A is the union of the isolated points of A with the accumulation points of A.
For the converse we note: If x is isolated, then x is definitely in A. If x is an accumulation point of A, then $x \in \overline{A}$ □

Definition 1.1.20. A subset A of $\mathbb{R}$ is said to be *dense* in $\mathbb{R}$ if its closure is equal to $\mathbb{R}$, i.e. $\overline{A} = \mathbb{R}$.

Proposition 1.1.21. *The set of rational numbers, $\mathbb{Q}$, is dense in $\mathbb{R}$.*

Proof. For an arbitray $x \in \mathbb{R}$ we consider a neighborhood U of x. Then we know that U contains the interval $(x-\varepsilon, x+\varepsilon)$ for a sufficiently small $\varepsilon > 0$. By an earlier result we have that there exists a rational number in $(x-\varepsilon, x+\varepsilon)$. □

We also have that the set of irrational numbers is dense in $\mathbb{R}$.

The property that $\mathbb{Q}$ has only countably elements, but still is dense in $\mathbb{R}$ is a very favorable property and occurs in various other situations. We say that $\mathbb{R}$ is separable.

$\mathbb{Q}$ is a dense subset of $\mathbb{R}$ with empty interior and thus the boundary of Q is all of $\mathbb{R}$. The same is true for the set of irrational numbers.

1.1.4. Supplementary material.

Theorem 1.19 (Nested Interval Theorem). *Let $\{I_j\}_{j=1}^{\infty}$ be a sequence of closed bounded intervals in $\mathbb{R}$, such that $I_j \subset I_{j+1}$ for all $j \in \mathbb{N}$. We assume in addtion that the lengths of the intervals $|I_j|$ tends to zero. Then $I := \cap_{j \in J} I_j = \{z\}$ for some $z \in \mathbb{R}$.*

Proof. Without loss of generality we assume $I_j = [a_j, b_j]$. Then the assumptions yield that $a_1 \leq a_2 \leq \cdots \leq b_2 \leq b_1$ and that for every $\varepsilon > 0$ there exist a $j \in \mathbb{N}$ such that $b_j - a_j \leq \varepsilon$.
We set $A := \{a_j : j \in \mathbb{N}\}$ and $B := \{b_j : j \in \mathbb{N}\}$, note that $a := \sup A < \infty$ and $b = \inf B < \infty$, and $a_j \leq a \leq b_j$ for $j \in \mathbb{N}$. Hence we have $[a, b] = \cap_{j=1}^{\infty} [a_j, b_j]$ and by the assumption on the shrinking of the interval lengths we get that $a = b = z$ for some $z \in \mathbb{R}$. □

CHAPTER 2

Normed spaces and innerproduct spaces

2.1. Normed spaces and innerproduct spaces

Vector spaces formalize the notion of linear combinations of objects that might be vectors in the plane, polynomials, smooth functions, sequences. Many problems in engineering, mathematics and science are naturally formulated and solved in this setting due to their linear nature. Vector spaces are ubiquitous for several reasons, e.g. as linear approximation of a non-linear object, or as building blocks for more complicated notions, such as vector bundles over topological spaces.

In this course vector spaces are equipped with additional structures in order to measure the distance between elements and formulate convergence of sequences of elements of vector spaces, or to provide quantitative and qualitative information on operators.

A set V is a vector space if it is possible to build linear combinations out of the elements in V. More formally, on V we have the operations of addition of vectors and multiplication by scalars. The scalars will be taken from a field $\mathbb{F}$, which is either the real numbers $\mathbb{R}$ or $\mathbb{C}$. In various situations $\mathbb{F}$ might also be a finite field or a field different from $\mathbb{R}$ and $\mathbb{C}$. If it is necessary we will refer to these vector spaces as real or complex vector spaces.

Developing an understanding of these vector spaces is one of the main objectives of this course. The axioms for a vector space specify the properties that addition of vectors and scalar multiplication.

Definition 2.1.1. A *vector space* over a field $\mathbb{F}$ is a set V together with the operations of addition $V \times V \to V$ and scalar multiplication $\mathbb{F} \times V \to V$ satisfying the following properties:

(1) Commutativity: $u + v = v + u$ for all $u, v \in V$ and $(\lambda\mu v) = \lambda(\mu v)$ for all $\lambda, \mu \in \mathbb{F}$;
(2) Associativity: $(u + v) + w = u + (v + w)$ for all $u, v, w \in V$;
(3) Additive identity: There exists an element $0 \in V$ such that $0 + v = v$ for all $v \in V$;
(4) Additive inverse: For every v ? V , there exists an element w ? V such that v+w = 0;
(5) Multiplicative identity: $1v = v$ for all $v \in V$;
(6) Distributivity: $\lambda(u+v) = \lambda u+\lambda v$ and $(\lambda+\mu)u = \lambda u+\mu u$ for all $u, v \in V$ and $\lambda, \mu \in \mathbb{F}$.

The elements of a vector space are called vectors. Given $v_1, ..., v_n$ be in V and $\lambda_1, ..., \lambda_n \in \mathbb{F}$ we call the vector

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n$$

a *linear combination.*

Our focus will be on three classes of examples.

EXAMPLES 2.1.2. We define some useful vector spaces.

- **Spaces of n-tuples:** The set of tuples $(x_1, ..., x_n)$ of real and complex numbers are vector spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ with respect to componentwise addition and scalar multiplication: $(x_1, ..., x_n) + (y_1, ..., y_n) = (x_1 + y_1, ..., x_n + y_n)$ and $\lambda(x_1, ..., x_n) = (\lambda x_1, ..., \lambda x_n)$.
- The space of polynomials of degree at most n, denoted by $\mathcal{P}_n$, where we define the operations of multiplication and addition coefficient-wise: For $p(x) = a_0 + a_1 x + \cdots a_n x^n$ and $q(x) = b_0 + b_1 x + \cdots b_n x^n$ we define

$$(p+q)(x) = (a_0+b_0)+(a_1+b_1)x+\cdots(a_n+b_n)x^n \text{ and } (\lambda p)(x) = \lambda a_0+\lambda a_1 x+\cdots\lambda a_n x^n$$

  for $\lambda \in \mathbb{F}$.

  The space of all polynomials $\mathcal{P}$ is the vector space of polynomials of arbitrary degrees.
- **Sequence spaces**: s denotes the set of sequences, c the set of all convergent sequences, c_0 the set of all convergent sequences tending to 0, c_f the set of all sequences with finitely many non-zero elements.
- **Function spaces:** The set of continuous functions $C(I)$ on an interval of $\mathbb{R}$, popular choices for I are $[0, 1]$ and $\mathbb{R}$. We define addition and scalar multiplication as follows: For $f, g \in C(I)$ and $\lambda \in \mathbb{F}$

$$(f+g)(x) = f(x) + g(x) \quad \text{and } (\lambda f)(x) = \lambda f(x).$$

  We denote by $C^{(n)}(I)$ the space of n-times continuously differentiable functions on I and the space $C^\infty(I)$ of smooth functions on I is the space of functions with infinitely many continuous derivatives. More generally, the set $\mathcal{F}(X)$ of functions from a set X to $\mathbb{F}$ is a vector space for the operations defined above. Note that $\mathcal{F}(\{1, 2, ..., n\})$ is just $\mathbb{F}^n$ and hence the first class of examples.

There are relations between the vector spaces in the aforementioned list. We start with clarifying their inclusion properties.

DEFINITION 2.1.3. A subset W of a vector space V is called a *subspace* if any linear combination of vectors of W is itself a vector in W.

If W is a subspace of V, then addition and scalar multiplication restricted to W, gives W the structure of a vector space.

Here are some examples of vector subspaces: $\mathcal{P}_n \subset \mathcal{P} \subset \mathcal{F}$, $C^\infty(I) \subset C^{(n)}(I) \subset C(I)$, $c_f \subset c_0 \subset c \subset s$. We define the linear span, span$W$, of a subset M of a vector space V to be the intersection of all subspaces of V containing M.

2.1.1. Normed spaces. The norm on a general vector space generalizes the notion of the length of a vector in $\mathbb{R}^2$ and $\mathbb{R}^3$.

DEFINITION 2.1.4. A *normed space* $(X, \|.\|)$ is a vector space X together with a function $\|.\| : X \to \mathbb{R}$, the *norm* on X, such that for all $x, y \in X$ and $\lambda \in \mathbb{R}$:

(1) *Positivity:* $0 \leq \|x\| < \infty$ and $\|x\| = 0$ if and only if $x = 0$;
(2) *Homogeneity:* $\|\lambda x\| = |\lambda| \|x\|$;
(3) *Triangle inequality:* $\|x + y\| \leq \|x\| + \|y\|$.

Normed spaces have a rich structure.

PROPOSITION 2.1.5. *Let $(X, \|.\|)$ be a normed space. Then $d : X \times X \to \mathbb{R}$ defined by $d(x, y) = \|x - y\|$ satisfies for all $x, y, z \in X$ (i) $d(x, y) \geq 0$ and $d(x, x) = 0$ if and only if $x = 0$ (positivity); (ii) $d(x, y) = d(y, x)$ (symmetry); (iii) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).*

The function $d(x, y) = \|x-y\|$ on the vector space X is an example of a distance function on X, aka as a metric. We will later discuss such distance functions on a general set.

PROOF. The properties (i)-(iii) are direct consequences of the axioms for a norm. In particular, (i) follows from property (1) of a norm, (ii) is derived from property (ii) of a norm for $\lambda = -1$ and (iii) is deduced from property (3) of a norm. □

The metric d on X is also compatible with the linear structure of a vector space:

- *Translation invariance:* $d(x + z, y + z) = d(x, y)$ for all $x, y, z \in X$;
- *Homogeneity:* $d(\lambda x, \lambda y) = |\lambda| d(x, y)$ for all $x, y \in X$ and scalars $\lambda \in \mathbb{R}$.

The metric d on X gives us a way to generalize intervals in $\mathbb{R}$, called balls.

DEFINITION 2.1.6. For $r > 0$ and $x \in X$ we define the *open ball* $B_r(x)$ of radius r and center x as the set

$$B_r(x) = \{y \in X : \|x - y\| < r\},$$

and the *closed ball* $\overline{B}_r(x)$ of radius r and center x as

$$\overline{B}_r(x) = \{y \in X : \|x - y\| \leq r\}.$$

The translation invariance and the homogeneity imply that the ball $B_r(x)$ is the image of the unit ball $B_1(0)$ centered at the origin under the affine mapping $f(y) = ry + x$.

The balls $B_r(x)$ have another peculiar feature. Namely, these are convex subsets of X.

DEFINITION 2.1.7. Let X be a vector space.

- For two points $x, y \in X$ the *interval* $[x, y]$ is the set of points $\{z|\, z = \lambda x + (1 - \lambda) y \ 0 \leq \lambda \leq 1\}$.
- A subset E of X is called *convex* if for any two points $x, y \in E$ the interval $[x, y]$ is also in E.

The notion of convexity is central to the theory of vector spaces and enters in an intricate manner in functional analysis, numerical analysis, optimization, etc. .

LEMMA 2.1. *Let $(X, \|.\|)$ be a normed vector space. Then the unit ball $B_1(0) = \{x \in X | \, \|x\| \leq 1\}$ is a convex set.*

PROOF. For $x, y \in B_1(0)$ we have that $\|\lambda x + (1-\lambda)y\| \leq |\lambda|\|x\| + |1-\lambda|\|y\| = 1$, because $\|x\|, \|y\|$ are both less than or equal to 1. Thus $\lambda x + (1-\lambda)y \in B_1(0)$. □

The real numbers with the absolute value is a normed space $(\mathbb{R}, |.|)$ and the open ball $B_r(x)$ is the open interval $(x-r, x+r)$ and $\overline{B}_r(x)$ is the closed interval $[x-r, x+r]$.

A fundamental class of metric spaces is $\mathbb{R}^n$ with the ℓ^p-norms.

DEFINITION 2.1.8. For $p \in [1, \infty)$ we define the ℓ^p-norm $\|.\|_p$ on $\mathbb{R}^n$ by assigning to $x = (x_1, ..., x_n) \in \mathbb{R}^n$ the number $\|x\|_p$:

$$\|x\|_p = (|x_1|^p + |x_2|^p + \cdots |x_n|^p)^{1/p}$$

. For $p = \infty$ we define the ℓ^∞-norm $\|.\|_\infty$ on $\mathbb{R}$ by

$$\|x\|_\infty = \max |x_1|, ..., |x_n|.$$

The notation for $\|.\|_\infty$ is justified by the fact that it is the limit of the $\|.\|_p$-norms.

LEMMA 2.2. *For $x \in \mathbb{R}^n$ we have that*

$$\|x\|_\infty = \lim_{p \to \infty} \|x\|_p.$$

Some inequalities enter the stage: Hölder's inequality and Young's inequality. For $p \in (1, \infty)$ we define its *conjugate* q as the number such that

$$\frac{1}{p} + \frac{1}{q} = 1.$$

If $p = 1$, then we define its conjugate q to be ∞ and if $p = \infty$ then $q = 1$.

LEMMA 2.3 (Young's inequality). *For $p \in (1, \infty)$ and q its conjugate we have*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q},$$

for $a, b \geq 0$.

PROOF. Consider the function $f(x) = x^{p-1}$ and integrate this with respect to x from zero to a. Now take the inverse of f given by $f^{-1}(y) = y^{q-1}$ and integrate it from zero to V. Then the sum of these two integrals always exceeds the product ab, but the integrals are a^p/p and b^q/q. Hence we have established the desired inequality. □

A consequence of Young's inequality is Hölder's inequality.

LEMMA 2.4. *Suppose $p \in (1, \infty)$ and $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ are vectors in $\mathbb{R}^n$. Then*

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

PROOF. Set $a_i = |x_i|/(\sum_{i=1}^n |x_i|^p)^{1/p}$ and $b_i = |y_i|/(\sum_{i=1}^n |y_i|^q)^{1/q}$. Then we have $\sum_i a_i^p = 1$ and $\sum_i b_i^q = 1$. By Young's inequality

$$\sum_{i=1}^n |x_i||y_i| \leq (\sum_{i=1}^n |x_i|^p)^{1/p}(\sum_{i=1}^n |y_i|^q)^{1/q}.$$

□

PROPOSITION 2.1.9. *The space $\mathbb{R}^n$ with the ℓ^p-norm $\|.\|_p$ is a normed space for $p \in [1, \infty]$.*

As an exercise I propose to draw the unit balls of $(\mathbb{R}^2, \|.\|_1)$, $(\mathbb{R}^2, \|.\|_2)$ and $(\mathbb{R}^2, \|.\|_\infty)$.

PROOF. First we show that ℓ^p is a vector space for $p \in [1, \infty)$: For $\lambda \in \mathbb{F}$ and $x \in \ell^p$ we have $\lambda x \in \ell^p$. One has to work a little bit to see that for $x, y \in \ell^p$ also $x + y \in \ell^p$:

$$\begin{aligned}
\|x+y\|_p^p &= \sum_{n=1}^\infty |x_n + y_n|^p \\
&\leq \sum_{n=1}^\infty |2\max\{|x_n|, |y_n|\}|^p \\
&= 2^p \sum_{n=1}^\infty |\max\{|x_n|, |y_n|\}|^p \\
&\leq 2^p(\sum_{n=1}^\infty |x_n|^p + \sum_{n=1}^\infty |y_n|^p) = 2^p(\|x\|_p^p + \|y\|_p^p) < \infty.
\end{aligned}$$

Positivity and homogeneity are consequences of the corresponding properties of the absolute value of a real number. The triangle inequality is the non-trivial assertion that we split up in three cases $p = 1$, $p = \infty$ and $p \in (1, \infty)$. Let $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ be points in $\mathbb{R}^n$.

(1) For $p = 1$ we have

$$\|x+y\|_1 = |x_1 + y_1| + \cdots + |x_n + y_n| \leq |x_1| + |y_1| + \cdots + |x_n| + |y_n| \leq \|x\|_1 + \|y\|_1$$

.

(2) For $p = \infty$ the argument is similar:

$$\begin{aligned}
\|x+y\|_\infty &= \max\{|x_1 + y_1|, ..., |x_n + y_n|\} \\
&= \max\{|x_1| + |y_1|, ..., |x_n| + |y_n|\} \\
&= \max\{|x_1|, ..., |x_n|\} + \max\{|y_1|, ..., |y_n|\} = \|x\|_\infty + \|y\|_\infty.
\end{aligned}$$

(3) The general case $p \in (1, \infty)$: The triangle inequality in this case is also known as *Minkowski's* inequality. We deduce it from Hölder's inequality

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

Dividing by $\|x+y\|_p^{1/q}$ and using $1 - 1/q = 1/p$ we arrive at Minkowski's inequality:

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p.$$

□

A natural generalization of the normed spaces $(\mathbb{R}^n, \|.\|_p)$ is to replace tuples of finite length with ones of infinite length $x = (x_1, x_2, ...)$ with $x_i \in \mathbb{R}$, i.e. $(\mathbb{R}^\infty, \|.\|_p)$. The standard notation for these normed spaces is $(\ell^p, \|.\|_p)$ because these are special classes of the Lebesgue spaces $L^p(\mathbb{N}, d\mu)$ for the counting measure. One often refers to these spaces as "little L^p"-spaces.

Example 2.1.10. For $1 \leq p < \infty$ the spaces $(\ell^p, \|.\|_p)$ are normed spaces of convergent sequences $x = (x_i)_i$ such that

$$\|x\|_p = |x_1|^p + |x_2|^p + \cdots < \infty,$$

and $(\ell^\infty, \|.\|_\infty)$ is the space of bounded sequences $(x_i)_i$ with respect to the norm

$$\|x\|_\infty = \sup\{|x_i| : i = 1, 2, ...\}.$$

We have the following inclusions:

$$\ell^1 \subset \ell^2 \subset \cdots \ell^\infty.$$

For example $(1/n)_n$ is in ℓ^p for $p \geq 2$, but not in ℓ^1.

Exercise 2.1.11. *Suppose $p, q \in [1, \infty]$. Show that for $p < q$ the space ℓ^p is a proper subspace of ℓ^∞.*

Let us view these vectors of infinite length as real-valued sequences. Then the assumption $\|x\|_p$ imposes conditions on the structure of the sequences. For example, $\|x\|_\infty = \sup_i |x_i|$ is finite if and only if x is a bounded sequence, and $\|x\|_1 = \sum_{i=1}^\infty |x_i|$ is finite if the sequence (x_i) is absolutely summable. The norms $\|.\|_p$ for $1 \leq p < \infty$ describe different notions of convergence, but $\|.\|_\infty$ does not impose convergence but just boundedness.

Proposition 2.1.12. *For $1 \leq p \leq \infty$ the spaces $(\ell^p, \|.\|_p)$ are normed spaces.*

The proof of the finite-dimensional setting extends to the infinite-dimensional setting because Hölder's inequality is valid for ℓ^p-norms.

Lemma 2.5 (Hölder's inequality). *For $1 < p < \infty$ and q its conjugate index, $x \in \ell^p$ and $y \in \ell^q$ we have*

$$\sum_{i=1}^{\infty} |x_i||y_i| \leq \|x\|_p \|y\|_q.$$

2.1.2. Innerproduct spaces. For vectors in $\mathbb{R}^3$ we have the 'dot product' aka 'scalar product' that assigns to a pair of vectors $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ the number

$$\langle x, y\rangle = x_1 y_1 + x_2 y_2 + x_3 y_3.$$

Pythagoras' theorem gives the length of $x = (x_1, x_2, x_3)$ as $\sqrt{x_1^2 + x_2^2 + x_3^2}$. Note that $\langle x, x\rangle = \sqrt{x_1^2 + x_2^2 + x_3^2}$. Innerproduct spaces are a generalization of these basic facts from Euclidean geometry to general vector spaces.

Definition 2.1.13. Let X be a vector space. An *innerproduct* on X is a map $\langle ., .\rangle : X \times X \to \mathbb{F}$, which has the following properties:

(1) (Linearity) For vectors $x_1, x_2, y \in X$ and scalars $\lambda_1, \lambda_2 \in \mathbb{F}$ we have $\langle \lambda_1 x_1 + \lambda_2 x_2, y\rangle = \lambda_1 \langle x_1, y\rangle + \lambda_2 \langle x_2, y\rangle$.
(2) (Symmetry) For vectors $x, y \in X$ we have $\langle x, y\rangle = \langle y, x\rangle$ for $\mathbb{F} = \mathbb{R}$ and $\langle x, y\rangle = \overline{\langle y, x\rangle}$ for $\mathbb{F} = \mathbb{C}$.
(3) (Positive definiteness) For any $x \in X$ we have $\langle x, x\rangle \geq 0$ and $\langle x, x\rangle = 0$ if and only if $x = 0$.

We call $(X, \langle ., .\rangle)$ an *innerproduct space* and denote by $\|x\| = \langle x, x\rangle^{1/2}$ the length of x.

We state a theorem of utmost importance about innerproduct spaces.

Theorem 2.6 (Cauchy-Schwarz). *Suppose X is an innerproduct space. Then for all $x, y \in X$ we have*

$$|\langle x, y\rangle| \leq \|x\|\|y\|.$$

We have $|\langle x, y\rangle| = \|x\|\|y\|$ if and only if $x = \lambda y$ for some $\lambda \in \mathbb{F}$.

Proof. Suppose $x \neq 0$ otherwise the inequality is trivial. Then we consider $z = \langle x, y\rangle x - \langle x, x\rangle y$. By the properties of an innerproduct we have

$$0 \leq \langle z, z\rangle = |\langle x, y\rangle|^2 \langle x, x\rangle - 2|\langle x, y\rangle|^2 \langle x, x\rangle + \langle x, x\rangle^2 \langle y, y\rangle,$$

hence we obtain

$$|\langle x, y\rangle|^2 \langle x, x\rangle \leq \langle x, x\rangle^2 \langle y, y\rangle$$

and after dividing through by the strictly positive number $\langle x, x\rangle$ we obtain the Cauchy-Schwarz inequality.

We have equality if and only if $z = 0$, which yields that $x = \lambda y$ for $\lambda = \langle x, x\rangle \langle x, y\rangle^{-1}$. □

As first consequence we deduce that innerproduct spaces $(X, \langle ., .\rangle)$ are normed spaces for $\|x\| = \langle x, x\rangle^{1/2}$.

Proposition 2.1.14. *For $(X, \langle ., .\rangle)$ the expression $\|x\| = \langle x, x\rangle^{1/2}$ defines a norm on X.*

PROOF. Homogeneity follows from the linearity of the innerproduct. The triangle inequality follows from Cauchy-Schwarz:

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|,$$

so the right side is $(\|x\| + \|y\|)^2$ and thus we have $\|x+y\| \leq \|x\| + \|y\|$. □

The sequence space ℓ^2 was the first example of an innerproduct space, studied by D. Hilbert in 1901 in his work on Fredholm operators.

EXAMPLE 2.1.15. The sequence space ℓ^2 is an innerproduct space for real-valued sequences $(x_i), (y_i)$

$$\langle x, y\rangle = \sum_{i=1}^{\infty} x_i y_i$$

and

$$\langle x, y\rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}$$

for complex-valued sequences.

The innerproduct $\langle ., .\rangle$ and its associated norm $\|.\| = \langle ., .\rangle^{1/2}$ are related by the *polarization identity*.

LEMMA 2.7 (Polarization identity). *Let $(X, \langle ., .\rangle)$ be an innerproduct space with norm $\|.\| = \langle ., .\rangle^{1/2}$.*

(1) *For a real innerproduct space we have $\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$ for all $x, y \in X$.*
(2) *For a complex innerproduct space we have $\langle x, y\rangle = \frac{1}{4}\sum_{k=1}^{4} i^k \|x + i^k y\|^2$.*

PROOF. The arguments are based on the homogeneity properties of innerproducts.

(1) $\|x+(-1)^k y\|^2 = \|x\|^2 + \|y\|^2 + (-1)^k \langle x, y\rangle$ for $k = 0, 1$. Adding these two identities yields the desired polarization identity.
(2) Left as an exercise.

□

Jordan and von Neumann gave an elementary characterizations of norms that arise from innerproducts.

THEOREM 2.8 (Jordan-von Neumann). *Suppose $(X, \|.\|)$ is a complex normed space. If the norm satisfies the parallelogram identity*

$$\|x-y\|^2 + \|x+y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad forall\ x, y \in X,$$

then X is a Hilbert space for the innerproduct

$$\langle x, y\rangle = \frac{1}{4}\sum_{k=1}^{4} i^k \|x + i^k y\|^2.$$

The proof of this useful result is elementary and will be given in the supplement to the chapter.

Innerproduct spaces are the infinite-dimensional counterparts of $(\mathbb{R}^n, \|.\|_2)$ and share many properties with these finite-dimensional spaces, in contrast to general normed spaces such as $C(I)$ with the sup-norm.

EXAMPLE 2.1.16. The supremums norm of $C[0,1]$ does not come from an innerproduct. Use the polarization identity to show this fact.

A way to address this issue is to change the norm. Namely, if one instead equips $C(I)$ with the 2-norm $\|.\|_2$ for functions, then one gets an innerproduct space.

LEMMA 2.9. *Let I be an interval of $\mathbb{R}$. Then the space of continuous complex-valued functions $C(I)$ is an innerproduct space for*

$$\langle f, g\rangle = \int_I f(x)\overline{g(x)}dx$$

for functions $f \in C(I)$ with finite norm

$$\|f\|_2 = \int_I |f(x)|^2 dx < \infty.$$

PROOF. We have $\langle \lambda f, g\rangle \int_I \lambda f(x)\overline{g(x)}dx = \lambda \int_I f(x)\overline{g(x)}dx = \lambda fg$ for $\lambda \in \mathbb{C}$, and $\langle f, g\rangle = \int_I f(x)\overline{g(x)}dx = \overline{\int_I \overline{f(x)}g(x)dx}$. Note that $|f(x)|^2$ is non-negative for $f \in C(I)$ and that it is zero for those $x \in I$ with $f(x) = 0$. By the properties of the integral we have shown the positivity of $\langle .,.\rangle$. $\square$

Historical note: The Cauchy-Schwarz inequality for $(C(\mathbb{R}), \langle .,.\rangle$ is due to Karl H. A. Schwarz in 1888 for continuous functions, and Cauchy for $\mathbb{R}^n$ with the Euclidean innerproduct.

Innerproducts yield a generalization of the notion of *orthogonality* of elements.

DEFINITION 2.1.17. Two elements x, y in an innerproduct space $(V, \langle .,,\rangle)$ are *orthogonal* to each other if $\langle x, y\rangle = 0$

The theorem of Pythagoras is true for any innerproduct space $(X, \langle .,.\rangle)$.

PROPOSITION 2.1.18 (Pythagoras's Theorem). *Let $(X, \langle .,.\rangle)$ be an innerproduct space. For two orthogonal elements $x, y \in X$ we have*

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

PROOF. The argument is based on the fact that $\langle x, x\rangle$ is a norm. By assumption we have $\langle x, y\rangle = 0$

$$\|x+y\|^2 = \|x\|^2 + 2\mathrm{Re}\,\langle x, y\rangle + \|y\|^2 = \|z\|^2 + \|y\|^2.$$

$\square$

As an example we consider some orthogonal vectors in $(C([0,1]), \langle .,.\rangle$. For $m \neq n$ we define the exponentials $e_m(x) = e^{2\pi i m x}$ and $e_n(x) = e^{2\pi i n x}$. Then

$$\langle e_m, e_n\rangle = \int_0^1 e^{2\pi i(m-n)x}dx = (2\pi i(m-n))^{-2}(e^{2\pi i(m-n)} - 1) = 0.$$

Note that $\langle e_n, e_n\rangle = 1$ for any $n \in \mathbb{Z}$. With the help of Kronecker's delta function we may express this as $\langle e_m, e_n\rangle = \delta_{m,n}$.

The theorem of Pythagoras is now at our disposal in any innerproduct spaces such as ℓ^2.

DEFINITION 2.1.19. A set of vectors $\{e_i\}_{i\in I}$ in an innerproduct space $(X, \langle .,, \rangle)$ is called an *orthogonal family* if $\langle e_i, e_j\rangle = 0$ for all $i \neq j$. In case that the orthogonal family $\{e_i\}_{i\in I}$ in V satisfies in addition $\|e_i\| = 1$ for any $i \in I$, then we refer to it as *orthonormal family.*

The set of vectors $\{e_i\}_{i\in I}$ is in general an infinite set. The exponentials $\{e^{2\pi nx}\}_{n\in\mathbb{Z}}$ is an orthonormal family in $C[0,1]$ with respect to $\langle .,.\rangle_2$ and is a system of utmost importance, e.g. it lies at the heart of Fourier analysis or more generally harmonic analysis.

Orthonormal families have an interesting property, known as *Bessel's inequality.*

PROPOSITION 2.1.20 (Bessel's inequality). *Suppose $\{e_i\}_{i\in I}$ is a countably infinite orthonormal family in an innerproduct space $(X, \langle .,.\rangle)$. Then for any $x \in X$ we have*

$$\sum_{i\in I} |\langle x, e_i\rangle|^2 \leq \|x\|^2.$$

Recall that a set I is *countably infinite* if there exists a bijection between I and the set of natural numbers $\mathbb{N}$, e.g. the set of integers $\mathbb{Z}$.

PROOF. It suffices to check the inequality for $I = \mathbb{N}$. Consider the vector $\tilde{x} = \sum_{i=1}^n \langle x, e_i\rangle e_i$. for each $n \in \mathbb{N}$. By the orthonormality of the set $\{e_i\}_{i\in I}$ we have

$$0 \leq \|x - \tilde{x}\|^2 = \|x\|^2 - \sum_{i=1}^n |\langle x, e_i\rangle|^2.$$

Thus the sequence of real numbers $(\sum_{i=1}^n |\langle x, e_i\rangle|^2)$ is bounded above and nondecreasing. Therefore it has a limit

$$\sum_{i=1}^\infty |\langle x, e_i\rangle|^2 \leq \|x\|^2.$$

□

The case of equality in Bessel's inequality characterizes an important properties of orthonormal systems and will be discussed in the chapter on Hilbert spaces.

For example Bessel's inequality for the set of exponentials $\{e^{2\pi inx}\}_{n\in\mathbb{Z}}$ in $(C[0,1], \langle .,.\rangle_2)$ is a statement about the Fourier coefficients of f

$$\widehat{f}(n) = \int_0^1 f(x)e^{-2\pi nx}dx,$$

then we have

$$\sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^2 \leq \|f\|_2^2.$$

Therefore we will refer to $(\langle x, e_i\rangle)_{i\in I}$ as the Fourier coefficients of $x \in X$ and of

$$\sum_{i\in I} \langle x, e_i\rangle e_i$$

as the *Fourier series* of x.

2.1.3. Bounded operators between normed spaces. Mappings between vector spaces are of interest in a wide range of applications. We restrict our focus to mappings that respect the vector space structure: linear mappings aka linear operators.

Definition 2.1.21. Let X, Y be vector spaces over the same scalar field $\mathbb{F}$. Then a mapping $T : X \to Y$ is *linear* if

$$T(x + \lambda y) = Tx + \lambda Ty$$

for all $x, y \in X$ and $\lambda \in \mathbb{F}$. We denote by $\mathcal{L}(X, Y)$ the set of all linear operators between X and Y.

Linear mappings are a special class of functions between two sets. Hence it has the structure of a vector space.Here are some examples of linear mappings for the classes of vector spaces of our interest.

(1) Linear mappings between $\mathbb{F}^n$ and $\mathbb{F}^m$ are given by $m \times n$ matrices A with entries in $\mathbb{F}$, $x \mapsto Ax$ for $x \in \mathbb{F}^n$.
(2) On the space of polynomials $\mathcal{P}_n$ of degree at most n we define the *differentiation operator* $Dp(x) = a_1 x + \cdots m a_n x^{n-1}$, the operator $p \mapsto \int p(x)dx$ and the evaluation operator $Tp(x) = p(0)$.
(3) Operators on sequence spaces: For an element of the vector space s, a sequence $x = (x_n)_n$, we define the *left shift* $Lx = (0, x_0, x_1, x_2, ...)$, the *right shift* $Rx = (x_1, x_2, ...)$ and the multiplication operator $T_a x = (a_0 x_0, a_1 x_1, ...)$ for a sequence $a = (a_0, a_1, ...) \in s$. On the vector space of convergent sequences c we define $Tx = \lim_n x_n$ for $x = (x_n) \in c$.
(4) Operators on function spaces: The set of continuous functions $C(I)$ on an interval of $\mathbb{R}$, popular choices for I are $[0, 1]$ and $\mathbb{R}$. For $f \in C(I)$ we define the *integral operator* $f \mapsto \int k(x, y)f(y)dx$ for a function k defined on $I \times I$, the kernel of the operator, and the evaluation operator $Tf(x) = f(a)$ for $a \in I$. For a differentiable continuous function f we are able to study the *differentiation operator* $Df(x) = f'(x)$.

Norms on these spaces provide a tool to understand the properties of these mappings via the notion of *operator norm* that measures the size of the measure of distortion of x induced by T: For normed spaces $(X, \|.\|_X)$, $(Y, \|.\|_Y)$ and a linear mapping $T : X \to Y$ we are interested in operators such that there exists a constant c such that

$$\|Tx\|_Y \le c\|x\|_X \quad forall\ x \in X.$$

Often we will omit the subscripts to ease the notation. The operators with a finite c are of particular relevance and are called *bounded operators*. We denote by $\mathcal{B}(X, Y)$ the set of all bounded linear operators from X to Y.

Definition 2.1.22. Let T be a linear operator between the normed spaces $(X, \|.\|_X)$ and $(Y, \|.\|_Y)$. The *operator norm* of T is defined by

$$\|T\| = \sup\{\frac{\|Tx\|_Y}{\|x\|_X} : \|x\|_X \neq 0\}.$$

Sometimes we denote the operator norm of T by $\|T\|_{\mathrm{op}}$.

Lemma 2.10. *For $T \in \mathcal{B}(X, Y)$ the following quantities are all equal to the operator norm $\|T\|$ of T:*

(1) $C_1 = \inf\{c \in \mathbb{R} : \|Tx\|_Y \leq c\|x\|_X\}$,
(2) $C_2 = \sup\{\frac{\|Tx\|_Y}{\|x\|_X} : \|x\|_X \leq 1\}$,
(3) $C_3 = \sup\{\frac{\|Tx\|_Y}{\|x\|_X} : \|x\|_X = 1\}$.

Proof. The argument is based on some inequalities:

(1) $C_2 \leq C_1$: By definition of C_1 we have $\|Tx\| \leq C_1\|x\|$. Hence for all $x \in \overline{B_1}(0)$ we have $\|Tx\| \leq C_1$ and thus we have $C_2 \leq C_1$.
(2) $C_3 \leq C_2$: For all $x \in \overline{B_1}(0)$ we have $\|Tx\| \leq C_2$. Pick an x with $\|x\| = 1$ and define the sequence of vectors $(x_n = (1 - 1/n)v)_n$ which all have $\|x_n\| \leq 1$ and hence $\|Tx_n\| \leq C_2$ for all $n \in \mathbb{N}$. Taking the limit gives $\|Tx\| \leq C_2$ and thus $C_3 \leq C_2$.
(3) $\|T\| \leq C_3$: By definition of C_3 we have $\|Tx\| \leq C_3$ for all x with $\|x\| = 1$. Take an arbitrary non-zero vector $x \in X$. Then $x/\|x\|$ has unit length and hence $\|T(\frac{x}{\|x\|})\| = \frac{\|Tx\|}{\|x\|} \leq C_3$, which establishes the desired inequality $\|T\| \leq C_3$.
(4) We have $\|Tx\|\|x\| \leq \|T\|$ for all $x \in X$. Hence $\|Tx\| \leq \|T\|\|x\|$ for all $x \in X$. Hence we have $C_1 \leq \|T\|$. Hence we have $C_1 \leq C_2 \leq C_3 \leq \|T\| \leq C_1$ and so the assertion is established.

□

These different expressions for the operator norm of a linear operator are elementary but nonetheless useful. Before we discuss some examples we note some properties of the operator norm.

Proposition 2.1.23. *For $S, T \in \mathcal{B}(X, Y)$ we have*

(1) $\|I\| = 1$ *for the identity operator* $I : X \to X$.
(2) $\|\lambda S + \mu T\| \leq |\lambda|\|S\| + |\mu|\|T\|$ *for* $\lambda, \mu \in F$.
(3) *Submultiplicativity:* $\|S \circ T\| \leq \|S\|\|T\|$.

Proof. (1) By the definition of the operator norm we have $\|I\| = 1$.
(2) The triangle inequality for norms yields the assertion.
(3) By definition we have

$$\|S \circ T\| = \sup\{\|STx\| : \|x\| = 1\} \leq \sup\{\|S\|\|Tx\| : \|x\| = 1\} = \|S\|\|T\|.$$

□

Proposition 2.1.24. *The vector space $\mathcal{B}(X, Y)$ of bounded operators between two normed spaces is a normed spaces wrt the operator norm.*

Proof. The preceding proposition implies the homogeneity property and the triangle inequality. The operator norm is clearly positive definite, and we have $\|T\| = 0$ if and only if $T = 0$ because it is defined in terms of a norm on Y. fined in terms of a norm on Y. □

We treat some of the operators defined above.

(1) Let D be the differentiation operator on $\mathcal{P}_n$: Then D is unbounded on $(\mathcal{P}_n, \|.\|_\infty)$, because for $p(x) = x^n$ $Dp(x) = nx^{n-1}$ and so $\|Dp\|_\infty = n$ for all $n \in \mathbb{N}$. The integration operator $p \mapsto \int p(x)dx$ has norm 1 for intervals of finite length.

(2) The left shift $Lx = (0, x_0, x_1, x_2, ...)$ has $\|L\| = 1$ and also the right shift $\|R\| = 1$. For the multiplication operator $T_a x = (a_0 x_0, a_1 x_1, ...)$ for a sequence $a = (a_0, a_1, ...) \in s$ we have $\|T_a\| = \|a\|_\infty$.

(3) The operator norm of the integral operator $C[a, b]$ with $\|.\|_\infty$ for an interval of finite length is $(b - a)^2 \|K\|_\infty$.

Some classes of operators on a normed space X: (i) *isometries* on X are linear operators T with $\|Tx\| = \|x\|$ for all $x \in X$, (ii) projections are linear operators P on X satisfying $P^2 = P$.

CHAPTER 3

Banach spaces and Hilbert spaces

3.1. Banach spaces and Hilbert spaces

We extend the topological notions introduced for the real line to general normed spaces and we focus on completeness in this section. Complete normed spaces are nowadays called Banach spaces, after the numerous seminal contributions of the Polish mathematician Stefan Banach to these objects. The class of complete innerproduct spaces are named after David Hilbert, who introduced the sequence space ℓ^2. His students made numerous contributions to the theory of innerproduct spaces, e.g. Erhard Schmidt, Hermann Weyl, Otto Toeplitz,... .

3.1.1. Completeness. We start with the generalization of open and closed intervals in $\mathbb{R}$ to general normed spaces.

Definition 3.1.1. Let $(X, \|.\|)$ be normed space.

(1) $B_r(x) = \{y \in X : \|y - x\| < r\}$ denotes the *open ball* of radius r around a point $x \in X$.
(2) $\overline{B_r}(x) = \{y \in X : \|y - x\| \leq r\}$ denotes the *closed ball* of radius r around a point $x \in X$.

For the sequence spaces ℓ^p open balls $B_r(x)$ around $x = (x_k)$ are all sequences $y = (y_k) \in \ell^p$ with $\|x - y\| < r$. In the setting of $(C(I), \|.\|)$ the ball $B_\varepsilon(f)$ are all continuous functionss g that are in an ε-strip of f.

Here are the the notions of a convergent sequence and Cauchy sequence in a normed space.

Definition 3.1.2. Let $(X, \|.\|)$ be a normed space.

(1) A sequence $(x_k)_{k\in\mathbb{N}}$ converges to $x \in X$ if for a given $\varepsilon > 0$ there exists a N such that $\|x - x_k\| < \varepsilon$ for $k \geq N$.
(2) A sequence $(x_k)_{k\in\mathbb{N}}$ is a *Cauchy sequence* if for any $\varepsilon > 0$ there exists a N such that $\|x_m - x_n\| < \varepsilon$ for all $m, n > N$.

This notion of sequences is a natural generalization of the one for real and complex numbers. Note that the elements of the sequences are vectors in a normed space. For example, a sequence in ℓ^2 is a sequence where the elements themselves are also sequences. The difference between the the normed space $\mathbb{Q}$ and the real numbers $\mathbb{R}$ viewed as normed space is that not all Cauchy sequences in $\mathbb{Q}$ converge to a rational number but that is the case for $\mathbb{R}$.

Definition 3.1.3. A normed space $(X, \|.\|)$ is called *complete* if every Cauchy sequence (x_k) in X has a limit x belonging to X. Moreover, a complete normed space is refered to as *Banach space* and a complete innerproduct space is known as *Hilbert space.*

THEOREM 3.1. *For $p \in [1,\infty]$ the normed space $(\mathbb{R}^n, \|.\|_p)$ is complete.*

PROOF. $\square$

The infinite-dimensional counterpart of the previous is also true, but its proof is more intricate.

THEOREM 3.2. *For $p \in [1,\infty]$ the normed spaces $(\ell^p, \|.\|_p)$ are complete.*

PROOF. $\square$

THEOREM 3.3. *For a finite interval $[a,b]$ the normed space $C[a,b]$ with respect to the sup-norm $\|.\|_\infty$ is complete.*

For the proof we have to discuss notions of convergence for sequences of functions.

DEFINITION 3.1.4. Let (f_n) be a sequence of functions on a set X.

- We say that (f_n) *converges pointwise* to a limit function f if for a given $\varepsilon > 0$ and $x \in X$ there exists an N so that
$$|f_n(x) - f(x)| < \varepsilon \quad forall\ n \geq N.$$
- We say that *converges uniformly* to a limit function f if for a given $\varepsilon > 0$ there exists an N so that
$$|f_n(x) - f(x)| < \varepsilon \quad forall\ n \geq N$$
holds for all $x \in X$.

There is a substantial difference between these two definitions. In pointwise convergence, one might have to choose a different N for each point $x \in X$. In the case of uniform convergence there is an N that holds for all $x \in X$. If one draws the graphs of a uniformly convergent sequence, then one realises that the definition amounts for a given $\varepsilon > 0$ to have a N so that the graphs of all the f_n for $n \geq N$, lie in an ε-band about the graph of f. In other words, the f_n's get uniformly close to f.

Note that uniform convergence implies pointwise convergence. Uniform convergence has an important property.

THEOREM 3.4. *Let (f_n) be a uniformly convergent sequence in $C(I)$ with limit f. Then the limit function f is continuous on I.*

PROOF. Let $y \in I$ and $\varepsilon > 0$ be given. By the uniform convergence of $f_n \to f$, there exists an N such that $n \geq N$ implies that
$$|f_n(x) - f(x)| \leq \varepsilon/3 \quad \text{for all } x \in I.$$
The continuity of f_N implies that there exists a $\delta > 0$ such that
$$|f_n(x) - f(y)| \leq \varepsilon/3 \quad \text{for } |x-y| \leq \delta.$$
We want to show that f is continuous. For all x such that $|x-y| < \delta$ we have that
$$\begin{aligned}|f(x) - f(y)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.\end{aligned}$$

$\square$

PROOF. $\square$

Theorem 3.5. *The normed space of bounded operators $(B(X,Y), \|.\|_{\mathrm{op}})$ is complete if and only if Y is a Banach space.*

The Banach space $(B(X,\mathbb{C}), \|.\|_{\mathrm{op}})$ is known as the *dual space* of X, denoted by X', and its elements are refered to as *functionals* on X.

Proof. Let (T_n) be a Cauchy sequence in $B(X,Y)$, so for any $\varepsilon > 0$ there exists a $N \in \mathbb{N}$ such that for all $m, n \geq \mathbb{N}$ we have $\|T_m - T_n\|_{\mathrm{op}} < \varepsilon$. Hence for any $x \in X$ we have

$$\|(T_m - T_n)x\|_Y \leq \|T_m - T_n\|_{\mathrm{op}}\|x\|_X < \varepsilon\|x\|_X.$$

Hence for all $x \in X$ the sequence $(T_n x)$ is a Cauchy sequence in Y. Since Y is a Banach space, it has a limit denoted by Tx, and thus we define $Tx = \lim_{n\to\infty} T_n x$. The limit operator T is linear and bounded.

$$\|Tx\|_Y \leq \sup_n \|T_n x\|_Y \leq \|x\|_X \sup_n \|T_n\|_{\mathrm{op}},$$

and thus we have $\|T\|_{\mathrm{op}} \leq \sup_n \|T_n\|_{\mathrm{op}}$, i.e. $T \in \mathrm{B}(X,Y)$.
We show that $\|T_n - T\|_{\mathrm{op}} \to 0$. We assume otherwise that $\|T_n - T\|_{\mathrm{op}}$ does not converge to 0. Then there exists an $\varepsilon > 0$ and a subsequence $(T_{n_k})_k$ of (T_n) such that

$$\|T_n - T\|_{\mathrm{op}} \geq \varepsilon \quad \text{for all k.}$$

Consequently, for every k there exists a $x_k \in X$ with $\|x_k\| = 1$ and

$$\|T_{n_k}(x_k) - T_m(x_k)\| \geq \varepsilon.$$

By assumption (T_n) is a Cauchy sequence, so one can choose a N_0 such that for all $m, n_k \geq N_0$ we have

$$\|T_{n_k}(x_k) - T_m(x_k)\| \leq \varepsilon/2$$

and this gives

$$\varepsilon \leq \|T_{n_k}(x_k) - T(x_k)\|_Y \leq \|T_{n_k}(x_k) - T_m(x_k)\|_Y + \|T_m(x_k) - T(x_k)\|_Y.$$

Hence for all $m \geq N_0$ we have

$$\|T_m(x_k) - T(x_k)\|_Y \geq \varepsilon/2.$$

That is a contradicition to the definition of T, thus we have $T_m(x_k) - T(x_k) \to 0$ in $(\mathcal{B}(X,Y), \|.\|_{\mathrm{op}})$. □

3.1.2. Topology of normed spaces. Definitions and properties of open and closed sets, sequences and other notions have natural counterparts in the setting of normed spaces. The motivation is once more an understanding of sequences of elements in normed spaces.

Definition 3.1.5.
(1) A set $U \subset X$ is a *neighborhood* of $x \in X$ if $B_r(x) \subset U$ for some $r > 0$.
(2) A set $O \subset X$ is *open* if every $x \in O$ has a neighborhood U contained in O.
(3) A set $C \subset X$ is *closed* if its complement $C^c = X \backslash F$ is open.

Note that the definition of open sets depends on the norm. In other words, open sets with respect to one norm need not be open with respect to another norm.

Lemma 3.6. *Let $(X, \|.\|)$ be normed space. Then $B_r(x)$ is open and $\overline{B_r}(x)$ is closed for $x \in X$ and $r > 0$.*

PROOF. The proof goes along the same lines as in the case of the real line. Suppose that $y \in B_r(x)$ and choose ε as $\varepsilon = r - d(x,y) > 0$. The triangle inequality yields that $B_\varepsilon(y) \subset B_r(x)$, i.e. $B_r(x)$ is open.
We show that $X \backslash \overline{B_r}(x)$ is open. For $y \in X \backslash \overline{B_r}(x)$ we set $\varepsilon = d(x,y) - r > 0$ and once more by the triangle inequality we deduce that $B_\varepsilon(y) \subset X \backslash \overline{B_r}(x)$. Hence $X \backslash \overline{B_r}(x)$ is open and $\overline{B_r}(x)$ is closed. □

DEFINITION 3.1.6. For a subset A of $(X, \|.\|)$ we introduce some notions.

(1) The *closure* of a subset A of X, denoted by $\overline{A}$, is the intersection of all closed sets containing A.
(2) The *interior* of a subset of A of X, denoted by $\mathrm{int}A$, is the union of all open subsets of X contained in A.
(3) The boundary of a subset A of X, denoted by $\mathrm{bd}A$, is the set $\overline{A} \backslash \mathrm{int}A$.

We continue with some defintions

DEFINITION 3.1.7. Let A be a subset of $(X, \|.\|)$.

(1) A point $x \in A$ is *isolated* in A if there exists a neighborhood U of x such that $U \cap A = \{x\}$.
(2) A point $x \in \mathbb{R}$ is said to be an *accumulation point* of A if every neighborhood of x contains points in $A \backslash \{x\}$.

DEFINITION 3.1.8. A subset A of $(X, \|.\|)$ is said to be *dense* in $\mathbb{R}$ if its closure is equal to X, i.e. $\overline{A} = X$. If the dense subset A is countable, then X is called *separable*.
In other words, a subset A of a normed space X is dense in X if for each $x \in X$ and each $\varepsilon > 0$ there exists a vector $y \in A$ such that

$$\|x - y\| < \varepsilon.$$

The relevance of a dense subset of a normed space is that it provides a way to approximate elements of the normed space by ones from the dense subset up to any given precision.

LEMMA 3.7. *Suppose A is a dense subspace of a normed space X. For any $x \in X$ there exists a sequence of elements $x_k \in A$ such that $\|x_k - x\| \to 0$ as $k \to \infty$.*

PROOF. For $x \in X$ there exists an x_k such that $\|x_k - x\| < 1/k$ for $k = 1, 2, ...$ By construction x_k converges to x. □

The next results have been proved in the section on real numbers and these are also true for normed spaces. The proofs of these results are along the same lines as the ones for the real line.

LEMMA 3.8. *Let $\{O_j : j \in J\}$ be a family of open sets of $(X, \|.\|)$.*

(1) *$\cap_{j=1}^n O_j$ is an open set for any $n \in \mathbb{N}$.*
(2) *$\cup_{j \in J} O_j$ is open for a general index set J.*

LEMMA 3.9. *Suppose A is a subset of $(X, \|.\|)$.*

(1) $\overline{A} = (\mathrm{Int}(A^c))^c$ *and* $\mathrm{int}(A) = (\overline{A^c})^c$
(2) $\mathrm{bd}A = \mathrm{bd}(A^c) = \overline{A} \cap \overline{A^c}$
(3) $\overline{A} = A \cup \mathrm{br}A = \mathrm{int}A \cup \mathrm{bd}A$

LEMMA 3.10. *Suppose A is a subset of $(X, \|.\|)$.*

(1) $\overline{A} = \{x \in X :$ *every neighborhood of x intersects A*$\}$
(2) $\mathrm{int}(A) = \{x \in X :$ *some neighborhood of x is contained in A*$\}$
(3) $\mathrm{bd}(A) = \{x \in X :$ *every neighborhood of x intersects A and its complement*$\}$

LEMMA 3.11. *A point x in a normed space $(X, \|.\|)$ is an accumulation point of A if and only if every neighborhood of x contains infinitely many points of A.*

APPENDIX A

Sets and functions

A.1. Sets and functions

In order to formalize our intution about collections of objects we use the framework of set theory. The relation between sets and their elements will be described by functions.

DEFINITION A.1.1. A *set* is a collection of distinct objects, its *elements*. If an object x is an element of a set X, we denote it by $x \in X$. If x is not an element of A, then we wrtie $x \neq X$.

A set is uniquely determined by its elements. Suppose X and Y are sets. Then they are identical, $X = Y$, if they have the same elements. More formalized, $X = Y$ if and only if for all $x \in X$ we have $x \in Y$, and for all $y \in Y$ we have $y \in X$.
The *empty set* is the set with no elements, denoted by $\emptyset$.

DEFINITION A.1.2. Suppose X and Y are sets. Then Y is a subset of X, denoted by $Y \subset X$, if for all $y \in Y$ we have $y \in X$.

If $Y \subseteq X$, one says that Y is contained in X. If $Y \subseteq X$ and $X \neq Y$, then Y is a proper subset of X and we use the notation $Y \subset X$.

Here are a few constructions of sets.

DEFINITION A.1.3. Let X and Y be sets.

- The *union* of X and Y, denoted by $X \cup Y$, is defined by
$$X \cup Y = z | \, z \in X \quad \text{or } z \in Y.$$
- The *intersection* of X and Y, denoted by $X \cap Y$, is defined by
$$X \cap Y = z | \, z \in X \quad \text{and } z \in Y.$$
- . The *difference set* of X from Y, denoted by $X \backslash Y$, is defined by
$$X \backslash Y = \{z \in X : z \in X \quad \textit{and } z \neq Y\}.$$
If all sets are contained in one set X, then the difference set $X \subset Y$ is called the *complement* of Y.
- The *Cartesian product* of X and Y, denoted by $X \times Y$, is the set
$$X \times Y = \{(x, y) | \, x \in X, y \in Y\},$$
i.e the set of all ordered pairs (x, y), with $x \in X$ and $y \in Y$.

Here are some basic properties of sets.

LEMMA A.1. *Let X, Y and Z be sets.*

(1) $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$ *and* $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$ *(distributition law)*
(2) $(X \cup Y)^c = X^c \cap Y^c$ *and* $(X \cap Y)^c = X^c \cup Y^c$ *(De Morgan's laws)*
(3) $X\backslash(Y \cup Z) = (X\backslash Y) \cap (X\backslash Z)$ *and* $X\backslash(Y \cap Z) = (X\backslash Y) \cup (X\backslash Z)$

Let X and Y be sets. A function with *domain* X and *codomain* Y, denoted by $f : X \to Y$, is a relation between the elements of X and Y satisfying the properties: for all $x \in X$, there is a unique $y \in Y$ such that $(x, y) \in f$, we denote it by: $f(x) = y$.

By definition, for each $x \in X$ there is exactly one $y \in Y$ such that $f(x) = y$. We say that y the *image* of x under f. The *graph* $G(f)$ of a function f is the subset of $X \times Y$ defined by

$$G(f) = \{(x, f(x))|\, x \in X\}.$$

The *range* of a function $f : X \to Y$, denoted by range(f), or $f(A)$, is the set of all $y \in Y$ that are the image of some $x \in X$:

$$\mathrm{range}(f) = \{y \in Y|\,\text{there exists}\, x \in X\,\text{such that}\, f(x) = y\}.$$

The *pre-image* of $y \in Y$ is the subset of all $x \in X$ that have y as their image. This subset is often denoted by $f^{?1}(y)$:

$$f^{?1}(y) = \{x \in X|\, f(x) = y\}.$$

Note that $f^{?1}(y) = \emptyset$ if and only if $y \in Y\backslash\mathrm{range}(f)$.

The following notions are central for the theory of functions.

Definition A.1.4. Let $f : X \to$ be a function.

(1) Then we call f *injective* or one-to-one if $f(x_1) = f(x_2)$ implies $x_1 = x_2$, i.e. no two elements of the domain have the same image.
(2) Then we call f *surjective* or *onto* if range$(f) = Y$, i.e. each $y \in Y$ is the image of at least one $x \in X$.
(3) Then we call f *bijective* if f is both injective and surjective.

Let $f : X \to Y$ and $g : Y \to Z$ be two functions so that the codomain of f coincides with the domain of g. Then we define the *composition*, denoted by $g \circ f$, as the function $g \circ f : X \to Z$, defined by $x \mapsto g(f(x))$.

For every set X, we define the *identity map*, denoted by id_X or id for short: $\mathrm{id}_X : X \to X$ is defined by $\mathrm{id}_X(x) = x$ for all $x \in X$. The identity mapis a bijection.

If f is a bijection, then it is invertible. Hence, the inverse relation is also a function, denoted by $f^{?1}$. It is the unique bijection $Y \to X$ such that $f^{?1} \circ f = \mathrm{id}_X$ and $f \circ f^{?1} = \mathrm{id}_Y$.

Lemma A.2. *Let* $f : X \to Y$ *and* $g : Y \to Z$ *by bijections. Then* $g \circ f$ *is also a bijection and* $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.