

LECTURE NOTES
METRIC SPACES (2018)

SERGEY MOZGOVOY

CONTENTS

1. Metric spaces	2
1.1. Definition and examples	2
1.2. Balls and bounded sets	5
1.3. Open and closed sets	6
1.4. Convergent sequences	8
1.5. Continuous maps	10
1.6. Complete metric spaces	12
1.7. Completion	14
1.8. Banach fixed point theorem	16
2. Topological spaces	18
2.1. Definition and examples	18
2.2. Subspace and product topologies	20
2.3. Hausdorff spaces	22
2.4. Connected topological spaces	23
2.5. Compact topological spaces	26
2.6. Compact metric spaces	28
2.7. Continuous functions on compact spaces	30
3. Normed vector spaces	32
3.1. Definition of a normed vector space	32
3.2. Examples of normed vector spaces	33
3.3. Bounded linear operators	34
3.4. Equivalent norms	36
3.5. Banach spaces	37
Appendix A. Some inequalities	40

1. METRIC SPACES

1.1. Definition and examples. The usual distance between points on the plane is given by

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

where $x = (x_1, x_2) \in \mathbb{R}^2$ and $y = (y_1, y_2) \in \mathbb{R}^2$. This distance satisfies in particular the triangle inequality

$$d(x, z) \leq d(x, y) + d(y, z), \quad x, y, z \in \mathbb{R}^2.$$

Using this example as a motivation we can axiomatize the properties of a distance as follows.

Definition 1.1. A **metric space** is a pair (X, d) , where X is a set and $d: X \times X \rightarrow \mathbb{R}$ is map such that for all $x, y, z \in X$

- (0) $d(x, y) \geq 0$,
- (1) $d(x, y) = 0 \iff x = y$,
- (2) $d(x, y) = d(y, x)$ (symmetry),
- (3) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

The map d is called a **metric** (or a distance function). The elements of X are called **points** of X .

Remark 1.2. Axiom 0 follows from other axioms. Indeed $0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y)$, hence $d(x, y) \geq 0$.

Example 1.3 (Euclidean spaces). Let

$$\mathbb{R}^n = \{ (x_1, \dots, x_n) \mid x_i \in \mathbb{R} \forall 1 \leq i \leq n \}.$$

This is a vector space over $\mathbb{R}$, where the sum of elements $x, y \in \mathbb{R}^n$ is defined by

$$x + y = (x_1 + y_1, \dots, x_n + y_n) \in \mathbb{R}^n$$

and the product of $x \in \mathbb{R}^n$ with a scalar $\lambda \in \mathbb{R}$ is defined by

$$\lambda x = (\lambda x_1, \dots, \lambda x_n) \in \mathbb{R}^n.$$

Define the **scalar product** on $\mathbb{R}^n$ by

$$(x, y) = x_1 y_1 + \dots + x_n y_n \in \mathbb{R}.$$

Define the **norm** on $\mathbb{R}^n$ by

$$\|x\| = \sqrt{(x, x)} = \sqrt{x_1^2 + \dots + x_n^2}.$$

Define the map $d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$d(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}.$$

To prove that d is a metric on $\mathbb{R}^n$ we will apply (see Theorem A.1)

Theorem 1.4 (Cauchy-Schwartz inequality). *For all $x, y \in \mathbb{R}^n$, we have*

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

Corollary 1.5 (Triangle inequality). *For all $x, y \in \mathbb{R}^n$, we have*

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

Proof. We have

$$\begin{aligned} \|x + y\|^2 &= (x + y, x + y) = (x, x) + 2(x, y) + (y, y) \\ &\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2, \end{aligned}$$

hence $\|x + y\| \leq \|x\| + \|y\|$. □

Corollary 1.6. *The map $d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ defined above is a metric on $\mathbb{R}^n$.*

Proof. Let us verify the axioms of a metric:

- (1) If $x = y$, then clearly $d(x, y) = \|x - y\| = 0$. Conversely, if $d(x, y) = \sqrt{\sum (x_i - y_i)^2} = 0 \implies (x_i - y_i)^2 = 0$ for all $i \implies x_i = y_i$ for all $i \implies x = y$.
- (2) $d(x, y) = \|x - y\| = \|y - x\| = d(y, x)$.

- (3) We have to prove $d(x, z) \leq d(x, y) + d(y, z)$ or equivalently $\|x - z\| \leq \|x - y\| + \|y - z\|$. But $\|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\|$. $\square$

Remark 1.7. The metric space $(\mathbb{R}^n, d)$ is called the **Euclidean space**. The metric d is called the **Euclidean metric** (or Euclidean distance).

Example 1.8. Define a metric d_∞ on $\mathbb{R}^2$ by

$$d_\infty(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}.$$

More generally, define for $x, y \in \mathbb{R}^n$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|, \quad d_\infty(x, y) = \|x - y\|_\infty = \max_{1 \leq i \leq n} |x_i - y_i|.$$

Then d_∞ is a metric on $\mathbb{R}^n$ different from the Euclidean metric.

Example 1.9. Define a metric d_1 on $\mathbb{R}^2$ by

$$d_1(x, y) = |x_1 - y_1| + |x_2 - y_2|.$$

More generally, define for $x, y \in \mathbb{R}^n$

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \quad d_1(x, y) = \|x - y\|_1 = \sum_{i=1}^n |x_i - y_i|.$$

Then d_1 is a metric on $\mathbb{R}^n$ different from the Euclidean metric.

Example 1.10. For any $p \geq 1$ and $x, y \in \mathbb{R}^n$, define

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad d_p(x, y) = \|x - y\|_p = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{1/p}.$$

To see that d_p satisfies the triangle inequality we will apply (see Theorem A.2)

Theorem 1.11 (Minkowski inequality). *For any $p \geq 1$ and $x, y \in \mathbb{R}^n$, we have*

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

We can write this inequality as $\|x + y\|_p \leq \|x\|_p + \|y\|_p$. Triangle inequality for d_p follows from

$$d_p(x, z) = \|x - z\|_p = \|(x - y) + (y - z)\|_p \leq \|x - y\|_p + \|y - z\|_p = d_p(x, y) + d_p(y, z).$$

Remark 1.12. For $p = 2$, we obtain $d_2(x, y) = \sqrt{\sum (x_i - y_i)^2}$ which is the Euclidean metric. For $p = 1$, we obtain $d_1(x, y) = \sum |x_i - y_i|$ which is the metric considered in the previous example.

Remark 1.13. One can show that for any $x \in \mathbb{R}^n$, we have $\lim_{p \rightarrow \infty} \|x\|_p = \max_i |x_i| = \|x\|_\infty$. This implies that $\lim_{p \rightarrow \infty} d_p(x, y) = d_\infty(x, y)$.

Example 1.14 (Discrete metric). Let X be any set. Define the discrete metric on X by

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

Example 1.15 (Hamming distance). This is a distance used in information theory and coding theory. Let X be a set and let $n \geq 1$. Define a metric on X^n by the rule

$$d(x, y) = \#\{1 \leq i \leq n \mid x_i \neq y_i\}, \quad x, y \in X^n.$$

This metric is called the Hamming distance. It is equal to the number of positions where the messages x and y are different. For $n = 1$ we obtain the discrete metric on X . For $X = \{0, 1\}$ we can write the distance also in the form $d(x, y) = \sum_i |x_i - y_i|$.

Example 1.16 (Levenshtein distance). This is a distance between strings used in computer science and linguistics. Let X be a finite set and let $X^* = \bigcup_{n \geq 0} X^n$ be the set of finite sequences $x = (x_1, \dots, x_n)$, $n \geq 0$, called words (or strings) in the alphabet X . Define the distance between two words to be the minimal number of 1-letter edits (insertion, deletion or substitution) needed to obtain one word from the other. Formally, let x, y be two words of length m, n respectively. For $0 \leq i \leq m$ and $0 \leq j \leq n$ define

$$\ell_{x,y}(i, j) = \begin{cases} \max\{i, j\} & ij = 0, \\ \min\{\ell_{x,y}(i-1, j) + 1, \ell_{x,y}(i, j-1) + 1, \ell_{x,y}(i-1, j-1) + 1_{x_i \neq y_j}\} & ij > 0 \end{cases}$$

Then define $d(x, y) = \ell_{x,y}(m, n)$. For example

$$\begin{aligned} d(\text{tip}, \text{top}) &= d(\text{ti}, \text{to}) = d(\text{t}, \text{t}) + 1 = 1 \\ d(\text{desert}, \text{dessert}) &= d(\text{de}, \text{des}) = d(\text{de}, \text{de}) + 1 = 1 \end{aligned}$$

Example 1.17 (Distance on graphs). A graph is a pair (V, E) , where V is a set, called the set of vertices, and E is a set of unordered pairs of elements in V , called the set of edges. An edge $e = \{v, v'\} \in E$ (also written as vv') is called an edge between vertices v and v' . Given a function $f: E \rightarrow \mathbb{R}_{\geq 0}$, we can interpret $f(e)$ for an edge $e = vv'$ as the time needed to travel between v and v' along e . For any path $p = e_1 \dots e_n$ between vertices v and v' , let $f(p) = \sum f(e_i)$. Define $d(v, v')$ as the minimal value of $f(p)$ for the paths p between v and v' (and $+\infty$ if there are no such paths). One can show that (V, d) is a metric space (if we also allow the value $+\infty$). For example, if $f(e) = 1$ for every $e \in E$, then $d(v, v')$ is equal to the number of edges along the shortest path between v and v' .

Example 1.18. We can interpret the space $\mathbb{R}^n$ as the set of all functions $f: \{1, \dots, n\} \rightarrow \mathbb{R}$. Let us generalize this construction to infinite domains. Given real numbers $a < b$, let $C[a, b]$ be the set of all continuous functions $f: [a, b] \rightarrow \mathbb{R}$. Define a metric $d = d_\infty$ on $C[a, b]$ by

$$d(f, g) = \sup_{x \in [a, b]} |f(x) - g(x)|, \quad f, g \in C[a, b].$$

Let us check the axioms of a metric

- (1) If $f = g$ then clearly $d(f, g) = 0$. Conversely, if $d(f, g) = 0$ then $|f(x) - g(x)| = 0$ for all $x \in [a, b]$. Therefore $f(x) = g(x)$ for all $x \in [a, b]$, hence $f = g$.
- (2) $d(f, g) = d(g, f)$ is clear.
- (3) We have to show that for any $f, g, h \in C[a, b]$, we have $d(f, g) \leq d(f, h) + d(h, g)$. For any $x \in [a, b]$ we have

$$|f(x) - g(x)| \leq |f(x) - h(x)| + |h(x) - g(x)| \leq d(f, h) + d(h, g).$$

Taking the supremum over all $x \in [a, b]$ we obtain the required inequality.

Example 1.19. Define a metric d_1 on $C[a, b]$ by

$$d_1(f, g) = \int_a^b |f(x) - g(x)| dx, \quad f, g \in C[a, b].$$

The axioms 1-2 of a metric are verified as before. The triangle inequality

$$d_1(f, g) \leq d_1(f, h) + d_1(h, g)$$

follows from the inequality

$$\int_a^b |u(x) + v(x)| dx \leq \int_a^b |u(x)| dx + \int_a^b |v(x)| dx$$

(substitute $f - h$ by u , $h - g$ by v and $f - g$ by $u + v$).

Example 1.20 (Subspace). Let (X, d) be a metric space and let $A \subset X$ be any subset. Then A together with the restriction of d to $A \times A$ is a metric space denoted by (A, d_A) and called a **subspace** of (X, d) . The metric d_A on A is called the **induced metric**.

Example 1.21. Let (X, d) and (Y, d) be two metric spaces. Then we can equip $X \times Y$ with different induced metrics. For example

$$\begin{aligned} d_1((x_1, y_1), (x_2, y_2)) &= d(x_1, x_2) + d(y_1, y_2) \\ d_2((x_1, y_1), (x_2, y_2)) &= \sqrt{d(x_1, x_2)^2 + d(y_1, y_2)^2} \end{aligned}$$

1.2. **Balls and bounded sets.** Let (X, d) be a metric space.

Definition 1.22. For any $x \in X$ and $r > 0$ define

- (1) the open ball $B(x, r)$ with center x and radius r by

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

- (2) the closed ball $\overline{B}(x, r)$ with center x and radius r by

$$\overline{B}(x, r) = \{y \in X \mid d(x, y) \leq r\}.$$

Example 1.23. Describe an open ball $B(0, 1)$ and a closed ball $\overline{B}(0, 1)$ in $\mathbb{R}^2$ with metrics d_2, d_1, d_∞ .

Example 1.24. Let X be a discrete metric space and $x \in X$. Describe $B(x, 1), \overline{B}(x, 1), B(x, 2)$.

Definition 1.25. A subset $A \subset X$ is called bounded if it is contained in some open ball, that is, if there exist $x \in X$ and $r > 0$ s.t. $A \subset B(x, r)$

Definition 1.26. Define the diameter of a subset $A \subset X$ to be

$$d(A) := \sup_{x, y \in A} d(x, y).$$

Example 1.27. The diameter of $B(0, 1)$ in the Euclidean space $\mathbb{R}^2$ is 2.

Remark 1.28. If $A \subset B$ then $d(A) \leq d(B)$. Indeed, for any $x, y \in A$ we have also $x, y \in B$. Therefore

$$d(x, y) \leq d(B).$$

Taking the supremum over all $x, y \in A$ we obtain $d(A) \leq d(B)$.

Lemma 1.29. The diameter of an open ball $B(x, r)$ is $\leq 2r$.

Proof. For any two points $y, z \in B(x, r)$ we have

$$d(y, z) \leq d(y, x) + d(x, z) < r + r = 2r.$$

This implies

$$d(B(x, r)) = \sup_{y, z \in B(x, r)} d(y, z) \leq 2r.$$

□

Lemma 1.30. A subset $A \subset X$ is bounded if and only if its diameter is finite.

Proof. If A is bounded then $A \subset B(x, r)$ for some $x \in X, r > 0$. Then

$$d(A) \leq d(B(x, r)) \leq 2r < \infty.$$

Conversely, assume that $d(A) < \infty$ and let $x \in A, r = d(A) + 1$. Then, for any $y \in A$,

$$d(x, y) \leq d(A) < r \implies y \in B(x, r).$$

This means that $A \subset B(x, r)$.

□

Example 1.31. (1) $y = x^2$ in $\mathbb{R}^2$ is unbounded

- (2) Let X have a discrete metric and assume that $\#X \geq 2$. Then $d(X) = 1$. The set X and any its subset are bounded.

1.3. Open and closed sets.

1.3.1. *Open sets.* Let (X, d) be a metric space.

Definition 1.32. Let $A \subset X$ be a subset.

- (1) A point $x \in A$ is called an internal (or interior) point if there exists $r > 0$ such that $B(x, r) \subset A$.
- (2) A subset $A \subset X$ is called an open set if every point of A is an internal point. That is, if for any $x \in A$ there exists $r > 0$ such that $B(x, r) \subset A$.
- (3) The set of all internal points of A is denoted by A° and is called the interior of A .

Lemma 1.33. Every open ball $B(x, r)$ in X is an open set in X .

Proof. For any $y \in B(x, r)$ we have to find $r' > 0$ such that $B(y, r') \subset B(x, r)$. As $y \in B(x, r) \implies d(x, y) < r \implies r - d(x, y) > 0$. Let $r' = r - d(x, y)$. Then, for any $z \in B(y, r')$,

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + r' = r.$$

This implies that $z \in B(x, r)$. Therefore $B(y, r') \subset B(x, r)$ as required. $\square$

Example 1.34. Any open interval $(a, b) \subset \mathbb{R}$ is an open set as we can represent it as an open Ball $B(x, r)$, where $x = \frac{a+b}{2}$, $r = \frac{b-a}{2}$.

Theorem 1.35. We have

- (1) The sets $\emptyset$ and X are open sets in X .
- (2) The union of any collection of open sets is an open set.
- (3) The intersection of any finite collection of open sets is an open set.

Proof. (1) Clear.

- (2) Let $(U_i)_{i \in I}$ be a collection of open sets in X . Let $U = \bigcup_{i \in I} U_i$. For any $x \in U$, there exists $i \in I$: $x \in U_i$. As U_i is open, $\exists r > 0$: $B(x, r) \subset U_i \subset U$. This implies that x is an internal point of U . This proves that U is open.
- (3) Let $(U_i)_{i \in I}$ be a finite collection of open sets in X . Let $U = \bigcap_{i \in I} U_i$. Given $x \in U$, we have $x \in U_i \forall i \in I$. For any $i \in I$ the set U_i is open, therefore $\exists r_i > 0$: $x \in B(x, r_i)$. Let $r = \min_{i \in I} r_i$ (we can take this minimum as I is finite). Then $B(x, r) \subset U_i$ for any $i \in I$. Therefore $B(x, r) \subset \bigcap_{i \in I} U_i = U$. This implies that x is an internal point of U . This proves that U is open. $\square$

Example 1.36. The union of open intervals in $\mathbb{R}$ is an open set. In particular, the set $(a, +\infty) = \bigcup_{n \geq 1} (a, a+n)$ is an open set in $\mathbb{R}$.

Remark 1.37. It is not true in general that the intersection of any collection of open sets is an open set. For example, consider a collection of open sets in $\mathbb{R}$ defined by

$$U_n = \left(-\frac{1}{n}, \frac{1}{n}\right), \quad n \geq 1.$$

Their intersection is

$$\bigcap_{n \geq 1} U_n = \bigcap_{n \geq 1} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\},$$

which is not an open set.

Lemma 1.38. For any subset $A \subset X$ its interior A° is an open set. It is the union of all open subsets of A .

Proof. For every $x \in A^\circ$, there exists $\varepsilon_x > 0$ such that $B(x, \varepsilon_x) \subset A$. Then

$$A^\circ \subset U = \bigcup_{x \in A^\circ} B(x, \varepsilon_x) \subset A.$$

The set U is open as a union of open balls. Therefore all points of U are internal in U , hence also in A . This implies that $U \subset A^\circ$ and we conclude that $A^\circ = U$ and it is open.

If $V \subset A$ is the union of all open subsets of A , then $A^\circ \subset V$. On the other hand all points of V are internal in A , hence $V \subset A^\circ$ and we conclude that $A^\circ = V$. $\square$

1.3.2. Closed sets.

Definition 1.39. Let $A \subset X$ be a subset.

- (1) A point $x \in X$ is called a limit point of A if for any $\varepsilon > 0$ there exists $y \in A \setminus \{x\}$ such that $d(x, y) < \varepsilon$.
- (2) A is called a closed set if it contains all of its limit points.
- (3) Define the closure of A to be the union of A and the set of all its limit points. It is denoted by $\overline{A}$.

Theorem 1.40. A subset $A \subset X$ is a closed set if and only if $X \setminus A$ is an open set.

Proof. Assume that A is a closed set. To prove that $X \setminus A$ is open, we have to show that for any $x \in X \setminus A$ there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subset X \setminus A$. If this is not the case then for any $\varepsilon > 0$ the set $B(x, \varepsilon)$ is not contained in $X \setminus A$, that is,

$$B(x, \varepsilon) \cap A \neq \emptyset.$$

This implies that there exists $y \in A$ such that $d(x, y) < \varepsilon$. This means that x is a limit point of A . As A is closed, we conclude that $x \in A$. But by our assumption $x \in X \setminus A$, a contradiction.

Conversely, assume that $X \setminus A$ is an open set. We have to show that every limit point x of A is contained in A . Assume that $x \in X \setminus A$. As $X \setminus A$ is open, there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subset X \setminus A$, that is, $B(x, \varepsilon) \cap A = \emptyset$. But this implies that x can not be a limit point of A , a contradiction. $\square$

Example 1.41. The closed interval $[a, b] \subset \mathbb{R}$ is a closed set. Indeed, its complement $\mathbb{R} \setminus [a, b] = (-\infty, a) \cup (b, \infty)$ is a union of open intervals and we know that they are open sets.

Theorem 1.42. Let (X, d) be a metric space. Then

- (1) The sets $\emptyset$ and X are closed sets in X .
- (2) The intersection of any collection of closed sets is a closed set.
- (3) The union of any finite collection of closed sets is an closed set.

Proof. (1) $X \setminus \emptyset = X$ is an open set $\implies \emptyset$ is closed. $X \setminus X = \emptyset$ is an open set $\implies X$ is closed.

- (2) Let $(F_i)_{i \in I}$ be a collection of closed sets. By the De Morgan's law

$$X \setminus \bigcap_{i \in I} F_i = \bigcup_{i \in I} (X \setminus F_i).$$

The sets $X \setminus F_i$ are open $\implies \bigcup_{i \in I} (X \setminus F_i)$ is open $\implies X \setminus \bigcap_{i \in I} F_i$ is open $\implies \bigcap_{i \in I} F_i$ is closed.

- (3) Let $(F_i)_{i \in I}$ be a finite collection of closed sets. By the De Morgan's law

$$X \setminus \bigcup_{i \in I} F_i = \bigcap_{i \in I} (X \setminus F_i).$$

The sets $X \setminus F_i$ are open $\implies \bigcap_{i \in I} (X \setminus F_i)$ is open $\implies X \setminus \bigcup_{i \in I} F_i$ is open $\implies \bigcup_{i \in I} F_i$ is closed. $\square$

Theorem 1.43. For any subset $A \subset X$ its closure $\overline{A}$ is a closed set. It is the intersection of all closed sets containing A .

Proof. Let x be a limit point of $\overline{A}$ such that $x \notin \overline{A}$. Then $\forall \varepsilon > 0 \exists y \in \overline{A}$ with $d(x, y) < \varepsilon/2$. If $y \in \overline{A} \setminus A$ then $\exists z \in A$ with $d(y, z) < \varepsilon/2$, hence

$$d(x, z) \leq d(x, y) + d(y, z) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

If $y \in A$ then we take $z = y$ and also obtain $d(x, z) < \varepsilon/2 < \varepsilon$. This implies that x is a limit point of A , hence $x \in \overline{A}$, a contradiction. We conclude that $\overline{A}$ is closed.

If F is any closed set containing A then all limit points of A are in F . Therefore $\overline{A} \subset F$. The set $\overline{A}$ is one of the closed sets containing A . Therefore $\overline{A} = \bigcap_{A \subset F - \text{cl}} F$. $\square$

Example 1.44. Consider $A = (0, 1) \subset \mathbb{R}$. Its closure is $[0, 1]$ and we have seen that it a closed set.

1.4. Convergent sequences. Let (X, d) be a metric space. A sequence $(x_n)_{n=1}^{\infty}$ of points in X is a collection of elements $x_1, x_2, \dots$ in X . We denote a sequence $(x_n)_{n=1}^{\infty}$ also by $(x_n)_{n \geq 1}$ or just (x_n) .

Definition 1.45. A sequence (x_n) of points in X is called convergent to $x \in X$ if $\forall \varepsilon > 0 \exists N > 0$ such that

$$d(x_n, x) < \varepsilon \quad \forall n \geq N.$$

The point x is called the limit of the sequence and is denoted by $\lim_{n \rightarrow \infty} x_n$. We also write $x_n \rightarrow x$ as $n \rightarrow \infty$.

Remark 1.46. The definition of convergence of a sequence (x_n) to x is equivalent to the statement that $d(x_n, x) \rightarrow 0, n \rightarrow \infty$.

Lemma 1.47. Every convergent sequence in a metric space has a unique limit.

Proof. Assume that $(x_n)_{n \geq 1}$ converges to $x, y \in X$. For any $\varepsilon > 0$ there exist $N, N' > 0$ such that $d(x_n, x) < \varepsilon$ for $n \geq N$ and $d(x_n, y) < \varepsilon$ for $n \geq N'$. Then for any $n \geq N, N'$ we have

$$d(x, y) \leq d(x, x_n) + d(x_n, y) < \varepsilon + \varepsilon = 2\varepsilon.$$

This inequality is true for any $\varepsilon > 0 \implies d(x, y) = 0 \implies x = y$. $\square$

Example 1.48. Consider a sequence $(x_n)_{n \geq 1}$ in $\mathbb{R}^2$ with $x_n = (\frac{1}{n}, \frac{n-1}{n})$. Then $\lim_{n \rightarrow \infty} x_n = x = (0, 1)$. Indeed,

$$d(x_n, x) = \sqrt{\left(\frac{1}{n} - 0\right)^2 + \left(\frac{n-1}{n} - 1\right)^2} = \sqrt{\frac{1}{n^2} + \frac{1}{n^2}} = \frac{\sqrt{2}}{n} \rightarrow 0, \quad n \rightarrow \infty.$$

This example shows that convergence in $\mathbb{R}^k$ can be studied coordinatewise. This is the content of the next lemma.

Lemma 1.49. A sequence $(x^{(n)})_{n \geq 1}$ in the Euclidean space $\mathbb{R}^k$ converges to $x \in \mathbb{R}^k \iff x_i^{(n)} \rightarrow x_i$ as $n \rightarrow \infty$, for all $1 \leq i \leq k$.

Proof. $\implies$ If $x^{(n)} \rightarrow x$ as $n \rightarrow \infty$, then $d(x^{(n)}, x) = \sqrt{\sum_i (x_i^{(n)} - x_i)^2} \rightarrow 0 \implies (x_i^{(n)} - x_i)^2 \rightarrow 0$ for all $1 \leq i \leq k \implies x_i^{(n)} \rightarrow x_i$ as $n \rightarrow \infty$.

$\Leftarrow$ If $x_i^{(n)} \rightarrow x_i$ as $n \rightarrow \infty$ for all i , then $(x_i^{(n)} - x_i)^2 \rightarrow 0$ for all i , hence $d(x^{(n)}, x) = \sqrt{\sum_i (x_i^{(n)} - x_i)^2} \rightarrow 0$. This means that $x^{(n)} \rightarrow x$ as $n \rightarrow \infty$. $\square$

Sequences are useful for the characterization of limit points of subsets.

Lemma 1.50. Let $A \subset X$ be a subset. A point $x \in X$ is a limit point of $A \subset X \iff$ there exists a sequence in $A \setminus \{x\}$ convergent to x .

Proof. $\implies$ Let $x \in X$ be a limit point of A . Then, for any $n \geq 1$ there exists $x_n \in A \setminus \{x\}$ such that $d(x_n, x) < 1/n$. This implies that the sequence $(x_n)_{n \geq 1}$ is contained in $A \setminus \{x\}$ and converges to x .

$\Leftarrow$ Assume that there exists a sequence $(x_n)_{n \geq 1}$ in $A \setminus \{x\}$ convergent to x . Then, for any $\varepsilon > 0$ there exists $n \geq 1$ such that $d(x_n, x) < \varepsilon$. This implies that x is a limit point of A . $\square$

Corollary 1.51. A subset $A \subset X$ is closed $\iff$ the limit of any convergent sequence of points in A is contained in A .

Proof. $\implies$ If A is closed and (x_n) is a sequence in A convergent to $x \in X$, then x is a limit point of A , hence $x \in A$. $\Leftarrow$ If x is a limit point of A , then there exists a sequence (x_n) in $A \setminus \{x\}$ convergent to x . By our assumption this limit is contained in A , hence $x \in A$ and A is closed set. $\square$

Example 1.52. Consider the space $C[0, 1]$ equipped with the metric d_1 . Consider the sequence of elements $x_n \in C[0, 1]$ defined by $x_n(t) = t^n$ for $t \in [0, 1]$. For every $t \in [0, 1)$, we have $\lim_{n \rightarrow \infty} x_n(t) = \lim_{n \rightarrow \infty} t^n = 0$. Therefore we can guess that the sequence (x_n) converges to the constant zero function $0 \in C[0, 1]$. This is indeed the case as

$$d_1(x_n, 0) = \int_0^1 |t^n - 0| dt = \int_0^1 t^n dt = \frac{t^{n+1}}{n+1} \Big|_0^1 = \frac{1}{n+1} \rightarrow 0, \quad n \rightarrow \infty.$$

On the other hand, let us equip $C[0, 1]$ with the metric d_∞ (called the supremum metric). Then

$$d_\infty(x_n, 0) = \sup_{t \in [0, 1]} |t^n - 0| = 1,$$

hence $x_n \not\rightarrow 0$ as $n \rightarrow \infty$ with respect to the metric d_∞ .

Example 1.53. Let us again equip $C[a, b]$ with the supremum metric d_∞ . Then a sequence of functions (x_n) in $C[a, b]$ converges to a function $x \in C[a, b]$ $\iff \forall \varepsilon > 0 \exists N > 0$:

$$d_\infty(x_n, x) = \sup_{t \in [a, b]} |x_n(t) - x(t)| \leq \varepsilon \quad \forall n \geq N$$

or equivalently

$$|x_n(t) - x(t)| \leq \varepsilon \quad \forall n \geq N \quad \forall t \in [a, b].$$

This leads us to the following definition.

Definition 1.54. Let $x(t)$ be a function and $(x_n(t))_{n \geq 1}$ be a sequence of functions on the interval $[a, b]$.

- (1) A sequence $(x_n(t))_{n \geq 1}$ is said to **converge pointwise** to a function $x(t)$ if

$$\forall t \in [a, b] \quad \forall \varepsilon > 0 \exists N > 0: \quad |x_n(t) - x(t)| < \varepsilon \quad \forall n \geq N$$

or equivalently, $\forall t \in [a, b]$ we have $\lim_{n \rightarrow \infty} x_n(t) = x(t)$.

- (2) A sequence $(x_n(t))$ is said to **converge uniformly** to a function $x(t)$ if

$$\forall \varepsilon > 0 \exists N > 0: \quad |x_n(t) - x(t)| < \varepsilon \quad \forall n \geq N \quad \forall t \in [a, b].$$

Remark 1.55.

- (1) Note that for pointwise convergence we can choose N depending on t , while for uniform convergence we have to choose N that works for all t simultaneously (uniformly).
- (2) Uniform convergence implies pointwise convergence.
- (3) We can see from the previous example that uniform convergence is equivalent to the convergence with respect to the metric d_∞ .

Example 1.56. Consider the sequence of functions $x_n(t) = t^n - t^{2n}$ in $C[0, 1]$. It converges pointwise to the zero function as $\lim_{n \rightarrow \infty} x_n(t) = \lim_{n \rightarrow \infty} (t^n - t^{2n}) = 0$ for every $t \in [0, 1]$. On the other hand this sequence does not converge uniformly to the zero function as

$$d_\infty(x_n, 0) = \sup_{t \in [0, 1]} (t^n - t^{2n}) = \frac{1}{4} \quad \forall n \geq 1.$$

Example 1.57. We have seen earlier that the sequence of functions $x_n(t) = t^n$ in $C[0, 1]$ does not converge to the zero function with respect to the metric d_∞ . But does it converge to any other function $x \in C[0, 1]$? If this is the case, then (x_n) converges uniformly to x , hence also pointwise. Then, for any $t \in [0, 1]$, we have

$$x(t) = \lim_{n \rightarrow \infty} x_n(t) = \lim_{n \rightarrow \infty} t^n = \begin{cases} 0, & 0 \leq t < 1 \\ 1, & t = 1 \end{cases}$$

But this function is not continuous, hence the sequence (x_n) does not converge in $(C[0, 1], d_\infty)$.

Lemma 1.58. Let (x_n) be a sequence in X convergent to $x \in X$ and let $y \in X$. Then

$$\lim_{n \rightarrow \infty} d(x_n, y) = d(x, y).$$

Proof. In the proof of this lemma we will need the following useful result

Lemma 1.59. For any points $x, y, z \in X$ we have

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

Proof. From the triangle inequalities $d(x, y) \leq d(x, z) + d(y, z)$ and $d(x, z) \leq d(x, y) + d(y, z)$, we obtain $d(x, y) - d(x, z) \leq d(y, z)$ and $d(x, z) - d(x, y) \leq d(y, z)$, hence the required statement. $\square$

Going back to the original lemma, we obtain

$$|d(x_n, y) - d(x, y)| \leq d(x_n, x) \rightarrow 0, \quad n \rightarrow \infty.$$

Therefore $d(x_n, y) \rightarrow d(x, y)$ as $n \rightarrow \infty$. $\square$

1.5. Continuous maps.

Definition 1.60. Let (X, d) , (Y, d') be two metric spaces. A map $f: X \rightarrow Y$ is called

- (1) **continuous at a point** $x \in X$ if $\forall \varepsilon > 0 \exists \delta > 0$:

$$d(x, y) < \delta \implies d'(f(x), f(y)) < \varepsilon.$$

- (2) **continuous** if it is continuous at every point of X .

Remark 1.61. Continuity at the point x can be written in the form

$$\forall \varepsilon > 0 \exists \delta > 0: \quad y \in B(x, \delta) \implies f(y) \in B(f(x), \varepsilon)$$

or in the form

$$\forall \varepsilon > 0 \exists \delta > 0: \quad f(B(x, \delta)) \subset B(f(x), \varepsilon).$$

Example 1.62. A constant function $f: X \rightarrow Y$ is continuous. Assume that $f(x) = z$ for all $x \in X$ and some $z \in Y$. For any $\varepsilon > 0$ choose arbitrary $\delta > 0$. Then $f(B(x, \delta)) = \{z\} \subset B(f(x), \varepsilon)$.

Example 1.63. Let X have the discrete metric. Then any map $f: X \rightarrow Y$ is continuous. For every $\varepsilon > 0$ we choose $\delta = 1$. Then $B(x, \delta) = B(x, 1) = \{x\}$ and the condition $f(x) \in B(f(x), \varepsilon)$ is obviously satisfied.

Theorem 1.64. A map $f: X \rightarrow Y$ is continuous $\iff$ for any sequence (x_n) in X convergent to $x \in X$, the sequence $(f(x_n))$ converges to $f(x)$.

Proof. $\implies$ Let f be continuous and let (x_n) be a sequence in X convergent to $x \in X$. Then $\forall \varepsilon > 0 \exists \delta > 0$

$$d(x, y) < \delta \implies d'(f(x), f(y)) < \varepsilon.$$

As $x_n \rightarrow x$, there exists $N > 0$

$$d(x, x_n) < \delta \quad \forall n \geq N.$$

This implies

$$d'(f(x), f(x_n)) < \varepsilon \quad \forall n \geq N,$$

hence $f(x_n) \rightarrow f(x)$ as $n \rightarrow \infty$.

$\Leftarrow$ Assume that f is not continuous at $x \in X$. This means that $\exists \varepsilon > 0 \forall \delta > 0 \exists y \in X$ such that

$$d(x, y) < \delta, \quad d'(f(x), f(y)) \geq \varepsilon.$$

For any $n \geq 1$, choosing $\delta = 1/n$, we can find $x_n \in X$ with $d(x, x_n) < 1/n$ and $d'(f(x), f(x_n)) \geq \varepsilon$. In this way we obtain a sequence (x_n) such that $x_n \rightarrow x$, but $f(x_n) \not\rightarrow f(x)$. A contradiction. $\square$

Theorem 1.65. A map $f: X \rightarrow Y$ is continuous $\iff$ for any open set $U \subset Y$ the preimage $f^{-1}(U) = \{x \in X \mid f(x) \in U\}$ is open in X .

Proof. $\implies$ Assume that $f: X \rightarrow Y$ is continuous and let $U \subset Y$ be open. For any point $x \in f^{-1}(U)$, the point $f(x) \in U$ is an internal point of U . Therefore $\exists \varepsilon > 0$ with $B(f(x), \varepsilon) \subset U$. By the continuity of f , there exists $\delta > 0$ such that

$$f(B(x, \delta)) \subset B(f(x), \varepsilon) \subset U.$$

Then $B(x, \delta) \subset f^{-1}(U) \implies x$ is an internal point of $f^{-1}(U)$. Therefore $f^{-1}(U)$ is an open set.

$\Leftarrow$ Let us show that f is continuous at any point $x \in X$. For every $\varepsilon > 0$, the set $U = B(f(x), \varepsilon)$ is open in Y . Therefore its preimage $f^{-1}(U)$ is open in X . In particular, the point $x \in f^{-1}(U)$ is an internal point. Therefore there exists $\delta > 0$ such that $B(x, \delta) \subset f^{-1}(U)$, that is,

$$f(B(x, \delta)) \subset U = B(f(x), \varepsilon).$$

This means that f is continuous at $x \in X$. Therefore f is continuous. $\square$

Lemma 1.66. Let $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ be continuous maps between metric spaces. Then their composition $gf: X \rightarrow Z$ is also a continuous map.

Proof. We have to show that for any open set $U \subset Z$ the preimage $(gf)^{-1}(U)$ is also open. We have

$$(gf)^{-1}(U) = f^{-1}(g^{-1}(U)).$$

If $U \subset Z$ is open, then $g^{-1}(U) \subset Y$ is open as g is continuous. Therefore $f^{-1}(g^{-1}(U))$ is open as f is continuous. This implies that $(gf)^{-1}(U)$ is open. $\square$

Definition 1.67. A map $f: (X, d) \rightarrow (Y, d')$ is called **Lipschitz continuous** if there exists $L \geq 0$ such that

$$d'(f(x), f(y)) \leq L \cdot d(x, y) \quad \forall x, y \in X.$$

The number L above is called a **Lipschitz constant** for f .

Theorem 1.68. If $f: X \rightarrow Y$ is Lipschitz continuous, then it is continuous.

Proof. Let $L > 0$ be a Lipschitz constant for f . For any $\varepsilon > 0$ choose $\delta = \varepsilon/L$. Then if $d(x, y) < \delta \implies d'(f(x), f(y)) \leq L \cdot d(x, y) < L\delta = \varepsilon$. Therefore f is continuous. $\square$

Example 1.69. Consider the function $f(t) = \sqrt{t}$ for $t \in [0, 1]$. This function is continuous, although it is not Lipschitz continuous. Indeed, if f would be Lipschitz continuous with a Lipschitz constant $L > 0$, then in particular,

$$\left| \sqrt{t} - 0 \right| \leq L \cdot |t - 0| = Lt,$$

hence $\sqrt{t} \geq \frac{1}{L}$ for $t > 0$. But this is obviously wrong for $\sqrt{t} = \frac{1}{2L}$.

Lemma 1.70. Let $f: [a, b] \rightarrow \mathbb{R}$ be a function having a continuous derivative. Then f is Lipschitz continuous with the Lipschitz constant $\max_{t \in [a, b]} |f'(t)|$.

Proof. The function $f': [a, b] \rightarrow \mathbb{R}$ is continuous, hence bounded. Let $L = \max_{t \in [a, b]} |f'(t)|$. By the mean value theorem, for any $a \leq x < y \leq b$, there exists $t \in (x, y)$ such that

$$\frac{f(y) - f(x)}{y - x} = f'(t).$$

Therefore

$$|f(y) - f(x)| = |f'(t)| \cdot |x - y| \leq L \cdot |x - y|$$

and f is Lipschitz continuous. $\square$

Definition 1.71. A map $f: (X, d) \rightarrow (Y, d')$ between two metric spaces is called an **isometry** if

$$d'(f(x), f(y)) = d(x, y) \quad \forall x, y \in X.$$

An isometry is called a **global isometry** if it is bijective.

Lemma 1.72. Let $f: (X, d) \rightarrow (Y, d')$ be an isometry. Then

- (1) f is continuous.
- (2) f is injective.
- (3) If f is a global isometry, then $f^{-1}: Y \rightarrow X$ is continuous.

Proof. (1) An isometry is Lipschitz continuous with a Lipschitz constant $L = 1$. Therefore it is also continuous.

(2) Assume that $x, y \in X$ satisfy $f(x) = f(y)$. Then

$$d(x, y) = d'(f(x), f(y)) = 0,$$

hence $x = y$. This means that f is injective.

(3) If f is a global isometry then $f^{-1}: Y \rightarrow X$ is also an isometry. Therefore by (1) it is continuous. $\square$

1.6. Complete metric spaces. Let (X, d) be a metric space.

Definition 1.73. A sequence $(x_n)_{n \geq 1}$ in X is called a **Cauchy sequence** (or a fundamental sequence) if $\forall \varepsilon > 0 \exists N > 0$:

$$d(x_m, x_n) < \varepsilon \quad \forall m, n \geq N.$$

Lemma 1.74. A convergent sequence in a metric space is a Cauchy sequence.

Proof. Let $(x_n)_{n \geq 1}$ be a sequence in X convergent to $x \in X$. Then $\forall \varepsilon > 0 \exists N > 0$

$$d(x_n, x) < \frac{\varepsilon}{2} \quad \forall n \geq N.$$

This implies that $\forall m, n \geq N$

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

hence $(x_n)_{n \geq 1}$ is a Cauchy sequence. $\square$

Example 1.75. There exist Cauchy sequences that are not convergent. For example, the sequence $(1/n)_{n \geq 1}$ is a Cauchy sequence which is not convergent in $(0, 1]$. It is a Cauchy sequence as it converges to 0 in $\mathbb{R}$. It is not convergent in $(0, 1]$ as the limit $\lim_{n \rightarrow \infty} \frac{1}{n} = 0 \notin (0, 1]$.

Definition 1.76. A metric space (X, d) is called a **complete metric space** if any Cauchy sequence in X converges in X .

Example 1.77. We obtain from the previous example that $(0, 1]$ is not complete.

Theorem 1.78. The space $\mathbb{R}$ is a complete metric space.

We will need some auxiliary statements.

Theorem 1.79. A bounded monotone sequence in $\mathbb{R}$ is convergent.

Proof. Without loss of generality we assume that (x_n) is a non-decreasing sequence bounded from above. By a property of real numbers there exists $x = \sup \{x_n \mid n \geq 1\}$. Then for any $\varepsilon > 0$ there exists $N \geq 1$ such that $x_N > x - \varepsilon$. Therefore for all $n \geq N$ we have $x - \varepsilon < x_N \leq x_n \leq x$, hence $|x_n - x| < \varepsilon$. This implies that (x_n) converges to x . $\square$

Theorem 1.80 (Bolzano-Weierstrass theorem). A bounded sequence in $\mathbb{R}$ has a convergent subsequence.

Proof. Let (x_n) be a bounded sequence. We will say that $n \geq 1$ is a peak if $x_n > x_m$ for all $m > n$. If there exist infinitely many peaks $n_1 < n_2 < \dots$, then we have $x_{n_1} > x_{n_2} > \dots$, hence we found a monotone bounded subsequence, which is convergent by the previous result. If there exist only finitely many peaks, then we can choose $N \geq 1$ such every $n \geq N$ is not a peak. As $n_1 = N$ is not a peak, there exists $n_2 > n_1$ with $x_{n_1} \leq x_{n_2}$. Continuing this process we find a bounded monotone subsequence $x_{n_1} \leq x_{n_2} \leq \dots$ which is convergent by the previous result. $\square$

Proof of Theorem 1.78. Let (x_n) be a Cauchy sequence. This sequence is bounded. Indeed, for any $\varepsilon > 0$ there exists $N \geq 1$ such that $d(x_m, x_n) \leq \varepsilon$ for all $m, n \geq N$. In particular, $d(x_N, x_n) \leq \varepsilon$ for $n \geq 1$. Choosing $r = \max \{\varepsilon, d(x_N, x_n) \mid 1 \leq n \leq N\}$, we obtain $x_n \in B(x_N, r + 1)$ for all $n \geq 1$.

By the Bolzano-Weierstrass theorem there exists a convergent subsequence (x_{n_k}) , with a limit, say $x \in \mathbb{R}$. For every $\varepsilon > 0$ there exists $N \geq 1$ such that $d(x_m, x_n) < \frac{\varepsilon}{2}$ for all $m, n \geq N$. On the other hand there exists $K \geq 1$ such that $n_K \geq N$ and $d(x_{n_k}, x) < \frac{\varepsilon}{2}$ for all $k \geq K$. This implies that for all $n \geq N$ we have $d(x_n, x) \leq d(x_n, x_{n_K}) + d(x_{n_K}, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. $\square$

Example 1.81. Let us show that the Euclidean space $\mathbb{R}^k$ is complete. Let $(x^{(n)})_{n \geq 1}$ be a Cauchy sequence in $\mathbb{R}^k$. Then for any $1 \leq i \leq k$, the sequence $(x_i^{(n)})_{n \geq 1}$ in $\mathbb{R}$ is a Cauchy sequence: $\forall \varepsilon > 0 \exists N > 0$

$$\forall m, n \geq N: \quad d(x^{(m)}, x^{(n)}) = \sqrt{\sum (x_i^{(m)} - x_i^{(n)})^2} < \varepsilon,$$

hence $|x_i^{(m)} - x_i^{(n)}| < \varepsilon$. But we know that $\mathbb{R}$ is complete, hence the sequence $(x_i^{(n)})_{n \geq 1}$ converges to some $x_i \in \mathbb{R}$. We conclude that $(x^{(n)})_{n \geq 1}$ converges to $x = (x_1, \dots, x_k)$.

Lemma 1.82. *Let (X, d) be a metric space and $A \subset X$ be a subset. Then*

- (1) *If X is complete and A is closed then (A, d) is complete.*
- (2) *If (A, d) is complete then A is closed.*

Proof. (1) Assume that (x_n) is a Cauchy sequence in A . Then it is Cauchy in $X \implies$ it is convergent to some $x \in X$. As A is closed, we conclude that $x \in A$. This means that (x_n) is convergent to a point in A . Therefore A is complete.

(2) Assume that A is complete. Let (x_n) be a sequence in A convergent to $x \in X$. We have to show that $x \in A$. As (x_n) is a convergent sequence, it is a Cauchy sequence. Therefore it converges to some $y \in A$. We conclude from the uniqueness of limits that $x = y \in A$. $\square$

Example 1.83. A subspace $[a, b] \subset \mathbb{R}$ is complete as $\mathbb{R}$ is complete and $[a, b]$ is closed. A subspace $\mathbb{Q} \subset \mathbb{R}$ is not complete as $\mathbb{Q}$ is not closed in $\mathbb{R}$ (we have seen that $\overline{\mathbb{Q}} = \mathbb{R}$).

Theorem 1.84. *The space $C[a, b]$ with the supremum metric d_∞ is complete.*

Proof. Let (x_n) be a Cauchy sequence in $C[a, b]$. Then, for any $t_0 \in [a, b]$, the sequence $(x_n(t_0))_{n \geq 1}$ of real numbers is also a Cauchy sequence: $\forall \varepsilon > 0 \exists N > 0 \forall m, n \geq N$:

$$(1) \quad |x_m(t_0) - x_n(t_0)| \leq \sup_{t \in [a, b]} |x_m(t) - x_n(t)| = d_\infty(x_m, x_n) < \varepsilon.$$

Therefore this sequence converges to some $x(t_0) \in \mathbb{R}$. In this way we obtain a function $x: [a, b] \rightarrow \mathbb{R}$. We have to show that (x_n) converges uniformly to x and that x is a continuous function. We can write the Cauchy condition in the form: $\forall \varepsilon > 0 \exists N > 0$

$$\forall m, n \geq N \forall t \in [a, b]: \quad |x_m(t) - x_n(t)| < \varepsilon.$$

Taking the limit $m \rightarrow \infty$ we obtain

$$\forall n \geq N \forall t \in [a, b]: \quad |x(t) - x_n(t)| \leq \varepsilon$$

hence the sequence of functions (x_n) converges uniformly to $x: [a, b] \rightarrow \mathbb{R}$. The continuity of x follows from the next lemma. $\square$

Lemma 1.85. *Assume that a sequence of functions $(x_n(t))$ in $C[a, b]$ converges uniformly to a function $x: [a, b] \rightarrow \mathbb{R}$. Then $x \in C[a, b]$.*

Proof. Let us show that $x: [a, b] \rightarrow \mathbb{R}$ is continuous at $t_0 \in [a, b]$. For any $\varepsilon > 0 \exists N > 0$

$$|x(t) - x_n(t)| < \frac{\varepsilon}{3} \quad \forall n \geq N \forall t \in [a, b].$$

We will need just $n = N$. As $x_N \in C[a, b]$, there exists $\delta > 0$

$$|x_N(t) - x_N(t_0)| < \frac{\varepsilon}{3} \quad \forall t \in [t_0 - \delta, t_0 + \delta].$$

Then $\forall t \in [t_0 - \delta, t_0 + \delta]$:

$$|x(t) - x(t_0)| \leq |x(t) - x_N(t)| + |x_N(t) - x_N(t_0)| + |x_N(t_0) - x(t_0)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

This implies that x is continuous at t_0 . Therefore x is continuous on $[a, b]$. $\square$

1.7. Completion. Assume that X is a metric space which is not complete. We would like to embed X into a larger metric space $X \subset X^*$ such that X^* is complete and such that the closure of X is equal to X^* .

Definition 1.86. A subset $A \subset X$ is called **dense** in X if $\overline{A} = X$.

Example 1.87.

- (1) Let $X = [0, 1]$ and $A = (0, 1) \subset X$. Then the closure of A is $[0, 1] = X$, hence A is dense in X .
- (2) The set $\mathbb{Q}$ is dense in $\mathbb{R}$. Indeed, any $x \in \mathbb{R}$ has a decimal representation

$$x = a_0.a_1a_2\dots$$

where $a_i \in \mathbb{Z}$ and $0 \leq a_i \leq 9$ for $i \geq 1$ (the digits). For any $n \geq 1$ the numbers

$$x_n = a_0.a_1\dots a_n = \sum_{i=0}^n \frac{a_i}{10^i}$$

are rational. Moreover, $\lim_{n \rightarrow \infty} x_n = x$ as

$$|x - x_n| = 0.0\dots 0a_{n+1}a_{n+2}\dots = \sum_{i \geq n+1} \frac{a_i}{10^i} \leq \sum_{i \geq n+1} \frac{9}{10^i} = \frac{1}{10^n} \rightarrow 0, \quad n \rightarrow \infty.$$

This implies that x is a limit point of $\mathbb{Q}$. Therefore, the closure of $\mathbb{Q}$ equals $\mathbb{R}$.

Lemma 1.88. Let $f, g: X \rightarrow Y$ be two continuous maps between metric spaces. Assume that there exists a dense subset $A \subset X$ such that $f(x) = g(x)$ for all $x \in A$. Then $f(x) = g(x)$ for all $x \in X$.

Proof. Let $F = \{x \in X \mid f(x) = g(x)\}$. We claim that F is closed in X . Indeed, let (x_n) be a sequence in F convergent to $x \in X$. By the continuity of f and g , we have $f(x_n) \rightarrow f(x)$ and $g(x_n) \rightarrow g(x)$ as $n \rightarrow \infty$. But $x_n \in F \implies f(x_n) = g(x_n) \implies$ the limits of these sequences coincide, hence $f(x) = g(x)$ and $x \in F$. This means that F is closed. But $A \subset F$, hence $X = \overline{A} \subset F$ and $F = X$. Therefore $f(x) = g(x) \forall x \in X$. $\square$

Definition 1.89. A **completion** of a metric space (X, d) is a complete metric space (X^*, d^*) together with an isometry $i: X \rightarrow X^*$ such that $\overline{i(X)} = X^*$ (that is, $i(X)$ is dense in X^*).

Remark 1.90. We have seen that an isometry is an injective map. Therefore we will often identify X with its image $i(X) \subset X^*$.

Example 1.91. Let $\mathbb{Q}$ be equipped with the Euclidean metric. Its completion is the Euclidean space $\mathbb{R}$ with the canonical inclusion $i: \mathbb{Q} \rightarrow \mathbb{R}$, $x \mapsto x$. Indeed

- (1) We know that $\mathbb{R}$ is complete.
- (2) The canonical inclusion is an isometry.
- (3) We have seen that $\mathbb{Q}$ is dense in $\mathbb{R}$.

Theorem 1.92. Every metric space has a completion (unique up to an isometry).

Proof. The idea of the construction is the following:

- (1) Define an equivalence relation on the set of all Cauchy sequences in X

$$(x_n) \sim (y_n) \iff \lim_{n \rightarrow \infty} d(x_n, y_n) = 0.$$

- (2) Define the completion X^* to be the set of all equivalence classes of Cauchy sequences.
- (3) Define the metric on X^* by $d([(x_n)], [(y_n)]) = \lim_{n \rightarrow \infty} d(x_n, y_n)$.
- (4) Construct an embedding $i: X \rightarrow X^*$ by sending $x \in X$ to the (equivalence class of) constant sequence (x_n) with $x_n = x$ for all $n \geq 1$.

Uniqueness: Assume that X has two completions $i: X \rightarrow X^*$ and $i': X \rightarrow X'$. Let us construct an isometry $\phi: X^* \rightarrow X'$ that preserves X . We will identify $x \in X$ with $i(x) \in X^*$ and $i'(x) \in X'$. For any point $x^* \in X^*$ there exists a sequence (x_n) in X that converges to x^* . The sequence (x_n) is a Cauchy sequence, therefore it converges to some $x' \in X'$. This point does not depend on the choice of the sequence (x_n) converging to x^* . We define $\phi(x^*) = x'$. For any $x \in X$, we have $\phi(x) = x$. To see that $\phi: X^* \rightarrow X'$ is an isometry, consider two Cauchy sequences $(x_n), (y_n)$ in X that converge to $x^*, y^* \in X^*$ and $x', y' \in X'$. Then

$$d^*(x^*, y^*) = \lim_{n \rightarrow \infty} d(x_n, y_n) = d'(x', y').$$

This implies that $d'(\phi(x^*), \phi(y^*)) = d^*(x^*, y^*)$. We can also show that ϕ is a bijection.

Existence:

1. Construction of X^* . Let $\bar{X}$ be the set of all Cauchy sequences in X . We call two sequences (x_n) , (y_n) equivalent if $\lim_{n \rightarrow \infty} d(x_n, y_n) = 0$. This is an equivalence relation (one can check reflexivity, symmetry, transitivity). Let X^* be the set of equivalence classes.

2. Metric on X^* . For any two equivalence classes $x^* = [(x_n)]$ and $y^* = [(y_n)]$ in X^* define

$$d(x^*, y^*) = \lim_{n \rightarrow \infty} d(x_n, y_n).$$

- (1) The sequence $(d(x_n, y_n))_{n \geq 1}$ is Cauchy and therefore converges in $\mathbb{R}$.
- (2) The limit $\lim_{n \rightarrow \infty} d(x_n, y_n)$ is independent of the choice of Cauchy sequences in the classes x^* and y^* .

To show that d defines a metric on X^* , we have to check the triangle inequality. Let $x^*, y^*, z^* \in X^*$ and let $(x_n), (y_n), (z_n)$ be their representatives. For any $n \geq 1$ we have

$$d(x_n, y_n) \leq d(x_n, z_n) + d(z_n, y_n).$$

Taking the limits, we obtain

$$d(x^*, y^*) \leq d(x^*, z^*) + d(z^*, y^*).$$

3. Embedding of X into X^* . For any $x \in X$, consider the constant sequence $x_n = x$ for $n \geq 1$. This is a Cauchy sequence that defines a point in X^* . This embedding $X \subset X^*$ is an isometry.

4. Density of X in X^* . Indeed, let $x^* \in X^*$ and let (x_n) be its representative. Then (x_n) is a Cauchy sequence $\implies \forall \varepsilon > 0 \exists N > 0$

$$\forall m, n \geq N: \quad d(x_m, x_n) < \varepsilon.$$

Therefore $\forall n \geq N$:

$$d(x^*, x_n) = \lim_{m \rightarrow \infty} d(x_m, x_n) \leq \varepsilon.$$

This implies that the sequence (x_n) in X converges to $x^* \in X^* \implies X$ is dense in X^* .

5. Completeness of X^* . Any Cauchy sequence (x_n) in X converges to $x^* \in X^*$ which is an equivalence class of (x_n) . Consider now a Cauchy sequence (x_n^*) in X^* . As X is dense in X^* , for any $n \geq 1$ we can choose $x_n \in X$ such that $d(x_n, x_n^*) < 1/n$. The sequence (x_n) is also a Cauchy sequence:

$$d(x_m, x_n) < d(x_m^*, x_n^*) + 1/m + 1/n.$$

Therefore it converges to some $x^* \in X^*$. This implies that (x_n^*) also converges to x^*

$$d(x_n^*, x^*) \leq d(x_n^*, x_n) + d(x_n, x^*) < 1/n + d(x_n, x^*) \rightarrow 0, \quad n \rightarrow \infty.$$

□

Example 1.93. We have seen that $\mathbb{R}$ is the completion of $\mathbb{Q}$. But we can also apply the above theorem in order to construct the set $\mathbb{R}$ of real numbers as the completion of $\mathbb{Q}$. According to the above construction $\mathbb{R}$ can be identified with the set of equivalence classes of Cauchy sequences in $\mathbb{Q}$. One can show that every such class x^* has a representative Cauchy sequence (x_n) of the form $x_n = \sum_{i=0}^n a_i/10^i$, where $a_i \in \mathbb{Z}$ and $0 \leq a_i \leq 9$ for $i \geq 1$. In this way we can represent x^* using the usual decimal representation $a_0.a_1a_2a_3\dots$. Note that two different decimal representations can induce equivalent Cauchy sequences, hence the same real number. For example, consider decimal representations $0.999\dots$ and $1.000\dots$ with the corresponding Cauchy sequences $x_n = \sum_{i=1}^n 9/10^i = 1 - 1/10^n$ and $y_n = 1 + \sum_{i=1}^n 0 = 1$. These sequences are equivalent as $d(x_n, y_n) = 1/10^n \rightarrow 0$ as $n \rightarrow \infty$, hence they represent the same real number.

1.8. Banach fixed point theorem.

Definition 1.94. Let (X, d) be a metric space. A map $f: X \rightarrow X$ is called a **contraction** if there exists $0 \leq \alpha < 1$ such that $\forall x, y \in X$

$$d(f(x), f(y)) \leq \alpha d(x, y).$$

Remark 1.95. A contraction is always continuous.

Definition 1.96. An element $x \in X$ is called a **fixed point** of a map $f: X \rightarrow X$ if $f(x) = x$.

Theorem 1.97 (Banach's fixed point theorem). *Let X be a complete metric space and $f: X \rightarrow X$ be a contraction. Then f has a unique fixed point $x \in X$ such that $f(x) = x$.*

Proof. Existence: Let $x_0 \in X$, $x_1 = f(x_0)$ and generally $x_{n+1} = f(x_n)$. The sequence (x_n) is Cauchy. Indeed, for any $k, n \geq 1$

$$\begin{aligned} d(x_n, x_{n+k}) &\leq d(x_n, x_{n+1}) + \cdots + d(x_{n+k-1}, x_{n+k}) = \sum_{i=0}^{k-1} d(x_{n+i}, x_{n+i+1}) \\ &= \sum_{i=0}^{k-1} d(f^{n+i}(x_0), f^{n+i}(x_1)) \leq \sum_{i=0}^{k-1} \alpha^{n+i} d(x_0, x_1) \leq \alpha^n d(x_0, x_1) \sum_{i=0}^{\infty} \alpha^i = \alpha^n \frac{d(x_0, x_1)}{1 - \alpha}. \end{aligned}$$

For any $\varepsilon > 0 \exists N > 0$ such that $\alpha^N \frac{d(x_0, x_1)}{1 - \alpha} < \varepsilon$. Then for any $m \geq n \geq N$

$$d(x_m, x_n) \leq \alpha^n \frac{d(x_0, x_1)}{1 - \alpha} \leq \alpha^N \frac{d(x_0, x_1)}{1 - \alpha} < \varepsilon.$$

This implies that (x_n) is a Cauchy sequence and therefore it converges to some $x \in X$. As f is continuous, we have

$$f(x) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = x,$$

that is, x is a fixed point.

Uniqueness: If $x, y \in X$ are two fixed points, then

$$d(x, y) = d(f(x), f(y)) < \alpha d(x, y).$$

Therefore $d(x, y) = 0$ and $x = y$. □

Example 1.98. Let $X = (0, 1)$ and $f: X \rightarrow X$ be given by $f(x) = \frac{1}{2}x$. Then f is a contraction, but it does not have fixed points (if $\frac{1}{2}x = x$ then $x = 0 \notin X$). The reason is that X is not complete. On the other hand, if $X = [0, 1]$, which is complete, and $f: X \rightarrow X$, $x \mapsto \frac{1}{2}x$, then f has a fixed point $x = 0$.

Example 1.99 (Picard Theorem). Let

$$f: [a, b] \times [c, d] \rightarrow \mathbb{R}$$

be a differentiable map. Consider an ordinary differential equation (ODE)

$$\frac{dy}{dt}(t) = f(t, y(t))$$

with an initial value condition

$$y(t_0) = y_0$$

for some $t_0 \in (a, b)$, $y_0 \in (c, d)$. We are going to show that this equation has a unique solution in a neighborhood $[t_0 - \varepsilon, t_0 + \varepsilon]$ of t_0 for some $\varepsilon > 0$. We can assume that $\exists L, M > 0$ such that

$$|f(t, y_1) - f(t, y_2)| \leq L |y_1 - y_2|, \quad t \in [a, b], y_1, y_2 \in [c, d],$$

$$|f(t, y)| \leq M, \quad t \in [a, b], y \in [c, d].$$

Let us choose $\varepsilon > 0$ such that

$$\varepsilon L < 1, \quad I = [t_0 - \varepsilon, t_0 + \varepsilon] \subset [a, b], \quad J = [y_0 - \varepsilon M, y_0 + \varepsilon M] \subset [c, d].$$

Consider the space

$$X = \{y \in C[t_0 - \varepsilon, t_0 + \varepsilon] \mid y(t) \in J \forall t \in I\}$$

equipped with the supremum metric. Consider the map

$$F: X \rightarrow X, \quad F(y)(t) = y_0 + \int_{t_0}^t f(s, y(s)) ds \quad y \in X, \quad t \in I.$$

Note that $F(y)$ is continuous and

$$|F(y)(t) - y_0| = \left| \int_{t_0}^t f(s, y(s)) ds \right| \leq \int_{t_0}^t M ds \leq \varepsilon M.$$

Therefore $F(y) \in X$. Note that $F(y) = y$ if and only if

$$y(t) = y_0 + \int_{t_0}^t f(s, y(s)) ds \quad \forall t \in I$$

if and only if $y'(t) = f(t, y(t))$ and $y(t_0) = y_0$, that is, y is a solution of our ODE. To show that F has a unique fixed point we will apply the previous theorem. The space X is complete as a closed subspace of a complete metric space $C[t_0 - \varepsilon, t_0 + \varepsilon]$. The map F is a contraction as

$$\begin{aligned} d(F(y_1), F(y_2)) &= \sup_{t \in I} |F(y_1)(t) - F(y_2)(t)| = \sup_{t \in I} \left| \int_{t_0}^t (f(s, y_1(s)) - f(s, y_2(s))) ds \right| \\ &\leq \sup_{t \in I} \int_{t_0}^t L |y_1(s) - y_2(s)| ds \leq \sup_{t \in I} \int_{t_0}^t L d_\infty(y_1, y_2) ds = L d_\infty(y_1, y_2) \sup_{t \in I} |t - t_0| = \varepsilon L d(y_1, y_2) \end{aligned}$$

and $\varepsilon L < 1$. This implies that we can apply the Banach fixed point theorem to F and find a unique $y \in X$ with $F(y) = y$. This proves existence and uniqueness of a solution of ODE.

2. TOPOLOGICAL SPACES

2.1. Definition and examples. Given a metric space (X, d) , we proved that the collection of all its open sets satisfies the following properties: $\emptyset$ and X are open sets, unions of open sets are open, finite intersections of open sets are open. A collection of subsets of X satisfying these conditions is called a topology on X . It turns out that many properties of the metric space depend only on the underlying topology. For this reason we will study topologies in more detail.

Definition 2.1. A **topological space** is a pair (X, τ) , where X is a set and τ is a collection of subsets of X such that

- (1) $\emptyset, X \in \tau$.
- (2) If $(U_i)_{i \in I}$ is a collection of sets in τ then the union $\bigcup_{i \in I} U_i$ is in τ .
- (3) If $(U_i)_{i \in I}$ is a finite collection of sets in τ then the intersection $\bigcap_{i \in I} U_i$ is in τ .

The collection τ is called a **topology** on X . The elements of τ are called **open sets**.

Example 2.2. Let (X, d) be a metric space and τ be the collection of all open sets in (X, d) . Then τ is a topology on X , called a **metric topology generated** by metric d .

Definition 2.3. A topological space (X, τ) is called **metrizable** if τ is generated by some metric on X .

Example 2.4. Let X be an arbitrary set and $\tau = \{\emptyset, X\}$. Then τ is a topology, called the **trivial topology**. This topology is not metrizable if $|X| \geq 2$. Indeed, if τ is generated by some metric d , consider $x, y \in X$ and let $r = d(x, y)$. Then $y \notin B(x, r)$, hence $B(x, r) \neq \emptyset, X$. Therefore $B(x, r) \notin \tau$, while it should be open. A contradiction.

Example 2.5. Let X be an arbitrary set and let τ consist of all subsets of X . Then τ is a topology, called the **discrete topology**. This topology is actually a metric topology generated by the discrete metric

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

Indeed, $B(x, 1) = \{x\}$ is open with respect to d for all $x \in X$. Then for any $A \subset X$, the set $A = \bigcup_{x \in A} \{x\}$ is also open with respect to d .

Example 2.6. Let $X = \{a, b\}$. Then the possible topologies on X are

- (1) $\tau = \{\emptyset, X\}$ (trivial topology).
- (2) $\tau = \{\emptyset, \{a\}, X\}$.
- (3) $\tau = \{\emptyset, \{b\}, X\}$.
- (4) $\tau = \{\emptyset, \{a\}, \{b\}, X\}$ (discrete topology).

Only the last topology is metrizable. One can show that a metric on a finite set always generates the discrete topology.

Definition 2.7. Let (X, τ) be a topological space.

- (1) A subset $A \subset X$ is called **closed** if its complement $X \setminus A$ is open.
- (2) Given a subset $A \subset X$, its **closure** $\bar{A}$ is defined to be the minimal closed set that contains A . It is the intersection of all closed sets that contain A .
- (3) Given a subset $A \subset X$, its **interior** A° is defined to be the maximal open set contained in A . It is the union of all open sets contained in A .

Example 2.8. If X is equipped with the trivial topology, then $\bar{\{x\}} = X$ for every $x \in X$. If X is equipped with the discrete topology, then $\bar{\{x\}} = \{x\}$ for every $x \in X$.

Example 2.9. Let $X = \{a, b\}$ and $\tau = \{\emptyset, \{a\}, X\}$. Then $\bar{\{a\}} = \{a\}$ and $\bar{\{b\}} = X$.

Definition 2.10. A map $f: (X, \tau) \rightarrow (Y, \tau')$ is called **continuous** if for every open set $U \subset Y$, the preimage $f^{-1}(U)$ is open in X . A map f is called a **homeomorphism** if f is bijective and both f, f^{-1} are continuous.

Example 2.11. Let X be a set with $|X| \geq 2$, τ_{disc} be the discrete topology on X and τ_{triv} be the trivial topology on X . Then the identity map

$$f: (X, \tau_{disc}) \rightarrow (X, \tau_{triv}), \quad x \mapsto x,$$

is bijective and continuous. On the other hand $g = f^{-1}(X, \tau_{\text{triv}}) \rightarrow (X, \tau_{\text{disc}})$ is not continuous. Indeed a one-point set $\{x\}$ is open in (X, τ_{disc}) , but its preimage $g^{-1}(\{x\}) = \{x\}$ is not open in (X, τ_{triv}) . We conclude that f is not a homeomorphism, although it is bijective and continuous.

Lemma 2.12. *If $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ are continuous maps between topological spaces, then $gf: X \rightarrow Z$ is also continuous.*

Proof. For any open set $U \subset Z$, we have $(gf)^{-1}(U) = f^{-1}(g^{-1}(U))$. Note that $g^{-1}(U)$ is open in Y , hence $f^{-1}(g^{-1}(U))$ is open in X . $\square$

Example 2.13. Let X be an arbitrary set and let τ consist of $\emptyset$ and all $U \subset X$ with $X \setminus U$ finite. This is a topology, called a **cofinite topology** (or Zariski topology). Let us verify the axioms.

- (1) It is clear that $\emptyset$ and X are in τ .
- (2) Let $(U_i)_{i \in I}$ be a collection of elements in τ . To show that the union $\bigcup_i U_i$ is in τ , we can assume that $U_i \neq \emptyset$, hence $X \setminus U_i$ is finite for all $i \in I$. Consider the complement

$$X \setminus \left(\bigcup_{i \in I} U_i \right) = \bigcap_{i \in I} (X \setminus U_i),$$

where we applied De Morgan's law. This set is finite as an intersection of finite sets. Therefore $\bigcup_i U_i \in \tau$.

- (3) Let $(U_i)_{i \in I}$ be a finite collection of elements in τ and let us show that $\bigcap_i U_i \in \tau$. If some set U_i is empty, then $\bigcap_i U_i = \emptyset \in \tau$. Therefore we can assume that $U_i \neq \emptyset$, hence $X \setminus U_i$ is finite for all $i \in I$. Consider the complement

$$X \setminus \left(\bigcap_{i \in I} U_i \right) = \bigcup_{i \in I} (X \setminus U_i),$$

where we again applied De Morgan's law. This set is a finite union of finite sets, hence is itself finite.

This implies that $\bigcap_{i \in I} U_i \in \tau$.

We conclude that τ is a topology on X . Note that if X is finite, then τ is a discrete topology and in particular metrizable. If X is infinite, then τ is not metrizable. Indeed, assume that τ is generated by a metric d . Let $x, y \in X$ be two distinct points and let $r = \frac{1}{2}d(x, y)$. Then open sets $U = B(x, r)$ and $V = B(y, r)$ satisfy $U \cap V = \emptyset$ and therefore

$$X = X \setminus (U \cap V) = (X \setminus U) \cup (X \setminus V).$$

But U, V are open and non-empty, hence $X \setminus U, X \setminus V$ are finite. This implies that X is finite as a union of two finite sets. A contradiction.

2.2. Subspace and product topologies.

Definition 2.14. Let (X, τ) be a topological space and $A \subset X$ be a subset. Then the collection

$$\tau_A = \{U \cap A \mid U \in \tau\}$$

is a topology on A , called the **subspace topology**. The topological space (A, τ_A) is called a **subspace** of (X, τ) .

Lemma 2.15. Let (X, τ) be a topological space and $A \subset X$ be a subset. Then

- (1) the inclusion map $i: A \rightarrow X$, $x \mapsto x$, is continuous.
- (2) if $f: (X, \tau) \rightarrow (Y, \tau')$ is a continuous map, then the restriction map $f|_A: A \rightarrow Y$, $x \mapsto f(x)$, is continuous.

Proof. (1) Let $U \subset X$ be open. Then its preimage $i^{-1}(U) = U \cap A$ is open in (A, τ_A) . Therefore $i: A \rightarrow X$ is continuous. (2) We can write $f|_A: A \rightarrow Y$ as a composition $A \xrightarrow{i} X \xrightarrow{f} Y$. The composition fi of two continuous maps is continuous. Therefore $f|_A$ is continuous. $\square$

Definition 2.16. Let (X, τ) , (Y, τ') be two topological spaces. We define the topology on $X \times Y$ to be the collection of all unions $\bigcup_{i \in I} U_i \times V_i$, where U_i are open subsets of X and V_i are open subsets of Y for $i \in I$. This topology is called the **product topology**.

Remark 2.17. Let us check that the above collection $\bar{\tau}$ of subsets in $X \times Y$ is indeed a topology.

- (1) The first axiom is satisfied as $\emptyset \times \emptyset$ and $X \times Y$ are in $\bar{\tau}$.
- (2) The second axiom (about unions of open sets) is automatically satisfied.
- (3) To prove the third axiom it is enough to consider an intersection of just two elements $\bigcup U_i \times V_i$ and $\bigcup U'_j \times V'_j$ from $\bar{\tau}$. We obtain

$$\begin{aligned} \left(\bigcup_{i \in I} U_i \times V_i \right) \cap \left(\bigcup_{j \in J} U'_j \times V'_j \right) &= \bigcup_{(i,j) \in I \times J} (U_i \times V_i) \cap (U'_j \times V'_j) \\ &= \bigcup_{(i,j) \in I \times J} (U_i \cap U'_j) \times (V_i \cap V'_j). \end{aligned}$$

The intersections $U_i \cap U'_j$ are open in X and the intersections $V_i \cap V'_j$ are open in Y . Therefore the above union is an element of $\bar{\tau}$.

This implies that $\bar{\tau}$ satisfies all axioms of a topology.

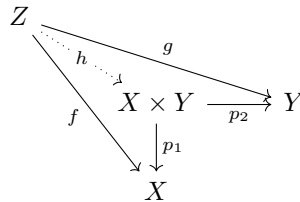
Theorem 2.18. Let X, Y be two topological spaces and let $X \times Y$ be equipped with a product topology. Then

- (1) The projection maps

$$p_1: X \times Y \rightarrow X, \quad (x, y) \mapsto x, \quad p_2: X \times Y \rightarrow Y, \quad (x, y) \mapsto y$$

are continuous.

- (2) Given continuous maps $f: Z \rightarrow X$, $g: Z \rightarrow Y$, there exists a unique continuous map $h: Z \rightarrow X \times Y$ that makes the diagram commute ($p_1 h = f$ and $p_2 h = g$).



Proof. 1. Let us show that $p_1: X \times Y \rightarrow X$ is continuous. For any open $U \subset X$ its preimage is $p_1^{-1}(U) = U \times Y$ which is an open subset of $X \times Y$.

2. It is clear that $h: Z \rightarrow X \times Y$ should satisfy

$$h(z) = (f(z), g(z)), \quad z \in Z.$$

We just have to show that this map is continuous. An open set in $X \times Y$ can be written as $\cup_{i \in I} U_i \times V_i$, where U_i are open in X and V_i are open in Y . Its preimage is equal to

$$\bigcup_{i \in I} h^{-1}(U_i \times V_i)$$

and it is enough to show that $h^{-1}(U \times V)$ is open for any open $U \subset X$ and any open $V \subset Y$. We have

$$h^{-1}(U \times V) = \{z \in Z \mid f(z) \in U, g(z) \in V\} = f^{-1}(U) \cap g^{-1}(V)$$

and the last set is open as an intersection of two open sets $f^{-1}(U)$ and $g^{-1}(V)$. This means that preimages of open sets in $X \times Y$ with respect to h are open in Z . Therefore $h: Z \rightarrow X \times Y$ is continuous. $\square$

Example 2.19. Let us show that the following three topologies on $\mathbb{R}^2$ coincide

- (1) Topology τ_2 generated by the Euclidean metric d_2 .
- (2) Topology τ_∞ generated by the supremum metric d_∞ .
- (3) The product topology $\bar{\tau}$ on $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$.

Similarly one can show that these topologies are equivalent to the topology τ_p generated by the metric d_p for every $p \geq 1$. To show that two topologies τ, τ' on a set X are equal one needs to prove that the identity map $f: (X, \tau) \rightarrow (X, \tau'), x \mapsto x$, is a homeomorphism. Indeed, continuity of f implies that $\tau' \subset \tau$ (if $U \in \tau' \implies U = f^{-1}(U) \in \tau$) and continuity of f^{-1} implies that $\tau \subset \tau'$. We conclude that $\tau = \tau'$. In our case we consider the identity map $f: (\mathbb{R}^2, \tau_2) \rightarrow (\mathbb{R}^2, \tau_\infty)$. Then

$$d_\infty(f(x), f(y)) = d_\infty(x, y) = \max_{i=1,2} |x_i - y_i| \leq \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} = d_2(x, y).$$

This implies that f is Lipschitz continuous, hence continuous. The inverse function $g = f^{-1}: (\mathbb{R}^2, \tau_\infty) \rightarrow (\mathbb{R}^2, \tau_2)$ satisfies

$$d_2(g(x), g(y)) = d_2(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \leq \sqrt{2} \cdot d_\infty(x, y).$$

This implies that g is Lipschitz continuous, hence continuous. We conclude that f is a homeomorphism, hence $\tau_2 = \tau_\infty$. Let us compare now τ_∞ and $\bar{\tau}$. The identity map $h: (\mathbb{R}^2, \tau_\infty) \rightarrow (\mathbb{R}^2, \bar{\tau})$ is continuous by the previous theorem (take $Z = (\mathbb{R}^2, \tau_\infty)$ and $X = Y = \mathbb{R}$). We need to show that its inverse $g = h^{-1}: (\mathbb{R}^2, \bar{\tau}) \rightarrow (\mathbb{R}^2, \tau_\infty)$ is also continuous. Every open set in $(\mathbb{R}^2, \tau_\infty)$ is a union of open balls with respect to the metric d_∞ . Such an open ball has a form

$$\begin{aligned} B(x, \varepsilon) &= \left\{ y \in \mathbb{R}^2 \mid d_\infty(x, y) = \max_{i=1,2} |x_i - y_i| < \varepsilon \right\} = \left\{ y \in \mathbb{R}^2 \mid y_i \in (x_i - \varepsilon, x_i + \varepsilon) \forall i \right\} \\ &= (x_1 - \varepsilon, x_1 + \varepsilon) \times (x_2 - \varepsilon, x_2 + \varepsilon) = U_1 \times U_2, \end{aligned}$$

where $U_i = (x_i - \varepsilon, x_i + \varepsilon)$ is open in $\mathbb{R}$. We conclude that $g^{-1}(B(x, \varepsilon)) = B(x, \varepsilon) = U_1 \times U_2$. But this set is open in the product topology $\bar{\tau}$, hence g is continuous. We conclude that h is a homeomorphism, hence $\tau_\infty = \bar{\tau}$.

Example 2.20. In contrast to the previous example, consider the identity map

$$\text{Id}: (C[0, 1], d_\infty) \rightarrow (C[0, 1], d_1).$$

This map is Lipschitz continuous, hence continuous. On the other hand its inverse is not continuous as there exists a sequence of functions $x_n(t) = t^n$ that converges to zero with respect to d_1 , but diverges with respect to d_∞ . This implies that the corresponding topologies satisfy $\tau_1 \subset \tau_\infty$, but they are not equal.

2.3. Hausdorff spaces. To motivate the introduction of Hausdorff topological spaces let us discuss first convergent sequences in topological spaces. In the case of metric spaces we required open balls for the definition of convergent sequences. In the case of topological spaces the analogue of open balls are open neighborhoods of a point.

Definition 2.21. Let (X, τ) be a topological space and $x \in X$.

- (1) An open set $U \subset X$ that contains x is called an **open neighborhood** of x .
- (2) A subset $V \subset X$ is called a **neighborhood** of x if there exists an open set $U \subset X$ such that $x \in U \subset V$.

Remark 2.22. Note that a subset $V \subset X$ is open $\iff$ for every point $x \in V$ there exists an open neighborhood $U \ni x$ contained in V . Indeed, if V is open, then we can take $U = V$ for every $x \in V$. Conversely, assume that for every $x \in V$, there exists open $U_x \ni x$ contained in V . Then $V = \bigcup_{x \in V} U_x$, hence V is open.

Definition 2.23. Let (X, τ) be a topological space. A sequence (x_n) in X is called **convergent** to $x \in X$ if for any neighborhood $U \ni x$ there exists $N > 0$ such that $x_n \in U$ for $n \geq N$.

Example 2.24. Let (X, τ) be a topological space with the trivial topology $\tau = \{\emptyset, X\}$ and let $x \in X$. Then the sequence $(x_n = x)_{n \geq 1}$ converges to any $y \in X$. Indeed, the only open set that contains y is X and the elements x_n are contained in X for all $n \geq 1$.

This example shows that arbitrary topological spaces can have rather undesirable properties. In order to ensure that sequences have at most one limit, we have to put some restrictions on the topological space.

Definition 2.25. A topological space (X, τ) is called a **Hausdorff space** (or a Hausdorff topological space) if for any $x, y \in X$ with $x \neq y$ there exist open sets $U \ni x$, $V \ni y$ such that $U \cap V = \emptyset$.

Example 2.26. Any metric space (X, d) is Hausdorff. For $x \neq y$, let $r = \frac{1}{2}d(x, y)$ and

$$U = B(x, r), \quad V = B(y, r).$$

Then U, V are open and $U \cap V = \emptyset$.

Example 2.27. Let (X, τ) have the trivial topology and let $|X| \geq 2$. Then X is not Hausdorff. Indeed, let $x, y \in X$ with $x \neq y$ and let $U \ni x$, $V \ni y$ be open. Then $U = V = X$, hence $U \cap V \neq \emptyset$.

Lemma 2.28. Let (X, τ) be a Hausdorff topological space. Then any convergent sequence in X has a unique limit.

Proof. Let (x_n) be a sequence convergent to $x, y \in X$ and assume that $x \neq y$. There exist open sets $U \ni x$, $V \ni y$ with $U \cap V = \emptyset$. As $x_n \rightarrow x, y$ for $n \rightarrow \infty$, there exist $N, N' > 0$ s.t.

$$x_n \in U \quad \forall n \geq N, \quad x_n \in V \quad \forall n \geq N'.$$

Then for any $n \geq \max\{N, N'\}$: $x_n \in U \cap V \implies U \cap V \neq \emptyset$ which contradicts to our assumption. Therefore $x = y$. $\square$

Lemma 2.29. Let (X, τ) be a Hausdorff topological space. For any point $x \in X$ the set $\{x\}$ is closed.

Proof. Let $U = X \setminus \{x\}$. By the Hausdorff condition, for any $y \in U$, there exists an open set $V_y \ni y$ such that $x \notin V_y$. Therefore $V_y \subset U$. It is clear that $U = \bigcup_{y \in U} V_y$ and this set is open as a union of open sets. This means that $\{x\} = X \setminus U$ is closed. $\square$

Example 2.30. Let X be an infinite set with the Zariski topology (closed sets are finite sets and X). Then every 1-point set $\{x\}$ is closed. However, X is not Hausdorff. Indeed, let $x \neq y$ and let $U \ni x$, $V \ni y$ be open sets such that $U \cap V = \emptyset$. Then

$$X = X \setminus (U \cap V) = (X \setminus U) \cup (X \setminus V)$$

is finite as a union of two finite sets. This contradicts to our assumption that X is infinite.

Exercise 2.31. A topological space X is Hausdorff if and only if the diagonal $\{(x, x) \mid x \in X\}$ is closed in $X \times X$.

2.4. Connected topological spaces.

Definition 2.32. A topological space (X, τ) is called **connected** if X can not be represented as a disjoint union $X = U \dot{\cup} V$ of two nonempty open subsets U, V . Equivalently, we require that if $U \subset X$ is both open and closed then $U = \emptyset$ or $U = X$.

Remark 2.33. A subset $A \subset X$ is called connected if the subspace (A, τ_A) is connected. This means that there don't exist open $U, V \subset X$ such that $A \cap U \neq \emptyset$, $A \cap V \neq \emptyset$ and $A \subset U \cup V$, $A \cap U \cap V = \emptyset$.

Example 2.34. The subset $A = [0, 1] \cup [2, 3] \subset \mathbb{R}$ is not connected. Note that the interval $[0, 1]$ is not open in $\mathbb{R}$, but it is open in A .

Proposition 2.35. Let $f: X \rightarrow Y$ be a continuous map between topological spaces and let X be connected. Then $f(X)$ is connected.

Proof. Assume that $f(X)$ is not connected. We can assume that $Y = f(X)$. We can represent $Y = U \dot{\cup} V$ with nonempty disjoint U, V . Then $X = f^{-1}(U) \dot{\cup} f^{-1}(V)$, where $f^{-1}(U), f^{-1}(V)$ are nonempty and open. This contradicts to the assumption that X is connected. $\square$

Corollary 2.36. Let $f: X \rightarrow Y$ be a continuous map between topological spaces and let Y have discrete topology. Then f is a constant map.

Proof. The image $f(X)$ is connected and has discrete topology. If $y \in f(X)$ then we can represent $f(X) = \{y\} \cup (f(X) \setminus \{y\})$, where both subsets on the right are open. As $f(X)$ is connected, we conclude that $f(X) = \{y\}$. $\square$

Lemma 2.37. Intervals $[a, b], (a, b), [a, b), (a, b] \subset \mathbb{R}$ are connected.

Proof. We will prove this result just for the interval (a, b) . For other intervals one can use similar arguments or apply the result for (a, b) together with the result about connected closures, to be proved later.

Assume that $(a, b) = U \dot{\cup} V$, where $U, V \neq \emptyset$ are both open and closed in (a, b) . Choose $c \in U$, $d \in V$ and assume that $c < d$ without loss of generality. Let $x = \sup U \cap [c, d]$. Then $x \in U$ as U is closed in (a, b) . We have $(x, d] \subset V$, hence its limit point x is contained in V . We conclude that $x \in U \cap V$, a contradiction. $\square$

Theorem 2.38 (Intermediate value theorem). Let $f: [a, b] \rightarrow \mathbb{R}$ be a continuous function. Then every value between $f(a)$ and $f(b)$ is attained by f .

Proof. Let $f(a) \leq f(b)$ without loss of generality and assume that $f(a) < y < f(b)$ is not in the image of f . But this would imply that $A = f([a, b])$ is not connected as

$$A = (A \cap (-\infty, y)) \cup (A \cap (y, +\infty))$$

and these two sets contain $f(a)$ and $f(b)$ respectively. We proved, however, that the image $f([a, b])$ of a connected set is connected. $\square$

Lemma 2.39. Let (X, τ) be a topological space and $A \subset X$ be a connected subset. Then the closure $\bar{A} \subset X$ is connected.

Proof. Restricting the topology to $\bar{A}$, we can assume that $\bar{A} = X$. Assume that $X = U \dot{\cup} V$, where U, V are open. Then

$$A = (U \cap A) \cup (V \cap A).$$

Without loss of generality $U \cap A \neq \emptyset$. As A is connected we conclude $U \cap A = A$ and $A \subset U$. As U is closed in X , we obtain $X = \bar{A} \subset U$ and therefore $U = X$. This implies that X is connected. $\square$

Lemma 2.40. If X, Y are connected topological spaces, then $X \times Y$ is also connected.

Proof. Assume that $X \times Y = U \dot{\cup} V$ with open U, V . Assume that U is nonempty and let $(x_0, y_0) \in U$. The subset $A = \{x_0\} \times Y$ is homeomorphic to Y , hence is connected. From

$$A = (U \cap A) \cup (V \cap A)$$

and $(x_0, y_0) \in U \cap A$, we obtain that $U \cap A = A$, hence $A \subset U$. This implies that, for every $y \in Y$, we have $(x_0, y) \in U$. Similarly to the above argument, we conclude that also $X \times \{y\} \in U$. This implies that

$$X \times Y = \bigcup_{y \in Y} X \times \{y\} \subset U,$$

hence $U = X \times Y$. Therefore $X \times Y$ is connected. $\square$

Example 2.41. The square $[0, 1] \times [0, 1] \subset \mathbb{R}^2$ is connected as $[0, 1] \subset \mathbb{R}$ is connected.

Example 2.42. The disk $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ is connected. Consider a continuous map (polar coordinates)

$$f: A = [0, 1] \times [0, 2\pi] \rightarrow \mathbb{R}^2, \quad (r, \theta) \mapsto (r \cos \theta, r \sin \theta).$$

Then A is connected (as a product of two intervals), hence $D = f(A)$ is also connected.

Proposition 2.43. Let (X, τ) be a topological space and let $(A_i)_{i \in I}$ be a collection of connected subsets such that $\cap_{i \in I} A_i \neq \emptyset$. Then $\cup_{i \in I} A_i$ is connected.

Proof. We can assume that $X = \cup_{i \in I} A_i$. Assume that $X = U \dot{\cup} V$, where U, V are open. Let $x \in \cap_{i \in I} A_i$. Without loss of generality we can assume that $x \in U$. For any $i \in I$, we have $x \in A_i \cap U$ and therefore $A_i \cap U \neq \emptyset$. As

$$A_i = (U \cap A_i) \cup (V \cap A_i)$$

and A_i is connected we conclude that $U \cap A_i = A_i$ and $A_i \subset U$. This implies $X = \cup_{i \in I} A_i \subset U$, that is, $U = X$. Therefore $X = \cup_{i \in I} A_i$ is connected. $\square$

Definition 2.44. A subset $A \subset X$ is called a **connected component** of X if it is connected and it is not contained in any larger connected subset.

Theorem 2.45. Let X be a topological space. Then

- (1) X is a disjoint union of its connected components.
- (2) Every connected component of X is closed.

Proof. (1) For any $x \in X$, let A be the union of all connected subsets of X that contain x . Then A is connected by the previous result. It is not contained in any larger connected subset B as this would imply that B is contained in the above union, hence $B \subset A$ and $A = B$. We conclude that A is a connected component. As every point $x \in X$ is contained in a connected component, we can cover X with connected components (represent it as a union). If two connected components A, B have a nonempty intersection then $A \cup B$ is connected by the previous result. By maximality of A, B , this implies that $A = B = A \cup B$. We conclude that X is a disjoint union of its connected components.

(2) If A is a connected component then $\bar{A}$ is connected and $A \subset \bar{A}$. By maximality of A we obtain $A = \bar{A}$, hence A is closed. $\square$

Remark 2.46. If X is a finite union $A_1 \cup \dots \cup A_n$ of its connected components, then all A_i are both open and closed. We proved already that A_i are closed. On the other hand $A_i = X \setminus \bigcup_{j \neq i} A_j = \bigcap_{j \neq i} (X \setminus A_j)$. This set is open as a finite intersection of open sets $X \setminus A_j$.

Example 2.47. Let $U \subset \mathbb{R}$ be an open set. Then U can be represented as a disjoint union $\bigcup_{i \in I} U_i$ of its connected components. We claim that every U_i is an open interval (possibly infinite), hence every open set $U \subset \mathbb{R}$ can be written as a disjoint union of open intervals.

Let $V \subset \mathbb{R}$ be an arbitrary connected subset and let $a = \inf V$ and $b = \sup V$ (possibly $a = -\infty$ and $b = +\infty$). Then $(a, b) \subset V$ as otherwise we would find a point $c \in (a, b) \setminus V$ and decompose $V = (V \cap (-\infty, c)) \cup (V \cap (c, +\infty))$ into a disjoint union of two non-empty open sets. This implies that V is one of the intervals (a, b) , $[a, b]$, $(a, b]$, $[a, b)$.

Assume now that V is a connected component of U and a, b are defined as before. If $a \in V$, then $A = (a - \varepsilon, a + \varepsilon) \subset U$ for some $\varepsilon > 0$, hence $A \cup V$ is connected and therefore $A \subset V$. But this implies that $\inf V < a$, a contradiction. We conclude that $a \notin V$ and similarly that $b \notin V$. Therefore $V = (a, b)$, hence every connected component of U is an open interval.

Definition 2.48. A topological space X is called **path-connected** if for any two points $x, y \in X$ there exists a continuous map $f: [0, 1] \rightarrow X$ such that $f(0) = x$, $f(1) = y$ (a path from x to y).

Lemma 2.49. If X is path-connected, then it is connected.

Proof. Assume that X is not connected. Then we can decompose $X = U \dot{\cup} V$, where $U, V \neq \emptyset$ are open subsets of X . Choose $x \in U$, $y \in V$ and a continuous map $f: [0, 1] \rightarrow X$ such that $f(0) = x$ and $f(1) = y$. Then we can decompose $[0, 1] = f^{-1}(U) \dot{\cup} f^{-1}(V)$. The sets $f^{-1}(U), f^{-1}(V)$ are open and non-empty as $f(0) = x \in U$ and $f(1) = y \in V$. This implies that $[0, 1]$ is not connected, a contradiction. $\square$

Example 2.50. There exist connected topological spaces that are not path-connected. For example, consider a subset $E \subset \mathbb{R}^2$, called a **deleted comb space**,

$$E = \left(\bigcup_{n \geq 1} \left\{ \frac{1}{n} \right\} \times [0, 1] \right) \cup \left([0, 1] \times \{0\} \right) \cup \{(0, 1)\}$$

One can show that the set $E' = E \setminus \{p\}$, where $p = (0, 1)$, is path-connected, hence also connected. Its closure (in E) is equal to E , hence E is also connected. On the other hand there is no path (contained in E) between p and E' , hence E is not path-connected.

2.5. Compact topological spaces.

Definition 2.51. Let (X, τ) be a topological space.

- (1) A collection $(U_i)_{i \in I}$ of open sets in X is called an open cover of X if $X = \bigcup_{i \in I} U_i$.
- (2) X is called a compact topological space if every open cover $(U_i)_{i \in I}$ of it contains a finite subcover, that is, there exists a finite subset $I' \subset I$ such that $X = \bigcup_{i \in I'} U_i$.
- (3) A subset $A \subset X$ is called compact if (A, τ_A) is a compact topological space.

Remark 2.52. The compactness of $A \subset X$ means that given an open cover of A , that is, a collection $(U_i)_{i \in I}$ of open sets in X such that $A \subset \bigcup_{i \in I} U_i$, there exists a finite subset $I' \subset I$ such that $A \subset \bigcup_{i \in I'} U_i$.

Example 2.53. The following topological spaces are compact.

- (1) Finite topological sets.
- (2) (X, τ) with trivial topology (that is, $\tau = \{\emptyset, X\}$).
- (3) A closed interval $[a, b] \subset \mathbb{R}$ (Heine-Borel Theorem).

Example 2.54. The following topological spaces are not compact.

- (1) Euclidean space $\mathbb{R}$. Consider the cover $(U_n)_{n \geq 1}$, where $U_n = (-n, n)$. It does not have a finite subcover.
- (2) The interval $(0, 1)$. Consider the cover $(U_n)_{n \geq 2}$, where $U_n = (\frac{1}{n}, 1)$. It does not have a finite subcover.

Theorem 2.55 (Heine-Borel Theorem on $\mathbb{R}$). *A closed interval $[a, b]$ is compact.*

Proof. Assume that $I_0 = [a, b]$ is not compact and let $(U_i)_{i \in S}$ be an infinite cover without a finite subcover. Divide I_0 into two equal subintervals. At least one of them does not have a finite subcover. Denote it by I_1 . Continuing this process we get intervals

$$I_0 \supset I_1 \supset I_2 \supset \dots$$

with I_n having length $\frac{b-a}{2^n}$. Let $x_n = \min I_n$. Then

$$x_0 \leq x_1 \leq x_2 \leq \dots$$

is an increasing sequence in $[a, b]$. Therefore it converges to some $x \in [a, b]$. For any $n \geq 1$ and $k \geq n$, we have $x_k \in I_k \subset I_n$. Therefore $x = \lim_{k \rightarrow \infty} x_k \in I_n$ as I_n is closed. As $(U_i)_{i \in S}$ is a cover of $[a, b]$, we can find $i \in S$ such that $x \in U_i$. Then there exists $\varepsilon > 0$ such that $(x - \varepsilon, x + \varepsilon) \subset U_i$. There exists $n > 0$ such that $\frac{b-a}{2^n} < \varepsilon$ and therefore $I_n \subset (x - \varepsilon, x + \varepsilon) \subset U_i$ (I_n contains x and has length $< \varepsilon$). This implies that $I_n \subset U_i$, contradicting to our assumption that I_n does not have a finite subcover. $\square$

Lemma 2.56. *Let (X, τ) be a compact topological space and let $A \subset X$ be closed. Then A is compact.*

Proof. Let $(U_i)_{i \in I}$ be an open cover of A , that is, $A \subset \bigcup_{i \in I} U_i$ and $U_i \subset X$ are open. The set $U = X \setminus A$ is open, so there is an open cover of X

$$X = \left(\bigcup_{i \in I} U_i \right) \cup U.$$

By the compactness of X , there exists a finite subset $I' \subset I$ such that $X = (\bigcup_{i \in I'} U_i) \cup U$. This implies $A \subset \bigcup_{i \in I'} U_i$, hence A is compact. $\square$

Theorem 2.57. *Let X, Y be compact topological spaces. Then $X \times Y$ (with product topology) is compact.*

Proof. Let $(W_i)_{i \in I}$ be an open cover of $X \times Y$. Refining the cover (W_i) , we can assume that $W_i = U_i \times V_i$, where U_i, V_i are open in X, Y respectively. For any $x \in X$, the space $\{x\} \times Y$ is compact and has an open cover $(U_i \times V_i)_{i \in I}$. Therefore there exists a finite subset $I_x \subset I$ such that $x \in U_i$ for all $i \in I_x$ and $Y = \bigcup_{i \in I_x} V_i$. Let $U_x = \bigcap_{i \in I_x} U_i \ni x$. The family $(U_x)_{x \in X}$ covers X , hence has a finite subcover $(U_{x_1}, \dots, U_{x_n})$. We claim that

$$(U_i \times V_i \mid i \in I_{x_j}, 1 \leq j \leq n)$$

is a finite cover of $X \times Y$. Indeed, for any $(x, y) \in X \times Y$, we have $x \in U_{x_j}$ for some $1 \leq j \leq n$. We also have $Y = \bigcup_{i \in I_{x_j}} V_i$, hence $y \in V_i$ for some $i \in I_{x_j}$. Then $x \in U_{x_j} \subset U_i$, hence $(x, y) \in U_i \times V_i$. This proves the above statement, hence the cover of $X \times Y$ has a finite subcover and we conclude that $X \times Y$ is compact. $\square$

Example 2.58. Let us show that every closed and bounded subset $A \subset \mathbb{R}^n$ is compact (part of the Heine-Borel theorem on $\mathbb{R}^n$ to be proved later). As A is bounded, we can embed $A \subset [-N, N]^n$ for some $N > 0$. We know that the interval $[-N, N]$ is compact. Therefore the product $[-N, N]^n$ is compact by the previous theorem. As $A \subset [-N, N]^n$ is closed, we conclude that it is also compact.

Proposition 2.59. *Let $f: X \rightarrow Y$ be a continuous map between topological spaces. If X is compact then $f(X)$ is compact.*

Proof. Let $(U_i)_{i \in I}$ be an open cover of $f(X)$, that is, $U_i \subset Y$ are open and $f(X) \subset \bigcup_{i \in I} U_i$. Then

$$X = \bigcup_{i \in I} f^{-1}(U_i),$$

where $f^{-1}(U_i)$ are open. By the compactness of X there exists a finite subset $I' \subset I$ such that

$$X = \bigcup_{i \in I'} f^{-1}(U_i).$$

This implies $f(X) \subset \bigcup_{i \in I'} U_i$, hence we found a finite subcover of $f(X)$. Therefore $f(X)$ is compact. $\square$

Example 2.60. Let us show that the unit circle $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ is compact. Consider a continuous map

$$f: [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \theta \mapsto (\cos \theta, \sin \theta).$$

Then $S^1 = f([0, 2\pi])$. As the closed interval $[0, 2\pi]$ is compact, we conclude that $S^1 = f([0, 2\pi])$ is also compact.

Proposition 2.61. *Let (X, τ) be a Hausdorff topological space and let $A \subset X$ be a compact subspace. Then A is closed in X .*

Proof. We need to show that $X \setminus A$ is open. Consider some $x \in X \setminus A$. For every $y \in A$ there exist open sets $U_y \ni x$, $V_y \ni y$ such that $U_y \cap V_y = \emptyset$. Then $A \subset \bigcup_{y \in A} V_y$ and by the compactness of A there exists a finite set $\{y_1, \dots, y_n\} \subset A$ such that $A \subset V_{y_1} \cup \dots \cup V_{y_n}$. We claim that

$$U = U_{y_1} \cap \dots \cap U_{y_n}$$

is an open neighborhood of x contained in $X \setminus A$. It is clear that U is open and $x \in \bigcap_i U_{y_i} = U$. Moreover, we know that $U_{y_i} \cap V_{y_i} = \emptyset$, hence $U \cap V_{y_i} = \emptyset$ and therefore

$$U \cap (V_{y_1} \cup \dots \cup V_{y_n}) = \emptyset.$$

But $A \subset V_{y_1} \cup \dots \cup V_{y_n}$, hence $U \cap A = \emptyset$. This implies that $U \subset X \setminus A$ and x is an internal point of $X \setminus A$. Therefore $X \setminus A$ is open. $\square$

Corollary 2.62. *Let $f: X \rightarrow Y$ be a continuous and bijective map between topological spaces. If X is compact and Y is Hausdorff, then f is a homeomorphism.*

Proof. We have to prove that $g = f^{-1}: Y \rightarrow X$ is continuous. To do this it is enough to show that for any closed subset $A \subset X$ the set $g^{-1}(A) = f(A)$ is closed. As X is compact and $A \subset X$ is closed, we obtain that A is also compact. Therefore $f(A) \subset Y$ is compact. By the assumption that Y is Hausdorff and the previous statement, we conclude that $f(A) \subset Y$ is closed. $\square$

Example 2.63. We have seen previously that the identity map

$$\text{Id}: (C[0, 1], d_\infty) \rightarrow (C[0, 1], d_1), \quad f \mapsto f$$

is Lipschitz continuous, hence also continuous. On the other hand we have seen that the inverse map is not continuous (there is a sequence of functions $x_n(t) = t^n$ convergent to zero with respect to d_1 , but divergent with respect to d_∞). This means that the previous statement can not be applied. As any topology generated by a metric is automatically Hausdorff, we conclude that $(C[0, 1], d_\infty)$ is not compact.

2.6. Compact metric spaces.

Lemma 2.64. *Let (X, d) be a metric space and $A \subset X$ be a compact subspace. Then A is closed and bounded (that is, $A \subset B(x, r)$ for some $x \in X$ and $r > 0$).*

Proof. (1) As X is a metric space, it is Hausdorff, hence every compact subset $A \subset X$ is closed in X . (2) For any $x \in X$, the family of open balls $(B(x, n))_{n \geq 1}$ is an open cover of A . By the compactness of A it contains a finite subcover. If N is the maximal radius of balls in this subcover, then $A \subset B(x, N)$, hence A is bounded. $\square$

Theorem 2.65 (Heine-Borel Theorem on $\mathbb{R}^n$). *A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Proof. We have seen that a compact set in a metric space is always closed and bounded. Conversely, assume that $A \subset \mathbb{R}^n$ is closed and bounded. Then there exists $N > 0$ such that $A \subset [-N, N]^n$. We know that $[-N, N]$ is compact in $\mathbb{R}$. Therefore $[-N, N]^n$ is compact. Therefore its closed subset A is compact. $\square$

Example 2.66. It is not true in general that a closed and bounded set is compact. For example, let X be an infinite set equipped with the discrete metric. Then X is closed and bounded as $X = B(x, 2)$ for any $x \in X$. But X is not compact. For example, the open cover $X = \bigcup_{x \in X} B(x, 1) = \bigcup_{x \in X} \{x\}$ does not have a finite subcover.

We will prove now some useful characterizations of compact metric spaces.

Definition 2.67. A metric space (X, d) is called **sequentially compact** if any sequence (x_n) in X has a convergent subsequence.

Definition 2.68. Let (X, d) be a metric space.

- (1) Given $\varepsilon > 0$, a subset $A \subset X$ is called an ε -**net** of X if $\forall x \in X \exists y \in A: d(x, y) < \varepsilon$. Equivalently, $X = \bigcup_{y \in A} B(y, \varepsilon)$.
- (2) A metric space X is called **totally bounded** if for any $\varepsilon > 0$ there exists a finite ε -net of X .

Theorem 2.69. *Let (X, d) be a metric space. Then the following conditions are equivalent*

- (1) (X, d) is compact.
- (2) (X, d) is sequentially compact.
- (3) (X, d) is complete and totally bounded.

Proof. (1) $\implies$ (2). Assume that (X, d) is compact. Let $(x_n)_{n \geq 1}$ be a sequence in X . Then it contains a subsequence convergent to $y \in X$ if and only if $\forall \varepsilon > 0$ and $\forall N > 0$

$$\exists n \geq N: \quad x_n \in B(y, \varepsilon).$$

Assuming that (x_n) does not contain a convergent subsequence, for any $y \in X$ there exist $\varepsilon_y > 0$ and $N_y > 0$ such that

$$x_n \notin B(y, \varepsilon_y) \quad \forall n \geq N_y.$$

The open balls $B(y, \varepsilon_y)$ with $y \in X$ form a cover of X . By the compactness of X there exists a finite subset $I \subset X$ s.t.

$$X = \bigcup_{y \in I} B(y, \varepsilon_y).$$

But then for any $n \geq \max \{N_y \mid y \in I\}$ we would have $x_n \notin B(y, \varepsilon_y)$ for $y \in I$ and therefore $x_n \notin \bigcup_{y \in I} B(y, \varepsilon_y) = X$, a contradiction.

(2) $\implies$ (3). Assume that X is sequentially compact. We will prove that it is complete and totally bounded.

Let (x_n) be a Cauchy sequence. We can find a convergent subsequence $(x_{n_k})_{k \geq 1}$ with a limit $x \in X$. For any $\varepsilon > 0$ there exists $N > 0$ such that $d(x_m, x_n) < \varepsilon/2$ for $m, n \geq N$. There exists $k > 0$ such that $n_k \geq N$ and $d(x_{n_k}, x) < \varepsilon/2$. Then for any $n \geq N$

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

This implies that the sequence (x_n) converges to x . This proves that X is complete.

Assume that X is not totally bounded. Then there exists $\varepsilon > 0$ s.t. for any finite set of points $\{x_1, \dots, x_n\}$ the balls $B(x_i, \varepsilon)$ don't cover X . Then we can choose a sequence $(x_n)_{n \geq 1}$ in such a way that

$$x_{n+1} \notin B(x_1, \varepsilon) \cup \dots \cup B(x_n, \varepsilon), \quad n \geq 1.$$

This implies that $d(x_m, x_n) \geq \varepsilon$ for $m \neq n$. This sequence can not contain a convergent subsequence (any subsequence is not Cauchy).

(3) $\implies$ (1). Assume that X is complete and totally bounded and let's prove that X is compact. Let $(U_i)_{i \in I}$ be an open cover that does not contain a finite subcover. By our assumption, for any $n \geq 1$, there exists a finite set $A_n \subset X$ such that

$$X = \bigcup_{x \in A_n} B(x, 1/n).$$

At least one of the sets $B(x, 1)$ with $x \in A_1$ can not be covered by a finite number of U_i . Let $x_1 \in A_1$ be the corresponding point. At least one of the sets

$$B(x_1, 1) \cap B(x, 1/2), \quad x \in A_2$$

can not be covered by a finite number of U_i . Let $x_2 \in A_2$ be the corresponding point. Continuing this process we find a sequence of points $x_n \in A_n$ such that

$$B(x_1, 1) \cap B(x_2, 1/2) \cap \dots \cap B(x_n, 1/n)$$

can not be covered by a finite number of U_i and in particular is non-empty. The sequence (x_n) is Cauchy. Indeed, for $m > n$ we have $B(x_m, 1/m) \cap B(x_n, 1/n) \neq \emptyset$, hence

$$d(x_m, x_n) < 1/m + 1/n < 2/n.$$

As X is complete, the sequence (x_n) converges to some $x \in X$. Choose $i \in I$ such that $x \in U_i$. Then $B(x, \varepsilon) \subset U_i$ for some $\varepsilon > 0$. There exists $n > 0$ such that $d(x_n, x) < \varepsilon/2$ and $\frac{1}{n} < \varepsilon/2$. Then $B(x_n, 1/n) \subset B(x, \varepsilon) \subset U_i$, hence

$$B(x_1, 1) \cap B(x_2, 1/2) \cap \dots \cap B(x_n, 1/n)$$

is covered by just one U_i , contradicting to our choice of the sequence (x_n) . $\square$

Example 2.70. Consider the space $C[0, 1]$ with the supremum metric d_∞ . The set $A = \bar{B}(0, 1) = \{f \in C[0, 1] \mid d_\infty(f, 0) \leq 1\}$ is closed and bounded, but it is not compact. By the previous theorem it is enough to show that A is not sequentially compact. The sequence $x_n(t) = t^n$ in A converges pointwise to the function $x(t) = \delta_1(t)$ which is 0 for $t \neq 1$ and 1 for $t = 1$. This function is not continuous, hence the sequence (x_n) does not have a convergent subsequence.

We can use Theorem 2.69 to give an alternative proof of the Heine-Borel theorem:

Theorem 2.71 (Heine-Borel theorem, see 2.55). *A closed interval $[a, b]$ is compact.*

Proof. We know that $[a, b]$ is complete. To show that it is totally bounded, for every $\varepsilon > 0$, choose $N > 1/\varepsilon$ and divide $[a, b]$ into N equal subintervals. The set A of the end-points of these subintervals forms a finite ε -net of $[a, b]$. We conclude that $[a, b]$ is totally bounded, hence compact. $\square$

Remark 2.72. In the same way we can show that an n -dimensional cube $[a, b]^n \subset \mathbb{R}^n$ is complete and totally bounded, hence compact. This can be used to prove Theorem 2.65.

We can also use Theorem 2.69 to give an alternative proof of the Bolzano-Weierstrass theorem:

Theorem 2.73 (Bolzano-Weierstrass theorem, see 1.80). *A bounded sequence in $\mathbb{R}^k$ has a convergent subsequence.*

Proof. Let (x_n) be a bounded sequence in $\mathbb{R}^k$. Let $A = \{x_n \mid n \geq 1\} \subset \mathbb{R}^k$. Then $\bar{A}$ is closed and bounded. Therefore $\bar{A}$ is compact, hence also sequentially compact. This implies that a sequence (x_n) in $\bar{A}$ has a convergent subsequence. $\square$

2.7. Continuous functions on compact spaces.

Remark 2.74. Recall that given a subset $A \subset \mathbb{R}$, we define

- (1) The **supremum** $\sup A \in \mathbb{R} \cup \{+\infty, -\infty\}$ to be the **least upper bound** of A , that is, an element $c \in \mathbb{R} \cup \{+\infty, -\infty\}$ such that $x \leq c$ for all $x \in A$ (upper bound of A) and c is the least element with this property, meaning that if $x \leq d$ for all $x \in A$, then $c \leq d$. By the supremum property of $\mathbb{R}$, there always exists $\sup A$. For example $\sup(0, 1) = 1$, $\sup \mathbb{R} = +\infty$, $\sup \emptyset = -\infty$. Given a map $f: X \rightarrow \mathbb{R}$, we define $\sup_{x \in X} f(x) = \sup f(X)$.
- (2) The **infimum** $\inf A \in \mathbb{R} \cup \{+\infty, -\infty\}$ to be the **greatest lower bound** of A , that is, an element $c \in \mathbb{R} \cup \{+\infty, -\infty\}$ such that $x \geq c$ for all $x \in A$ (lower bound of A) and c is the greatest element with this property, meaning that if $x \geq d$ for all $x \in A$, then $c \geq d$. There always exists $\inf A$. For example $\inf(0, 1) = 0$, $\inf \mathbb{R} = -\infty$, $\inf \emptyset = +\infty$. Given a map $f: X \rightarrow \mathbb{R}$, we define $\inf_{x \in X} f(x) = \inf f(X)$.
- (3) The **maximum** $\max A$ to be an element $c \in A$ such that $x \leq c$ for all $x \in A$. For example $\max(0, 1)$ is undefined, $\max \mathbb{R}$ is undefined, $\max \emptyset$ is undefined, $\max[0, 1] = 1$. Note that if A has the maximum, then $\max A = \sup A$. Given a map $f: X \rightarrow \mathbb{R}$, we define $\max_{x \in X} f(x) = \max f(X)$. We say that f attains its maximum at $x_0 \in X$ if $f(x_0) = \sup_{x \in X} f(x)$ (then necessarily $f(x_0) = \max_{x \in X} f(x)$).
- (4) The **minimum** $\min A$ to be an element $c \in A$ such that $x \geq c$ for all $x \in A$. Note that if A has the minimum, then $\min A = \inf A$. Given a map $f: X \rightarrow \mathbb{R}$, we define $\min_{x \in X} f(x) = \min f(X)$. We say that f attains its minimum at $x_0 \in X$ if $f(x_0) = \inf_{x \in X} f(x)$ (then necessarily $f(x_0) = \inf_{x \in X} f(x)$).

Theorem 2.75. Let (X, τ) be a compact topological space and let $f: X \rightarrow \mathbb{R}$ be a continuous function. Then the set of values $f(X)$ is bounded and there exist points $x_0, x_1 \in X$ such that

$$f(x_0) = \inf_{x \in X} f(x), \quad f(x_1) = \sup_{x \in X} f(x).$$

Proof. Let $A = f(X) \subset \mathbb{R}$. Then A is compact, hence closed and bounded. Therefore

$$M = \sup_{x \in X} f(x) = \sup A < +\infty$$

and $M \in A$. This implies that there exists $x_1 \in X$ with $M = f(x_1)$. The proof for the infimum is similar. $\square$

Remark 2.76. Let us generalize the space of continuous functions $C[a, b]$. Given a topological space X , define $C(X)$ to be the set of continuous functions $f: X \rightarrow \mathbb{R}$. If X is compact, define the supremum metric d_∞ on $C(X)$

$$d_\infty(f, g) = \sup_{x \in X} |f(x) - g(x)|.$$

This metric is well-defined as the function $h: X \rightarrow \mathbb{R}$, $x \mapsto |f(x) - g(x)|$ is continuous, hence bounded by the above result.

Remark 2.77. Even more generally, given topological spaces X, Y , define $C(X, Y)$ to be the set of continuous maps $f: X \rightarrow Y$. If X is compact and (Y, d_Y) is a metric space, define the supremum metric on $C(X, Y)$

$$d_\infty(f, g) = \sup_{x \in X} d_Y(f(x), g(x)).$$

This metric is well-defined as the function $h: X \rightarrow \mathbb{R}$, $x \mapsto d_Y(f(x), g(x))$ is continuous, hence bounded by the above result. One can show that the generated topology on $C(X, Y)$ coincides with the **compact-open** topology. This is the minimal topology that contains the sets $W(K, U) = \{f \in C(X, Y) \mid f(K) \subset U\}$ for all compact $K \subset X$ and open $U \subset Y$.

Example 2.78. If X is finite and has the discrete topology, then $C(X) \simeq \mathbb{R}^n$, where $n = |X|$, hence a subset $A \subset C(X)$ is compact if it is closed and bounded (by the Heine-Borel theorem). But this is not true in general. For example, the subset $A = \bar{B}(0, 1) \subset C[0, 1]$ is closed and bounded, but not compact (see Example 2.70). The next result is a useful characterization of compact subsets of $C(X)$.

Theorem 2.79 (Arzela-Ascoli Theorem). Let X be a compact metric space and $C(X)$ be the space of continuous maps $f: X \rightarrow \mathbb{R}$ with the supremum metric. A subset $A \subset C(X)$ is compact if and only if

- (1) A is closed and bounded.

(2) A is **equicontinuous**, that is, $\forall \varepsilon > 0 \exists \delta > 0$

$$\forall f \in A: \quad d(x, y) < \delta \implies |f(x) - f(y)| < \varepsilon.$$

Definition 2.80. A map $f: X \rightarrow Y$ between two metric spaces is called **uniformly continuous** if $\forall \varepsilon > 0 \exists \delta > 0$:

$$d_X(x, y) < \delta \implies d_Y(f(x), f(y)) < \varepsilon.$$

Remark 2.81. For a map $f: X \rightarrow Y$ between metric spaces one can show that

- (1) Lipschitz continuity implies uniform continuity.
- (2) Uniform continuity implies continuity.

Example 2.82. Let us construct a continuous map that is not uniformly continuous. Consider the map $f: (0, 1) \rightarrow \mathbb{R}$, $x \mapsto 1/x$. It is continuous, but not uniformly continuous. Indeed, let $\varepsilon > 0$ and assume $\exists \delta > 0$ such that

$$d(x, y) < \delta \implies d(f(x), f(y)) < \varepsilon.$$

Let $y = \delta$ and $0 < x < \delta$. Then $d(x, y) = |x - y| < \delta$ and

$$d(f(x), f(y)) = |1/x - 1/\delta| \rightarrow \infty, \quad x \rightarrow 0.$$

This means that $d(f(x), f(y)) > \varepsilon$ for $x > 0$ sufficiently small.

Example 2.83. Let us construct a uniform continuous map $f: X \rightarrow Y$ that is not Lipschitz continuous. Consider the map $f: [0, 1] \rightarrow \mathbb{R}$, $x \mapsto \sqrt{x}$. This map is continuous on a compact, hence also uniformly continuous by the next theorem. We can also show this directly: for any $0 < \varepsilon < 1$, choose $\delta = \varepsilon^2$. Assume that $x \leq y$ and $|x - y| < \delta$. If $\sqrt{y} < \varepsilon$, then $|\sqrt{x} - \sqrt{y}| < \varepsilon$ as required. If $\sqrt{y} \geq \varepsilon$, then

$$\delta > |y - x| = |\sqrt{x} + \sqrt{y}| \cdot |\sqrt{x} - \sqrt{y}| \geq \varepsilon |\sqrt{x} - \sqrt{y}|,$$

hence $|\sqrt{x} - \sqrt{y}| < \delta/\varepsilon = \varepsilon$ as required. On the other hand f is not Lipschitz continuous:

$$\frac{|f(x) - f(y)|}{|x - y|} = \frac{1}{|\sqrt{x} + \sqrt{y}|} \rightarrow \infty, \quad x, y \rightarrow 0.$$

Theorem 2.84 (Uniform continuity theorem). *Let $f: X \rightarrow Y$ be a continuous map between metric spaces. If X is compact, then f is uniformly continuous.*

Proof. Assume the contrary. Then there exists $\varepsilon > 0$ such that for all $n \geq 1 \exists x_n, y_n \in X$:

$$d_X(x_n, y_n) < 1/n, \quad d_Y(f(x_n), f(y_n)) \geq \varepsilon.$$

As X is compact, there exists a subsequence $(x_{n_k})_{k \geq 1}$ of $(x_n)_{n \geq 1}$ convergent to some $x \in X$. By the above inequality, the sequence (y_{n_k}) also converges to x . Then $(f(x_{n_k}))$ and $(f(y_{n_k}))$ converge to $f(x)$, hence $\exists N > 0 \forall k \geq N$:

$$d_Y(f(x_{n_k}), f(x)) < \varepsilon/2, \quad d_Y(f(y_{n_k}), f(x)) < \varepsilon/2.$$

$\implies d_Y(f(x_{n_k}), f(y_{n_k})) < \varepsilon$, a contradiction. □

3. NORMED VECTOR SPACES

3.1. Definition of a normed vector space. Recall that a vector space over a field K (we will need only $K = \mathbb{R}$ and $K = \mathbb{C}$) is a set X together with two operations

$$\begin{aligned} \text{addition: } X \times X &\rightarrow X, & (x, y) &\mapsto x + y \\ \text{scalar multiplication: } K \times X &\rightarrow X, & (\lambda, x) &\mapsto \lambda x \end{aligned}$$

satisfying the axioms:

- (1) $(X, +)$ is an abelian group,
- (2) $\lambda(\mu x) = (\lambda\mu)x$,
- (3) $1x = x$,
- (4) $\lambda(x + y) = \lambda x + \lambda y$,
- (5) $(\lambda + \mu)x = \lambda x + \mu x$

for any $x, y \in X$ and $\lambda, \mu \in K$.

Definition 3.1. A normed vector space is a pair $(X, \|\cdot\|)$, where X is a vector space (over $K = \mathbb{R}$ or $K = \mathbb{C}$) and $\|\cdot\|$ is a map

$$\|\cdot\| : X \rightarrow \mathbb{R}, \quad x \mapsto \|x\|$$

satisfying the axioms

- (1) $\|x\| = 0 \iff x = 0$.
- (2) $\|\lambda x\| = |\lambda| \cdot \|x\|$ for all $\lambda \in K, x \in X$.
- (3) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$ (triangle inequality).

The map $\|\cdot\|$ is called a norm on X and $\|x\|$ is called the norm of x .

Remark 3.2. For any $x \in X$ we have

$$\|-x\| = \|(-1) \cdot x\| = |-1| \cdot \|x\| = \|x\|.$$

We also have $\|x\| \geq 0$. Indeed

$$0 = \|x + (-x)\| \leq \|x\| + \|-x\| = 2\|x\|,$$

so $\|x\| \geq 0$. We can think of $\|x\|$ as the length of a vector x .

Lemma 3.3. Let $(X, \|\cdot\|)$ be a normed vector space. Then the map

$$d: X \times X \rightarrow \mathbb{R}, \quad d(x, y) = \|x - y\|, \quad x, y \in X$$

defines a metric on X .

Proof. For any $x, y, z \in X$ we have

- (1) $d(x, y) = 0 \iff \|x - y\| = 0 \iff x - y = 0 \iff x = y$.
- (2) $d(x, y) = \|x - y\| = \|y - x\| = d(y, x)$.
- (3) $d(x, y) = \|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y)$.

□

In this way any normed vector space can be considered as a metric space (or a topological space). Many examples of metric spaces that we considered before arise in this way.

3.2. Examples of normed vector spaces.

Example 3.4. The Euclidean norm on $\mathbb{R}^n$

$$\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$$

for $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

Example 3.5. For any $p \geq 1$ there is a norm $\|\cdot\|_p$ on $\mathbb{R}^n$ given by

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$$

for $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. The triangle inequality axiom follows from the Minkowski inequality. There is also the ∞ -norm (or the supremum norm) on $\mathbb{R}^n$

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

Example 3.6. Let $C[a, b]$ be the set of all continuous functions $f: [a, b] \rightarrow \mathbb{R}$. Define addition and scalar multiplication on $C[a, b]$ pointwise:

For any $x, y \in C[a, b]$ define $x + y \in C[a, b]$ by

$$(x + y)(t) = x(t) + y(t), \quad t \in [a, b].$$

For any $\lambda \in \mathbb{R}$, $x \in C[a, b]$ define $\lambda x \in C[a, b]$ by

$$(\lambda x)(t) = \lambda \cdot x(t), \quad t \in [a, b].$$

The set $C[a, b]$ equipped with these operations is a vector space. For any $p \geq 1$ there is a norm $\|\cdot\|_p$ on $C[a, b]$ given by

$$\|x\|_p = \left(\int_a^b |x(t)|^p dt \right)^{1/p}.$$

There is also the ∞ -norm (or the supremum norm) on $C[a, b]$

$$\|x\|_\infty = \sup_{t \in [a, b]} |x(t)| = \max_{t \in [a, b]} |x(t)| < \infty$$

where in the last equality we used the fact that $[a, b]$ is compact.

Example 3.7. For any $p \geq 1$, let ℓ_p be the set of sequences $(x_n)_{n \geq 1}$ in $\mathbb{R}$ such that $\sum_{n \geq 1} |x_n|^p < \infty$. For any such sequence, define $\|x\|_p = \left(\sum_{n \geq 1} |x_n|^p \right)^{1/p}$. Define addition and scalar multiplication of sequences componentwise:

$$(x + y)_n = x_n + y_n, \quad (\lambda x)_n = \lambda \cdot x_n.$$

It is clear that if $x \in \ell_p$ and $\lambda \in \mathbb{R}$, then $\lambda x \in \ell_p$. Let us show that if $x, y \in \ell_p$, then $x + y \in \ell_p$. By the Minkowski inequality, for any $m \geq 1$,

$$\left(\sum_{n=1}^m |x_n + y_n|^p \right)^{1/p} \leq \left(\sum_{n=1}^m |x_n|^p \right)^{1/p} + \left(\sum_{n=1}^m |y_n|^p \right)^{1/p} \leq \|x\|_p + \|y\|_p.$$

Taking the limit as $m \rightarrow \infty$, we obtain

$$\left(\sum_{n \geq 1} |x_n + y_n|^p \right)^{1/p} \leq \|x\|_p + \|y\|_p < \infty,$$

hence $x + y \in \ell_p$. This implies that ℓ_p is a vector space. The map $\|\cdot\|_p$ defines a norm on ℓ_p .

Similarly, define ℓ_∞ to be the set of all sequences $(x_n)_{n \geq 1}$ in $\mathbb{R}$ such that $\sup_{n \geq 1} |x_n| < \infty$ (that is, bounded sequences). Then ℓ_p is a vector space with addition and scalar multiplication defined as before. It has a norm

$$\|x\|_\infty = \sup_{n \geq 1} |x_n| \quad \forall x \in \ell_\infty.$$

3.3. Bounded linear operators.

Definition 3.8. Let $(X, \|\cdot\|)$, $(Y, \|\cdot\|)$ be two normed vector spaces over K .

- (1) A map $A: X \rightarrow Y$ is called a **linear operator** (or a linear map) if

$$A(x + y) = Ax + Ay, \quad A(\lambda x) = \lambda Ax$$

for any $x, y \in X$ and $\lambda \in K$.

- (2) A linear operator $A: X \rightarrow Y$ is called **bounded** if there exists $M \geq 0$ such that

$$\|Ax\| \leq M \|x\| \quad \forall x \in X.$$

- (3) A linear operator $A: X \rightarrow Y$ is called **continuous** if it is continuous with respect to the metrics on X and Y induced by the norms.

Lemma 3.9. A linear operator $A: X \rightarrow Y$ between two normed vector spaces is bounded if and only if it is continuous.

Proof. (1) $\implies$ (2). If $A: X \rightarrow Y$ is bounded with a constant $M > 0$, then it is Lipschitz continuous

$$d_Y(Ax, Ay) = \|Ax - Ay\| = \|A(x - y)\| \leq M \|x - y\| = M d_X(x, y),$$

hence continuous.

(2) $\implies$ (1). Let $A: X \rightarrow Y$ be continuous. If A is not bounded, then for any $n \geq 1$, there exists $x_n \in X$ such that $\|Ax_n\| > n \|x_n\|$. Let $y_n = \frac{x_n}{n \|x_n\|}$. Then $\|y_n\| = 1/n$ and

$$\|Ay_n\| = \frac{\|Ax_n\|}{n \|x_n\|} > 1.$$

This implies that $y_n \rightarrow 0$, while $Ay_n \not\rightarrow 0$ as $n \rightarrow \infty$. This contradicts to the continuity of A . $\square$

Definition 3.10. Let X, Y be normed vector spaces. Define $L(X, Y)$ to be the set of all bounded linear operators from X to Y . The set $L(X, Y)$ has a structure of a vector space:

- (1) Given $A, B \in L(X, Y)$, define $A + B \in L(X, Y)$ by

$$(A + B)x = Ax + Bx, \quad x \in X.$$

If A is bounded by $M > 0$ and B is bounded by $N > 0$, then $A + B$ is also bounded:

$$\|(A + B)x\| = \|Ax + Bx\| \leq \|Ax\| + \|Bx\| \leq M \|x\| + N \|x\| = (M + N) \|x\|.$$

- (2) Given $\lambda \in K$ and $A \in L(X, Y)$ define $\lambda A \in L(X, Y)$ by

$$(\lambda A)x = \lambda \cdot Ax.$$

It is clear that if A is bounded, then λA is bounded.

Define $L(X) = L(X, X)$ and define the **dual space** $X^* = L(X, \mathbb{R})$.

Definition 3.11. Define the **operator norm** $\|A\|$ of a bounded linear operator $A \in L(X, Y)$ to be the minimal number $M \geq 0$ satisfying $\|Ax\| \leq M \|x\|$ for all $x \in X$. That is,

$$\|A\| = \inf \{ M \geq 0 \mid \|Ax\| \leq M \|x\| \quad \forall x \in X \}.$$

Lemma 3.12. For any $A \in L(X, Y)$, we have

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|.$$

Proof. For any $x \neq 0$, consider the vector $y = x / \|x\|$. Then $\|y\| = 1$ and $\frac{\|Ax\|}{\|x\|} = \frac{\|Ay\|}{\|y\|} = \|Ay\|$. This implies the second equality of the lemma.

Let $a = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$. Then $\frac{\|Ax\|}{\|x\|} \leq a$ for all $x \neq 0$, hence $\|Ax\| \leq a \|x\|$ for all $x \in X$. Therefore $\|A\| \leq a$. Conversely, we have $\|Ax\| \leq \|A\| \cdot \|x\|$, hence $\frac{\|Ax\|}{\|x\|} \leq \|A\|$ for all $x \neq 0$. Therefore $a = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \leq \|A\|$. We conclude that $\|A\| = a$. $\square$

Example 3.13. Consider a linear map $A: (\mathbb{R}^n, \|\cdot\|_1) \rightarrow (\mathbb{R}^n, \|\cdot\|_\infty)$ given by a matrix $(a_{ij})_{i,j=1}^n$. Then

$$\|Ax\|_\infty = \max_i \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \max_i \left(\max_j |a_{ij}| \right) \sum_j |x_j| = \max_{i,j} |a_{ij}| \cdot \|x\|_1.$$

This implies that $\|A\| \leq \max_{i,j} |a_{ij}|$. To prove the inverse inequality, consider $x = e_j \in \mathbb{R}^n$ for $1 \leq j \leq n$. Then

$$\|Ax\|_\infty = \|(a_{1j}, a_{2j}, \dots, a_{nj})\|_\infty = \max_i |a_{ij}|.$$

Therefore

$$\|A\| \geq \frac{\|Ax\|_\infty}{\|x\|_1} = \max_i |a_{ij}|.$$

This is true for all $1 \leq j \leq n$, hence $\|A\| = \max_{i,j} |a_{ij}|$.

Theorem 3.14. For any $A, B \in L(X, Y)$ and $\lambda \in K$, we have

- (1) $\|A\| = 0 \iff A = 0$.
- (2) $\|\lambda A\| = |\lambda| \|A\|$.
- (3) $\|A + B\| \leq \|A\| + \|B\|$.

Therefore $L(X, Y)$ equipped with the operator norm is itself a normed vector space.

Proof. (1) Assume that $\|A\| = 0$. Then $\forall x \in X: \|Ax\| \leq 0 \|x\| = 0 \implies \|Ax\| = 0 \implies Ax = 0$. Therefore $A = 0$. Conversely, if $A = 0$, then

$$\|A\| = \sup_{x \neq 0} \|Ax\| / \|x\| = \sup_{x \in X} 0 = 0.$$

(2)

$$\|\lambda A\| = \sup_{x \neq 0} \frac{\|\lambda Ax\|}{\|x\|} = |\lambda| \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = |\lambda| \cdot \|A\|.$$

(3) For any $x \in X$

$$\|(A + B)x\| = \|Ax + Bx\| \leq \|Ax\| + \|Bx\| \leq \|A\| \|x\| + \|B\| \|x\|.$$

Therefore $\|A + B\| \leq \|A\| + \|B\|$. □

Example 3.15. Let $X = (\mathbb{R}^n, \|\cdot\|_1)$. We will show that the dual space $X^* = L(X, \mathbb{R})$ is given by $(\mathbb{R}^n, \|\cdot\|_\infty)$. A linear map $f: \mathbb{R}^n \rightarrow \mathbb{R}$ can be written in the form

$$f(x) = \sum_{i=1}^n f_i x_i \quad x \in \mathbb{R}^n$$

for some $(f_1, \dots, f_n) \in \mathbb{R}^n$ (we will denote this vector also by f). Then

$$\|f(x)\| = \left| \sum f_i x_i \right| \leq \max_i |f_i| \cdot \sum |x_i| = \|f\|_\infty \cdot \|x\|_1,$$

hence $\|f\| = \sup_{x \neq 0} \|f(x)\| / \|x\|_1 \leq \|f\|_\infty$. On the other hand, if $\|f\|_\infty = |f_i|$ for some i , then for $x = e_i$ (element of the standard basis) we have $\|f(x)\| / \|x\|_1 = |f_i|$, hence $\|f\| = \sup_{x \neq 0} \|f(x)\| / \|x\|_1 = \|f\|_\infty$.

Similarly one can show that the dual space of $(\mathbb{R}^n, \|\cdot\|_\infty)$ is given by $(\mathbb{R}^n, \|\cdot\|_1)$. More generally, for any $p, q \geq 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$, the dual space of $(\mathbb{R}^n, \|\cdot\|_p)$ is given by $(\mathbb{R}^n, \|\cdot\|_q)$.

Lemma 3.16. Let $A \in L(X, Y)$ and $B \in L(Y, Z)$. Then BA is bounded and $\|BA\| \leq \|B\| \cdot \|A\|$.

Proof. For any $x \in X$ we have

$$\|BAx\| \leq \|B\| \cdot \|Ax\| \leq \|B\| \cdot \|A\| \cdot \|x\|.$$

This implies $\|BA\| \leq \|B\| \cdot \|A\|$. □

Remark 3.17. The previous result implies that the normed vector space $L(X) = L(X, X)$ has a multiplication given by composition of operators. This means that $L(X)$ is equipped with a structure of an algebra. Its identity element is the identity operator.

3.4. Equivalent norms.

Definition 3.18. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a vector space X are called **equivalent** if there exist $C, C' > 0$ such that

$$\|x\|_2 \leq C \|x\|_1, \quad \|x\|_1 \leq C' \|x\|_2, \quad \forall x \in X.$$

Theorem 3.19. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a vector space X are equivalent if and only if they generate the same topology on X .

Proof. Consider the identity operator

$$A: (X, \|\cdot\|_1) \rightarrow (X, \|\cdot\|_2), \quad x \mapsto x.$$

Then the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if and only if A and its inverse are bounded linear operators:

- (1) If $\|x\|_2 \leq C \|x\|_1 \implies \|Ax\|_2 \leq C \|x\|_1 \implies \|A\| \leq C$.
- (2) If A is bounded, then $\|x\|_2 = \|Ax\|_2 \leq \|A\| \|x\|_1$ and we can take $C = \|A\|$.
- (3) The argument for A^{-1} is the same.

On the other hand the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ generate the same topology if and only if A and its inverse are continuous (the preimage of an open set is open). But we know that for linear operators continuity and boundedness are equivalent. $\square$

Example 3.20. Consider the vector space $X = C[0, 1]$ with the norms

$$\|x\|_1 = \int_0^1 |x(t)| dt, \quad \|x\|_\infty = \sup_{t \in [0, 1]} |x(t)|.$$

These norms are not equivalent. The sequence $x_n(t) = t^n$ converges to zero with respect to $\|\cdot\|_1$, but it diverges with respect to $\|\cdot\|_\infty$ (it converges pointwise to $t \mapsto \delta_{1t}$ which is not continuous).

Theorem 3.21. Let X be a finite dimensional vector space. Then all norms on X are equivalent.

Proof. Fixing a basis $(e_1, \dots, e_n)$ of X , we can assume that $X = \mathbb{R}^n$. We have to prove that any norm $\|\cdot\|$ on X is equivalent to the fixed norm

$$\|x\|_1 = \sum_{k=1}^n |x_k|, \quad x = (x_1, \dots, x_n).$$

Consider the identity operator $A: (X, \|\cdot\|_1) \rightarrow (X, \|\cdot\|)$. Then

$$\|Ax\| = \|x\| = \|x_1 e_1 + \dots + x_n e_n\| \leq \sum |x_i| \cdot \|e_i\| \leq C \sum |x_i| = C \|x\|_1,$$

where $C = \max_i \|e_i\|$. Therefore A is bounded and continuous. Define

$$a = \inf_{x \neq 0} \frac{\|Ax\|}{\|x\|_1} = \inf_{\|x\|_1=1} \|Ax\|.$$

The last minimum is attained as the set $\{x \in \mathbb{R}^n \mid \|x\|_1 = 1\}$ is compact and the map $(\mathbb{R}^n, \|\cdot\|_1) \rightarrow \mathbb{R}$, $x \mapsto \|Ax\|$ is continuous by the previous argument. This implies that $a > 0$. We have $\|x\| / \|x\|_1 \geq a$ for all $x \neq 0$, hence

$$\|x\|_1 \leq \frac{1}{a} \|x\|.$$

Therefore the norms are equivalent. $\square$

Theorem 3.22. Let $A: X \rightarrow Y$ be a linear operator between normed vector spaces. If X is finite-dimensional, then A is continuous.

Proof. Choosing a basis $(e_1, \dots, e_n)$ of X , we can assume that $X = \mathbb{R}^n$. As all norms on X are equivalent, we can assume that the norm on X is given by $\|x\|_1 = \sum_{i=1}^n |x_i|$. Then

$$\begin{aligned} \|Ax\| &= \left\| A \left(\sum x_i e_i \right) \right\| = \left\| \sum x_i A e_i \right\| \leq \sum \|x_i A e_i\| \\ &= \sum |x_i| \|A e_i\| \leq C \sum |x_i| = C \|x\|_1, \end{aligned}$$

where $C = \max_i \|A e_i\|$. This implies that A is bounded. $\square$

3.5. Banach spaces.

Definition 3.23. A **Banach space** is a normed vector space $(X, \|\cdot\|)$ which is complete (with respect to the metric $d(x, y) = \|x - y\|$).

Example 3.24. Any finite dimensional normed vector space is Banach. We have seen that the Euclidean space $\mathbb{R}^n$ is complete. Any other norm on $\mathbb{R}^n$ is equivalent to the Euclidean norm.

Example 3.25. We have seen that the space of continuous functions $C[a, b]$ with the supremum metric is complete. Therefore $C[a, b]$ with the supremum norm

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

is a Banach space.

Example 3.26. Consider the space $C[0, 1]$ equipped with the norm $\|x\|_1 = \int_0^1 |x(t)| dt$. Let us show that it is non-Banach. Consider the function $f(t) = 0$ for $t \in [0, \frac{1}{2}]$ and $f(t) = 1$ for $t \in (\frac{1}{2}, 1]$. This is a non-continuous function which can be interpolated by the continuous functions (for $n \geq 2$)

$$x_n(t) = \begin{cases} 0 & t \leq \frac{1}{2} \\ n(t - \frac{1}{2}) & \frac{1}{2} \leq t \leq \frac{1}{2} + \frac{1}{n} \\ 1 & t \geq \frac{1}{2} + \frac{1}{n} \end{cases}$$

We have $\|f - x_n\|_1 \leq 1/n$, hence the sequence (x_n) is Cauchy. On the other hand if it converges to some continuous function g with respect to the norm $\|\cdot\|_1$, then $\|f - g\|_1 = 0$. This would imply that $g(t) = 0$ for $t < \frac{1}{2}$ and $g(t) = 1$ for $t > \frac{1}{2}$, contradicting the continuity of g .

Theorem 3.27. The space ℓ_p is Banach.

Proof. Let $(x^{(n)})_{n \geq 1}$ be a Cauchy sequence in ℓ_p . Then $\forall \varepsilon > 0 \exists N > 0$:

$$(2) \quad \left\|x^{(m)} - x^{(n)}\right\|_p = \left(\sum_{i \geq 1} |x_i^{(m)} - x_i^{(n)}|^p\right)^{1/p} < \varepsilon \quad \forall m, n \geq N.$$

Then for any $k \geq 1$

$$|x_k^{(m)} - x_k^{(n)}| \leq \left(\sum_{i \geq 1} |x_i^{(m)} - x_i^{(n)}|^p\right)^{1/p} < \varepsilon.$$

Therefore the sequence $(x_k^{(n)})_{n \geq 1}$ in $\mathbb{R}$ is Cauchy, hence converges to some $x_k \in \mathbb{R}$. Consider the sequence $x = (x_k)_{k \geq 1}$. Taking the limit $m \rightarrow \infty$ in (2), we obtain

$$\left\|x - x^{(n)}\right\|_p \leq \varepsilon.$$

This implies that $x - x^{(n)} \in \ell_p$, hence $x = (x - x^{(n)}) + x^{(n)} \in \ell_p$. This also implies that $(x^{(n)})_{n \geq 1}$ converges to x . $\square$

Theorem 3.28. Let X, Y be two normed vector spaces and assume that Y is a Banach space. Then the space of bounded linear operators $L(X, Y)$ is a Banach space.

Proof. Let $(A_n)_{n \geq 1}$ be a Cauchy sequence in $L(X, Y)$. For any $x \in X$ and for any $\varepsilon > 0$ there exists $N > 0$ s.t.

$$\|A_m - A_n\| < \varepsilon \quad \forall m, n \geq N.$$

Therefore

$$\|A_m x - A_n x\| \leq \|A_m - A_n\| \|x\| \leq \varepsilon \|x\|, \quad m, n \geq N$$

and this implies that $(A_n x)$ is a Cauchy sequence in Y . Therefore it is convergent to some $Ax \in Y$. Let us show that $A: X \rightarrow Y$ is linear. For any $x, y \in X$, we have

$$A(x + y) = \lim_{n \rightarrow \infty} A_n(x + y) = \lim_{n \rightarrow \infty} (A_n x + A_n y) = \lim_{n \rightarrow \infty} A_n x + \lim_{n \rightarrow \infty} A_n y = Ax + Ay.$$

Similarly one proves that $A(\lambda x) = \lambda Ax$ for any $A \in \mathbb{R}$, $x \in X$. Let us show that A is bounded. For any $\varepsilon > 0$ there exists $N > 0$ such that

$$\|A_m - A_n\| < \varepsilon \quad \forall m, n \geq N.$$

Therefore

$$\|A_m x - A_n x\| \leq \|A_m - A_n\| \|x\| = \varepsilon \|x\|$$

and

$$\|Ax - A_n x\| \leq \varepsilon \|x\|.$$

This implies that $A - A_n$ is bounded and therefore A is bounded. We also obtain that

$$\|A - A_n\| \leq \varepsilon \quad \forall n \geq N.$$

This implies that $A_n \rightarrow A$ as $n \rightarrow \infty$. $\square$

Remark 3.29. Given a normed vector space $(X, \|\cdot\|)$ over $\mathbb{R}$, we defined its dual space to be $X^* = L(X, \mathbb{R})$. Its elements are continuous linear maps $f: X \rightarrow \mathbb{R}$, called linear functionals. Note that X^* is a Banach space (as $\mathbb{R}$ is complete), although X can be non-Banach. This implies that X^{**} is not necessarily isomorphic to X (although there is a canonical map $X \rightarrow X^{**}$).

Definition 3.30. Let X be a normed vector space and $(x_n)_{n \geq 1}$ be a sequence in X .

- (1) The (formal) series $\sum_{n \geq 1} x_n$ is called **convergent** to $x \in X$ if the sequence of partial sums $(s_n)_{n \geq 1}$ with $s_n = \sum_{k=1}^n x_k$ converges to x . We denote the limit $\lim_{n \rightarrow \infty} s_n$ by $\sum_{n \geq 1} x_n$.
- (2) The series $\sum_{n \geq 1} x_n$ is called **absolutely convergent** if the series $\sum_{n \geq 1} \|x_n\|$ converges in $\mathbb{R}$.

Theorem 3.31. Let X be a Banach space and $\sum_{n=1}^{\infty} x_n$ be an absolutely convergent series in X . Then $\sum_{n=1}^{\infty} x_n$ is convergent.

Proof. For any $n \geq 1$, let

$$s_n = \sum_{k=1}^n x_k \in X, \quad a_n = \sum_{k=1}^n \|x_k\| \in \mathbb{R}.$$

By assumption, the sequence $(a_n)_{n \geq 1}$ is convergent and in particular Cauchy. Therefore $\forall \varepsilon > 0 \exists N > 0$

$$|a_m - a_n| < \varepsilon \quad \forall m, n \geq N.$$

If $n > m \geq N$, then

$$\|s_n - s_m\| = \|x_{m+1} + \cdots + x_n\| \leq \sum_{k=m+1}^n \|x_k\| = a_n - a_m < \varepsilon,$$

hence the sequence $(s_n)_{n \geq 1}$ in X is Cauchy. As X is complete, we conclude that $(s_n)_{n \geq 1}$ is convergent. $\square$

Theorem 3.32. Let X be a Banach space and $A \in L(X)$. Then the series

$$e^A = \sum_{k \geq 0} \frac{A^k}{k!}$$

is convergent.

Proof. As X is Banach, the space $L(X) = L(X, X)$ equipped with an operator norm is also Banach. To prove the theorem, we have to show that the above series is absolutely convergent. We have

$$\sum_{k \geq 0} \|A^k / k!\| = \sum_{k \geq 0} \frac{1}{k!} \|A^k\| \leq \sum_{k \geq 0} \frac{1}{k!} \|A\|^k = e^{\|A\|},$$

where we used the fact that $\|AB\| \leq \|A\| \cdot \|B\|$ (and therefore $\|A^k\| \leq \|A\|^k$). This means that the series $\sum_{k \geq 0} \frac{A^k}{k!}$ in $L(X)$ is absolutely convergent and therefore convergent. $\square$

Definition 3.33. Let X be a normed vector space and let $I \in L(X)$ be the identity operator defined by $I(x) = x$ for $x \in X$. An operator $A \in L(X)$ is called **invertible** in $L(X)$ if there exists an operator $B \in L(X)$ such that $AB = BA = I$. The operator B is called the inverse of A and is denoted by A^{-1} .

Theorem 3.34. *Let X be a Banach space and let $A \in L(X)$ with $\|A\| < 1$. Then the operator $I - A$ is invertible in $L(X)$ and its inverse is $\sum_{k \geq 0} A^k$.*

Proof. The normed vector space $L(X)$ is Banach. To show that $\sum_{k \geq 0} A^k$ is convergent, we have to show that it is absolutely convergent. We have

$$\sum_{k \geq 0} \|A^k\| \leq \sum_{k \geq 0} \|A\|^k = \frac{1}{1 - \|A\|} < \infty.$$

This means that $\sum_{k \geq 0} A^k$ is absolutely convergent and therefore convergent. Let $s_n = \sum_{k=0}^n A^k$ and $B = \lim_{n \rightarrow \infty} s_n = \sum_{k \geq 0} A^k$. Then

$$\begin{aligned} (I - A)B &= (I - A) \lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} (I - A)s_n \\ &= \lim_{n \rightarrow \infty} ((I + A + \cdots + A^n) - (A + A^2 + \cdots + A^{n+1})) = \lim_{n \rightarrow \infty} (I - A^{n+1}) = I \end{aligned}$$

as $\|A^n\| \leq \|A\|^n \rightarrow 0$ as $n \rightarrow \infty$. Similarly one can show that $B(I - A) = I$. Therefore $I - A$ is invertible and $B = \sum_{k \geq 0} A^k$ is its inverse. $\square$

Corollary 3.35. *Let X be a Banach space and let $A \in L(X)$ be invertible in $L(X)$. Then for any $B \in L(X)$ with $\|B\| < \|A^{-1}\|^{-1}$. The operator $A + B$ is invertible.*

Proof. It is enough to show that the operator $A^{-1}(A + B) = I + C$ (where $C = A^{-1}B$) is invertible. We have

$$\|C\| = \|A^{-1}B\| \leq \|A^{-1}\| \cdot \|B\| < \|A^{-1}\| \cdot \|A^{-1}\|^{-1} = 1.$$

Therefore by the above theorem $I + C$ is invertible. $\square$

Assume that we have a bounded linear operator $A \in L(X)$ that is bijective. Then its inverse map $A^{-1}: X \rightarrow X$ is automatically linear, but it is not bounded in general (hence A is not invertible in $L(X)$ in general). However, one can show that A^{-1} is bounded if X is a Banach space.

Theorem 3.36 (Banach-Schauder Theorem). *If X, Y are Banach spaces and $A \in L(X, Y)$ is surjective then it maps open sets in X to open sets in Y .*

Corollary 3.37. *Let X be a Banach space and let $A \in L(X)$ be bijective. Then the linear operator $A^{-1}: X \rightarrow X$ is bounded. Therefore A is invertible in $L(X)$.*

Proof. By the above theorem A maps open sets to open sets. This means that the preimages of open sets with respect to A^{-1} are open. Therefore A^{-1} is continuous, or equivalently, bounded. $\square$

Example 3.38. Let X be the set of sequences $(x_n)_{n \geq 1}$ in $\mathbb{R}$ that have only finitely many nonzero elements. We equip X with the componentwise addition and scalar multiplication and with the norm $\|x\| = \sup_{n \geq 1} |x_n|$. Then the linear operator

$$A: X \rightarrow X, \quad x \mapsto (x_n/n)_{n \geq 1}$$

is bounded and bijective, while its inverse $A^{-1}: X \rightarrow X, x \mapsto (nx_n)_{n \geq 1}$ is not bounded. This implies that A is not invertible in $L(X)$. By the previous result we conclude that X is not Banach.

APPENDIX A. SOME INEQUALITIES

Theorem A.1 (Cauchy-Schwarz inequality). *For all $x, y \in \mathbb{R}^n$, we have*

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

Proof. We have

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n (x_i y_j - x_j y_i)^2 &= \sum_{i=1}^n x_i^2 \sum_{j=1}^n y_j^2 + \sum_{i=1}^n y_i^2 \sum_{j=1}^n x_j^2 - 2 \sum_{i=1}^n x_i y_i \sum_{j=1}^n x_j y_j \\ &= \|x\|^2 \|y\|^2 + \|y\|^2 \|x\|^2 - 2(x, y)(x, y) = 2\|x\|^2 \|y\|^2 - 2(x, y)^2. \end{aligned}$$

As the first expression above is non-negative, we obtain

$$2\|x\|^2 \|y\|^2 - 2(x, y)^2 \geq 0,$$

hence $|(x, y)| \leq \|x\| \|y\|$. □

Theorem A.2 (Minkowski's inequality). *For any $p \geq 1$, we have*

$$\left(\sum_{i=1}^n |x_i + y_i|^p \right)^{1/p} \leq \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} + \left(\sum_{i=1}^n |y_i|^p \right)^{1/p}$$

First, let us prove two auxiliary inequalities.

Theorem A.3 (Young's inequality). *If $p, q > 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$, then*

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}, \quad x, y \geq 0.$$

Proof. As $f(x) = -\log(x)$ is convex (convex-downward) and $\frac{1}{p} + \frac{1}{q} = 1$, we obtain

$$f\left(\frac{1}{p}x^p + \frac{1}{q}y^q\right) \leq \frac{1}{p}f(x^p) + \frac{1}{q}f(y^q)$$

hence

$$\log\left(\frac{1}{p}x^p + \frac{1}{q}y^q\right) \geq \frac{1}{p}\log(x^p) + \frac{1}{q}\log(y^q) = \log(xy)$$

and $xy \leq \frac{1}{p}x^p + \frac{1}{q}y^q$. □

Remark A.4. A convex (convex-downward) function is a function $f: [a, b] \rightarrow \mathbb{R}$ satisfying

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

for $0 < t < 1$ and $a \leq x < y \leq b$. If f is twice differentiable and $f'' \geq 0$, then f is convex.

Theorem A.5 (Hölder inequality). *If $p, q > 1$ satisfy $1/p + 1/q = 1$, then*

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

Proof. We can assume that $x_i, y_i \geq 0$. Also we can rescale our vectors to get $\sum x_i^p = \sum y_i^q = 1$. Then we obtain

$$\sum x_i y_i \leq \frac{1}{p} \sum x_i^p + \frac{1}{q} \sum y_i^q = \frac{1}{p} + \frac{1}{q} = 1 = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \cdot \left(\sum_{i=1}^n |y_i|^q \right)^{1/q}.$$

□

Remark A.6. For $p = 1$, we would get $q = \infty$ and the corresponding inequality should have the form

$$\sum |x_i y_i| \leq \|x\|_1 \cdot \|y\|_\infty = \sum |x_i| \cdot \max_i |y_i|$$

which is obviously true.

Remark A.7. For $p = q = 2$, we obtain the Cauchy-Schwarz inequality

$$(x, y) = \sum x_i y_i \leq \|x\| \cdot \|y\|.$$

Proof of Minkowski's inequality. If $p = 1$, the inequality follows from the triangle inequality for real numbers $|x + y| \leq |x| + |y|$. If $p > 1$, let $q > 1$ be defined by $1/p + 1/q = 1$. We can assume that $x_i, y_i \geq 0$. Then

$$\begin{aligned} \sum_{i=1}^n (x_i + y_i)^p &= \sum_{i=1}^n (x_i + y_i)^{p-1} x_i + \sum_{i=1}^n (x_i + y_i)^{p-1} y_i \\ &\leq \left(\sum_{i=1}^n (x_i + y_i)^{(p-1)q} \right)^{1/q} \left(\left(\sum_{i=1}^n x_i^p \right)^{1/p} + \left(\sum_{i=1}^n y_i^p \right)^{1/p} \right) \end{aligned}$$

From $1/p + 1/q = 1$ we obtain $(p-1)q = p$, hence we can divide both sides by $(\sum_{i=1}^n (x_i + y_i)^p)^{1/q}$ and get

$$\left(\sum_{i=1}^n (x_i + y_i)^p \right)^{1/p} \leq \left(\sum_{i=1}^n x_i^p \right)^{1/p} + \left(\sum_{i=1}^n y_i^p \right)^{1/p}$$

□

Theorem A.8 (Jensen's inequality). *Let f be a convex function, let $x_1, \dots, x_n$ be elements in its domain and let $p_1, \dots, p_n \geq 0$ satisfy $\sum p_i = 1$. Then*

$$f\left(\sum p_i x_i\right) \leq \sum p_i f(x_i).$$

Intuitive explanation. Put particles with mass p_i at the points $(x_i, f(x_i))$ of the plain. Then the center of gravity $(\sum p_i x_i, \sum p_i f(x_i))$ will be over the curve, hence $\sum p_i f(x_i) \geq f(\sum p_i x_i)$. □

Proof. By induction on n

$$\begin{aligned} f\left(\sum_{i=1}^{n+1} p_i x_i\right) &= f\left((1 - p_{n+1}) \sum_{i=1}^n \frac{p_i}{1 - p_{n+1}} x_i + p_{n+1} x_{n+1}\right) \\ &\leq (1 - p_{n+1}) f\left(\sum_{i=1}^n \frac{p_i}{1 - p_{n+1}} x_i\right) + p_{n+1} f(x_{n+1}) \leq \sum_{i=1}^n p_i f(x_i) + p_{n+1} f(x_{n+1}). \end{aligned}$$

□

Remark A.9. We can consider the sequence (p_i) as a collection of probabilities and the sequence $x = (x_i)$ as a random variable. Then we define expectations $E[x] = \sum p_i x_i$, $E[f(x)] = \sum p_i f(x_i)$ and we can write Jensen's inequality in the form

$$f(E[x]) \leq E[f(x)].$$