LECTURE NOTES
MA2223: METRIC SPACES (2014)

SERGEY MOZGOVOY

Contents

1. Metric spaces

1.1. **Definition of a metric space.** Define the distance between two elements $x, y \in \mathbb{R}$ to be

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

Define the distance between two elements $x, y \in \mathbb{R}^2$ to be

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

where $x = (x_1, x_2)$, $y = (y_1, y_2)$. The properties of these distance functions suggest the following general definition.

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

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

Date: October 22, 2014.

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

1.1.1. *Euclidean spaces.* Let

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

This is a vector space over $\mathbb{R}$, where the sum of two elements $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ is

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

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

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

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 $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by

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

In order to prove that this map is a metric on $\mathbb{R}^n$ we will need the following result.

Theorem 1.2 (Cauchy-Schwarz inequality)**.** *For any $x, y \in \mathbb{R}^n$, we have*

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

Proof. We have

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

Note that the first expression above is always non-negative. Therefore

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

and $(x,y)^2 \le \|x\|^2 \|y\|^2$. Taking the square roots we obtain $|(x,y)| \le \|x\| \|y\|$. □

Corollary 1.3. *The map $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined above is a metric on $\mathbb{R}^n$.*

Proof. We will check all the axioms from the definition of a metric space.

(1) $d(x,y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} \ge 0$ is clear.

(2) If $d(x,y) = 0$ then

$$\sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} = 0$$

and $(x_1 - y_1)^2 + \dots + (x_n - y_n)^2 = 0$. This implies that $(x_i - y_i)^2 = 0$ for any i. Therefore $x_i = y_i$ for any i. This means that $x = y$.

Conversely, if $x = y$ then $x_i = y_i$ for every i and therefore

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} = 0.$$

(3)

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

(4) We have to prove $d(x,z) \leq d(x,y) + d(y,z)$ or equivalently

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

Let $a = x - y$, $b = y - z$. Then $x - z = a + b$ and we have to prove

$$\begin{aligned} &\|a + b\| \leq \|a\| + \|b\| \\ \Leftrightarrow &\|a + b\|^2 \leq (\|a\| + \|b\|)^2 \\ \Leftrightarrow &(a + b, a + b) \leq (a, a) + (b, b) + 2\|a\|\|b\| \\ \Leftrightarrow &(a, a) + (b, b) + 2(a, b) \leq (a, a) + (b, b) + 2\|a\|\|b\| \\ \Leftrightarrow &(a, b) \leq \|a\|\|b\|. \end{aligned}$$

The last inequality follows from the Cauchy-Schwarz inequality.

□

Remark 1.4. *The metric space $(\mathbb{R}^n, d)$ is called the Euclidean space. The metric d is called the Euclidean metric (or Euclidean distance).*

1.2. **Examples of metric spaces.**

Example 1.5. 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 a metric d_∞ on $\mathbb{R}^n$ by

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

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

Example 1.6. Define a metric d_1 on $\mathbb{R}^2$ by

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

Note that this metric is different from the Euclidean metric on $\mathbb{R}^2$. Similarly we can define a metric on $\mathbb{R}^n$ by

$$d_1(x,y) = \sum_{i=1}^n |x_i - y_i| = |x_1 - y_1| + \dots + |x_n - y_n|.$$

Example 1.7. For any $p \ge 1$, define a metric d_p on $\mathbb{R}^n$ by

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

To see that it satisfies the triangle inequality we will use (without proof)

Theorem 1.8 (Minkowski inequality)**.** *Let $p \ge 1$ and $a, b \in \mathbb{R}^n$. Then*

$$\left(\sum_{i=1}^n |a_i + b_i|^p\right)^{1/p} \le \left(\sum_{i=1}^n |a_i|^p\right)^{1/p} + \left(\sum_{i=1}^n |b_i|^p\right)^{1/p}.$$

Triangle inequality $d_p(x,z) \le d_p(x,y) + d_p(y,z)$ can be written as

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

Let $a_i = x_i - y_i$ and $b_i = y_i - z_i$. Then $x_i - z_i = a_i + b_i$ and the previous inequality can be written as

$$\left(\sum_{i=1}^n |a_i + b_i|^p\right)^{1/p} \le \left(\sum_{i=1}^n |a_i|^p\right)^{1/p} + \left(\sum_{i=1}^n |b_i|^p\right)^{1/p}$$

which is precisely the Minkowski inequality.

Remark 1.9. *Note that*

$$d_2(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

is the Euclidean metric. Note that

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

is a metric considered in the previous example.

Example 1.10. Let X be any set. Define the discrete metric on X by

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

Example 1.11. Given a set S, let 2^S be the set of all subsets of S, called the power set of S. Assume that S is finite. We would like to define a metric on 2^S. Any two elements $A, B \in 2^S$ are subsets of S. It is natural to define a distance between them as a difference between these subsets, that is, to define

$$d(A,B) = \#(A \backslash B) + \#(B \backslash A).$$

We can identify the power set 2^S with the set of all functions $f : S \to \{0,1\}$, where a subset $A \subset S$ is identified with its characteristic function $1_A : S \to \{0,1\}$

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

The number $d(A,B)$ counts those elements $x \in S$ which are contained in one of A, B but not in the other. This means that it counts $x \in S$ where $1_A(x) \neq 1_B(x)$. Therefore

$$d(A,B) = \sum_{x \in S} |1_A(x) - 1_B(x)|.$$

More generally, for any two functions $f : S \to \mathbb{R}$, $g : S \to \mathbb{R}$ we can define

$$d(f,g) = \sum_{x \in S} |f(x) - g(x)|.$$

This gives a metric on the set $\mathbb{R}^S$ of all functions $f : S \to \mathbb{R}$. Note that this metric coincides with a metric d_1 on $\mathbb{R}^n$.

Example 1.12. Let $C[a,b]$ be the set of all continuous functions $f : [a,b] \to \mathbb{R}$. Define a metric $d = d_\infty$ on $C[a,b]$ by

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

Let us check the axioms of a metric

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

$$\sup_{x \in [a,b]} |f(x) - g(x)| \leq \sup_{x \in [a,b]} |f(x) - h(x)| + \sup_{x \in [a,b]} |h(x) - g(x)|.$$

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 $x_0 \in [a,b]$ we obtain the required inequality.

Example 1.13. Define a metric d_1 on $C[a,b]$ by

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

The axioms 1-3 of a metric a 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 1.14 (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.

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

Definition 1.15. 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 1.16. 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 1.17. Let X be a discrete metric space and $x \in X$. Describe $B(x, 1)$, $\overline{B}(x, 1)$, $B(x, 2)$.

Definition 1.18. 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 1.19. Define the diameter of a subset $A \subset X$ to be
$$d(A) := \sup_{x,y \in A} d(x, y).$$

Example 1.20. The diameter of $B(0, 1)$ in the Euclidean space $\mathbb{R}^2$ is 2.

Remark 1.21. *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 1.22. *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 1.23. *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 $y \in A$,
$$d(x, y) \leq d(A) < r \Rightarrow y \in B(x, r).$$
This means that $A \subset B(x, r)$. □

Example 1.24. (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.

1.4. **Open and closed sets.**

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

Definition 1.25. 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 1.26. *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 $y \in B(x, r) \Rightarrow d(x, y) < r \Rightarrow r - d(x, y) > 0$. 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. □

Example 1.27. 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 1.28. *We have*

(1) *The sets $\varnothing$ 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.

□

Example 1.29. 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 1.30. *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*

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

Their intersection is

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

which is not an open set.

Lemma 1.31. *Let $A \subset X$ be a subset. Then its interior A° is an open set.*

Proof. For any $x \in A^\circ$, there exists $r_x > 0$ such that $B(x, r_x) \subset A$. We know that $B(x, r_x)$ is an open set. This means that any point of $B(x, r_x)$ is an internal point of $B(x, r_x)$ and therefore also of A. This implies that $B(x, r_x) \subset A^\circ$ (all points are internal). Therefore

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

As a union of open sets it is again an open set. □

1.4.2. *Closed sets.*

Definition 1.32. 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 $y \in A \backslash \{x\}$ 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 1.33. *A subset $A \subset X$ is a closed set if and only if $X \backslash A$ is an open set.*

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

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

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 $x \in X \backslash A$, a contradiction.

Conversely, assume that $X \backslash A$ is an open set. We have to show that every limit point x of A is contained in A. Assume that $x \in X \backslash A$. As $X \backslash A$ is open, there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subset X \backslash A$, that is, $B(x, \varepsilon) \cap A = \varnothing$. But this implies that x can be a limit point of A, a contradiction. □

Example 1.34. The closed interval $[a, b] \subset \mathbb{R}$ is a closed set. Indeed, its complement $\mathbb{R} \backslash [a, b] = (-\infty, a) \cup (b, \infty)$ is a union of open intervals and we know that they are open sets.

Theorem 1.35. *Let (X, d) be a metric space. Then*

(1) *The sets $\varnothing$ 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) $X \backslash \varnothing = X$ is an open set $\Rightarrow \varnothing$ is closed. $X \backslash X = \varnothing$ is an open set $\Rightarrow X$ is closed.

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

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

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

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

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

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

□

Theorem 1.36. *Let $A \subset X$ be a subset. Then its closure $\overline{A}$ is a closed set. It is an intersection of all closed sets containing A.*

Proof. Let x be a limit point of $\overline{A}$. We have to show that $x \in \overline{A}$. Assume that $x \notin \overline{A}$. Then for any $\varepsilon > 0$ there exists $y \in \overline{A}$ such that $d(x, y) < \varepsilon/2$. If $y \in \overline{A} \backslash A$ then there exists $z \in A$ such that $d(y, z) < \varepsilon/2$. Then

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

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, that is, $x \in \overline{A}$. A contradiction.

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 $\overline{A} = \bigcup_{A \subset F - \text{ cl}} F$. □

Example 1.37. Consider $A = (0, 1) \subset \mathbb{R}$. Its closure is $[0, 1]$ and we know that it a closed set.

1.5. **Convergent sequences.** Let (X,d) be a metric space. A sequence $(x_n)_{n\geq 1}$ of points in X is a collection of elements $x_1, x_2, \dots$ in X.

Definition 1.38. A sequence $(x_n)_{n\geq 1}$ of points in X is said to converge to $x \in X$ if for any $\varepsilon > 0$ there exists $N > 0$ such that

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

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

Remark 1.39. *The definition of convergence of a sequence (x_n) to x is equivalent to the statement that $d(x_n, x) \to 0$, $n \to \infty$.*

Lemma 1.40. *Every convergent sequence in a metric space has a unique limit.*

Proof. Assume that $(x_n)_{n\geq 1}$ converges to x and y. For any $\varepsilon > 0$ there exist $M > 0$ and $N > 0$ such that $d(x_n, x) < \varepsilon$ for $n \geq M$ and $d(x_n, y) < \varepsilon$ for $n \geq N$. Without loss of generality we can assume that $M \leq N$. Then for any $n \geq N$ we have

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

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

Example 1.41. Consider a sequence $(x_n)_{n\geq 1}$ in $\mathbb{R}^2$ with $x_n = \left(\frac{1}{n}, \frac{n-1}{n}\right)$. Then $\lim_{n\to\infty} x_n = x = (0,1)$. Indeed,

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

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

Lemma 1.42. *A sequence $(x^{(n)})_{n\geq 1}$ in the Euclidean space $\mathbb{R}^k$ converges to $x \in \mathbb{R}^k$ if and only if, for any $1 \leq i \leq k$, we have $x_i^{(n)} \to x_i$, $n \to \infty$.*

Proof. Assume that $x^{(n)} \to x$, $n \to \infty$. Then

$$d(x^{(n)}, x) = \sqrt{(x_1^{(n)} - x_1)^2 + \cdots + (x_k^{(n)} - x_k)^2} \to 0.$$

This implies that for any $1 \leq i \leq k$: $(x_i^{(n)} - x_i)^2 \to 0$. Therefore $x_i^{(n)} \to x_i$, $n \to \infty$.

Conversely, if for any $1 \leq i \leq k$: $x_i^{(n)} \to x_i$, $n \to \infty$, then $(x_i^{(n)} - x_i)^2 \to 0$ and also

$$d(x^{(n)}, x) = \sqrt{(x_1^{(n)} - x_1)^2 + \cdots + (x_k^{(n)} - x_k)^2} \to 0.$$

This means that $x^{(n)} \to x$, $n \to \infty$. □

Example 1.43. Let $C[a,b]$ be equipped with a supremum metric d_∞. Then a sequence of functions $(x_n(t))_{n\geq 1}$ in $C[a,b]$ converges to $x(t) \in C[a,b]$ if and only if $\forall \varepsilon > 0\ \exists N > 0\ \forall n \geq N$:

$$d(x_n, x) = \sup_{t\in[a,b]} |x_n(t) - x(t)| < \varepsilon.$$

This means that $\forall t \in [a,b]$, we have

$$|x_n(t) - x(t)| < \varepsilon.$$

This leads us to the following definition.

Definition 1.44. Let $(x_n(t))_{n\geq 1}$ be sequence of functions on the interval $[a,b]$ and let $x(t)$ be also a function on $[a,b]$. Then

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

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

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

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

that is, $\forall t\in[a,b]:\ \lim_{n\to\infty}x_n(t)=x(t)$.

Note that the uniform convergence implies the pointwise convergence. The uniform convergence is equivalent to the convergence in the metric space $(C[a,b],d_\infty)$.

Example 1.45. Consider a sequence of functions $x_n(t)=t^n$ in $C[0,1]$ with metric d_∞. Assume that $\lim_{n\to\infty}x_n=x$ for some $x\in C[a,b]$. Then $x_n(t)$ converges uniformly to $x(t)$ and therefore also pointwise. Then, for any $t\in[0,1]$, we have

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

This function is not continuous. Therefore the sequence (x_n) is not convergent in $(C[0,1],d_\infty)$. On the other hand, it is convergent to the zero function with respect to the metric d_1. Indeed,

$$d_1(x_n,0)=\int_0^1|t^n-0|\,dt=\int_0^1 t^n dt=\left.\frac{t^{n+1}}{n+1}\right|_0^1=\frac{1}{n+1}\to 0,\qquad n\to\infty.$$

Therefore $x_n\to 0$, $n\to\infty$.

Lemma 1.46. *Assume that a sequence $(x_n)_{n\geq 1}$ in X converges to x. Then, for any point $y\in X$, we have*

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

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

Lemma 1.47. *For any points $x,y,z\in X$ we have*

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

Proof. Indeed, we have by the triangle inequality

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

Therefore

$$d(x,y)-d(x,z)\leq d(y,z),\qquad d(x,z)-d(x,y)\leq d(y,z)$$

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

Now let us prove the original lemma. We have

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

But $d(x_n,x)\to 0$ as $n\to\infty$. $\Rightarrow$ $|d(x_n,y)-d(x,y)|\to 0$ as $n\to\infty$ $\Rightarrow$ $d(x_n,y)\to d(x,y)$ as $n\to\infty$. □

Lemma 1.48. *Let $A\subset X$ be a subset. A point $x\in X$ is a limit point of A if and only if there exists a sequence in $A\backslash\{x\}$ convergent to x.*

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

Conversely, assume that there exists a sequence $(x_n)_{n\geq 1}$ in $A\backslash\{x\}$ 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. □

Lemma 1.49. *Let (X,d) be a metric space. A subset $A\subset X$ is closed in X if and only if the limit of any convergent sequence of points in A is contained in A.*

Proof. "$\Rightarrow$ ": Assume that A is closed and let (x_n) be a sequence of point in A convergent to $x\in X$. Then x is a limit point of A. Therefore $x\in A$.

"$\Leftarrow$": Let x be a limit point of A. Then there exists a sequence of points in $A\backslash\{x\}$ convergent to x. By our assumption the limit of this sequence is contained in A, that is $x\in A$. This implies that A is closed set. □

1.6. **Continuous maps.**

Definition 1.50. Let (X, d), (Y, d') be two metric spaces. A map $f : X \to Y$ is called

(1) continuous at a point $x_0 \in X$ if $\forall \varepsilon > 0\ \exists \delta > 0$:
$$\forall x \in X \text{ s.t. } d(x_0, x) < \delta : \quad d'(f(x_0), f(x)) < \varepsilon.$$
(2) continuous if it is continuous at every point of X.

Remark 1.51. *The definition of continuity at a point x_0 can be written equivalently as*
$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in B(x_0, \delta) : \quad f(x) \in B(f(x_0), \varepsilon)$$
or in the form
$$\forall \varepsilon > 0\ \exists \delta > 0 : \qquad f(B(x_0, \delta)) \subset B(f(x_0), \varepsilon).$$

Lemma 1.52. *A map $f : X \to Y$ is continuous if and only if for any sequence (x_n) in X convergent to $x_0 \in X$, the sequence $(f(x_n))$ converges to $f(x_0)$.*

Proof. $\Rightarrow$: Assume that f is continuous and $x_n \to x_0$, $n \to \infty$. Let $\varepsilon > 0$. As f is continuous at x_0, $\exists \delta > 0$:
$$d(x_0, x) < \delta \ \Rightarrow\ d'(f(x_0), f(x)) < \varepsilon.$$
As $\lim_{n\to\infty} x_n = x_0$,
$$\exists N > 0\ \forall n \geq N :\ d(x_0, x_n) < \delta.$$
This implies that
$$\exists N > 0\ \forall n \geq N :\ d'(f(x_0), f(x_n)) < \varepsilon.$$
This means that $f(x_n) \to f(x_0)$ as $n \to \infty$.

$\Leftarrow$: Assume that f is not continuous at $x_0 \in X$. Then
$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x \in X :\ d(x_0, x) < \delta,\ d'(f(x_0), f(x)) \geq \varepsilon.$$
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 $x_n \to x$, but $f(x_n) \not\to f(x_0)$. A contradiction. □

Theorem 1.53. *A map $f : (X, d) \to (Y, d')$ between two metric spaces is continuous if and only if for any open set $U \subset Y$ its preimage $f^{-1}(U)$ is open in X.*

Proof. $\Rightarrow$: Assume that $f : X \to Y$ 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 $\exists \varepsilon > 0 :\ B(f(x), \varepsilon) \subset U$. By the continuity of f, there exists $\delta > 0$ such that
$$f(B(x, \delta)) \subset B(f(x), \varepsilon) \subset U.$$
This means that $B(x, \delta) \subset f^{-1}(U)$, that is, x is an internal point of $f^{-1}(U)$. Therefore $f^{-1}(U)$ is an open set.

$\Leftarrow$: Assume that the preimage of any open set in Y is open in X. Let $x \in X$ and $\varepsilon > 0$. 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,
$$f(B(x, \delta)) \subset U = B(f(x), \delta).$$
This meas that f is continuous at $x \in X$. Therefore f is continuous. □

Lemma 1.54. *Let $f : X \to Y$ and $g : Y \to Z$ be continuous maps between metric spaces. Then their composition $gf : X \to Z$ 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. Note that
$$(gf)^{-1}(U) = f^{-1}(g^{-1}(U)).$$
The set $g^{-1}(U)$ is open as g is continuous. The set $f^{-1}(g^{-1}(U))$ is open as f is continuous. □

Definition 1.55. A map $f : (X, d) \to (Y, d')$ between two metric spaces is called an isometry if

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

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

Lemma 1.56. *Let $f : (X, d) \to (Y, d')$ be an isometry. Then*

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

Proof. (1) Let us show that f is continuous at some point $x \in X$. For any $\varepsilon > 0$, let $\delta = \varepsilon$. Then, for any $y \in X$ with $d(x, y) < \delta$, we have

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

This implies that f is continuous at x.

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

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

which implies $x = y$. This means that f is injective.

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

□

Definition 1.57. A subset $A \subset X$ is called dense if its closure $\overline{A} = X$.

Example 1.58. (1) Let $X = [0, 1]$ have the Euclidean metric and let $A = (0, 1)$. Then the closure of A is $[0, 1] = X$. Therefore A is dense in X.

(2) The set $\mathbb{Q}$ is dense in $\mathbb{R}$. Indeed, given $x \in \mathbb{R}$, we can write it in a decimal notation as

$$x = x_0.x_1x_2x_3 \ldots$$

where x_0 is an integer part of x and $x_1, x_2, \ldots$ are the digits forming the fractional part of x. For any $n \geq 1$ the numbers

$$y_n = x_0.x_1 \ldots x_n = \sum_{i=0}^{n} \frac{x_i}{10^i}$$

are rational. Moreover, $\lim_{n\to\infty} y_n = x$ as

$$|x - y_n| = 0.0 \ldots 0x_{n+1}x_{n+2} \cdots = \sum_{i\geq n+1} \frac{x_i}{10^i} < \sum_{i\geq n+1} \frac{10}{10^i}$$

$$= \frac{1}{10^n}\left(1 + \frac{1}{10} + \ldots\right) = \frac{1}{10^n} \cdot \frac{1}{1 - 1/10} \to 0, \quad n \to \infty.$$

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

Lemma 1.59. *Let $f, g : X \to Y$ 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 $x \in X$. 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 a closed set. Indeed, let (x_n) be a sequence in F convergent to $x \in X$. By the continuity of f and g, we have $f(x_n) \to f(x)$ and $g(x_n) \to g(x)$ as $n \to \infty$. But $x_n \in F \Rightarrow f(x_n) = g(x_n)$. Therefore the limits of these sequences coincide, that is $f(x) = g(x)$. This implies that $x \in F$. Therefore F is closed. By our assumption $A \subset F$. Hence $X = \overline{A} \subset F$, that is $F = X$ and

$$f(x) = g(x), \qquad \forall x \in X.$$

□

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

Definition 1.60. A sequence $(x_n)_{n\geq 1}$ in X is called a Cauchy (or fundamental) sequence if $\forall \varepsilon > 0$ $\exists N > 0$:

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

Lemma 1.61. *Any 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 $\forall \varepsilon > 0$ $\exists N > 0$

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

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

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

This means that $(x_n)_{n\geq 1}$ is a Cauchy sequence. □

Example 1.62. The sequence $(1/n)_{n\geq 1}$ in $(0, +\infty)$ is a Cauchy sequence, but it is not convergent in $(0, +\infty)$. It is a Cauchy sequence as it is convergent to 0 in $\mathbb{R}$. It is not convergent in $(0, +\infty)$ as the limit $\lim_{n\to\infty} \frac{1}{n} = 0 \notin (0, +\infty)$.

Definition 1.63. A metric space (X, d) is called a complete metric space if any Cauchy sequence in X is convergent in X.

Example 1.64. We know that any Cauchy sequence in $\mathbb{R}$ is convergent. Therefore $\mathbb{R}$ is a complete metric space.

Example 1.65. Let us show that the Euclidean space $\mathbb{R}^k$ is complete. Let $(x^{(n)})_{n\geq 1}$ be a Cauchy sequence. Then for any $1 \leq i \leq k$, the sequence $(x_i^{(n)})_{n\geq 1}$ in $\mathbb{R}$ is a Cauchy sequence: $\forall \varepsilon > 0$ $\exists N > 0$

$$\forall m, n \geq N : \qquad d(x^{(m)}, x^{(n)}) = \sqrt{\sum_{i=1}^{n} (x_i^{(m)} - x_i^{(n)})^2} < \varepsilon$$

and therefore $\left| x_i^{(m)} - x_i^{(n)} \right| < \varepsilon$. Therefore the sequence $(x_i^{(n)})_{n\geq 1}$ is convergent to some $x_i \in \mathbb{R}$. This implies that $(x^{(n)})_{n\geq 1}$ converges $x = (x_1, \ldots, x_k)$.

Example 1.66. Let us show that the space $C[a, b]$ with a supremum (or uniform) metric d_∞ is complete. Let $(x_n)_{n\geq 1}$ 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: $\forall \varepsilon > 0$ $\exists N > 0$ $\forall m, n \geq N$:

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

Therefore this sequence converges to some $x(t_0) \in \mathbb{R}$. In this way we obtain a function $x : [a, b] \to \mathbb{R}$. We have to show that $(x_n)_{n\geq 1}$ converges uniformly to x and that x is a continuous function. We can write the Cauchy condition in the form: $\forall \varepsilon > 0$ $\exists N > 0$

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

Taking a limit $m \to \infty$ we obtain

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

which means the sequence of functions $(x_n)_{n\geq 1}$ converges uniformly to $x : [a, b] \to \mathbb{R}$. The continuity of x follows from the following lemma.

Lemma 1.67. *Assume that a sequence of functions $(x_n(t))$ in $C[a, b]$ converges uniformly to a function $x : [a, b] \to \mathbb{R}$. Then $x \in C[a, b]$.*

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

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

Taking $n = N$ we obtain

$$\forall t \in [a,b] : \qquad |x(t) - x_N(t)| \leq \frac{\varepsilon}{3}.$$

From $x_N \in C[a,b] \Rightarrow \exists \delta > 0$

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

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 $[a,b]$. □

Lemma 1.68. *Let (X,d) be a metric space, $A \subset X$ be a subset and (A, d_A) be a metric space with an induced metric. Then*

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

Proof. 1. Assume that (x_n) is a Cauchy sequence in A. Then it is Cauchy in $X \Rightarrow$ 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 $x = y \in A$. □

Example 1.69. 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 $\overline{\mathbb{Q}} = \mathbb{R}$).

Definition 1.70. Let (X,d) be a metric space. Define a completion of (X,d) to be a complete metric space (X^*, d^*) together with an isometry $i : X \to X^*$ such that $i(X)$ is dense in X^*.

Remark 1.71. *We have seen that an isometry is an injective map. Therefore we can consider X as a subspace of X^*.*

Example 1.72. Let $\mathbb{Q}$ be equipped with the Euclidean metric. Its completion is the Euclidean space $\mathbb{R}$ with a canonical inclusion $i : \mathbb{Q} \to \mathbb{R}$:

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

Theorem 1.73. *Every metric space has a completion (unique up to an isometry).*

Proof. Uniqueness. Assume that X has two completions $i : X \to X^*$ and $i' : X \to X'$. Let us construct an isometry $\varphi : X^* \to X'$ 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 $\varphi(x^*) = x'$. For any $x \in X$, we have $\varphi(x) = x$. To see that $\varphi : X^* \to X'$ is an isometry, consider two Cauchy sequences (x_n), y_n that converge to $x^*, y^* \in X^*$ and $x', y' \in X'$. Then

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

This implies that $d'(\varphi(x^*), \varphi(y^*)) = d^*(x^*, y^*)$.

Existence. Let $\overline{X}$ be the set of all Cauchy sequences in X. We call two sequences (x_n), (y_n) equivalent if $\lim d(x_n, y_n) = 0$. This is an equivalence relation (reflexivity, symmetry, transitivity). Let X^* be the set of equivalence classes.

Metric on X^*. For any two classes x^*, y^* choose representatives, that is, Cauchy sequences (x_n), (y_n), and define

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

(1) The sequence $(d(x_n, y_n))_{n\geq 1}$ is Cauchy and therefore convergent.
(2) The limit $\lim_{n\to\infty} d(x_n, y_n)$ does not depend on the choice of 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

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

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.

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

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

Therefore $\forall n \geq N$:

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

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

X^* *is complete.* 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^*

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

□

1.7.1. *Banach fixed point theorem.*

Definition 1.74. Let (X, d) be a metric space. A map $f : X \to X$ 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 1.75. *A contraction is always continuous.*

Theorem 1.76 (Banach's fixed point theorem)**.** *Let X be a complete metric space and $f : X \to X$ be a contraction. Then f a unique fixed point ($x \in X$ such that $f(x) = x$).*

Proof. 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 $k, n \geq 1$

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

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

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

that is, x is a fixed point.

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

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

Therefore $d(x,y) = 0$ and $x = y$. □

Example 1.77. Let $X = (0,1)$ and $f : X \to X$ 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 $f : X \to X$, $x \mapsto \frac{1}{2}x$, then f has a fixed point $x = 0$.

Example 1.78 (Picard Theorem)**.** Let

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

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

$$|f(t,y_1) - f(t,y_2)| \le L\,|y_1 - y_2|\,, \qquad t \in [a,b],\ y_1, y_2 \in [c,d],$$
$$|f(t,y)| \le M, \qquad t \in [a,b],\ y \in [c,d].$$

Let us choose $\varepsilon > 0$ such that

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

Consider the space

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

equipped with a supremum metric. Consider a map

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

Note that $F(y)$ is continuous and

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

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

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

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

$$d(F(y_1), F(y_2)) = \sup_{t\in I} |F(y_1)(t) - F(y_2)(t)| = \sup_{t\in I} \left|\int_{t_0}^{t} (f(s, y_1(s)) - f(s, y_2(s)))\right| ds$$
$$\le \sup_{t\in I} \int_{t_0}^{t} L\,|y_1(s) - y_2(s)|\,ds \le \sup_{t\in I} \int_{t_0}^{t} Ld(y_1, y_2)ds = Ld(y_1, y_2) \sup_{t\in I} |t - t_0| = \varepsilon Ld(y_1, y_2)$$

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 the 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: $\varnothing$ 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) $\varnothing, X \in \tau$.
(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.

Definition 2.2. Let (X,τ) be a topological space.

(1) A subset $A \subset X$ is called closed if its complement $X\backslash A$ is open.
(2) The closure of a subset $A \subset X$ is the minimal closed set $\overline{A}$ that contains A. It is the intersection of all closed sets that contain A.

Definition 2.3. A map $f : (X,\tau) \to (Y,\tau')$ is called continuous if for any 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.

Remark 2.4. *In the same way as in the case of metric space one can prove that if* $f : (X,\tau) \to (Y,\tau')$ *and* $g : (Y\ \tau') \to (Z,\tau'')$ *are continuous maps then* $gf : (X,\tau) \to (Z,\tau'')$ *is continuous.*

Example 2.5. 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 2.6. A topological space (X,τ) is called metrizable if τ is generated by some metric on X.

Example 2.7. Let X be an arbitrary set and $\tau = \{\varnothing, X\}$. Then τ is a topology, called the trivial topology.

For example let $X = \{x,y\}$ and τ be the trivial topology. Then τ is not metrizable. Indeed, assume that τ is a metric topology for some metric d. Let $r = d(x,y)$. Then $B(x,r) = \{x\}$ is an open set. But $\{x\}$ is not an element of τ. A contradiction.

Example 2.8. Let X be an arbitrary set and let τ consist of all subsets of X. Then τ is a topology, called the discrete topology. Note that this is actually a metric topology for the discrete metric defined by

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

Indeed, we have seen that every subset of X is open with respect to this metric.

Example 2.9. Let X be an arbitrary set and τ consists of $\varnothing$ and all subsets $U \subset X$ with a finite complement $X\backslash U$. This is a topology on X. Indeed, let $(U_i)_{i\in I}$ be a collection of elements in τ. We will assume that neither of them is empty (we can remove them, if there are any) and therefore $X\backslash U_i$ are finite. To see that the union is in τ, we consider the complement

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

where we applied the De Morgan's rule. This set is finite as an intersection of finite sets. Therefore $\bigcup_{i\in I} U_i \in \tau$.

Now assume that $(U_i)_{i\in I}$ is a finite collection of elements in τ. We can assume that neither of them is empty as otherwise the intersection is automatically empty. To see that the intersection is in τ, we consider the complement

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

This set is finite as a finite union of finite sets. Therefore $\bigcap_{i\in I}U_i\in\tau$. This proves that τ is a topology on X. It is called sometimes the Zariski topology of X.

If X is finite then τ is a discrete topology. If X is infinite, then τ is not metrizable. Assume that there exists a metric d that generates τ. Let $x,y\in X$ be two distinct points and $r=d(x,y)/2$. Then open sets $U=B(x,r)$ and $V=B(y,r)$ satisfy $U\cap V=\varnothing$ and therefore

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

But as U, V are open and non-empty, their complements are finite and the union of these complements is finite. This contradicts to the assumption that X is inifinite.

2.2. **Subspace and product topologies.**

Definition 2.10. 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 2.11. *Let (X, τ) be a topological space and $A \subset X$ be a subset. Then*

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

2. We can write $f_A : A \to Y$ as a composition $A \xrightarrow{i} X \xrightarrow{f} Y$. The composition fi of two continuous maps is continuous. Therefore $f|_A$ is continuous. □

Definition 2.12. 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 is an open subset of X and V_i is an open subset of Y. This topology is called the product topology.

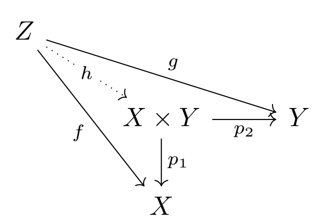
Remark 2.13. *Let us check that the above collection $\overline{\tau}$ of subsets in $X \times Y$ is indeed a topology. The first axiom is satisfied as $\varnothing \times \varnothing$ and $X \times Y$ are in $\overline{\tau}$. The second axiom is automatically satisfied. To prove the third axiom it is enough to consider the intersection of just two elements $\bigcup U_i \times V_i$ and $\bigcup U'_j \times V'_j$ from $\overline{\tau}$. We obtain*

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

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

Theorem 2.14. *Let X, Y be two topological spaces and let $X \times Y$ be equipped with a product topology. Then*

(1) *The projection maps*
$$p_1 : X \times Y \to X, \quad (x, y) \mapsto x, \qquad p_2 : X \times Y \to Y, \quad (x, y) \mapsto y$$
are continuous.
(2) *Given continuous maps $f : Z \to X$, $g : Z \to Y$, there exists a unique continuous map $h : Z \to X \times Y$ that makes the diagram commute ($p_1 h = f$ and $p_2 h = g$).*



Proof. 1. Let us show that $p_1 : X \times Y \to X$ 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 $h : Z \to X \times Y$ should satisfy $h(z) = (f(x), g(y))$. 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 $h : Z \to X \times Y$ is continuous. $\square$

2.3. **Hausdorff topological 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 2.15. 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 $x \in V \subset X$ is called a neighborhood of x if there exists open set $U \subset X$ such that $x \in U \subset V$.

Definition 2.16. Let (X, τ) be a topological space. A sequence (x_n) in X is said to converge to $x \in X$ if for any nbd $U \ni x$ there exists $N > 0$ such that $x_n \in U$ for $n \geq N$.

Example 2.17. Let X be a set and $\tau = \{\varnothing, X\}$ be the trivial topology on it. 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 2.18. A topological space (X, τ) is called Hausdorff if for any $x, y \in X$ with $x \neq y$ there exist open sets $U \ni x$, $V \ni y$ such that $U \cap V = \varnothing$.

Example 2.19. Any metric space $(X, d))$ is Hausdorff. For $x \neq y$, let $r = \frac{1}{2} d(x, y)$ and

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

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

Example 2.20. If X is a set with more than one element, then X equipped with the trivial topology is not Hausdorff. Indeed, let $x, y \in X$ and $x \neq y$. If $U \ni x$ and $V \ni y$ are open, then $U = V = X$ and $U \cap V \neq 0$.

Lemma 2.21. *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 $U \cap V = \varnothing$. As $x_n \to x, y$ for $n \to \infty$, there exist $N, N' > 0$ s.t.

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

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

Lemma 2.22. *Let (X, τ) be a Hausdorff topological space. For any point $x \in X$ the set $\{x\}$ is closed.*

Proof. Let $U = X \backslash \{x\}$. 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 $U = \cup_{y \in U} V_y$ and this set is open as a union of open sets. This means that $\{x\} = X \backslash U$ is closed. □

Example 2.23. Let X be an infinite set with a Zariski topology. For any point $x \in X$ the set $X \backslash \{x\}$ is open because is complement is finite. Therefore the set $\{x\}$ is closed. However, X is not Hausdorff. Indeed, let $x \neq y$ be points in X and assume that there exist open sets $U \ni x$, $V \ni y$ such that $U \cap V = \varnothing$. Then

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

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

2.4. **Connected topological spaces.**

Definition 2.24. 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, if $U \subset X$ is both open and closed then $U = \varnothing$ or $U = X$. A subset $A \subset X$ is called connected if (A, τ_A) is connected.

Example 2.25. 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 2.26. *Let $(X, \tau), (Y, \tau')$ be topological spaces and let (X, τ) be connected. If $f : X \to Y$ is continuous 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. □

Corollary 2.27. *Let (X, τ) be a connected topological space and let (Y, τ') have discrete topology. Then any continuous map $f : X \to Y$ is constant.*

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

Lemma 2.28. *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 the other intervals one can use similar arguments or apply the result for (a,b) together with a result on the connected closures, to be proved later. Assume that $(a,b) = U \dot\cup V$, where U, V are open, nonempty subsets. Let $c \in U$ and assume, without loss of generality, that $(c,b) \cap V \neq \varnothing$. Let $d = \inf[c,b) \cap V$. As V is closed in $[c,b)$ and d is a limit point of V, we conclude that $d \in V$. Therefore $c \neq d$ and $[c,d) \subset U$ (d is the minimal element in $[c,b) \cap V$). This implies that d is a limit point of U and as U is closed in (a,b), we conclude that $d \in U$. Therefore $d \in U \cap V$, a contradiction. □

Lemma 2.29. *Let (X, τ) be a topological space and $A \subset X$ be a connected subset. Then the closure $\overline{A} \subset X$ is connected.*

Proof. Restricting the topology to $\overline{A}$, we can assume that $\overline{A} = X$. 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 $U \cap A \neq \varnothing$. As A is connected we conclude $U \cap A = A$ and $A \subset U$. As U is closed in X, we obtain $X = \overline{A} \subset U$ and therefore $U = X$. This implies that X is connected. □

Lemma 2.30. *If $(X, \tau), (Y, \tau')$ are connected then $X \times Y$ is 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 connected. From

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

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

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

that is, $U = X \times Y$. Therefore $X \times Y$ is connected. □

Proposition 2.31. *Let (X, τ) be a topological space and let $(A_i)_{i \in I}$ be a collection of connected subsets such that $\cap_{i \in I} A_i \neq \varnothing$. 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 $A_i \cap U \neq \varnothing$. 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. □

Definition 2.32. A subset of X is called a connected component if it is connected and it is not contained in any larger connected subset.

Theorem 2.33. *Any connected component of X is a closed set and X is a disjoint union of its connected components.*

Proof. If A is a connected component then $\overline{A}$ is connected and $A \subset \overline{A}$. By maximality of A, we obtain $A = \overline{A}$. Therefore A is closed.

Let $x \in X$ and let A be the union of all connected subsets of X containing x. Then A is connected (by the previous result) and it is not contained in any larger connected subset. This means that A is a connected component. If two connected components A, B have a nonempty intersection then $A \cup B$ is connected. By the maximality of A, B this implies that $A = A \cup B = B$. As every point of X is contained in some connected component, we obtain that X is a disjoint union of its connected components. □

Example 2.34. Let $U \subset \mathbb{R}$ be an open set. We can represent U as a disjoint union $\cap_{i\in I} U_i$ of its connected components. If $V \subset \mathbb{R}$ is connected and $a = \inf V$, $b = \sup V$, then $(a, b) \subset V$ (otherwise we could find a point $a < c < b$, $c \notin V$ and decompose $V = (V \cap (-\infty, c)) \cup (V \cap (c, +\infty)))$. This implies that V is an interval. If $V \subset U$ is a connected component, then $V = (a, b)$ (if $a \in V$, then $(a - \varepsilon, a + \varepsilon) \subset V$ for some $\varepsilon > 0$ and therefore $(a - \varepsilon, b) \subset V$, a contradiction; similarly for b). This means that any open set $U \subset \mathbb{R}$ is a disjoint union of open intervals.

2.5. **Compact topological spaces.**

Definition 2.35. 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 its open cover $(U_i)_{i\in I}$ 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 2.36. *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 2.37. The following topological spaces are compact.

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

Example 2.38. The following topological spaces are not compact.

(1) The interval $(0,1)$. One can consider the cover $(U_n)_{n\ge2}$, where $U_n=\left(\frac1n,1\right)$. It does not have a finite cubcover.
(2) Euclidean space $\mathbb{R}$. One can consider the cover $(U_n)_{n\ge1}$, where $U_n=(-n,n)$. It does not have a finite subcover.

Theorem 2.39 (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
$$x_0\le x_1\le x_2\le\dots$$
is an increasing sequence in $[a,b]$. Therefore it is convergent to some $x\in[a,b]$. For any $n\ge1$ and $k\ge n$, we have $x_k\in I_k\subset I_n$. Therefore
$$x=\lim_{k\to\infty}x_k\in I_n$$
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. □

Lemma 2.40. *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, $U_i \subset X$ are open and $A \subset \bigcup_{i\in I} U_i$. The set $U = X\backslash A$ is open, so there is an open cover of X

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

By the compactness of X, there exists a finite subset $I' \subset I$ such that

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

Then $A \subset \bigcup_{i\in I'} U_i$. This implies that A is compact. □

Theorem 2.41. *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. Therefore there exists a finite subset $I_x \subset I$ such that $x \in U_i$ for $i \in I_x$ and $(V_i)_{i\in I_x}$ covers Y. Let $U_x = \bigcap_{i\in I_x} U_i \ni x$. The family $(U_x)_{x\in X}$ covers X and we can find a finite subcover $(U_{x_1}, \dots, U_{x_n})$. Then

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

is a finite cover of $X \times Y$. Indeed, let $(x, y) \in X \times Y$ and $x \in U_{x_j}$. There exists $i \in I_{x_j}$ such that $y \in V_i$. Then

$$(x, y) \in U_{x_j} \times V_i \subset U_i \times V_i.$$

This means that we found a finite subcover of $X \times Y$ and therefore $X \times Y$ is compact. □

Proposition 2.42. *Let $f : X \to Y$ 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 $f(X) \subset \cup_{i\in I} U_i$. 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 that

$$f(X) \subset \bigcup_{i\in I'} U_i$$

i.e. we found a finite subcover of $f(X)$. This means that $f(X)$ is compact. □

Proposition 2.43. *Let (X, τ) be a Hausdorff topological space and let $A \subset X$ be a compact subspace. Then A is closed in X.*

Proof. Let $x \in X\backslash A$. For any point $y \in A$ there exist nbds $U_y \ni x$, $V_y \ni y$ such that $U_y \cap V_y = \varnothing$. 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

$$A \subset V_{y_1} \cup \dots \cup V_{y_n}.$$

Let

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

be a new nbd of x. Then

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

(otherwise $U \cap V_{y_i} \neq \varnothing$ for some i, but we know that $U_{y_i} \cap V_{y_i} = \varnothing$). This implies that $U \cap A = \varnothing$, so $U \subset X\backslash A$. We proved that for any $x \in X\backslash A$ there exists an open nbd $U \ni x$ such that $U \subset X\backslash A$. This implies that $X\backslash A$ is open. □

2.6. Compact metric spaces.

Lemma 2.44. *Let (X, d) be a metric space and $A \subset X$ be a compact subspace. Then A is closed and bounded (i.e. $\sup \{d(x, y) \mid x, y \in A\} < \infty$).*

Proof. As X is a metric space, it is Hausdorff and therefore any compact subspace $A \subset X$ is closed in X.

Let $x \in X$. Then the family of balls $(B(x, n))_{n \geq 1}$ is an open cover of A. By the compactness of A it contains a finite subcover. Let N be the maximal index in this subcover. Then $A \subset B(x, N)$. Therefore A is bounded. □

Theorem 2.45 (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. □

Definition 2.46. A metric space (X, d) is called sequentially compact if any sequence $(x_n)_{n \geq 1}$ in X has a convergent subsequence.

Definition 2.47. Let (X, d) be a metric space.

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

Theorem 2.48. *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. $1 \Rightarrow 2$. 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 $\forall N > 0$

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

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

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

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.

$2 \Rightarrow 3$. 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

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

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).

$3 \Rightarrow 1$. 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 $I_n \subset X$ such that

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

At least one of the sets

$$B(x, 1), \qquad x \in I_1$$

can not be covered by a finite number of U_i. Let it be $B(x_1, 1)$ with $x_1 \in I_1$. At least one of the sets

$$B(x_1, 1) \cap B(x, 1/2), \qquad x \in I_2$$

can not be covered by a finite number of U_i. Let it be $B(x_1, 1) \cap B(x_2, 1/2)$ with $x_2 \in I_2$. Continuing this process we find a sequence of points $(x_n)_{n\geq 1}$ such that

$$B(x_1, 1) \cap B(x_2, 1/2) \cap \cdots \cap B(x_n, 1/n), \qquad n \geq 1$$

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

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

As X is complete, the sequence (x_n) converges to some $x \in X$. Taking limits in the first inequality for $m \to \infty$, we get

$$d(x, x_n) \leq 1/n.$$

Let $i \in I$ be s.t. $x \in U_i$. Then there exists $n \geq 1$ such that $B(x, 2/n) \subset U_i$. Then

$$B(x_n, 1/n) \subset B(x, 2/n) \subset U_i$$

(if $d(x_n, y) < 1/n$ then $d(x, y) \leq d(x, x_n) + d(x_n, y) < 2/n$) and therefore

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

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

Theorem 2.49 (Bolzano-Weierstrass theorem)**.** *A bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.*

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

2.7. **Continuous functions on the compact spaces.**

Proposition 2.50. *Let (X,τ) be a compact topological space and let $f: X\to\mathbb{R}$ be a continuous function. Then there exist points $x_m, x_M\in X$ such that*

$$f(x_m)=\inf_{x\in X} f(x),\qquad f(x_M)=\sup_{x\in X} f(x).$$

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

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

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

Definition 2.51. Let (X,d), (Y,d') be two metric spaces. A map $f: X\to Y$ is called uniformly continuous if $\forall\varepsilon>0\ \exists\delta>0$:

$$d(x,y)<\delta\quad\Rightarrow\quad d'(f(x),f(y))<\varepsilon,\qquad x,y\in X.$$

Remark 2.52. *If $f: X\to Y$ is uniformly continuous then f is continuous. The converse is not true in general.*

Example 2.53. Consider a map $f:(0,+\infty)\to(0,+\infty)$, $x\mapsto 1/x$. It is continuous, but not uniformly continuous. Indeed, let $\varepsilon>0$ and assume $\exists\delta>0$ such that

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

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

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

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

Example 2.54. If $f: X\to Y$ is a Lipschitz continuous map between two metric spaces (that is $\exists L>0$ such that $d'(f(x),f(y))\le Ld(x,y)$ for any $x,y\in X$) then f is uniformly continuous.

Theorem 2.55 (Uniform continuity theorem)**.** *Let $(X,d),(Y,d')$ be two metric spaces and assume that X is compact. If $f: X\to Y$ is continuous then it is uniformly continuous.*

Proof. Assume the contrary. $\Rightarrow\exists\varepsilon>0\ \forall n\ge 1\ \exists x_n,y_n\in X$:

$$d(x_n,y_n)<1/n,\qquad d'(f(x_n),f(y_n))\ge\varepsilon.$$

As X is compact, there exists a subsequence $(x_{n_k})_{k\ge1}$ of $(x_n)_{n\ge1}$ 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 $f(x)$. $\Rightarrow\exists N>0\ \forall k\ge N:$

$$d'(f(x_{n_k}),f(x))<\varepsilon/2,\qquad d'(f(y_{n_k}),f(x))<\varepsilon/2.$$

$\Rightarrow d'(f(x_{n_k}),f(y_{n_k}))<\varepsilon$, a contradiction. □

Remark 2.56. *If $f: X\to Y$ is continuous and X is compact then f is not necessarily Lipschitz continuous. For example, consider the function $f(x)=\sqrt{x}$ on $[0,1]$. Then*

$$\frac{|f(x)-f(y)|}{|x-y|}=\frac{1}{\left|\sqrt{x}+\sqrt{y}\right|}\to\infty,\qquad x,y\to 0.$$

Theorem 2.57 (Arzela-Ascoli Theorem)**.** *Let (X,d) be a compact metric space and $C(X)$ be the space of continuous maps $f: X\to\mathbb{R}$ with a 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 $\forall\varepsilon>0\ \exists\delta>0$*

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

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

$$\text{addition:} \quad X \times X \to X, \quad (x, y) \mapsto x + y$$
$$\text{scalar multiplication:} \quad K \times X \to X, \quad (\lambda, x) \mapsto \lambda x$$

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

$$\|\cdot\| : X \to \mathbb{R}, \qquad x \mapsto \|x\|$$

satisfying the axioms

(1) $\|x\| = 0 \Leftrightarrow x = 0$.
(2) $\|\lambda x\| = |\lambda| \cdot \|x\|$ for any $\lambda \in \Bbbk$, $x \in V$.
(3) $\|x + y\| \leq \|x\| + \|y\|$ for any $x, y \in X$ (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*

$$d : X \times X \to X, \qquad d(x, y) = \|x - y\| , \quad x, y \in X$$

defines a metric on X.

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

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

$\square$

In this way any normed vector space can be considered as a metric space. Most of the examples of metric spaces that we considered 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 $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

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

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

Example 3.6. Let $C[a, b]$ be the set of all continuous functions $f : [a, b] \to \mathbb{R}$. 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

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

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

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

$C[a, b]$ 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 on $C[a, b]$

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

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 $x \in l_p$, define

$$\|x\|_p = \left(\sum_{n \geq 1} |x_n|^p \right)^{1/p}.$$

Define addition and scalar multiplication on l_p componentwise:

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

Let us show that if $x, y \in l_p$ then $x + y \in l_p$. 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 $m \to \infty$, we obtain

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

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

Similarly, we define l_∞ to be the set of all sequences $(x_n)_{n \geq 1}$ of real numbers such that

$$\sup_{n \geq 1} |x_n| < \infty.$$

We equip l_∞ with a norm

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

3.3. **Bounded linear operators.**

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

(1) A map $A : X \to Y$ is called a linear operator if
$$A(x+y) = Ax + Ay, \qquad A(\lambda x) = \lambda Ax$$
for any $x, y \in X$ and $\lambda \in K$.

(2) A linear operator $A : X \to Y$ is called bounded if there exists $M > 0$ such that
$$\|Ax\| \leq M \|x\| \qquad \forall x \in X.$$

(3) A linear operator $A : X \to Y$ 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 $A : X \to Y$ between two normed vector spaces is bounded if and only if it is continuous.*

Proof. $\Rightarrow$. Assume that $A : X \to Y$ is bounded with a constant $M > 0$. For any $\varepsilon > 0$ let $\delta = \varepsilon/M$. If $x, y \in X$ and $d(x, y) < \delta$, then
$$d'(Ax, Ay) = \|Ax - Ay\| = \|A(x-y)\| \leq M \|x - y\| = Md(x, y) < M\delta = \varepsilon.$$
Therefore $A : X \to Y$ is continuous.

$\Leftarrow$. Let $A : X \to Y$ 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
$$\|Ay_n\| = \frac{\|Ax_n\|}{n\|x_n\|} > \frac{n\|x_n\|}{n\|x_n\|} = 1.$$
This implies that $y_n \to 0$, while $Ay_n \not\to 0$ as $n \to \infty$. This contradicts to the continuity of A. □

Let $L(X, Y)$ denote the set of all bounded linear operators from X to Y and let $L(X) = L(X, X)$. The set $L(X, Y)$ is a vector space:

(1) Given $A, B \in L(X, Y)$, define $A + B \in L(X, Y)$ by
$$(A+B)x = Ax + Bx, \qquad x \in X.$$
If A is bounded by $M > 0$ and B is bounded by $N > 0$ then
$$\|(A+B)x\| = \|Ax + Bx\| \leq \|Ax\| + \|Bx\| \leq M\|x\| + M\|x\| = (M+N)\|x\|,$$
that is, $A + B$ is bounded by $M + N$.

(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.

Definition 3.10. Define the operator norm 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,
$$\|A\| = \inf \{M \geq 0 \mid \|Ax\| \leq M\|x\| \;\; \forall x \in X\}.$$

Lemma 3.11. *For any $A \in L(X, Y)$, we have*
$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|.$$

Proof. Let
$$a = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|.$$
Then $\frac{\|Ax\|}{\|x\|} \leq a$ for all $x \neq 0$ and therefore $\|Ax\| \leq a\|x\|$ for all $x \in X$. This implies that $\|A\| \leq a$.

Conversely, for any $\varepsilon > 0$ there exists $x \neq 0$ such that
$$a - \varepsilon < \frac{\|Ax\|}{\|x\|} \leq \frac{\|A\|\|x\|}{\|x\|} = \|A\|.$$
As $\varepsilon > 0$ is arbitrary, we obtain $a \leq \|A\|$. □

Example 3.12. Consider a linear map $A : (\mathbb{R}^n, d_1) \to (\mathbb{R}^n, d_\infty)$ given by a matrix $(a_{ij})_{i,j=1}^n$. Let us find its norm.

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

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

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

Therefore

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

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

Lemma 3.13. *For any $A, B \in L(X, Y)$ and $\lambda \in K$, we have*

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

Proof. 1. Assume that $\|A\| = 0$. Then $\forall x \in X$: $\|Ax\| \leq 0 \, \|x\| = 0 \Rightarrow \|Ax\| = 0 \Rightarrow Ax = 0$. 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 $\|A + B\| \leq \|A\| + \|B\|$. □

Remark 3.14. *The last result implies that $L(X, Y)$ equipped with an operator norm is itself a normed vector space.*

Lemma 3.15. *Let $A \in L(X, Y)$ and $B \in L(Y, Z)$. Then $\|BA\| \leq \|B\| \cdot \|A\|$.*

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

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

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

3.4. **Equivalent norms.**

Definition 3.16. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a vector space X are called equivalent if there exist $C, C' > 0$ such that

$$\|x\|_2 \le C \|x\|_1, \qquad \|x\|_1 \le C' \|x\|_2, \qquad \forall x \in X.$$

Theorem 3.17. *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

$$A : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2), \qquad x \mapsto x.$$

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

(1) If $\|x\|_2 \le C \|x\|_1 \Rightarrow \|Ax\|_2 \le C \|x\|_1 \Rightarrow \|A\| \le C$.
(2) If A is bounded, then $\|x\|_2 = \|Ax\|_2 \le \|A\| \|x\|_1$ 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. □

Example 3.18. Consider $X = C[0, 1]$ with norms

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

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

Theorem 3.19. *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

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

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

$$\|Ax\| = \|x\| = \|x_1e_1 + \cdots + x_ne_n\|_1 \le \sum |x_k| \cdot \max_k \|e_k\|_1 = C \|x\|_1,$$

where $C = \max_k \|e_k\|_1$. Therefore A is bounded and continuous. Define

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

The last minimum is attained as the set

$$\{x \in \mathbb{R}^n \mid \|x\|_1 = 1\}$$

is compact and the map $(\mathbb{R}^n, \|\cdot\|_1) \to \mathbb{R}$, $x \mapsto \|Ax\|$ is continuous by the previous argument. Therefore $a > 0$. We have $\|x\| \ge a \|x\|_1$ for any $x \in \mathbb{R}^n$. This implies

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

Therefore the norms are equivalent. □

Theorem 3.20. *Let $A : X \to Y$ 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

$$\|x\| = \sum_{i=1}^{n} |x_i|, \qquad x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Then

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

where $C = \max_i \|Ae_i\|$. This implies that A is bounded. □

3.5. **Banach spaces.**

Definition 3.21. A Banach space is a normed vector space $(X, \|\cdot\|)$ which is complete (with respect to the metric $d(x,y) = \|x-y\|$).

Example 3.22. 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 3.23. 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.

Theorem 3.24. *The space l_p is Banach.*

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

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

Then for any $k \geq 1$

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

Therefore the sequence $(x_k^{(n)})_{n\geq 1}$ in $\mathbb{R}$ is Cauchy and therefore convergent to some element x_k. Let $x = (x_k)_{k\geq 1}$.

For any $K > 0$ we have

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

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

therefore $x - x^{(n)} \in l_p$ and $x \in l_p$. This also implies convergence.

□

Theorem 3.25. *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.

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

Therefore

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

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 $A : X \to Y$ is linear. For any $x, y \in X$, we have

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

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

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

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

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

This implies that $A_n \to A$ as $n \to \infty$. □

Remark 3.26. *Given a normed vector space $(X, \|\cdot\|)$ over $\mathbb{R}$, define its dual space to be $X^* = L(X, \mathbb{R})$. Its elements are continuous linear functions $f : X \to \mathbb{R}$, called linear functionals. Note that X^* is a Banach space (as $\mathbb{R}$ is complete), although X can be non-banach.*

Definition 3.27. Let X be a normed vector space and $(x_n)_{n\geq 1}$ be a sequence in X.

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

Theorem 3.28. *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

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

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

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

If $n > m \geq N$, then

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

and therefore 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. □

Theorem 3.29. *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

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

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

Definition 3.30. Let X be a normed vector space. The operator $I \in L(X)$ defined by $I(x) = x$ for $x \in X$ is called an identity operator. 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 3.31. *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 it 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

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

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 $B = \lim_{n\to\infty} s_n = \sum_{k\geq 0} A^k$. Then

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

as $\|A^n\| \leq \|A\|^n \to 0$ as $n \to \infty$. 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. □

Corollary 3.32. *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 $\|B\| < \left\|A^{-1}\right\|^{-1}$. 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

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

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

Theorem 3.33 (Banach-Shauder 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 3.34. *Let X be a Banach space and let $A \in L(X)$ be bijective. Then the linear operator $A^{-1} : X \to X$ is bounded, that is, 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. □