
Michael Papadimitrakis

Metric spaces

Contents

1	Metric spaces.	1
1.1	Metrics, neighborhoods, open sets, closed sets.	1
1.2	Limits and continuity of functions.	5
1.3	Sequences.	8
1.4	Compactness.	10
1.5	Connectedness.	15

Chapter 1

Metric spaces.

1.1 Metrics, neighborhoods, open sets, closed sets.

Definition. Let X be a non-empty set. We call **metric** on X every function $d : X \times X \rightarrow \mathbb{R}$ with the following properties:

- (i) $d(x, y) \geq 0$ for every $x, y \in X$.
- (ii) For every $x, y \in X$: $d(x, y) = 0$ if and only if $x = y$.
- (iii) $d(x, y) = d(y, x)$ for every $x, y \in X$.
- (iv) $d(x, y) \leq d(x, z) + d(z, y)$ for every $x, y, z \in X$.

We say that the pair (X, d) is a **metric space** or that “the set X is equipped with the metric d ” or we just say “the set X with the metric d ”. The value of $d(x, y)$ is called **distance** between x, y .

A metric space consists of two things: a non-empty set X and a metric $d : X \times X \rightarrow \mathbb{R}$ which measures distances between the elements of X . When we have a non-empty set X we may talk about the *metric space* X only when there is a preassigned specific metric d on the set X .

Example 1.1.1. The cartesian product $\mathbb{R}^d = \mathbb{R} \times \cdots \times \mathbb{R}$ with $d \geq 2$ factors is the set of all ordered d -tuples (x_1, \dots, x_d) of real numbers. Using orthogonal axes, we identify \mathbb{R}^2 with a plane and \mathbb{R}^3 with the space. If $d = 1$, we consider $\mathbb{R}^1 = \mathbb{R}$ and we identify \mathbb{R}^1 with a line.

If we denote $\|x\| = (x_1^2 + \cdots + x_d^2)^{1/2}$ the **euclidean norm** of $x = (x_1, \dots, x_d)$, then the **euclidean distance** between $x, y \in \mathbb{R}^d$ is $\|x - y\| = ((x_1 - y_1)^2 + \cdots + (x_d - y_d)^2)^{1/2}$.

It is well known that the function $d : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, defined by $d(x, y) = \|x - y\|$, satisfies all properties of a metric and it is called **euclidean metric** on \mathbb{R}^d .

In everything that follows we shall consider \mathbb{R}^d equipped with the euclidean metric. In case we want to use a different metric on \mathbb{R}^d we shall state this explicitly and give a description of the specific metric to be used.

Definition. Let (X, d) be a metric space. If $x \in X$, $r > 0$, we call **r -neighborhood** of x or **neighborhood** with **center** x and **radius** r the set

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

It is obvious that every r -neighborhood contains at least its center.

Example 1.1.2. In \mathbb{R}^2 (with the euclidean metric) $N_x(r)$ is usually denoted $D_x(r)$ and it is the *open disc* with center x and radius r : $D_x(r) = \{y \mid \|y - x\| < r\}$. The corresponding *closed disc* is $\overline{D}_x(r) = \{y \mid \|y - x\| \leq r\}$ and the corresponding *circle* is $C_x(r) = \{y \mid \|y - x\| = r\}$.

In particular, the open disc, the closed disc and the circle with center 0 and radius 1 are denoted \mathbb{D} , $\overline{\mathbb{D}}$ and \mathbb{T} , respectively.

Example 1.1.3. In \mathbb{R}^d (with the euclidean metric) $N_x(r)$ is usually denoted $B_x(r)$ and it is the d -dimensional open ball with center x and radius r : $B_x(r) = \{y \mid \|y - x\| < r\}$. The corresponding d -dimensional closed ball is $\overline{B}_x(r) = \{y \mid \|y - x\| \leq r\}$ and the corresponding $(d-1)$ -dimensional sphere is $S_x(r) = \{y \mid \|y - x\| = r\}$.

The closed ball with center 0 and radius 1 is usually denoted \mathbb{B}^d and the sphere with center 0 and radius 1 is usually denoted \mathbb{S}^{d-1} .

Thus, $\mathbb{B}^1 = [-1, 1]$ and $\mathbb{S}^0 = \{-1, 1\}$. Also, $\mathbb{B}^2 = \overline{\mathbb{D}}$ and $\mathbb{S}^1 = \mathbb{T}$.

Proposition 1.1. Let (X, d) be a metric space and $x, y \in X$, $x \neq y$. Then there is $r > 0$ so that $N_x(r) \cap N_y(r) = \emptyset$.

Proof. Take $r = \frac{1}{2} d(x, y) > 0$. If $z \in N_x(r) \cap N_y(r)$, i.e. $d(z, x) < r$ and $d(z, y) < r$, then

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

and we arrive at a contradiction. Therefore $N_x(r) \cap N_y(r) = \emptyset$. \square

Definition. Let (X, d) be a metric space, $A \subseteq X$ and $x \in X$.

We say that x is an **interior point** of A if some neighborhood of x is contained in A .

We say that x is a **boundary point** of A if every neighborhood of x intersects both A and A^c .

We say that x is a **limit point** of A if every neighborhood of x intersects A .

We say that x is an **accumulation point** of A if every neighborhood of x intersects A at a point different from x .

We also define

$$A^\circ = \{x \in X \mid x \text{ is an interior point of } A\},$$

$$\partial A = \{x \in X \mid x \text{ is a boundary point of } A\},$$

$$\overline{A} = \{x \in X \mid x \text{ is a limit point of } A\}.$$

The sets A° , ∂A and \overline{A} are called **interior**, **boundary** and **closure** of A , respectively.

If $A \subseteq X$, the complement of A with respect to X is denoted A^c .

Proposition 1.2. Let (X, d) be a metric space and $A \subseteq X$. Then

$$(i) \partial A = \partial(A^c).$$

$$(ii) A^\circ \subseteq A \subseteq \overline{A}.$$

$$(iii) \overline{A} \setminus A^\circ = \partial A.$$

$$(iv) A^\circ = A \setminus \partial A.$$

$$(v) \overline{A} = A \cup \partial A.$$

Proof. (i) From the definition of a boundary point it is clear that the boundary points of A are the same as the boundary points of A^c . In other words, the sets ∂A and $\partial(A^c)$ have the same elements.

(ii) If $x \in A^\circ$, then there is a neighborhood of x which is contained in A and hence $x \in A$. Also, if $x \in A$, then every neighborhood of x intersects A and hence $x \in \overline{A}$.

(iii) Let $x \in \overline{A} \setminus A^\circ$. Since $x \in \overline{A}$, every neighborhood of x intersects A . Since $x \notin A^\circ$, there is no neighborhood of x which is contained in A and hence every neighborhood of x intersects A^c . Therefore, $x \in \partial A$. Conversely, let $x \in \partial A$. Then every neighborhood of x intersects A and hence $x \in \overline{A}$. Also every neighborhood of x intersects A^c which means that there is no neighborhood of x which is contained in A and hence $x \notin A^\circ$. Thus $x \in \overline{A} \setminus A^\circ$.

(iv) and (v) are straightforward corollaries of (ii) and (iii). \square

Example 1.1.4. We consider \mathbb{R}^2 and a relatively simple curve Γ which divides the plane in three subsets: the set A_1 of the points on one side of Γ , the set A_2 of points on the other side of Γ and the set of points of Γ . For instance Γ can be a circle or an ellipse or a line or a closed polygonal line (the circumference of a rectangle, for instance). Just looking at these shapes on the plane, we understand that $A_1^\circ = A_1$, $\partial A_1 = \Gamma$ and $\overline{A_1} = A_1 \cup \Gamma$. We have analogous results for A_2 and also $\Gamma^\circ = \emptyset$, $\partial \Gamma = \Gamma$ and $\overline{\Gamma} = \Gamma$.

Example 1.1.5. Let Γ be a relatively simple surface in \mathbb{R}^3 which divides the space in the set A_1 of the points on one side of Γ , the set A_2 of points on the other side of Γ and the set of points of Γ . For instance Γ can be a plane or a spherical surface or the surface of a parallelopiped. Then, as in the last example, $A_1^\circ = A_1$, $\partial A_1 = \Gamma$ and $\overline{A_1} = A_1 \cup \Gamma$. There are similar results for A_2 and also $\Gamma^\circ = \emptyset$, $\partial \Gamma = \Gamma$ and $\overline{\Gamma} = \Gamma$.

Definition. Let (X, d) be a metric space and $A \subseteq X$.

We say that A is **open** if it consists only of its interior points.

We say that A is **closed** if it contains all its limit points.

In other words, A is open if and only if $A = A^\circ$, and A is closed if and only if $A = \overline{A}$. It is clear from proposition 1.2 that a set is open if and only if it contains none of its boundary points and that a set is closed if and only if it contains all its boundary points.

Example 1.1.6. In examples 1.1.4 and 1.1.5 the sets A_1, A_2 are open and the sets $A_1 \cup \Gamma, A_2 \cup \Gamma$ and Γ are closed.

Proposition 1.3. Let (X, d) be a metric space. Every r -neighborhood is open.

Proof. Let $x \in X, r > 0$. We take any $y \in N_x(r)$ and we shall prove that there is $s > 0$ so that $N_y(s) \subseteq N_x(r)$, i.e. that y is an interior point of $N_x(r)$. This will imply that $N_x(r)$ is open. We have $d(y, x) < r$ and we take $s = r - d(y, x) > 0$. If $w \in N_y(s)$, then $d(w, x) \leq d(w, y) + d(y, x) < s + d(y, x) = r$ and thus $w \in N_x(r)$. Therefore $N_y(s) \subseteq N_x(r)$. \square

Proposition 1.4. Let (X, d) be a metric space and $A \subseteq X$. Then A is closed if and only if A^c is open.

Proof. Since A and A^c have the same boundary points, we have the following successive equivalent statements: $[A \text{ is closed}] \Leftrightarrow [A \text{ contains all boundary points of } A] \Leftrightarrow [A \text{ contains all boundary points of } A^c] \Leftrightarrow [A^c \text{ contains no boundary point of } A^c] \Leftrightarrow [A^c \text{ is open}]$. \square

The complement of the complement of a set is the set itself and hence: A is open if and only if A^c is closed.

Proposition 1.5. Let (X, d) be a metric space and $A \subseteq X$. Then A° is the largest open set contained in A and \overline{A} is the smallest closed set containing A .

Proof. (i) Let $x \in A^\circ$. Then there is $r > 0$ so that $N_x(r) \subseteq A$. We take any $y \in N_x(r)$. Since $N_x(r)$ is open, there is some $s > 0$ so that $N_y(s) \subseteq N_x(r)$ and hence $N_y(s) \subseteq A$. Therefore $y \in A^\circ$. We see that $N_x(r) \subseteq A^\circ$ and so x is an interior point of A° . Thus, every point of A° is an interior point of A° and hence A° is an open set contained in A .

Now let B be an open set contained in A . If $x \in B$, then there is $r > 0$ so that $N_x(r) \subseteq B \subseteq A$ and hence $x \in A^\circ$. Therefore $B \subseteq A^\circ$.

(ii) Let x be a limit point of \overline{A} . We take any $r > 0$ and then $N_x(r)$ intersects \overline{A} . Let $y \in N_x(r) \cap \overline{A}$. Since $N_x(r)$ is open, there is some $s > 0$ so that $N_y(s) \subseteq N_x(r)$. Since $y \in \overline{A}$, $N_y(s)$ intersects A and hence $N_x(r)$ also intersects A . Therefore, every $N_x(r)$ intersects A and hence $x \in \overline{A}$. We see that every limit point of \overline{A} belongs to \overline{A} and thus \overline{A} is a closed set containing A .

Finally, let B be a closed set containing A . If $x \in \overline{A}$, then every $N_x(r)$ intersects A and hence intersects B . Therefore $x \in \overline{B}$ and, since B is closed, $x \in B$. Thus $\overline{A} \subseteq B$. \square

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

- (i) The union of any open subsets of X is open.
- (ii) The intersection of finitely many open subsets of X is open.
- (iii) The intersection of any closed subsets of X is closed.
- (iv) The union of finitely many closed subsets of X is closed.

Proof. (i) If x belongs to the union U of certain open sets, then x belongs to one of these sets, say A . Since A is open, there is $r > 0$ so that $N_x(r) \subseteq A \subseteq U$. Therefore every point of U is an interior point of U and then U is open.

(ii) Let $F = A_1 \cap \dots \cap A_n$, where A_k is open for every k . If $x \in F$, then $x \in A_k$ for every k . Thus, there are $r_1, \dots, r_n > 0$ so that $N_x(r_k) \subseteq A_k$ for every k . We take $r = \min\{r_1, \dots, r_n\} > 0$. Then $N_x(r) \subseteq N_x(r_k) \subseteq A_k$ for every k and hence $N_x(r) \subseteq F$. Therefore every point of F is an interior point of F and then F is open.

(iii) and (iv) are immediate consequences of (i) and (ii), of proposition 1.4 and of the laws of de Morgan: $(\bigcap A)^c = \bigcup A^c$ and $(\bigcup A)^c = \bigcap A^c$. \square

Definition. Let X be a non-empty set and d_1, d_2 be metrics on X . We say that the two metrics are **equivalent** if the metric spaces (X, d_1) and (X, d_2) have the same open sets: every A which is open in (X, d_1) is also open in (X, d_2) and conversely.

Proposition 1.4 says that the closed sets in any metric space are the complements of the open sets. Therefore, the metrics d_1, d_2 on X are equivalent if and only if the metric spaces (X, d_1) and (X, d_2) have the same closed sets.

Proposition 1.7. Let X be non-empty and d_1, d_2 be metrics on X . We denote $N_x^{d_1}(r)$ and $N_x^{d_2}(r)$ the neighborhoods of x in the metric spaces (X, d_1) and (X, d_2) , respectively. The following are equivalent.

(i) d_1, d_2 are equivalent.

(ii) For every $x \in X$ and every $\epsilon > 0$ there is $\delta > 0$ so that $N_x^{d_1}(\delta) \subseteq N_x^{d_2}(\epsilon)$ and, conversely, for every $x \in X$ and every $\epsilon > 0$ there is $\delta > 0$ so that $N_x^{d_2}(\delta) \subseteq N_x^{d_1}(\epsilon)$.

Proof. (i) \Rightarrow (ii) Let $x \in X$ and $\epsilon > 0$. The neighborhood $N_x^{d_2}(\epsilon)$ is open in the metric space (X, d_2) . Since (X, d_1) and (X, d_2) have the same open sets, $N_x^{d_2}(\epsilon)$ is also open in (X, d_1) . Because $x \in N_x^{d_2}(\epsilon)$, there is $\delta > 0$ so that $N_x^{d_1}(\delta) \subseteq N_x^{d_2}(\epsilon)$. The converse is similar.

(ii) \Rightarrow (i) Let A be open in (X, d_1) . We shall prove that A is also open in (X, d_2) .

We take any $x \in A$. Since A is open in (X, d_1) , there is $\epsilon > 0$ so that $N_x^{d_1}(\epsilon) \subseteq A$. Then there is $\delta > 0$ so that $N_x^{d_2}(\delta) \subseteq N_x^{d_1}(\epsilon)$ and thus $N_x^{d_2}(\delta) \subseteq A$. Therefore every element of A is an interior point of A in (X, d_2) and hence A is open in (X, d_2) . The converse is similar. \square

Exercises.

1.1.1. (i) We define three functions $d : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ by $d(x, y) = (x - y)^2$, $d(x, y) = |x - y|^{1/2}$ and $d(x, y) = \frac{|x-y|}{1+|x-y|}$. Which of these d is a metric on \mathbb{R} ?

(ii) For every $x = (x_1, x_2), y = (y_1, y_2)$ in \mathbb{R}^2 we set $d(x, y) = ((x_1 - y_1)^2 + 4(x_2 - y_2)^2)^{1/2}$. Is d a metric on \mathbb{R}^2 ?

(iii) Let $d(x, y) = |x_1 - y_1|$ for every $x = (x_1, x_2, x_3), y = (y_1, y_2, y_3)$ in \mathbb{R}^3 . Is d a metric on \mathbb{R}^3 ?

1.1.2. Which of the following are open or closed subsets of \mathbb{R} ?

$\mathbb{N}, \mathbb{Q}, \{1/n \mid n \in \mathbb{N}\}, \{0\} \cup \{1/n \mid n \in \mathbb{N}\}, [0, 1] \cup \{1 + 1/n \mid n \in \mathbb{N}\}.$

Find their interiors, their closures and their boundaries.

1.1.3. Which of the following are open or closed subsets of \mathbb{R}^2 ?

$\{(x_1, x_2) \mid x_1 > 0\}, \{(x_1, 0) \mid a \leq x_1 \leq b\}, \{(x_1, 0) \mid a < x_1 < b\}, \{(x_1, x_2) \mid x_1 x_2 \leq 1\}, \{(x_1, x_2) \mid x_1 x_2 > 1\}, \{(1/n, 0) \mid n \in \mathbb{N}\}, [0, 1] \times (\{0\} \cup \{1/n \mid n \in \mathbb{N}\}).$

Find their interiors, their closures and their boundaries.

1.1.4. Which of the following are open or closed subsets of \mathbb{R}^3 ?

$\{(x_1, x_2, x_3) \mid x_1 > 0\}, \{(x_1, 0, 0) \mid a < x_1 < b\}, \{(x_1, 0, 0) \mid a \leq x_1 \leq b\}, \{(x_1, x_2, 0) \mid a \leq x_1 \leq b, c \leq x_2 \leq d\}, \{(x_1, x_2, x_3) \mid x_1 + x_2 + x_3 > 1\}, \{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 < x_3\}.$

Find their interiors, their closures and their boundaries.

1.1.5. Let $x \cdot y = x_1y_1 + \cdots + x_dy_d$ be the usual euclidean inner product in \mathbb{R}^d . Let $a \in \mathbb{R}^d$, $a \neq 0$ and $a \in \mathbb{R}$. The set $\Gamma = \{x \in \mathbb{R}^d \mid a \cdot x = a\}$ is a hyperplane of \mathbb{R}^d . The open halfspaces of \mathbb{R}^d determined by Γ are $A_1 = \{x \in \mathbb{R}^d \mid a \cdot x > a\}$ and $A_2 = \{x \in \mathbb{R}^d \mid a \cdot x < a\}$ and the corresponding closed halfspaces are $B_1 = \{x \in \mathbb{R}^d \mid a \cdot x \geq a\}$ and $B_2 = \{x \in \mathbb{R}^d \mid a \cdot x \leq a\}$. Find the interiors, the closures and the boundaries of Γ , A_1 , A_2 , B_1 and B_2 .

1.1.6. In \mathbb{R}^d , the general open or closed orthogonal parallelepiped with edges parallel to the coordinate axes is $(a_1, b_1) \times \cdots \times (a_d, b_d)$ or $[a_1, b_1] \times \cdots \times [a_d, b_d]$, respectively. Prove that the first set is open and the second is closed.

1.1.7. Let (X, d) be a metric space.

- (i) Prove that both X and \emptyset are open and closed subsets of X .
- (ii) If $A \subseteq X$, prove that ∂A is closed.
- (iii) Prove that every finite subset of X is closed.
- (iv) If $A \subseteq B \subseteq X$, prove that $A^\circ \subseteq B^\circ$ and $\overline{A} \subseteq \overline{B}$.
- (v) If $A \subseteq X$ is open and $B \subseteq X$ is closed, prove that $A \setminus B$ is open and $B \setminus A$ is closed.

1.1.8. Let X be any non-empty set and $d : X \times X \rightarrow \mathbb{R}$ be the function defined by $d(x, x) = 1$ for every $x \in X$ and by $d(x, y) = 0$ for every $x, y \in X$ with $x \neq y$.

- (i) Prove that d is a metric on X . This metric is called **discrete metric**.
- (ii) Prove that every $A \subseteq X$ (with the discrete metric) is open and closed. Prove that $A^\circ = \overline{A} = A$ and $\partial A = \emptyset$ for every $A \subseteq X$.

1.1.9. Let (X, d) be a metric space. We define $d' : X \times X \rightarrow \mathbb{R}$ by $d'(x, y) = \frac{d(x, y)}{d(x, y) + 1}$. Prove that d' is a metric on X and that d, d' are equivalent.

1.1.10. Let X be non-empty, d_1, d_2 be equivalent metrics on X and $A \subseteq X$. Prove that in both metric spaces, (X, d_1) and (X, d_2) , A has the same interior points, the same boundary points and the same limit points.

1.2 Limits and continuity of functions.

Definition. Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$, $f : A \rightarrow Y$, $x_0 \in X$ be an accumulation point of A and $y_0 \in Y$. We say that y_0 is a **limit** of f at x_0 , and denote

$$y_0 = \lim_{x \rightarrow x_0} f(x),$$

if for every $\epsilon > 0$ there is $\delta > 0$ so that $f(x) \in N_{y_0}(\epsilon)$ for every $x \in N_{x_0}(\delta) \cap A$, $x \neq x_0$ or, equivalently, if for every $\epsilon > 0$ there is $\delta > 0$ so that $\rho(f(x), y_0) < \epsilon$ for every $x \in A$ with $0 < d(x, x_0) < \delta$.

This definition is the direct generalization of the well known definition of the limit of a function in case both metric spaces (X, d) and (Y, ρ) are the euclidean space \mathbb{R} .

Proposition 1.8. Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$, $f : A \rightarrow Y$ and $x_0 \in X$ be an accumulation point of A . If f has a limit at x_0 , then this limit is unique.

Proof. Let $y'_0 = \lim_{x \rightarrow x_0} f(x)$ and $y''_0 = \lim_{x \rightarrow x_0} f(x)$, where $y'_0, y''_0 \in Y$. We assume $y'_0 \neq y''_0$ and then proposition 1.1 implies that there is $\epsilon > 0$ so that $N_{y'_0}(\epsilon) \cap N_{y''_0}(\epsilon) = \emptyset$. Then there is $\delta > 0$ so that $f(x) \in N_{y'_0}(\epsilon)$ and $f(x) \in N_{y''_0}(\epsilon)$ for every $x \in N_{x_0}(\delta) \cap A$, $x \neq x_0$, and we arrive at a contradiction. \square

Proposition 1.8 allows us to talk about **the** limit of a function at a point.

Definition. Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$, $f : A \rightarrow Y$ and $x_0 \in A$. We say that f is **continuous** at x_0 if for every $\epsilon > 0$ there is $\delta > 0$ so that $f(x) \in N_{f(x_0)}(\epsilon)$ for every $x \in N_{x_0}(\delta) \cap A$ or, equivalently, if for every $\epsilon > 0$ there is $\delta > 0$ so that $\rho(f(x), f(x_0)) < \epsilon$ for every $x \in A$ with $d(x, x_0) < \delta$.

If $x_0 \in A$ is not an accumulation point of A , i.e. if it is an **isolated point** of A , then we may easily see that f is automatically continuous at x_0 . On the other hand, if $x_0 \in A$ is an accumulation point of A , then f is continuous at x_0 if and only if $\lim_{x \rightarrow x_0} f(x) = f(x_0)$.

Definition. Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$ and $f : A \rightarrow Y$. We say that f is **continuous** in A if it is continuous at every point of A .

Proposition 1.9. Let (X, d) , (Y, ρ) and (Z, τ) be metric spaces, $A \subseteq X$, $B \subseteq Y$, $x_0 \in A$, $f : A \rightarrow B$ and $g : B \rightarrow Z$. If f is continuous at x_0 and g is continuous at $y_0 = f(x_0)$, then $g \circ f : A \rightarrow Z$ is continuous at x_0 .

Proof. We take $\epsilon > 0$ and then there is $\delta' > 0$ so that

$$\tau(g(y), g(y_0)) < \epsilon \quad \text{for every } y \in B \text{ with } \rho(y, y_0) < \delta'. \quad (1.1)$$

Then there is $\delta > 0$ so that

$$\rho(f(x), y_0) = \rho(f(x), f(x_0)) < \delta' \quad \text{for every } x \in A \text{ with } d(x, x_0) < \delta. \quad (1.2)$$

From (1.2) and from (1.1) with $y = f(x)$ we get that for every $x \in A$ with $d(x, x_0) < \delta$ we have $\tau(g(f(x)), g(f(x_0))) < \epsilon$. Thus $g \circ f : A \rightarrow Z$ is continuous at x_0 . \square

Proposition 1.10. Let (X, d) be a metric space, $A \subseteq X$, $x_0 \in A$, $f, g : A \rightarrow \mathbb{R}$ be continuous at x_0 and $\lambda, \mu \in \mathbb{R}$. Then:

(i) $\lambda f + \mu g : A \rightarrow \mathbb{R}$ and $fg : A \rightarrow \mathbb{R}$ are continuous at x_0 .

(ii) If $B = \{x \in A \mid g(x) \neq 0\}$ and $g(x_0) \neq 0$, then $\frac{1}{g} : B \rightarrow \mathbb{R}$ is continuous at x_0 .

Proof. (i) We take any $\epsilon > 0$ and then there is $\delta > 0$ so that $|f(x) - f(x_0)| < \frac{\epsilon}{2(|\lambda|+1)}$ and $|g(x) - g(x_0)| < \frac{\epsilon}{2(|\mu|+1)}$ for every $x \in A$ with $d(x, x_0) < \delta$. This implies that

$$\begin{aligned} |(\lambda f(x) + \mu g(x)) - (\lambda f(x_0) + \mu g(x_0))| &\leq |\lambda||f(x) - f(x_0)| + |\mu||g(x) - g(x_0)| \\ &\leq |\lambda|\frac{\epsilon}{2(|\lambda|+1)} + |\mu|\frac{\epsilon}{2(|\mu|+1)} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

for every $x \in A$ with $d(x, x_0) < \delta$ and hence $\lambda f + \mu g : A \rightarrow \mathbb{R}$ is continuous at x_0 .

We then take any $\epsilon > 0$ and we set

$$\epsilon_1 = \min \left\{ \left(\frac{\epsilon}{3} \right)^{1/2}, \frac{\epsilon}{3(|f(x_0)|+1)}, \frac{\epsilon}{3(|g(x_0)|+1)} \right\} > 0.$$

Then there is $\delta > 0$ so that $|f(x) - f(x_0)| < \epsilon_1$ and $|g(x) - g(x_0)| < \epsilon_1$ for every $x \in A$ with $d(x, x_0) < \delta$. This implies that

$$\begin{aligned} |f(x)g(x) - f(x_0)g(x_0)| &\leq |f(x) - f(x_0)||g(x) - g(x_0)| + |f(x_0)||g(x) - g(x_0)| \\ &\quad + |g(x_0)||f(x) - f(x_0)| \\ &\leq \epsilon_1^2 + |f(x_0)|\epsilon_1 + |g(x_0)|\epsilon_1 < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \end{aligned}$$

for every $x \in A$ with $d(x, x_0) < \delta$ and hence $fg : A \rightarrow \mathbb{R}$ is continuous at x_0 .

(ii) We take any $\epsilon > 0$ and then there is $\delta > 0$ so that $|g(x) - g(x_0)| < \min\{\frac{|g(x_0)|}{2}, \frac{g(x_0)^2}{2}\epsilon\}$ for every $x \in A$ with $d(x, x_0) < \delta$. This implies that

$$|g(x)| = |g(x_0) + (g(x) - g(x_0))| \geq |g(x_0)| - |g(x) - g(x_0)| > |g(x_0)| - \frac{|g(x_0)|}{2} = \frac{|g(x_0)|}{2}$$

and hence

$$\left| \frac{1}{g(x)} - \frac{1}{g(x_0)} \right| = \frac{|g(x_0) - g(x)|}{|g(x)||g(x_0)|} \leq \frac{2|g(x) - g(x_0)|}{g(x_0)^2} < \epsilon$$

for every $x \in B$ with $d(x, x_0) < \delta$. Therefore $\frac{1}{g} : B \rightarrow \mathbb{R}$ is continuous at x_0 . \square

The proof of proposition 1.11 is almost identical to the previous proof.

Proposition 1.11. *Let (X, d) be a metric space, $A \subseteq X$, $f, g : A \rightarrow \mathbb{R}$, $x_0 \in X$ be an accumulation point of A , $\lim_{x \rightarrow x_0} f(x) = y_0 \in \mathbb{R}$, $\lim_{x \rightarrow x_0} g(x) = z_0 \in \mathbb{R}$ and $\lambda, \mu \in \mathbb{R}$. Then:*

- (i) $\lim_{x \rightarrow x_0} (\lambda f + \mu g) = \lambda y_0 + \mu z_0$ and $\lim_{x \rightarrow x_0} fg = y_0 z_0$.
- (ii) If $z_0 \neq 0$, then x_0 is an accumulation point of $B = \{x \in A \mid g(x) \neq 0\}$ and $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \frac{y_0}{z_0}$.

Combining propositions 1.9 and 1.10 and starting from very simple examples of continuous functions, we can produce more complicated ones.

Example 1.2.1. In \mathbb{R}^d we define the **k -projection** $\pi_k : \mathbb{R}^d \rightarrow \mathbb{R}$ by $\pi_k(x) = x_k$ for every $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. Every π_k is continuous, since $|\pi_k(x) - \pi_k(y)| \leq \|x - y\|$ for every $x, y \in \mathbb{R}^d$. Therefore, if $g : \mathbb{R} \rightarrow \mathbb{R}$ is continuous, then $f : \mathbb{R}^d \rightarrow \mathbb{R}$ defined by $f(x) = g(x_k)$ for every $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ is continuous, since $f = g \circ \pi_k$.

Thus **polynomial functions** $p(x_1, \dots, x_d) = Ax_1^{a_1} \dots x_d^{a_d} + Bx_1^{b_1} \dots x_d^{b_d} + \dots$ where all exponents are non-negative integers, all coefficients are real numbers and the sum is finite, are continuous functions. **Rational functions**, i.e. quotients of polynomial functions, are also continuous (except at the points where their denominator vanishes) as well as functions which are simple combinations of exponential or trigonometric or other simple continuous functions of the coordinates.

Proposition 1.12. *Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$ and $f : A \rightarrow Y$. Then the following are equivalent.*

- (i) f is continuous in A .
- (ii) For every open $W \subseteq Y$ there is an open $U \subseteq X$ so that $f^{-1}(W) = U \cap A$.
- (iii) For every closed $F \subseteq Y$ there is a closed $G \subseteq X$ so that $f^{-1}(F) = G \cap A$.

Proof. (i) \Rightarrow (ii) Let $x \in f^{-1}(W)$, i.e. $f(x) \in W$. Since W is open, there is $\epsilon_x > 0$ so that $N_{f(x)}(\epsilon_x) \subseteq W$. Since f is continuous, there is $\delta_x > 0$ so that $f(y) \in N_{f(x)}(\epsilon_x) \subseteq W$, and hence $y \in f^{-1}(W)$, for every $y \in N_x(\delta_x) \cap A$. Therefore, $N_x(\delta_x) \cap A \subseteq f^{-1}(W)$. Now we consider the set $U = \bigcup_{x \in f^{-1}(W)} N_x(\delta_x)$. Then U is a union of open sets and so it is open. We also have

$$U \cap A = \bigcup_{x \in f^{-1}(W)} (N_x(\delta_x) \cap A) \subseteq f^{-1}(W).$$

On the other hand it is clear that for every $x \in f^{-1}(W)$ we have $x \in N_x(\delta_x) \cap A$ and hence $x \in U \cap A$. Thus, $f^{-1}(W) \subseteq U \cap A$.

(ii) \Rightarrow (i) Take any $x_0 \in A$ and any $\epsilon > 0$. Then $N_{f(x_0)}(\epsilon)$ is open in Y and so there is an open $U \subseteq X$ so that $f^{-1}(N_{f(x_0)}(\epsilon)) = U \cap A$. Then $x_0 \in U \cap A$ and, since U is open, there is $\delta > 0$ so that $N_{x_0}(\delta) \subseteq U$. Now, for every $x \in N_{x_0}(\delta) \cap A$ we have $x \in U \cap A$ and hence $x \in f^{-1}(N_{f(x_0)}(\epsilon))$ i.e. $f(x) \in N_{f(x_0)}(\epsilon)$. Therefore, f is continuous at every $x_0 \in A$.

The equivalence (i) \Leftrightarrow (iii) is a consequence of the equivalence (i) \Leftrightarrow (ii) and of the general identity $f^{-1}(W^c) = (f^{-1}(W))^c \cap A$. \square

Definition. The metric spaces (X, d) and (Y, ρ) are called **homeomorphic** if there is $f : X \rightarrow Y$ which is one-to-one in X and onto Y and so that f is continuous in X and $f^{-1} : Y \rightarrow X$ is continuous in Y .

It is trivial to prove that the relation of homeomorphism between metric spaces is an equivalence relation. It is also trivial to see, based for instance on proposition 1.7, that, if d_1 and d_2 are two metrics on the non-empty set X , then the two metrics are equivalent if and only if the identity function between (X, d_1) and (X, d_2) is a homeomorphism.

Exercises.

1.2.1. Prove that $\{x \in \mathbb{R}^d \setminus \{0\} \mid e^{-\|x\|} + \sin \|x\| > 0\}$ is an open subset of \mathbb{R}^d .
Is $\{x \in \mathbb{R}^d \setminus \{0\} \mid \|x\| - \|x\|^3 \leq 3\}$ a closed subset of \mathbb{R}^d ?

1.2.2. Let X, Y be non-empty sets, $A \subseteq X$, $x_0 \in A$ and $f : A \rightarrow Y$. Let d_1, d_2 be equivalent metrics on X and ρ_1, ρ_2 be equivalent metrics on Y . Prove that f is continuous at x_0 with respect to d_1 and ρ_1 if and only if it is continuous at x_0 with respect to d_2 and ρ_2 .

1.2.3. Let (X, d) be a non-empty set with the discrete metric (exercise 1.1.8), (Y, ρ) be any metric space, $A \subseteq X$ and $f : A \rightarrow Y$. Prove that f is continuous in A .

1.3 Sequences.

The next definition is the generalization of the analogous definition in the euclidean space \mathbb{R} .

Definition. Let (X, d) be a metric space, $x \in X$ and let (x_n) be a sequence in X . We say that (x_n) **converges** to x in (X, d) or that x is a **limit** of (x_n) in (X, d) , and denote

$$x_n \rightarrow x \quad \text{or} \quad \lim_{n \rightarrow +\infty} x_n = x,$$

if for every $\epsilon > 0$ there is n_0 so that $x_n \in N_x(\epsilon)$ for every $n \geq n_0$ or, equivalently, if for every $\epsilon > 0$ there is n_0 so that $d(x_n, x) < \epsilon$ for every $n \geq n_0$.

It is clear that $x_n \rightarrow x$ in the metric space (X, d) if and only if $d(x_n, x) \rightarrow 0$ in \mathbb{R} .

Proposition 1.13. Let (X, d) be a metric space and let (x_n) be a sequence in X . If (x_n) has a limit, then this limit is unique.

Proof. Let $x_n \rightarrow x'$ and $x_n \rightarrow x''$ and assume that $x' \neq x''$. We know that there is $\epsilon > 0$ so that $N_{x'}(\epsilon) \cap N_{x''}(\epsilon) = \emptyset$. Then there is n_0 so that $x_n \in N_{x'}(\epsilon)$ and $x_n \in N_{x''}(\epsilon)$ for every $n \geq n_0$ and this is impossible. \square

Because of proposition 1.13, we can talk about **the** limit of a sequence.

The next proposition reduces convergence in the euclidean space \mathbb{R}^d to convergence in \mathbb{R} .

Proposition 1.14. Let $x_n = (x_{n,1}, \dots, x_{n,d}) \in \mathbb{R}^d$ for every n and $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. The following are equivalent.

- (i) $x_n \rightarrow x$ in \mathbb{R}^d .
- (ii) $x_{n,k} \rightarrow x_k$ in \mathbb{R} for every $k = 1, \dots, d$.

Proof. (i) \Rightarrow (ii) A consequence of $|x_{n,k} - x_k| \leq \|x_n - x\|$.

(ii) \Rightarrow (i) A consequence of $\|x_n - x\| \leq |x_{n,1} - x_1| + \dots + |x_{n,d} - x_d|$. \square

We shall now see the close relation between the notion of convergence of sequences and certain notions we have encountered already: the notion of limit point, the notion of closed set (and, indirectly, of open set) and, finally, the notions of the limit and continuity of a function.

Proposition 1.15. Let (X, d) be a metric space, $A \subseteq X$ and $x \in X$. Then x is a limit point of A if and only if there is a sequence (x_n) in A so that $x_n \rightarrow x$.

Proof. Let x be a limit point of A . We take any $n \in \mathbb{N}$ and then $N_x(\frac{1}{n})$ contains at least one point of A , i.e. there is $x_n \in A$ so that $d(x_n, x) < \frac{1}{n}$. Thus, the sequence (x_n) is in A and $x_n \rightarrow x$. Conversely, let (x_n) be a sequence in A so that $x_n \rightarrow x$. We take any $\epsilon > 0$ and then there is n_0 so that $x_n \in N_x(\epsilon)$ for every $n \geq n_0$. Thus $N_x(\epsilon)$ intersects A and so x is a limit point of A . \square

Example 1.3.1. Let us prove that the closure of the open ball $B_{x_0}(r)$ in \mathbb{R}^d is the corresponding closed ball $\overline{B}_{x_0}(r)$.

Assume that x is a limit point of $B_{x_0}(r)$. Then there is a sequence (x_n) in $B_{x_0}(r)$ so that $x_n \rightarrow x$, i.e. $\|x_n - x\| \rightarrow 0$. Then from $\|x_n - x_0\| - \|x - x_0\| \leq \|x_n - x\|$ we find $\|x_n - x_0\| \rightarrow \|x - x_0\|$. Since $\|x_n - x_0\| < r$ for every n , we get $\|x - x_0\| \leq r$ and so $x \in \overline{B}_{x_0}(r)$. Thus, $\overline{B_{x_0}(r)} \subseteq \overline{B}_{x_0}(r)$.

Conversely, take $x \in \overline{B_{x_0}}(r)$, i.e. $\|x - x_0\| \leq r$. For each $n \in \mathbb{N}$ we consider $x_n = \frac{1}{n}x_0 + (1 - \frac{1}{n})x$. Then $\|x_n - x_0\| = (1 - \frac{1}{n})\|x - x_0\| < r$ and hence $x_n \in B_{x_0}(r)$ for every n . Moreover, $x_n \rightarrow x$ and so $x \in \overline{B_{x_0}}(r)$. Therefore, $\overline{B_{x_0}}(r) \subseteq \overline{B_{x_0}}(r)$.

Proposition 1.16. *Let (X, d) be a metric space and $A \subseteq X$. The following are equivalent.*

- (i) A is closed.
- (ii) Every x , which is the limit of a sequence in A , belongs to A .

Proof. (i) \Rightarrow (ii) Take any x which is the limit of a sequence in A . Proposition 1.15 implies that x is a limit point of A and, since A is closed, $x \in A$.

(ii) \Rightarrow (i) Take any limit point x of A . Proposition 1.15 implies that there is a sequence in A with limit x and hence x belongs to A . Thus A contains all its limit points and so it is closed. \square

Propositions 1.17 and 1.18 are generalizations of analogous propositions for \mathbb{R} .

Proposition 1.17. *Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$, $x_0 \in A$ and $f : A \rightarrow Y$. The following are equivalent.*

- (i) f is continuous at x_0 .
- (ii) For every (x_n) in A with $x_n \rightarrow x_0$ we have $f(x_n) \rightarrow f(x_0)$.

Proof. (i) \Rightarrow (ii) Take (x_n) in A with $x_n \rightarrow x_0$. We take any $\epsilon > 0$ and then there is $\delta > 0$ so that

$$\rho(f(x), f(x_0)) < \epsilon \quad \text{for every } x \in A \text{ with } d(x, x_0) < \delta. \quad (1.3)$$

Then there is n_0 so that

$$d(x_n, x_0) < \delta \quad \text{for every } n \geq n_0. \quad (1.4)$$

Now (1.4) and (1.3) with $x = x_n$ imply that for every $n \geq n_0$ we have $\rho(f(x_n), f(x_0)) < \epsilon$. Therefore $f(x_n) \rightarrow f(x_0)$.

(ii) \Rightarrow (i) Assume that f is not continuous at x_0 . Then there is $\epsilon > 0$ so that for every $\delta > 0$ there is $x \in A$ such that $d(x, x_0) < \delta$ and $\rho(f(x), f(x_0)) \geq \epsilon$. Hence for every $n \in \mathbb{N}$ there is $x_n \in A$ with $d(x_n, x_0) < \frac{1}{n}$ and $\rho(f(x_n), f(x_0)) \geq \epsilon$. Then (x_n) is in A and $x_n \rightarrow x_0$ but $f(x_n) \not\rightarrow f(x_0)$ and we arrived at a contradiction. \square

The proof of proposition 1.18 is almost identical to the proof of proposition 1.17.

Proposition 1.18. *Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$, x_0 be an accumulation point of A , $y_0 \in Y$ and $f : A \rightarrow Y$. The following are equivalent.*

- (i) $\lim_{x \rightarrow x_0} f(x) = y_0$.
- (ii) For every (x_n) in $A \setminus \{x_0\}$ with $x_n \rightarrow x_0$ we have $f(x_n) \rightarrow y_0$.

Definition. Let (X, d) be a metric space and (x_n) be a sequence in X . We say that (x_n) is a **Cauchy sequence** if for every $\epsilon > 0$ there is n_0 so that $d(x_n, x_m) < \epsilon$ for every $n, m \geq n_0$.

Proposition 1.19. *Let (X, d) be a metric space and (x_n) be a sequence in X . If (x_n) converges to some element of X , then it is a Cauchy sequence.*

Proof. Let $x_n \rightarrow x$. If $\epsilon > 0$, then there is n_0 so that $d(x_n, x) < \frac{\epsilon}{2}$ for every $n \geq n_0$. Therefore, $d(x_n, x_m) \leq d(x_n, x) + d(x_m, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ for every $n, m \geq n_0$ and hence (x_n) is a Cauchy sequence. \square

Definition. Let (X, d) be a metric space and $A \subseteq X$. We say that A is **complete** if every Cauchy sequence in A converges to some element of A .

Proposition 1.20. *Let X be non-empty and let d_1, d_2 be metrics on X . The following are equivalent.*

- (i) The metrics d_1, d_2 are equivalent.
- (ii) The metric spaces (X, d_1) and (X, d_2) have the same convergent sequences.

Proof. (i) \Rightarrow (ii) Let $x_n \rightarrow x$ in (X, d_1) . We shall prove that $x_n \rightarrow x$ also in (X, d_2) . Let $\epsilon > 0$. Proposition 1.7 implies that there is $\delta > 0$ so that $N_x^{d_1}(\delta) \subseteq N_x^{d_2}(\epsilon)$. Since $x_n \rightarrow x$ in (X, d_1) , there is n_0 so that $x_n \in N_x^{d_1}(\delta)$, and hence $x_n \in N_x^{d_2}(\epsilon)$, for every $n \geq n_0$. Thus $x_n \rightarrow x$ in (X, d_2) .

The converse is similar.

(ii) \Rightarrow (i) Let $A \subseteq X$ be closed in (X, d_1) . We shall see that A is closed also in (X, d_2) .

We assume that (x_n) is in A and $x_n \rightarrow x$ in (X, d_2) . Then $x_n \rightarrow x$ also in (X, d_1) and, since A is closed in (X, d_1) , we get $x \in A$. Thus A is closed in (X, d_2) .

The converse is similar. □

Exercises.

1.3.1. Let $x_n \rightarrow x$ and $y_n \rightarrow y$ in (X, d) . Prove that $d(x_n, y_n) \rightarrow d(x, y)$ in \mathbb{R} .

1.3.2. Let (X, d) be a metric space, $A \subseteq X$ and $x \in X$. Prove that x is a boundary point of A if and only if there are sequences (x'_n) in A and (x''_n) in A^c so that $x'_n \rightarrow x$ and $x''_n \rightarrow x$.

1.3.3. We consider sequences (x_n) and (y_n) in \mathbb{R}^d and (λ_n) in \mathbb{R} . If $x_n \rightarrow x$, $y_n \rightarrow y$ in \mathbb{R}^d and $\lambda_n \rightarrow \lambda$ in \mathbb{R} , prove that $x_n + y_n \rightarrow x + y$ and $\lambda_n x_n \rightarrow \lambda x$ in \mathbb{R}^d .

1.3.4. Using sequences, prove that $\{\frac{1}{n} \mid n \in \mathbb{N}\}$ is not a closed subset of \mathbb{R} while $\{0\} \cup \{\frac{1}{n} \mid n \in \mathbb{N}\}$ is a closed subset of \mathbb{R} .

1.3.5. Using sequences, prove that closed balls, hyperplanes and closed halfspaces in \mathbb{R}^d are closed subsets of \mathbb{R}^d .

1.3.6. Let (X, d) be a non-empty set with the discrete metric (exercise 1.1.8). Prove that a sequence (x_n) in X converges if and only if it is constant after some value of its index n .

1.4 Compactness.

Definition. Let X be non-empty, $M \subseteq X$ and let Σ be a collection of subsets of X .

We say that Σ is a **covering** of M if $M \subseteq \bigcup_{A \in \Sigma} A$. If, moreover, Σ is finite, we say that it is a **finite covering** of M .

Let Σ and Σ' be coverings of M . If $\Sigma' \subseteq \Sigma$, then we say that Σ is **larger** than Σ' and that Σ' is **smaller** than Σ .

Let (X, d) be a metric space and $M \subseteq X$. If Σ is a covering of M and all $A \in \Sigma$ are open sets, then Σ is called **open covering** of M .

Definition. Let (X, d) be a metric space and $M \subseteq X$. We say that M is **compact** if for every open covering Σ of M there is a finite covering Σ' of M which is smaller than Σ .

Example 1.4.1. Let (X, d) be a metric space and $M = \{x_1, \dots, x_n\} \subseteq X$. We take any open covering Σ of M . Then every $x_k \in M$ belongs to some $A_k \in \Sigma$ and hence $M \subseteq A_1 \cup \dots \cup A_n$. Thus, $\Sigma' = \{A_1, \dots, A_n\}$ is a finite covering of M with $\Sigma' \subseteq \Sigma$. Hence M is compact.

Definition. Let (X, d) be a metric space and $M \subseteq X$. We say that M is **bounded** if there is $x_0 \in X$ and $r > 0$ so that $M \subseteq N_{x_0}(r)$.

Example 1.4.2. A set M in \mathbb{R}^d is bounded if and only if it is contained in some orthogonal parallelepiped with edges parallel to the coordinate axes.

Proposition 1.21. Let (X, d) be a metric space and $M \subseteq X$. If M is compact, then it is bounded and closed.

Proof. We take any $x_0 \in X$ and we consider the collection $\Sigma = \{N_{x_0}(n) \mid n \in \mathbb{N}\}$. Then Σ is an open covering of M , and so there is a covering Σ' of M which is smaller than Σ , i.e. there are n_1, \dots, n_N so that $M \subseteq N_{x_0}(n_1) \cup \dots \cup N_{x_0}(n_N)$. If $r = \max\{n_1, \dots, n_N\}$, then $M \subseteq N_{x_0}(r)$ and hence M is bounded.

Now we take any $x_0 \in M^c$. We consider the sets $A_n = \{x \in X \mid d(x, x_0) > \frac{1}{n}\}$ and the collection $\Sigma = \{A_n \mid n \in \mathbb{N}\}$. Then Σ is an open covering of M , and hence there is a finite covering Σ' of M which is smaller than Σ . I.e. there are n_1, \dots, n_N so that $M \subseteq A_{n_1} \cup \dots \cup A_{n_N}$. If $n = \max\{n_1, \dots, n_N\}$, then we have $M \subseteq A_n$ and hence $N_{x_0}(\frac{1}{n}) \subseteq M^c$. We proved that every $x_0 \in M^c$ is an interior point of M^c . Thus M^c is open and so M is closed. \square

Proposition 1.22. *Let (X, d) be a metric space and $N \subseteq M \subseteq X$. If M is compact and N is closed, then N is compact.*

Proof. We take any open covering Σ of N . Then $\Sigma_1 = \{N^c\} \cup \Sigma$ is an open covering of M . Since M is compact, there is a finite covering Σ'_1 of M which is smaller than Σ_1 . I.e. there are $A_1, \dots, A_n \in \Sigma_1$ so that $M \subseteq A_1 \cup \dots \cup A_n$.

If N^c is one of A_1, \dots, A_n , say $N^c = A_n$, then $M \subseteq A_1 \cup \dots \cup A_n$ becomes $M \subseteq A_1 \cup \dots \cup A_{n-1} \cup N^c$, and hence $N \subseteq A_1 \cup \dots \cup A_{n-1}$. Thus, $\Sigma' = \{A_1, \dots, A_{n-1}\}$ is a finite covering of N which is smaller than Σ . If N^c is not one of A_1, \dots, A_n , then $\Sigma' = \{A_1, \dots, A_n\}$ is a finite covering of N which is smaller than Σ .

In any case there is a finite covering of N which is smaller than Σ . \square

Proposition 1.23. *Let (X, d) be a metric space and $M, N \subseteq X$ so that $M \cap N = \emptyset$. If M is compact and N is closed, then there is $\epsilon > 0$ so that $d(x, y) \geq \epsilon$ for every $x \in M$ and $y \in N$.*

Proof. For every $x \in M$ we have $x \in N^c$ and, since N^c is open, there is $\epsilon_x > 0$ so that $N_x(\epsilon_x) \subseteq N^c$ and hence $N_x(\epsilon_x) \cap N = \emptyset$. This implies

$$d(x, y) \geq \epsilon_x \quad \text{for every } x \in M \text{ and } y \in N. \quad (1.5)$$

The collection $\{N_x(\frac{\epsilon_x}{2}) \mid x \in M\}$ is an open covering of M and, since M is compact, there are $x_1, \dots, x_n \in M$ so that $M \subseteq N_{x_1}(\frac{\epsilon_{x_1}}{2}) \cup \dots \cup N_{x_n}(\frac{\epsilon_{x_n}}{2})$. We set $\epsilon = \min\{\frac{\epsilon_{x_1}}{2}, \dots, \frac{\epsilon_{x_n}}{2}\} > 0$. If $x \in M$, there is $k = 1, \dots, n$ so that $x \in N_{x_k}(\frac{\epsilon_{x_k}}{2})$ and (1.5) implies that for every $y \in N$ we have

$$d(x, y) \geq d(y, x_k) - d(x, x_k) \geq \epsilon_{x_k} - \frac{\epsilon_{x_k}}{2} = \frac{\epsilon_{x_k}}{2} \geq \epsilon.$$

Therefore, $d(x, y) \geq \epsilon$ for every $x \in M$ and $y \in N$. \square

The next theorem is a generalization of the well known result for sequences of nested closed and bounded intervals in \mathbb{R} : if $[a_1, b_1] \supseteq [a_2, b_2] \supseteq \dots \supseteq [a_n, b_n] \supseteq \dots$, then there is x which belongs to every $[a_n, b_n]$ and if, moreover, $b_n - a_n \rightarrow 0$, then this x is unique.

Definition. Let (X, d) be a metric space and $M \subseteq X$. We define the **diameter** of M to be

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

Theorem 1.1. *Let (X, d) be a metric space and K_1, K_2, \dots be a sequence of non-empty compact subsets of X so that $K_{n+1} \subseteq K_n$ for every n . Then there is some element which belongs to all K_n . If, moreover, $\text{diam } K_n \rightarrow 0$, then the common element of all K_n is unique.*

Proof. We assume that $\bigcap_{n=1}^{+\infty} K_n = \emptyset$. Then the collection $\Sigma = \{K_n^c \mid n \in \mathbb{N}\}$ is an open covering of K_1 . Since K_1 is compact, there are n_1, \dots, n_N so that $K_1 \subseteq K_{n_1}^c \cup \dots \cup K_{n_N}^c$. We take $n = \max\{n_1, \dots, n_N\}$, and then $K_1 \subseteq K_n^c$. This is wrong, because $K_n \subseteq K_1$ and $K_n \neq \emptyset$.

Now, let $\text{diam } K_n \rightarrow 0$. If x, y belong to all K_n , then $0 \leq d(x, y) \leq \text{diam } K_n$ for every n and hence $d(x, y) = 0$. \square

The important theorem 1.2 describes the notion of compactness in terms of sequences.

Theorem 1.2. *Let (X, d) be a metric space and $M \subseteq X$. The following are equivalent.*

(i) *M is compact.*

(ii) *Every sequence in M has at least one subsequence which converges to an element of M .*

Proof. (i) \Rightarrow (ii) We take an arbitrary sequence (x_n) in M .

Assume that for every $x \in M$ there is a neighborhood $N_x(\epsilon_x)$ of x , which contains only finitely many terms of (x_n) . Then $\Sigma = \{N_x(\epsilon_x) \mid x \in M\}$ is an open covering of M and hence there are x_1, \dots, x_n so that $M \subseteq N_{x_1}(\epsilon_{x_1}) \cup \dots \cup N_{x_n}(\epsilon_{x_n})$. Each of these neighborhoods contains only finitely many terms of (x_n) . Therefore, M also contains only finitely many terms of (x_n) and we arrive at a contradiction.

Therefore there is $x_0 \in M$ so that for every $\epsilon > 0$ the neighborhood $N_{x_0}(\epsilon)$ contains infinitely many terms of (x_n) . Thus, there is $n_1 \geq 1$ so that $x_{n_1} \in N_{x_0}(1)$. Then there is $n_2 > n_1$ so that $x_{n_2} \in N_{x_0}(\frac{1}{2})$. We continue inductively and we find a subsequence (x_{n_k}) of (x_n) so that $x_{n_k} \in N_{x_0}(\frac{1}{k})$ or, equivalently, $d(x_{n_k}, x_0) < \frac{1}{k}$ for every k . Therefore $x_{n_k} \rightarrow x_0$.

(ii) \Rightarrow (i) *Step 1.* Let $\epsilon > 0$. Then there are $x_1, \dots, x_n \in M$ so that $M \subseteq N_{x_1}(\epsilon) \cup \dots \cup N_{x_n}(\epsilon)$. Assume that this is not true. We take any $x_1 \in M$. Then $M \not\subseteq N_{x_1}(\epsilon)$ and so there is $x_2 \in M$ with $x_2 \notin N_{x_1}(\epsilon)$. Then $M \not\subseteq N_{x_1}(\epsilon) \cup N_{x_2}(\epsilon)$ and so there is $x_3 \in M$ with $x_3 \notin N_{x_1}(\epsilon) \cup N_{x_2}(\epsilon)$. Then $M \not\subseteq N_{x_1}(\epsilon) \cup N_{x_2}(\epsilon) \cup N_{x_3}(\epsilon)$ and so there is $x_4 \in M$ with $x_4 \notin N_{x_1}(\epsilon) \cup N_{x_2}(\epsilon) \cup N_{x_3}(\epsilon)$. We continue inductively and we see that there is a sequence (x_n) in M so that $d(x_n, x_m) \geq \epsilon$ for every n, m with $n \neq m$. But this does not allow the existence of a convergent subsequence of (x_n) and we arrive at a contradiction.

Step 2. We take any open covering Σ of M . Then there is $\epsilon > 0$ so that for every $x \in M$ the neighborhood $N_x(\epsilon)$ is contained in some $A \in \Sigma$.

Assume that there is no $\epsilon > 0$ with this property. I.e. for every $\epsilon > 0$ there is $x \in M$ so that $N_x(\epsilon)$ is not contained in any $A \in \Sigma$. Thus, for every $n \in \mathbb{N}$ there is $x_n \in M$ so that $N_{x_n}(\frac{1}{n})$ is not contained in any $A \in \Sigma$. Now, there is a subsequence (x_{n_k}) of (x_n) so that $x_{n_k} \rightarrow x_0$ for some $x_0 \in M$. Then $x_0 \in A_0$ for some $A_0 \in \Sigma$. Since A_0 is open, there is $\delta > 0$ so that $N_{x_0}(\delta) \subseteq A_0$. We take n_k large enough so that $d(x_{n_k}, x_0) < \frac{\delta}{2}$ and $\frac{1}{n_k} < \frac{\delta}{2}$. Then for every $x \in N_{x_{n_k}}(\frac{1}{n_k})$ we have

$$d(x, x_0) \leq d(x, x_{n_k}) + d(x_{n_k}, x_0) < \frac{1}{n_k} + d(x_{n_k}, x_0) < \frac{\delta}{2} + \frac{\delta}{2} = \delta$$

and hence $N_{x_{n_k}}(\frac{1}{n_k}) \subseteq N_{x_0}(\delta) \subseteq A_0$. We arrive at a contradiction, because $N_{x_{n_k}}(\frac{1}{n_k})$ is not contained in any $A \in \Sigma$.

Step 3. We take an arbitrary open covering Σ of M . According to step 2, there is $\epsilon > 0$ so that for every $x \in M$ we have that $N_x(\epsilon)$ is contained in some $A \in \Sigma$. According to step 1, there are $x_1, \dots, x_n \in M$ so that $M \subseteq N_{x_1}(\epsilon) \cup \dots \cup N_{x_n}(\epsilon)$. Now let $N_{x_k}(\epsilon) \subseteq A_k \in \Sigma$ for each $k = 1, \dots, n$. Then $M \subseteq N_{x_1}(\epsilon) \cup \dots \cup N_{x_n}(\epsilon) \subseteq A_1 \cup \dots \cup A_n$ and hence $\Sigma' = \{A_1, \dots, A_n\}$ is a finite covering of M which is smaller than Σ . \square

Proposition 1.24. *Every closed orthogonal parallelepiped in \mathbb{R}^d with edges parallel to the coordinate axes is compact.*

Proof. Let $M = [a_1, b_1] \times \dots \times [a_d, b_d]$. We consider an arbitrary open covering Σ of M and we assume that there is no finite covering Σ' of M which is smaller than Σ .

We split every edge $[a_k, b_k]$ in the two subintervals $[a_k, \frac{a_k+b_k}{2}]$ and $[\frac{a_k+b_k}{2}, b_k]$. This induces a splitting of M in 2^d orthogonal parallelepipeds, each of which has dimensions equal to one half of the dimensions of M . We observe that for at least one of these parallelepipeds, call it M_1 , there is no finite covering which is smaller than Σ . Otherwise, for each of these parallelepipeds there would exist a finite covering which is smaller than Σ , and hence the (finite) union of these finite coverings would be a finite covering of M which is smaller than Σ . Similarly, we split M_1 in 2^d orthogonal parallelepipeds for at least one of which, call it M_2 , there is no finite covering smaller than Σ . We continue inductively and we end up with a sequence (M_l) of orthogonal parallelepipeds with the

following properties:

- (i) For every l there is no finite covering of M_l which is smaller than Σ .
- (ii) $M \supseteq M_1 \supseteq \dots \supseteq M_{l-1} \supseteq M_l \supseteq \dots$. This means that, if $M_l = [a_{l,1}, b_{l,1}] \times \dots \times [a_{l,d}, b_{l,d}]$, then for every $k = 1, \dots, d$ we have

$$a_k \leq a_{1,k} \leq \dots \leq a_{l-1,k} \leq a_{l,k} \leq \dots \leq b_{l,k} \leq b_{l-1,k} \leq \dots \leq b_{1,k} \leq b_k.$$

- (iii) For every $k = 1, \dots, d$ and $l \geq 1$ we have $b_{l,k} - a_{l,k} = \frac{b_k - a_k}{2^l}$ and hence $b_{l,k} - a_{l,k} \rightarrow 0$.
- (iv) For every $l \geq 2$ we have $\text{diam } M_l = \frac{\text{diam } M}{2^l}$ and hence $\text{diam } M_l \rightarrow 0$.

From (ii) we have that for every $k = 1, \dots, d$ the sequence $(a_{l,k})$ is increasing and bounded above and that the sequence $(b_{l,k})$ is decreasing and bounded below and hence both sequences converge to two limits which, because of (iii), coincide. We set

$$x_k = \lim_{l \rightarrow +\infty} a_{l,k} = \lim_{l \rightarrow +\infty} b_{l,k}.$$

Then $x = (x_1, \dots, x_d)$ belongs to every M_l . Since Σ is a covering of M , there is some $A_0 \in \Sigma$ so that $x \in A_0$. Now, A_0 is open and hence there is $\epsilon_0 > 0$ so that $N_x(\epsilon_0) \subseteq A_0$. Now, (iv) implies that there is l_0 so that $\text{diam } M_{l_0} < \epsilon_0$. Then, since $x \in M_{l_0}$, for every $y \in M_{l_0}$ we have $\|y - x\| \leq \text{diam } M_{l_0} < \epsilon_0$ and hence $y \in N_x(\epsilon_0)$. Thus $M_{l_0} \subseteq N_x(\epsilon_0) \subseteq A_0$ and so $\Sigma' = \{A_0\}$ is a finite covering of M_{l_0} which is smaller than Σ and we arrive at a contradiction with (i). \square

Bolzano-Weierstrass theorem. Every bounded sequence in \mathbb{R}^d has at least one convergent subsequence.

Proof. If (x_n) is any bounded sequence in \mathbb{R}^d , then there is a closed orthogonal parallelepiped M with edges parallel to the coordinate axes so that (x_n) is in M . Now, M is compact and hence there is a subsequence of (x_n) which converges (to an element of M). \square

The next theorem is the most useful result for the determination of compact sets.

Theorem 1.3. Let $M \subseteq \mathbb{R}^d$. Then M is compact if and only if it is bounded and closed.

First proof. Because of proposition 1.21, we have to prove only one direction.

Let M be closed and bounded. We take any (x_n) in M . Since M is bounded, (x_n) is also bounded and the Bolzano-Weierstrass theorem implies that there is a subsequence (x_{n_k}) so that $x_{n_k} \rightarrow x$ for some $x \in \mathbb{R}^d$. Since M is closed and (x_{n_k}) is in M , we have that $x \in M$. Hence every sequence in M has a subsequence which converges to an element of M and theorem 1.2 implies that M is compact.

Second proof. Again, proposition 1.21 proves one direction.

Since M is bounded, there is a closed orthogonal parallelepiped N with edges parallel to the coordinate axes so that $M \subseteq N$. Proposition 1.24 implies that N is compact and, since M is closed, proposition 1.22 implies that M is compact. \square

Example 1.4.3. Every closed ball is a compact subset of \mathbb{R}^d .

Theorem 1.3 says that the converse of proposition 1.21 is true in \mathbb{R}^d . This is not the case though in an arbitrary metric space.

Theorem 1.4. The metric space \mathbb{R}^d is complete.

Proof. Let (x_n) be a Cauchy sequence in \mathbb{R}^d . Then we easily see that (x_n) is bounded. Indeed, there is n_0 so that $\|x_n - x_m\| < 1$ for every $n, m \geq n_0$. This implies that $\|x_n - x_{n_0}\| < 1$ for every $n \geq n_0$ and hence $\|x_n\| \leq \|x_{n_0}\| + 1$ for every $n \geq n_0$. Therefore,

$$\|x_n\| \leq \max\{\|x_1\|, \dots, \|x_{n_0-1}\|, \|x_{n_0}\| + 1\} \quad \text{for every } n.$$

The Bolzano-Weierstrass theorem implies that there is a subsequence (x_{n_k}) so that $x_{n_k} \rightarrow x$ for some x . Now, we have that $\|x_k - x_{n_k}\| \rightarrow 0$, because (x_n) is a Cauchy sequence, and hence

$$\|x_k - x\| \leq \|x_k - x_{n_k}\| + \|x_{n_k} - x\| \rightarrow 0.$$

Therefore, $x_k \rightarrow x$. \square

Proposition 1.25. *Let (X, d) and (Y, ρ) be metric spaces, $M \subseteq X$ and $f : M \rightarrow Y$. If f is continuous in M and M is compact, then $f(M)$ is compact.*

Proof. Let T be an open covering of $f(M)$. Proposition 1.12 implies that for every $B \in T$ there is an open $A_B \subseteq X$ so that

$$f^{-1}(B) = A_B \cap M. \quad (1.6)$$

Since $f(M) \subseteq \bigcup_{B \in T} B$, we have $M \subseteq \bigcup_{B \in T} f^{-1}(B) \subseteq \bigcup_{B \in T} A_B$, i.e. the collection $\Sigma = \{A_B \mid B \in T\}$ is an open covering of M . Since M is compact, there are $B_1, \dots, B_n \in T$ so that $M \subseteq A_{B_1} \cup \dots \cup A_{B_n}$. This together with (1.6) imply

$$M \subseteq (A_{B_1} \cup \dots \cup A_{B_n}) \cap M = (A_{B_1} \cap M) \cup \dots \cup (A_{B_n} \cap M) = f^{-1}(B_1) \cup \dots \cup f^{-1}(B_n),$$

and hence $f(M) \subseteq B_1 \cup \dots \cup B_n$. Therefore $\{B_1, \dots, B_n\}$ is a finite covering of $f(M)$ which is smaller than T . \square

Proposition 1.26. *Every non-empty compact subset of \mathbb{R} has a maximal and a minimal element.*

Proof. Let $M \subseteq \mathbb{R}$ be non-empty and compact. Since M is non-empty and bounded, $u = \sup M$ is in \mathbb{R} . Then for every $\epsilon > 0$ there is $x \in M$ so that $u - \epsilon < x \leq u$. Therefore u is a limit point of M and, since M is closed, $u \in M$. Hence u is the maximal element of M .

The proof for the existence of a minimal element is similar. \square

Proposition 1.27 generalizes the familiar analogous proposition for continuous $f : [a, b] \rightarrow \mathbb{R}$.

Proposition 1.27. *Let (X, d) be a metric space, $M \subseteq X$ and $f : M \rightarrow \mathbb{R}$. If f is continuous on M and M is compact, then f is bounded and has a maximum and a minimum value.*

Proof. Proposition 1.25 implies that $f(M) \subseteq \mathbb{R}$ is compact. Now proposition 1.26 says that $f(M)$ is bounded and has a maximal and a minimal element. \square

Definition. *Let (X, d) and (Y, ρ) be metric spaces, $A \subseteq X$ and $f : A \rightarrow Y$. We say that f is **uniformly continuous** in A if for every $\epsilon > 0$ there is $\delta > 0$ so that $\rho(f(x'), f(x'')) < \epsilon$ for every $x', x'' \in A$ with $d(x', x'') < \delta$.*

Theorem 1.5. *Let (X, d) and (Y, ρ) be metric spaces, $M \subseteq X$ and $f : M \rightarrow Y$. If f is continuous in M and M is compact, then f is uniformly continuous in M .*

Proof. Let $\epsilon > 0$. Since f is continuous in M , for every $x \in M$ there is $\delta_x > 0$ so that

$$\rho(f(y), f(x)) < \frac{\epsilon}{2} \quad \text{for every } y \in M \text{ with } d(y, x) < \delta_x. \quad (1.7)$$

The collection $\{N_x(\frac{\delta_x}{2}) \mid x \in M\}$ is an open covering of M and, since M is compact, there are $x_1, \dots, x_n \in M$ so that

$$M \subseteq N_{x_1}(\frac{\delta_{x_1}}{2}) \cup \dots \cup N_{x_n}(\frac{\delta_{x_n}}{2}). \quad (1.8)$$

We define $\delta = \min\{\frac{\delta_{x_1}}{2}, \dots, \frac{\delta_{x_n}}{2}\} > 0$ and we take any $x', x'' \in M$ with $d(x', x'') < \delta$. Because of (1.8), there is $k = 1, \dots, n$ so that $x' \in N_{x_k}(\frac{\delta_{x_k}}{2})$ and hence $d(x', x_k) < \frac{\delta_{x_k}}{2} < \delta_{x_k}$. This implies that $d(x'', x_k) \leq d(x'', x') + d(x', x_k) < \delta + \frac{\delta_{x_k}}{2} \leq \delta_{x_k}$ and from (1.7) we have $\rho(f(x'), f(x_k)) < \frac{\epsilon}{2}$ and $\rho(f(x''), f(x_k)) < \frac{\epsilon}{2}$. Thus

$$\rho(f(x'), f(x'')) \leq \rho(f(x'), f(x_k)) + \rho(f(x''), f(x_k)) < \epsilon.$$

We proved that for every $x', x'' \in M$ with $d(x', x'') < \delta$ we have $\rho(f(x'), f(x'')) < \epsilon$. Therefore, f is uniformly continuous in M . \square

Exercises.

1.4.1. Prove that $\{(x_1, x_2) \mid x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 1\}$ is a compact subset of \mathbb{R}^2 and that $\{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 \leq x_3 \leq 1\}$ is a compact subset of \mathbb{R}^3 .

1.4.2. (i) Consider the subset $A = \{(x_1, x_2) \mid x_1^2 + x_2^2 \leq 1\}$ of \mathbb{R}^2 . Does the function $f(x_1, x_2) = e^{x_1+x_2}$ have a maximum and a minimum value in A ?

(ii) Consider the subset $A = \{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 + x_3 \leq 1, |x_3| \leq 2\}$ of \mathbb{R}^3 . Does the function $f(x_1, x_2, x_3) = e^{x_1+x_3} \sin(x_1 x_2)$ have a maximum and a minimum value in A ?

1.4.3. (i) Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ with $f(x) \rightarrow 0$ when $\|x\| \rightarrow +\infty$. This means, by definition, that for every $\epsilon > 0$ there is $R > 0$ so that $|f(x)| < \epsilon$ for every $x \in \mathbb{R}^d$ with $\|x\| > R$.

If there is $x_0 \in \mathbb{R}^d$ so that $f(x_0) \geq 0$, prove that f has a maximum value.

(ii) Prove that $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ with $f(x_1, x_2) = x_1 e^{-x_1^2 - x_2^2}$ has a maximum and a minimum value and find them.

1.4.4. Let (X, d) be a metric space, $x \in X$, (x_n) be a sequence in X so that $x_n \neq x$ for every n and $x_n \rightarrow x$. Prove that $\{x_n \mid n \in \mathbb{N}\}$ is not compact and that $\{x\} \cup \{x_n \mid n \in \mathbb{N}\}$ is compact.

1.4.5. Let (X, d) be a metric space and $M_1, \dots, M_n \subseteq X$. If M_1, \dots, M_n are compact, prove that $M_1 \cup \dots \cup M_n$ is compact.

1.4.6. Let (X, d) be a metric space and $A, B \subseteq X$. If A is compact and B is closed, prove that $A \cap B$ is compact.

1.4.7. Let (X, d) be a metric space, $x_0 \in X$ and M, N be non-empty compact subsets of X .

(i) Prove that there are $x', y' \in M$ so that $d(x', y') = \text{diam } M$.

(ii) Prove that there is $x' \in M$ so that $d(x_0, x') = \inf\{d(x_0, x) \mid x \in M\}$.

(iii) Prove that there are $x' \in M$ and $y' \in N$ so that $d(x', y') = \inf\{d(x, y) \mid x \in M, y \in N\}$.

1.4.8. Let $x_0 \in \mathbb{R}^d$, $M \subseteq \mathbb{R}^d$ be non-empty and closed and $N \subseteq \mathbb{R}^d$ be non-empty and compact.

(i) Prove that there is $x' \in M$ so that $\|x_0 - x'\| = \inf\{\|x_0 - x\| \mid x \in M\}$.

(ii) Prove that there are $x' \in M$ and $y' \in N$ so that $\|x' - y'\| = \inf\{\|x - y\| \mid x \in M, y \in N\}$.

1.4.9. Let M be a bounded subset of \mathbb{R}^d . Prove that \overline{M} and ∂M are compact.

1.4.10. Let (X, d) be a non-empty set with the discrete metric (exercise 1.1.8). Prove that $M \subseteq X$ is compact if and only if it is a finite set.

1.5 Connectedness.

Definition. Let (X, d) be a metric space and $A \subseteq X$. We say that B, C form a **decomposition** of A if (i) $B \cup C = A$, (ii) $B \neq \emptyset$, $C \neq \emptyset$, (iii) none of B, C contains a limit point of the other.

It is clear that (iii) is equivalent to $\overline{B} \cap C = \emptyset$ and $B \cap \overline{C} = \emptyset$.

Example 1.5.1. In \mathbb{R}^2 we consider the closed discs $B = \overline{D}_{(0,0)}(1)$, $C = \overline{D}_{(3,0)}(1)$ and their union $A = B \cup C$. It is clear that B, C form a decomposition of A .

If we consider the open discs $B = D_{(0,0)}(1)$, $C = D_{(2,0)}(1)$ and $A = B \cup C$, then the discs B, C are tangent but, again, they form a decomposition of A .

If we take the closed disc $B = \overline{D}_{(0,0)}(1)$, the open disc $C = D_{(2,0)}(1)$ and $A = B \cup C$, then the discs B, C are tangent and they *do not* form a decomposition of A . Indeed, B contains the limit point $(1, 0)$ of C .

Definition. Let (X, d) be a metric space and $A \subseteq X$. We say that A is **connected** if there is no decomposition of A , i.e. there is no pair of sets B, C with the properties (i)-(iii) of the previous definition.

Example 1.5.2. The first two sets A of example 1.5.1 are not connected since each of them admits a specific decomposition. But we cannot decide at this moment if the third set A of example 1.5.1 is connected or not. We know that the specific B, C related to this A do not form a decomposition of A . *To decide that A is connected we must prove that, not only the specific pair, but an arbitrary pair does not form a decomposition of A .*

Example 1.5.3. It is obvious that \emptyset as well as any $\{x\}$ is a connected set.

Lemma 1.1. *Let (X, d) be a metric space and $A, B, C \subseteq X$ with $B \cap C = \emptyset$ and assume that none of B, C contains a limit point of the other. If A is connected and $A \subseteq B \cup C$, then either $A \subseteq B$ or $A \subseteq C$.*

Proof. We define $B_1 = A \cap B$ and $C_1 = A \cap C$. Clearly, $B_1 \cup C_1 = A$ and $B_1 \cap C_1 = \emptyset$. Now let $x \in B_1$. Then $x \in B$, and x is not a limit point of C . Then there is $r > 0$ so that $N_x(r) \cap C = \emptyset$ and, since $C_1 \subseteq C$, we get $N_x(r) \cap C_1 = \emptyset$. Thus x is not a limit point of C_1 . We conclude that B_1 does not contain any limit point of C_1 . Similarly, C_1 does not contain any limit point of B_1 .

If $B_1 \neq \emptyset$ and $C_1 \neq \emptyset$, then B_1, C_1 form a decomposition of A and this contradicts the connectedness of A . Hence, either $B_1 = \emptyset$ or $C_1 = \emptyset$ and thus either $A \subseteq C$ or $A \subseteq B$, respectively \square

Proposition 1.28. *Let (X, d) be a metric space and Σ be a collection of connected subsets of X all of which have a common point. Then $\bigcup_{A \in \Sigma} A$ is connected.*

Proof. We set $U = \bigcup_{A \in \Sigma} A$ and we shall prove that U is connected. Let x_0 be the common point of all $A \in \Sigma$.

We assume that U is not connected. Then there are B, C which form a decomposition of U . Since $x_0 \in U$, we have $x_0 \in B$ or $x_0 \in C$. Assume that $x_0 \in B$ (the proof is the same if $x_0 \in C$). For every $A \in \Sigma$ we have $A \subseteq U$ and hence $A \subseteq B \cup C$. According to lemma 1.1, every $A \in \Sigma$ is contained either in B or in C . But if any $A \in \Sigma$ is contained in C , it cannot contain x_0 which is in B . Therefore every $A \in \Sigma$ is contained in B and hence $U \subseteq B$. This implies that $C = \emptyset$ and we arrived at a contradiction. \square

Proposition 1.29. *Let (X, d) be a metric space and $A, D \subseteq X$ so that $A \subseteq D \subseteq \overline{A}$. If A is connected, then D is connected.*

Proof. Let D not be connected. Then there are B, C which form a decomposition of D . Since $A \subseteq D$, we have $A \subseteq B \cup C$. Lemma 1.1 implies that $A \subseteq B$ or $A \subseteq C$. Let $A \subseteq B$. (The proof is similar if $A \subseteq C$.) Now, every point of D is a limit point of A and hence a limit point of B (since $A \subseteq B$). Therefore no point of D belongs to C (since C does not contain limit points of B) and this is wrong since $C \neq \emptyset$. \square

Proposition 1.30. *Let $(X, d), (Y, \rho)$ be metric spaces, $A \subseteq X$ and $f : A \rightarrow Y$. If f is continuous in A and A is connected, then $f(A)$ is connected.*

Proof. Assume that $f(A)$ is not connected. Then there are B', C' which form a decomposition of $f(A)$. We consider the inverse images of B', C' in A , i.e. the sets

$$B = f^{-1}(B') = \{x \in A \mid f(x) \in B'\}, \quad C = f^{-1}(C') = \{x \in A \mid f(x) \in C'\}.$$

It is clear that $B \cup C = A$, $B \cap C = \emptyset$, $B \neq \emptyset$, $C \neq \emptyset$.

Now, let B contain a limit point b of C . Then there is a sequence (c_n) in C so that $c_n \rightarrow b$. Since f is continuous at b , we get $f(c_n) \rightarrow f(b)$. The sequence $(f(c_n))$ is in C' and thus $f(b)$ is a limit point of C' . But $f(b) \in B'$ and we arrive at a contradiction, because B' does not contain any limit point of C' . Hence B does not contain any limit point of C . Similarly, C does not contain a limit point of B . Thus, B, C form a decomposition of A and this is wrong since A is connected. \square

Definition. Let (X, d) be a metric space, $x, y \in X$ and $r > 0$. Every finite set $\{z_0, \dots, z_n\} \subseteq X$ with $z_0 = x$, $z_n = y$ and $d(z_k, z_{k-1}) < r$ for every $k = 1, \dots, n$ is called **r -succession of points** which joins x, y . If, moreover, $z_k \in A$ for every $k = 0, \dots, n$, we say that the r -succession of points is in A .

Theorem 1.6. Let (X, d) be a metric space and K be a compact subset of X . Then K is connected if and only if for every $x, y \in K$ and every $r > 0$ there is an r -succession of points in K which joins x, y .

Proof. Assume K is connected. We take any $x, y \in K$ and any $r > 0$ and let there be no r -succession of points in K which joins x, y . We define the sets

$$B = \{b \in K \mid \text{there is an } r\text{-succession of points in } K \text{ which joins } x, b\},$$

$$C = \{c \in K \mid \text{there is no } r\text{-succession of points in } K \text{ which joins } x, c\}.$$

It is clear that $B \cup C = K$, $B \neq \emptyset$ (since $x \in B$) and $C \neq \emptyset$ (since $y \in C$).

Assume that B contains a limit point b of C . Then (since $b \in B$) there is an r -succession of points in K which joins x, b and, also, (since b is a limit point of C) there is $c \in C$ so that $d(c, b) < r$. If to the r -succession of points of K which joins x, b we attach c (as a final point after b), then we get an r -succession of points in K which joins x, c . This is wrong since $c \in C$. Hence B does not contain any limit point of C .

Now assume that C contains a limit point c of B . Then (since c is a limit point of B) there is $b \in B$ so that $d(c, b) < r$ and (since $b \in B$) there is an r -succession of points in K which joins x, b . If to the r -succession of points in K which joins x, b we attach c (as a final point after b), then we get an r -succession of points in K which joins x, c . This is wrong since $c \in C$. Hence C does not contain any limit point of B .

We conclude that B, C form a decomposition of K and this is wrong since K is connected.

Therefore there is an r -succession of points in K which joins x, y .

Conversely, assume that for every $x, y \in K$ and every $r > 0$ there is an r -succession of points in K which joins x, y .

We assume that K is not connected. Then there are B, C which form a decomposition of K .

Let x be a limit point of B . Since $B \subseteq K$, x is a limit point of K and, since K is closed, we get $x \in K$. Now, $x \notin C$ (because C does not contain any limit point of B) and we get that $x \in B$. Thus B contains all its limit points and it is closed. Finally, since $B \subseteq K$ and K is compact, B is also compact. Similarly C is also compact.

Now B, C are compact and disjoint and proposition 1.23 implies that there is $r > 0$ so that $d(b, c) \geq r$ for every $b \in B$ and $c \in C$. Since $B \neq \emptyset$, $C \neq \emptyset$, we consider $b' \in B$ and $c' \in C$. Then it is easy to see that there is no r -succession of points in K which joins b', c' , and we arrive at a contradiction. Indeed, assume that there is an r -succession $\{z_0, \dots, z_n\}$ in K so that $z_0 = b'$, $z_n = c'$ and $d(z_k, z_{k-1}) < r$ for every $k = 1, \dots, n$. Since $z_0 \in B$, $z_n \in C$, it is clear that there is k so that $z_{k-1} \in B$, $z_k \in C$. Then $d(z_k, z_{k-1}) < r$ contradicts that we have $d(b, c) \geq r$ for every $b \in B, c \in C$. \square

Proposition 1.31. A set $I \subseteq \mathbb{R}$ is connected if and only if it is an interval.

Proof. Let I be connected. If I is not an interval, there are $x_1, x_2 \in I$ and $x \notin I$ so that $x_1 < x < x_2$. Then the sets $B = I \cap (-\infty, x)$ and $C = I \cap (x, +\infty)$ form a decomposition of I and we have a contradiction. Thus I is an interval.

Conversely, let I be an interval. If I has only one element, then it is connected. If $I = [a, b]$ with $a < b$, then $[a, b]$ is compact and if we take any x, y in $[a, b]$ and any $r > 0$, it is clear that we can find an r -succession of points in $[a, b]$ which joins x and y . Thus $[a, b]$ is connected. If I is an interval of any other type, we can find a sequence of intervals $I_n = [a_n, b_n]$ which increase and their union is I . Then each I_n is connected and proposition 1.28 implies that I is also connected. \square

Now we have the following corollary of propositions 1.30 and 1.31.

Proposition 1.32. *Let (X, d) be a metric space, $A \subseteq X$ and $f : A \rightarrow \mathbb{R}$ be continuous on A . If A is connected, then f has the intermediate value property in A .*

Proof. $f(A)$ is a connected subset of \mathbb{R} and hence it is an interval. Not let u_1, u_2 be values of f in A , i.e. u_1, u_2 belong to the interval $f(A)$. Then every u with $u_1 < u < u_2$ also belongs to the interval $f(A)$. I.e. every number between the values u_1, u_2 of f in A is also a value of f in A . \square

A special case of proposition 1.32 is the well known *intermediate value theorem* which says that if $f : I \rightarrow \mathbb{R}$ is continuous in the interval $I \subseteq \mathbb{R}$, then it has the intermediate value property in I .

Definition. *Let (X, d) be a metric space, $I \subseteq \mathbb{R}$ be an interval and $\gamma : I \rightarrow X$ be continuous on I . We say that γ is a **curve** in (X, d) . The set $\gamma^* = \gamma(I) = \{\gamma(t) \mid t \in I\}$ is called **trajectory** of the curve γ . If $\gamma^* \subseteq A \subseteq X$, we say that the curve γ is in A .*

Propositions 1.30 and 1.31 imply that the trajectory of any curve in (X, d) is a connected subset of X . Also, if the interval I (the domain of definition of the curve) is closed and bounded (hence compact), then proposition 1.25 implies that the trajectory of the curve is a compact subset of X .

Example 1.5.4. Every linear segment $[x, y]$ in \mathbb{R}^d is the trajectory of the curve $\gamma : [a, b] \rightarrow \mathbb{R}^d$ given by $\gamma(t) = \frac{b-t}{b-a}x + \frac{t-a}{b-a}y$ for $a \leq t \leq b$.

A polygonal line consisting of two successive linear segments, i.e. $[x, y] \cup [y, z]$, is also the trajectory of a curve: we may take $a < b < c$ and $\gamma : [a, c] \rightarrow \mathbb{R}^d$ given by $\gamma(t) = \frac{b-t}{b-a}x + \frac{t-a}{b-a}y$ when $a \leq t \leq b$ and by $\gamma(t) = \frac{c-t}{c-b}y + \frac{t-b}{c-b}z$ when $b \leq t \leq c$. This function γ is continuous, since its formulas in the subintervals $[a, b]$ and $[b, c]$ of $[a, c]$ coincide at b .

In a similar manner we may see that a general polygonal line consisting of n successive linear segments is the trajectory of a curve.

Definition. *Let (X, d) be a metric space and $A \subseteq X$. We say that A is **arcwise connected** if for every two points of A there is a curve in A which joins these two points.*

Proposition 1.33. *Let (X, d) be a metric space and $A \subseteq X$. If A is arcwise connected, then it is connected.*

Proof. We fix any $x_0 \in A$. For every $x \in A$ there is a curve γ_x in A which joins x_0 and x . Then $\gamma_x^* \subseteq A$ and hence $\bigcup_{x \in A} \gamma_x^* \subseteq A$. Conversely, since every $x \in A$ is contained in the trajectory γ_x^* , we have that $A \subseteq \bigcup_{x \in A} \gamma_x^*$. Therefore $A = \bigcup_{x \in A} \gamma_x^*$. Now, every γ_x^* is connected and since all γ_x^* have the point x_0 in common, we conclude that A is connected. \square

Example 1.5.5. Every ring is a connected subset of \mathbb{R}^2 .

Example 1.5.6. Every convex set $A \subseteq \mathbb{R}^d$ is arcwise connected and hence connected. Indeed if we take any two points in A the linear segment which joins them is contained in A . For instance, balls and orthogonal parallelepipeds are connected subsets of \mathbb{R}^d .

Example 1.5.7. A set $A \subseteq \mathbb{R}^d$ is called **star-shaped** if there is a specific point $x_0 \in A$ so that for every $x \in A$ the linear segment $[x_0, x]$ is contained in A . Every such x_0 is called *center* of the star-shaped set A . The center of the star-shaped set A may not be unique, but this does not mean that every point of A is a center of it.

It is clear that a star-shaped A is arcwise connected and hence connected. Indeed, every two points of A can be joined with a polygonal line in A consisting of two successive linear segments: one segment from one of the points to the center x_0 and the other segment from x_0 to the second point.

Example 1.5.8. The set $A = \overline{D}_{(0,0)}(1) \cup D_{(2,0)}(1)$ in examples 1.5.1 and 1.5.2 is connected, since it is star-shaped with center 1.

Theorem 1.7. *Let A be an open subset of \mathbb{R}^d . Then A is connected if and only if it is arcwise connected.*

Proof. If A is arcwise connected, proposition 1.33 implies that it is connected.

Conversely, let A be connected. We take $x, y \in A$ and we assume that there is no polygonal line in A which joins x, y . We define the sets

$$B = \{b \in A \mid \text{there is a polygonal line in } A \text{ which joins } x, b\},$$

$$C = \{c \in A \mid \text{there is no polygonal line in } A \text{ which joins } x, c\}.$$

It is clear that $B \cup C = A$, $B \neq \emptyset$ (since $x \in B$) and $C \neq \emptyset$ (since $y \in C$).

We assume that B contains some limit point b of C . Then (since $b \in B$) there is a polygonal line in A which joins x, b . Since A is open, there is $r > 0$ so that $N_b(r) \subseteq A$ and (since b is a limit point of C) there is $c \in N_b(r) \cap C$. If to the polygonal line in A which joins x, b we attach (as last) the linear segment $[b, c]$ (which is contained in $N_b(r)$ and hence in A), we get a polygonal line in A which joins x, c . This is wrong, since $c \in C$. Thus B does not contain any limit point of C .

Now we assume that C contains a limit point c of B . Since A is open, there is $r > 0$ so that $N_c(r) \subseteq A$. Then (since c is a limit point of B) there is $b \in N_c(r) \cap B$. As before, (since $b \in B$) there is a polygonal line in A which joins x, b and, if to this we attach the linear segment $[b, c]$ (which is contained in $N_c(r)$ and hence in A), we get a polygonal line in A which joins x, c . This is wrong, since $c \in C$. Thus C does not contain any limit point of B .

We conclude that B, C form a decomposition of A and we arrive at a contradiction because A is connected.

Therefore, there is a polygonal line in A which joins x, y . □

Definition. *Let (X, d) be a metric space and $A \subseteq X$. We say that $C \subseteq A$ is a **connected component** of A if C is connected and has the following property: if $C \subseteq C' \subseteq A$ and C' is connected, then $C = C'$.*

In other words, C is a connected component of A if it is a connected subset of A and there is no strictly larger connected subset of A .

Let us see a characteristic property of connected components. Let C be a connected component of A and let B be any connected subset of A so that $C \cap B \neq \emptyset$. Then $C \cup B$ is connected (being the union of connected sets with a common point) and $C \subseteq C \cup B \subseteq A$. Since C is a connected component of A , we get $C \cup B = C$ and hence $B \subseteq C$. In other words, *a connected component of A swallows every connected subset of A intersecting it.*

Let C_1, C_2 be *distinct* connected components of A and assume that $C_1 \cap C_2 \neq \emptyset$. Since C_1 is a connected subset of A and intersects the connected component C_2 of A , we get $C_1 \subseteq C_2$. Symmetrically, $C_2 \subseteq C_1$ and hence $C_1 = C_2$. We arrive at a contradiction and we conclude that $C_1 \cap C_2 = \emptyset$. Thus, *different connected components of A are disjoint.*

Proposition 1.34. *Let (X, d) be a metric space and $A \subseteq X$. Then A is the union of its (mutually disjoint) connected components.*

Proof. We shall prove that every point of A belongs to a connected component of A . We take $x \in A$ and define C_x to be the union of all connected subsets B of A which contain x . (Such a set is $\{x\}$.) I.e.

$$C_x = \bigcup \{B \mid B \text{ is connected } \subseteq A \text{ and } x \in B\}.$$

Now C_x is a subset of A and contains x . It is also connected, since it is the union of connected sets B with x as a common point. If $C_x \subseteq C' \subseteq A$ and C' is connected, then C' is one of the connected subsets B of A which contain x and hence $C' \subseteq C_x$. Thus $C_x = C'$. Therefore C_x is a connected component of A and contains x . □

It is obvious that A is connected if and only if A is the only connected component of A .

Example 1.5.9. In \mathbb{R}^2 we consider the discs $B = D_{(0,0)}(1)$ and $C = D_{(3,0)}(1)$ and the set $A = B \cup C$. The discs B, C are connected subsets of A . Lemma 1.1 implies that any connected subset of A is contained either in B or in C . I.e. there is no connected subset of A strictly larger than either B or C . Therefore the discs B and C are the connected components of A .

Example 1.5.10. We take $\mathbb{Z} \subseteq \mathbb{R}$ and any $n \in \mathbb{Z}$.

Then $\{n\}$ is a connected set. Let $\{n\} \subseteq C' \subseteq \mathbb{Z}$ and $C' \neq \{n\}$. Then $C' = \{n\} \cup (C' \setminus \{n\})$ and it is clear that the sets $\{n\}$ and $C' \setminus \{n\}$ form a decomposition of C' . Thus C' is not connected and hence $\{n\}$ is a connected component of \mathbb{Z} .

Therefore \mathbb{Z} has infinitely many connected components, each of them being a singleton.

Proposition 1.35. Let (X, d) be a metric space and $A \subseteq X$. If A is closed, then every connected component of A is closed.

Proof. Let C be a connected component of A . Since $C \subseteq A$ and A is closed, we get $C \subseteq \overline{C} \subseteq A$. Proposition 1.29 implies that \overline{C} is connected and, since C is a connected component of A , we get that $C = \overline{C}$. Therefore C is closed. \square

Proposition 1.36. Let A be an open subset of \mathbb{R}^d . Every connected component of A is open.

Proof. Let C be a connected component of A and $x \in C$. Then $x \in A$ and, since A is open, there is $r > 0$ so that $N_x(r) \subseteq A$. Since $N_x(r)$ is a connected subset of A and intersects the connected component C of A , we see that $N_x(r) \subseteq C$. Thus x is an interior point of C . \square

Propositions 1.34 and 1.36 imply that every open subset of \mathbb{R}^d is the union of disjoint open connected sets.

Exercises.

1.5.1. Say which of the following subsets of \mathbb{R}^2 are connected and find their connected components. The complement of a circle or of an open triangular line or of a closed triangular line. Also: $\{(\frac{1}{n}, 0) | n \in \mathbb{N}\}$, $\{(0, 0)\} \cup \{(\frac{1}{n}, 0) | n \in \mathbb{N}\}$, $[(0, 0), (1, 0)] \cup \bigcup_{n=1}^{+\infty} [(\frac{1}{n}, 0), (\frac{1}{n}, \frac{1}{n})]$, $\bigcup_{n=1}^{+\infty} \{x \in \mathbb{R}^2 | \|x\| = 1 + \frac{1}{n}\}$, $\{(x, y) | x, y \in \mathbb{Q}\}$.

1.5.2. Prove that the following sets are connected $\{(x, \sin x) | x \in \mathbb{R}\}$, $\{(x, \sin \frac{1}{x}) | 0 < x \leq 1\}$, $\{(x, \sin \frac{1}{x}) | 0 < x \leq 1\} \cup [(0, -1), (0, 1)]$ in \mathbb{R}^2 .

1.5.3. (i) Find a simple example of two connected sets in \mathbb{R}^2 whose intersection is not connected.
(ii) Find a simple example of a connected set A in \mathbb{R}^2 so that ∂A is not connected.
(iii) Find a simple example of a connected set A in \mathbb{R}^2 so that A° is not connected.

1.5.4. Let $d \geq 2$, $U \subseteq \mathbb{R}^d$ be a connected open set and $a_1, \dots, a_n \in U$. Prove that $U \setminus \{a_1, \dots, a_n\}$ is connected and open.

1.5.5. Take a hyperplane L in \mathbb{R}^d and the two open halfspaces of \mathbb{R}^d which are determined by L . If a curve γ in \mathbb{R}^d joins a point of one halfspace and a point of the other halfspace, prove that the trajectory of γ intersects L .

1.5.6. Let (X, d) be a metric space, $A_n \subseteq X$ be connected and $A_n \cap A_{n+1} \neq \emptyset$ for every n . Prove that $\bigcup_{n=1}^{+\infty} A_n$ is connected.

1.5.7. Let $B \subseteq \mathbb{R}^d$. If B is open and closed prove that either $B = \emptyset$ or $B = \mathbb{R}^d$.

1.5.8. Let (X, d) be a metric space and $A \subseteq X$ be connected (not necessarily compact). Prove that for every $r > 0$ and every $x, y \in A$ there is an r -succession of points in A which joins x, y .

1.5.9. Let (X, d) be a metric space and $A \subseteq X$. Prove that A is connected if and only if the only continuous functions $f : A \rightarrow \mathbb{R}$ with $f(A) \subseteq \mathbb{Z}$ are the constant functions.

1.5.10. Let $A \subseteq \mathbb{R}^d$ be open and connected and let every point of $B \subseteq A$ be an isolated point of B . Prove that $A \setminus B$ is connected.

1.5.11. Let (X, d) be a metric space.

(i) Let $A_n \subseteq X$ be compact so that $A_{n+1} \subseteq A_n$ for every $n \in \mathbb{N}$ and so that every two points of A_n can be joined by some $\frac{1}{n}$ -succession of points in A_n . Prove that $\bigcap_{n=1}^{+\infty} A_n$ is connected.

(ii) Let $F \subseteq X$ be compact and let $x, y \in F$ belong to different connected components of F . Prove that there is a decomposition B, C of F so that $x \in B$ and $y \in C$.

1.5.12. Let (X, d) be a metric space. We say that (X, d) is **locally connected** if for every $x \in X$ and every $r > 0$ there is an open connected U so that $x \in U \subseteq N_x(r)$.

Prove that (X, d) is locally connected if and only if for every open $A \subseteq X$ all the connected components of A are open.

1.5.13. Let (X, d) be a metric space. We say that (X, d) is **locally arcwise connected** if for every $x \in X$ and every $r > 0$ there is an open arcwise connected U so that $x \in U \subseteq N_x(r)$.

If (X, d) is locally arcwise connected, prove that every open $A \subseteq X$ is connected if and only if it is arcwise connected.

1.5.14. Let (X, d) be a non-empty set with the discrete metric (exercise 1.1.8). Prove that $M \subseteq X$ is connected if and only if it has at most one element.