INSTITUT FOR MATEMATIK OG DATALOGI
SYDDANSK UNIVERSITET

Lecture Notes on Metric and Topological Spaces

Niels Jørgen Nielsen

2010

Introduction

The theory of metric and topological spaces is fundamental in most areas of mathematics and its applications. The study of concepts like convergence or continuity is essential in mathematics and this theory puts these concepts into a general framework.

Most students know the basic properties of continuous functions defined on the real line, the plane or the space and the basic tool for defining and studying such functions is the ability of measuring distances. We would like to put this into a more general setting. Assume e.g. that we have a set A of objects (this could be functions) each of which can be put into a system of equations to give a solution. This creates a set B of solutions. The perfect situation is of course that if two objects in A are "close" the corresponding two solutions in B are also "close".

The obstacle here is of course what we mean by "close". If we in some way could define a distance function both on A and B, then we could also define what it means that two elements in the given set are "close". This is the background of defining a *metric space* which is a set equipped with a distance function, also called a metric, which enables us to measure the distance between two elements in the set.

It is however not always possible to find a reasonable metric (reasonable relative to our specific purpose) on a given set, but then the more general concept of a *topological space* can often help. In such a space we use other ways to define what is means that two elements in the space are close.

One of the mostly used tools in mathematics is *compactness*. A compact space is a topological space where we loosely speaking have the ability to reduce a situation involving infinitely many objects of a certain type to a situation only involving finitely many objects.

The ability to perform compactness arguments is essential for anyone who wants to study mathematics and its applications.

We now wish to discuss the arrangement of these notes in greater detail.

Section 0 contains some preliminaries on the concept of a function and some properties of the natural numbers, the real numbers, and the complex numbers. Some readers might know the material presented here. Section 1 is devoted to inner product spaces which is a special case of metric spaces (and also cover the most important examples). This is done because many readers probably have studied these spaces in a course on linear algebra. In section 2 we define metric spaces and investigate their basic properties while we in Section 3 prove some important theorems on continuous functions defined on metric spaces; these results have also applications to the classical situation of real valued functions defined on an interval of $\mathbb{R}$. In Section 4 we extend our theory to general topological spaces and investigate continuous functions on topological spaces in Section 5. Section 6 is concerned with compact topological spaces, their properties, and how continuous functions behave on such spaces. We also give a proof of the classical theorem of Heine–Borel which characterizes the compact subsets of $\mathbb{R}^n$. Section 7 contains the most important results on connected subsets of topological spaces.

Though these notes are mainly on metric spaces, we have taken the attitude that if a notion is the same for metric spaces and the more general topological spaces, then we use the more general setting. Similarly, if a proof of a result is more or less the same for metric and topological spaces, we give the proof in full generalty.

0 Notation and preliminaries

In these notes we shall use the following notation on the different sets of numbers:

- $\mathbb{N}$ denotes the set of natural numbers.
- $\mathbb{Z}$ denotes the set of integers.
- $\mathbb{Q}$ denotes the set of rational numbers.
- $\mathbb{R}$ denotes the set of real numbers.
- $\mathbb{C}$ denotes the set of complex numbers.

0.1 The concept of a function

Let X and Y be two sets. A *function* from X to Y is a *relation* which to every $x \in X$ relates a uniquely determined element in Y which is called $f(x)$. X is called the *domain of defintion* for f and Y is called the *co–domain* of f. That f is a function from X to Y, we shall write as $f : X \to Y$. Once we have decided what X and Y should be, we will often just talk about the function f. If we change either the domain of definition, the relation f, or the co–domain, we get a new function. Let us now look on an example:

Example 0.1 *Let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by $f(x) = x^2$ for all $x \in \mathbb{R}$. We could also look on the function $g : [0,1] \to \mathbb{R}$ defined by $g(x) = x^2$ for all $x \in [0,1]$ or on the function $h : \mathbb{R} \to [0,\infty[$ defined by $h(x) = x^2$ for all $x \in \mathbb{R}$. Though the relation is the same, the three functions f, g, and h are different.*

The *image set* of a function $f : X \to Y$ (or just the image of f) is denoted by $f(X)$ and is the subset of Y defined by:

$$f(X) = \{f(x) \mid x \in X\}.$$

Note that we can consider f as a function from X to $f(X)$ which is sometimes handy. Likewise, if $A \subseteq X$ the *image of A by f* is the subset $f(A)$ of Y defined by:

$$f(A) = \{f(x) \mid x \in A\}$$

A function $f : X \to Y$ is said to be *injective* if it satisfies the condition

$$\forall x, y \in X : f(x) = f(y) \Rightarrow x = y,$$

or equivalently

$$\forall x, y \in X : x \neq y \Rightarrow f(x) \neq f(y).$$

An injective function f is also called *one–to–one* and we shall often say f *is 1–1* when we mean that f is injective. f is called an *injection*.

A function $f : X \to Y$ is said to be *surjective when* $f(X) = Y$. Instead of saying that f is surjective we shall often say that f is *onto* and call f a *surjection*.

Whether a given f gives an injective or a surjective function depends heavily on which domain of definition or co–domain we use as the following simple example shows:

Example 0.2 *If $f : \mathbb{R} \to \mathbb{R}$ is the function defined by f(x) = x^2 for all $x \in \mathbb{R}$, then f is neither injective nor surjective. If we change the function to (we shall still call it f) $f : [0, \infty[\to \mathbb{R}$ with $f(x) = x^2$ for all $x \in [0, \infty[$, then f becomes injective, but not surjective. If we finally consider $f : [0, \infty[\to [0, \infty[$, then f becomes both injective and surjective.*

The function $f : X \to Y$ is said to be *bijective* if it is both injective and surjective. We then call f a *bijection*. f is a bijection if and only if

$$\forall y \in Y \quad \exists! x \in X : y = f(x),$$

where "$\exists! x$" means that there exists exactly one x.

For a bijective f this statement can be used to define a function from Y to X, named $f^{-1} : Y \to X$ by the statement

$$\forall y \in Y \quad f^{-1}(y) = x \Leftrightarrow (x \in X) \wedge (f(x) = y).$$

f^{-1} is called the inverse function to f.

Unfortunately the symbol f^{-1} is also used when f is not bijective. If $f : X \to Y$ is any function and $V \subseteq Y$, we put

$$f^{-1}(V) = \{x \in X \mid f(x) \in V\}.$$

The set $f^{-1}(V)$ is called *the inverse image* of the set V by the function f.

In most cases this ambiguity does not cause problems as the following exercise hopefully shows.

Exercise 0.3 *Let $f : X \to Y$ be a bijective function and call its inverse $g : Y \to X$ in this exercise while we reserve the symbol f^{-1} to mean the inverse image by f. Prove that if $V \subseteq Y$, then $g(V) = f^{-1}(V)$. this means in particular that if $y \in Y$, then $g(y) = f^{-1}(\{y\})$.*

The inverse image construction is very essential and the next theorem gives a list of its basic properties.

Theorem 0.4 *Let $f : X \to Y$ be a function and let there for all i in an index set I be given a $V_i \subseteq Y$. Further let $V \subseteq Y$ and $U \subseteq X$. Then the following statements hold:*

$$
\begin{aligned}
f^{-1}(\emptyset) &= \emptyset && (0.1)\\
f^{-1}(Y) &= X && (0.2)\\
f^{-1}(Y \setminus V) &= X \setminus f^{-1}(V) && (0.3)\\
f^{-1}(\bigcup_{i\in I} V_i) &= \bigcup_{i\in I} f^{-1}(V_i) && (0.4)\\
f^{-1}(\bigcap_{i\in I} V_i) &= \bigcap_{i\in I} f^{-1}(V_i) && (0.5)\\
f(f^{-1}(V)) &\subseteq V && (0.6)\\
f^{-1}(f(U)) &\supseteq U && (0.7)
\end{aligned}
$$

Exercise 0.5 *Prove each of the statements in Theorem 0.4. Give examples where the inclusions in the last two staments are sharp.*

We used indexing in the formulation of the previous theorem so let us end this subsection by formalizing this concept.

Definition 0.6 *Let I and A be two sets. A is said to indexed by I if there exists a bijection $f : I \to A$. Usually we shall put $x_i = f(i)$ for all $i \in I$ and write $A = \{x_i \mid i \in I\}$ without mentioning f.*

Indexing a set is a very practical notion; just look on the formulas in the previous theorem which would have been more complicated to write if we had not used indexing. Note also that there is no loss of generality to use it because any set can be indexed. Indeed, let A be any set and let $Id : A \to A$ be the identity function, that is $Id(x) = x$ for all $x \in A$. Id is clearly a bijection so we have indexed A by itself.

0.2 Sequences, countable sets and properties of $\mathbb{N}$

We start with the following definition:

Definition 0.7 *If A is an arbitrary set, then a sequence in A is a function $f : \mathbb{N} \to A$. Usually we will put $x_n = f(n) \in A$ for all $n \in \mathbb{N}$ and talk about the sequence (x_n) without mentioning f.*

Note that it is not required that the function f in Definition 0.7 is injective so some of the x_n's might be equal. It could happen that the $\{x_n \mid n \in \mathbb{N}\}$ is finite. Consider e.g. the sequence $(x_n) \subseteq \mathbb{R}$ defined by setting $x_n = 1$ when n is even and $x_n = -1$ when n is odd.

Definition 0.8 *A set A is called countable if there is a bijection $f : \mathbb{N} \to A$. In other words A is indexed by the natural numbers and we can with $x_n = f(n)$ for all $n \in \mathbb{N}$ write $A = \{x_n \mid n \in \mathbb{N}\}$.*

It is proved in [4, Exercises 10,12] that $\mathbb{Z}$ and $\mathbb{Q}$ are countable. An infinite set which is not countable will be called *uncountable*. It will be proved later in these notes that $\mathbb{R}$ is uncountable.

A set is called *at most countable* if it is either finite or countable. Some authors also use the word *enumerable* instead of countable.

Let us here mention one very important property of $\mathbb{N}$ which is the basis of the induction principle (proofs by induction).

The well ordering axiom for $\mathbb{N}$

Every non–empty subset A of $\mathbb{N}$ has a minimal (smallest) element $\min A$.

We now get the fundamental theorem on the induction principle:

Theorem 0.9 *Let $P(n)$ be an open statement involving one variable n from $\mathbb{N}$. If*

(i) The statement $P(1)$ is true,

(ii) The statement ($\forall n \in \mathbb{N} : P(n) \Rightarrow P(n+1)$) is true,

then the statement $P(n)$ is true for all $n \in \mathbb{N}$ (or written with quatifiers: $\forall n \in \mathbb{N} \quad P(n)$).

Proof: We shall prove the theorem by contradiction so assume that $P(n)$ is not true for all $n \in \mathbb{N}$ whis means that the set

$$A = \{n \in \mathbb{N} \mid P(n) \text{ is untrue}\} \neq \emptyset.$$

By the well ordering axiom for $\mathbb{N}$ the set A has a smallest element m. (i) implies that $m > 1$ so that $m - 1 \in \mathbb{N}$ and hence by the choice of m $P(m-1)$ is true. Applying (ii) we get that then also $P(m)$ is true which is a contradiction. □

Proofs using Theorem 0.9 are called *induction proofs* or *proofs by induction*.

0.3 Properties of $\mathbb{R}$

Some of the basic properties of $\mathbb{R}$ which we shall discuss in this part might be well known to some of the readers.

Let $S \subseteq \mathbb{R}$ and let $m, M \in \mathbb{R}$. M is called an *upper bound* for S if $x \leq M$ for all $x \in S$. Similarly m is called an lower bound for S if $m \leq x$ for all $x \in S$. S is called *bounded from above* if there exists an upper bound for S and similarly S is called *bounded from below* if it has a lower bound. S is called *bounded* if it is bounded both from below and from above. Hence S is bounded if and only if it is a subset of an interval $[m, M]$ of finite length.

We now need the following definition:

Definition 0.10 *(Least upper bound) Let $S \subseteq \mathbb{R}$ and $M_0 \in \mathbb{R}$. M_0 is called a least upper bound for S, if*

(i) M_0 is an upper bound for S and

(ii) For every upper bound M for S we have that $M_0 \leq M$.

Note that M_0 is uniquely determined if it exists. Indeed, if M_0 and M_1 are least upper bounds for S, then appying (ii) on M_0, respectively M_1 we get that $M_0 \leq M_1$ and $M_1 \leq M_0$; that is $M_0 = M_1$.

If S is not bounded from above, it has of course no least upper bound since (i) above cannot be fulfilled. If $S = \emptyset$, every real number is an upper bound for S and hence (ii) cannot be fulfilled. Hence $\emptyset$ does not have a least upper bound.

It is a fundamental property of the real numbers that the empty set and the subsets of $\mathbb{R}$ which are not bounded from above are the only subsets of $\mathbb{R}$ without a least upper bound. In these notes we shall take this property as an axiom, but we will not make an axiomatic construction of $\mathbb{R}$.

The completeness axiom of the real numbers

Every non–empty subset $S \subseteq \mathbb{R}$ which is bounded from above has a least upper bound. This number is uniquely determined and is called the supremum of S and written as $\sup S$.

It is important to notice that $\sup S$ need not belong to S. If actually $\sup S \in S$, we shall write $\max S$ instead of $\sup S$. If S is unbounded from above and hence does not have a least upper bound, it is often convenient to put $\sup S = \infty$.

Similar to Definition 0.10 we can define:

Definition 0.11 *(Greatest lower bound) Let $S \subseteq \mathbb{R}$ and let $m_0 \in \mathbb{R}$. m_0 is called a greatest lower bound for S if*

(i) m_0 is a lower bound for S and

(ii) For every lower bound m for S we have that $m \leq m_0$

Similar to the above it follows that m_0 is uniquely determined if it exists.

If $S \subseteq \mathbb{R}$, we put $-S = \{-x \in \mathbb{R} \mid x \in S\}$. It is easily verified that m is a lower bound for S if and only $-m$ is an upper bound for $-S$ and therefore we get from the completeness axiom:

Corollary 0.12 *Every non–empty subset $S \subseteq \mathbb{R}$ which is bounded from below has a greatest lower bound. It is uniquely determined and is called the infimum of S and written* $\inf S$.

It follows of course that if S is bounded from below, then $-S$ is bounded from above and $\inf S = -\sup(-S)$. If $\inf S \in S$, we shall write $\min S$ instead of $\inf S$. If S is not bounded from below it is convenient to put $\inf S = -\infty$.

As an application we can prove

Theorem 0.13 *(Archimedes' axiom)*

The set $\mathbb{N}$ of natural numbers is not bounded from above.

Proof: We prove the theorem by contradiction so assume that $\mathbb{N}$ is bounded from above. From the completeness axiom for $\mathbb{R}$ we get that $a = \sup \mathbb{N}$ exists. If now $n \in \mathbb{N}$ is arbitrary, then

$n+1 \in \mathbb{N}$ and therefore $n+1 \leq a$ which gives that $n \leq a-1$. Since this holds for all n, $a-1$ is an upper bound for $\mathbb{N}$, but this contradicts that a was the least upper bound. $\square$

The reason for the name of this theorem is that in some textbooks it is taken as one of the axioms defining $\mathbb{N}$.

To make the reader more familar with the notions of $\sup$ and $\inf$ we prove the next theorem which will be used several times in these notes, in particular in Section 6.

Theorem 0.14 *Let $([a_n, b_n])$ be a sequence of closed and bounded intervals so that $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$ for all $n \in \mathbb{N}$. There exists $a, b \in \mathbb{R}$ with $a \leq b$ so that*

$$\cap_{n=1}^{\infty}[a_n, b_n] = [a, b]. \tag{0.8}$$

Proof: Since $[a_n, b_n] \subseteq [a_1, b_1]$ for all $n \in \mathbb{N}$ we get that b_1 is a upper bound for $\{a_n \mid n \in \mathbb{N}\}$ and that a_1 is a lower bound for $\{b_n \mid n \in \mathbb{N}\}$. Therefore we can put

$$a = \sup\{a_n \mid n \in \mathbb{N}\}$$

and

$$b = \inf\{b_n \mid n \in \mathbb{N}\}$$

and we claim that these numbers are the required ones. To see this we first observe that since the sequence of intervals is decreasing, we get that for all $n \leq m$ $a_n \leq a_m < b_m \leq b_n$. Looking on all the possible inequalities here we actually get that for all $n, m \in \mathbb{N}$ $a_n \leq b_m$. If we keep m fixed arbitrarily, this inequality shows that b_m is an upper bound for the set $\{a_n \mid n \in \mathbb{N}\}$ and therefore $a \leq b_m$. Since this works for all m, we conclude that a is a lower bound for $\{b_m \mid m \in \mathbb{N}\}$ and so $a \leq b$. We have left to prove (0.8). Since a is an upper bound for (a_n) and b is a lower bound for (b_n), we immediatelty get that $a_n \leq a \leq b \leq b_n$ which shows that the inclusion "$\supseteq$" in (0.8) holds. To get the other inclusion we let $x \in \cap_{n=1}^{\infty}[a_n, b_n]$ be arbitrary and get that $a_n \leq x \leq b_n$ for all $n \in \mathbb{N}$. This inequality shows that x is and upper bound for (a_n) and a lower bound for (b_n) and hence $a \leq x \leq b$, i.e $x \in [a, b]$. $\square$

Note that it can happen that $a = b$ so that the interval $[a, b]$ reduces to a point. In the exercise hours we will prove a much more general result [4, Exercise 5].

0.4 Properties of $\mathbb{C}$

We recall that $\mathbb{C}$ is the plane where we have put $i = (0, 1)$ so that every $z \in \mathbb{C}$ can be written as $z = a + ib$, where $a, b \in \mathbb{R}$. a is called the *real part of* z and b is called the *imaginary part* of z and we write $a = \Re z$ and $b = \Im z$. We add two complex numbers by adding their real and imaginary parts separately. Multiplication is defined so that $i^2 = -1$ and if $z = a_1 + ib_1$ and

$z_2 = a_2 + ib_2$, where $a_1, a_2, b_1, b_2 \in \mathbb{R}$, then we multiply z_1 with z_2 as usual, just remembering that $i^2 = 1$, that is

$$z_1 z_2 = (a_1 + ib_1)(a_2 + ib_2) = (a_1 a_2 - b_1 b_2) + i(b_1 a_2 + a_1 b_2).$$

If again $z = a + ib$ with $a, b \in \mathbb{R}$, the number $\overline{z} = a - ib$ is called the *complex conjugate* of z. We easily get the following formulas:

$$z + \overline{z} = 2a = 2\Re z$$

$$z - \overline{z} = 2ib = 2i\Im z.$$

The number $|z| = (a^2 + b^2)^{\frac{1}{2}}$ is called the the *absolute value* of z. It satisfies the following conditions:

(i) $z\overline{z} = |z|^2$.

(ii) $|z| \geq 0$ and $|z| = 0 \Leftrightarrow z = 0$.

(iii) For all $z_1, z_2 \in \mathbb{C}$ $|z_1 z_2| = |z_1||z_2|$.

(iv) For all $z_1, z_2 \in \mathbb{N}$ $|z_1 + z_2| \leq |z_1| + |z_2|$

The following calculation shows (i):

$$z\overline{z} = (a + ib)(a - ib) = a^2 - (ib)^2 = a^2 + b^2 = |z|^2.$$

(ii) is obvious and (iii) follows easily from (i) with $z = z_1 z_2$. To obtain (iv) we first observe that if z is written as above, then $|z| = (a^2 + b^2)^{\frac{1}{2}} \geq |a| = |\Re z|$. If now $z_1, z_2 \in \mathbb{C}$, we use (i) to get:

$$\begin{aligned} |z_1 + z_2|^2 = (z_1 + z_2)(\overline{z_1} + \overline{z_2}) &= \\ |z_1|^2 + |z_2|^2 + z_1\overline{z_2} + z_2\overline{z_1}. \end{aligned}$$

Since $\overline{z_1\overline{z_2}} = \overline{z_1} z_2$, we get from the above that

$$z_1\overline{z_2} + \overline{z_1} z_2 = 2\Re(z_1\overline{z_2}) \leq |z_1||z_2|,$$

and inserting this we get that:

$$\begin{aligned} |z_1 + z_2|^2 = |z_1|^2 + |z_2|^2 + 2\Re(z_1\overline{z_2}) \leq \\ |z_1|^2 + |z_2|^2 + 2|z_1||z_2| = (|z_1| + |z_2|)^2 \end{aligned}$$

and this shows (iv) which is called the triangle inequality.

In these notes we will develop a theory of topological spaces and will not use $\mathbb{C}$ very much since it turns out that as topological spaces $\mathbb{C}$ and $\mathbb{R}^2$ are identical. However, there are a few places where it is convenient to use complex numbers.

0.5 From Logic: The Axiom of Choice

Here we discuss an axiom from Set Theory, called the *Axiom of Choice* which turns out to be independent of the other more or less intuitive axioms from that theory. If we formulate it in ordinary human language it sounds plausible and looks more like a trivial theorem:

The Axiom of Choice

If $\{A_i \mid i \in I\}$ *is an arbitrary family of non–empty sets, then we can choose a point in each of the* A_i*'s.*

If we have to formulate this in a stringent mathematical way, we have to think about what it means "to choose". This means that one has to give a rule how to choose so the mathematical correct way of formulating the above is:

The Axiom of Choice

If $\{A_i \mid i \in I\}$ *is a family of non–empty sets, then there exists a function* $f : I \to \cup_{i\in I} A_i$ *so that* $f(i) \in A_i$ *for all* $i \in I$.

The Axiom of Choice was first formulated in 1904 by Ernst Zermelo (1871–1956) and in 1940 it was proved by Kurt Gödel (1906–1978) that it is consistent with the other axioms from Set Theory (i.e the usual more or less "intuitive" axioms). Consistent means here that if we take the usual axioms and add the axiom of choice, then we get a logical system which does not contain contradictions. In 1963 it was proved by Paul Cohen (1934–2007) that the axiom of choice is independent of the other axioms.

In the formulation above the axiom looks plausible and one has to think quite a bit to realize that the existence of the function f does not follow trivially from the other axioms. It is however much harder to accept some of the consequences of the axiom because some of them really contradict our intuition. We can for example mention that the axiom of choice implies that a ball in $\mathbb{R}^3$ can be divided into finitely many peaces which you can put together to two balls, each having the same volume as the original one! This follows from the fact that the axiom implies that there exists sets which cannot be given a volume and that such sets (the shape of which is beyond our imagination) can be put together to something nice.

After a heavy debate in the 1930's about the axiom of choice it is now accepted by most mathematicians and in modern mathematics, a set theory which consists of the usual axioms and the axiom of choice, is now used. This means that the axiom of choice is used in proofs without any reference to it. This is also the case for these notes (see e.g.. the proof of Theorem 4.19 below).

The reason for the acceptance is probably that with a axiom system including the axiom of choice as the foundation, one gets a mathematics the results of which are applicable to the existing world. In particular in physics where Mathematical Analysis is used heavily and in this area of mathematics one cannot avoid using the axiom of choice.

In the notes [6] written by this author there is a chapter on different equivalent formulations of the axiom of choice. Let us end this section giving a simple equivalent formulation which cannot be found in [6].

If X and Y are sets and $g : X \to Y$ is a function, then g is said to have a right inverse h if there exists a function $h : Y \to X$ so that $g(h(y)) = y$ for all $y \in Y$. We have the following theorem:

Theorem 0.15 *The following statements are equivalent.*

(i) The Axiom of Choice.

(ii) Every surjection has a right inverse.

Proof: $(i) \Rightarrow (ii)$: Assume (i) and let X and Y be given sets and $g : X \to Y$ be a surjection. Since g is surjective, $g^{-1}(\{y\}) \neq \emptyset$ for all $y \in Y$ and therefore the axiom of choice gives a map $f : Y \to \cup_{y\in Y} g^{-1}(\{y\}) \subseteq X$ so that $f(y) \in g^{-1}(\{y\})$ for all $y \in Y$. Clearly $g(f(y)) = y$ for all $y \in Y$ so that f is a right inverse of g.

$(ii) \Rightarrow (i)$: Assume (ii) and let $\{A_i \mid i \in I\}$ be an arbitrary family of non–empty sets. We consider the sets $A_i \times \{i\} \subseteq A_i \times I$ and define $X = \cup_{i\in I}(A_i \times \{i\}) \subseteq \cup_{i\in I} A_i \times I$. We note that the sets $A_i \times \{i\}$ are mutually disjoint and therefore X is often called the disjoint union of the A_i's; we also need the function $p : \cup_{i\in I} A_i \times I \to \cup_{i\in I} A_i$ defined by $p(x, i) = x$ for all $(x, i) \in \cup_{i\in I} A_i \times I$. We now define a function $g : X \to I$ by

$$g(x, i) = i \quad \text{for all } i \in I \text{ and all } (x, i) \in A_i \times \{i\}.$$

Note that g is well defined since the $A_i \times \{i\}$'s are mutually disjoint. By (ii) g has a right inverse $h : I \to X$ which means that $g(h(i)) = i$ for all $i \in I$. If we look on the definition of g, we see that necessarily $h(i) \in A_i \times \{i\}$ for all $i \in I$. The function $f : I \to \cup_{i\in I} A_i$ defined by $f = p \circ h$ clearly has the property that $f(i) \in A_i$ for all $i \in I$ so that (i) is satisfied. □

1 Normed vector spaces and inner product spaces

Some of the concepts we shall introduce in this section are well known from the course "Linear Algebra" and some of the results we shall prove are also well known from that course. If X is a vector space over $\mathbb{C}$, we shall call X a complex vector space and if X is a vector space over $\mathbb{R}$, we shall call X a real vector space. If it does not matter in the contents whether X is a real or complex vector space, we shall just write "vector space".

We start by introducing the concept of length of a vector in an abstract vector space, namely the concept of a norm.

Definition 1.1 *Let X be a vector space. A function $\|\cdot\| : X \to \mathbb{R}$ is called a norm if it satisfies the following conditions:*

N1 $\|x\| \geq 0$ *for all* $x \in X$.

N2 $\|tx\| = |t|\|x\|$ *for all* $x \in X$ *and all* $t \in \mathbb{C}$*, if* X *is a complex vector space (all* $t \in \mathbb{R}$ *if* X *is a real vector space).*

N3 $\|x + y\| \leq \|x\| + \|y\|$ *for all* $x \in X$ *and all* $y \in Y$*.*

N4 $\forall x \in X : \|x\| = 0 \Leftrightarrow x = 0$*.*

Remark: If we in N2 put $x = 0$ and $t = 0$, we get that $\|0\| = 0$ so the important implication in N4 is "$\Rightarrow$". N3 is called the triangle inequality.
If $\| \cdot \|$ is a norm on X, we shall call the pair $(X, \| \cdot \|)$ a *normed space*. If it is clear from the contents what the norm on X is, we shall simply talk about the normed space X.

The most obvious examples of normed spaces are of course $\mathbb{R}$ and $\mathbb{C}$ equipped with the absolute value of a number as the norm.

The most important normed spaces are $\mathbb{C}^n$ and $\mathbb{R}^n$ (where $n \in \mathbb{N}$) equipped with their Euclidian norm $\| \cdot \|_2$ defined as follows: If $\mathbf{x} \in \mathbb{C}^n$ with $\mathbf{x} = (x_1, x_2, \cdots, x_n)$, then

$$\|\mathbf{x}\|_2 = (\sum_{k=1}^{n} |x_k|^2)^{\frac{1}{2}} \tag{1.1}$$

and similarly for $\mathbf{x} \in \mathbb{R}^n$. It requires however a bit of work and some additional results to see that $\| \cdot \|_2$ satisfies the triangle inequality when $n > 1$, so we start with a simpler example and postpone the case of $\| \cdot \|_2$ to a little later.

Example 1.2 *Let* $n \in \mathbb{N}$ *and let* $X = \mathbb{C}^n$ *or* $X = \mathbb{R}^n$*. If* $\mathbf{x} \in X$ *with* $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ *where* $x_k \in \mathbb{C}$ *for all* $1 \leq k \leq n$ *(*$x_k \in \mathbb{R}$ *if* $X = \mathbb{R}^n$*), then we define* $\|\mathbf{x}\|_1$ *by*

$$\|\mathbf{x}\|_1 = \sum_{k=1}^{n} |x_k|.$$

We have to show that $\| \cdot \|_1$ *is a norm on* X*. Since obviously it satisfies N1 and N2, we concentrate on N3 and N4. To see that N3 is satisfied we let* $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in X$ *and* $\mathbf{y} = (y_1, y_2, \cdots, y_n) \in X$ *and obtain*

$$\|\mathbf{x} + \mathbf{y}\|_1 = \sum_{k=1}^{n} |x_k + y_k| \leq$$
$$\sum_{k=1}^{n} |x_k| + \sum_{k=1}^{n} |y_k| = \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1$$

where we have used that $|x_k + y_k| \leq |x_k| + |y_k|$ *for all* $1 \leq k \leq n$*. This shows that N3 holds. If* $\|\mathbf{x}\|_1 = 0$*, then*

$$\sum_{k=1}^{n} |x_k| = \|\mathbf{x}\|_1 = 0$$

and since all terms in the sum are non–negative, we get that $x_k = 0$ *for all* $1 \leq k \leq n$ *which means that* $\mathbf{x} = 0$*. Hence N4 is also satisfied.*

A more complicated example is the following:

Example 1.3 *let S be a set and define $B(S,\mathbb{C})$ to be the set of all bounded functions $f : S \to \mathbb{C}$. If $f, g \in B(S,\mathbb{C})$ and $t \in \mathbb{C}$, we define the functions $f + g; S \to \mathbb{C}$ and $tf : S \to \mathbb{C}$ by*

$$(f + g)(s) = f(s) + g(s) \quad \textit{for all } s \in S$$

$$(tf)(s) = tf(s) \quad \textit{for all } s \in S.$$

$$\|f\|_\infty = \sup\{|f(s)| \mid s \in S\}$$

We wish to show that $B(S,\mathbb{C})$ is a vector space and that $\|\cdot\|_\infty$ is a norm on $B(S,\mathbb{C})$ (do not think about the subscript ∞; it is simply a tradition to use that subscript when a norm is defined by a supremum). Hence let $t \in \mathbb{C}$, $f, g \in B(S,\mathbb{C})$ be given. Since $|f(s)| \leq \|f\|_\infty$ for all $s \in S$, we get that $|tf(s)| \leq |t|\|f\|_\infty$ which shows that tf is a bounded function, i.e. $tf \in B(S,\mathbb{C})$. It is also clear that

$$\|tf\|_\infty = \sup_{s\in S} |t||f(s)| = |t| \sup_{s\in S} |f(s)| = |t|\|f\|_\infty$$

so that N2 is satisfied.

Since also $|g(s)| \leq \|g\|_\infty$ for all $s \in S$ we get that

$$|f(s) + g(s)| \leq |f(s)| + |g(s)| \leq \|f\|_\infty + \|g\|_\infty \tag{1.2}$$

for all $s \in S$ which gives that $f + g \in B(S,\mathbb{C})$. It now follows that $B(S,\mathbb{C})$ is a vector space under the addition and scalar multiplication given above. In addition (1.2) shows that $\|f\|_\infty + \|g\|_\infty$ is an upper bound for the set $\{|f(s) + g(s)| \mid s \in S\}$ and since by definition $\|f + g\|_\infty$ is the least upper bound for that set we must have

$$\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$$

which shows that N3 is satisfied. If $\|f\|_\infty = 0$, then

$$|f(s)| \leq \|f\|_\infty = 0 \quad \textit{for all } s \in S$$

which shows that $f(s) = 0$ for all $s \in S$ so that f is the 0–function, that is the 0–vector in $B(S,\mathbb{C})$. Since trivially N1 is satisfied, we have proved that $\|\cdot\|_\infty$ is a norm on $B(S,\mathbb{C})$. The space $B(S,\mathbb{R})$ can of course be treated similarly.

We now define inner products and inner product spaces which should be known from Linear Algebra.

Definition 1.4 *Let H be a complex vector space. A function $(\cdot,\cdot)\colon H \times H \to \mathbb{C}$ is called an inner product on H if*

$$\forall x, y \in H\colon (x, y) = \overline{(y, x)} \tag{1.3}$$

$$\forall x, y, z \in H\colon (tx + y, z) = t(x, z) + (y, z) \tag{1.4}$$

$$\forall x \in H\colon (x, x) \geq 0 \tag{1.5}$$

$$(x, x) = 0 \implies x = 0. \tag{1.6}$$

Remark: If H is a real vector space, then $(x, y) \in \mathbb{R}$ for all $x, y \in H$, and (1.3) then looks like

$$\forall x, y \in H\colon (x, y) = (y, x). \tag{1.7}$$

If H is a vector space with inner product $(\cdot, \cdot)$, we put $\|x\| = \sqrt{(x, x)}$ for all $x \in H$. Note that $\|x\| \geq 0$ for all $x \in H$ and that $\|x\| = 0$ implies that $x = 0$. The pair $(H, (\cdot, \cdot))$ is called *an inner product space*.

Example 1.5 *Let $n \in \mathbb{N}$. If $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{C}^n$ and $\mathbf{y} \in \mathbb{C}^n$ we define*

$$< \mathbf{x}, \mathbf{y} >= \sum_{k=1}^{n} x_k \overline{y_k}$$

or if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$< \mathbf{x}, \mathbf{y} >= \sum_{k=1}^{n} x_k y_k.$$

$< \cdot, \cdot >$ is an inner product on $\mathbb{C}^n$ ($\mathbb{R}^n$) and for all $\mathbf{x} \in \mathbb{C}^n$ (for all $\mathbf{x} \in \mathbb{R}^n$) we have

$$\|\mathbf{x}\|_2 =< \mathbf{x}, \mathbf{x} >^{\frac{1}{2}}$$

where $\|\cdot\|_2$ is defined in (1.1). In the course Linear Algebra $< \mathbf{x}, \mathbf{y} >$ was called the dot–product between $\mathbf{x}$ and $\mathbf{y}$ and denoted $\mathbf{x} \cdot \mathbf{y}$. We shall not use the dot–notation here since the dot has a tendency of disappearing in writing.

Exercise 1.6 *Prove the statements in Example 1.5.*

We shall now prove an important inequality for inner products which actually also will give that $\|\cdot\|$ defined above is a norm.

Theorem 1.7 *(Cauchy-Schwartz' inequality)*

If H is a vector space with inner product $(\cdot, \cdot)$, we have:

$$|(x, y)| \leq \|x\| \|y\| \quad \textit{for alle } x \in H. \tag{1.8}$$

Proof: If $t \in \mathbb{R}$, $x \in H$ og $y \in H$ are arbitrary, we get that

$$\begin{aligned} 0 &\leq (tx + y, tx + y) = t^2\|x\|^2 + t(x, y) + t(y, x) + \|y\|^2 \\ &= t^2\|x\|^2 + t[(x, y) + \overline{(x, y)}] + \|y\|^2 \\ &= \|x\|^2 t^2 + 2\Re(x, y)t + \|y\|^2. \end{aligned} \tag{1.9}$$

Put now $p(t) = \|x\|^2 t^2 + 2\Re(x, y)t + \|y\|^2$ for all $t \in \mathbb{R}$. p is then a polynomium of degree 2 which (1.9) satisfies $p(t) \geq 0$ for all $t \in \mathbb{R}$. Therefore p has a most one root which implies that the discriminand D of p is non–positive, hence

$$0 \geq D = 4[\Re(x, y)]^2 - 4\|x\|^2\|y\|^2. \tag{1.10}$$

Dividing by 4 and taking the square root this gives

$$|\Re(x,y)| \le \|x\|\|y\|. \tag{1.11}$$

If H is a real vector space, (1.11) is the same as (1.8) and we have finished the proof in that case. If H is a complex vector space, we can continue as follows:

Put $\alpha = \frac{|(x,y)|}{(x,y)}$ if $(x,y) \neq 0$ and $\alpha = 1$, if $(x,y) = 0$. Since $|\alpha| = 1$ and $|(x,y)| \ge 0$, (1.11) gives:

$$|(x,y)| = \alpha(x,y) = (\alpha x, y) = \Re(\alpha x, y) \le \|\alpha x\|\|y\| = |\alpha|\|x\|\|y\| = \|x\|\|y\|, \tag{1.12}$$

which exactly is (1.8). $\square$

We are now able to prove:

Theorem 1.8 *If H is a complex or real vector space with inner product $(\cdot,\cdot)$ and we define $\|x\| = (x,x)^{\frac{1}{2}}$ for all $x \in H$, then $\|\cdot\|$ is a norm on H.*

Proof: We have to show that $\|\cdot\|$ satisfies N1–N4. N1 and N4 follows directly from (1.5) and (1.6). To prove N2 we let $t \in \mathbb{C}$ and $x \in H$ be given arbitrarily. A little computation shows that

$$\|tx\|^2 = (tx, tx) = t\bar{t}\|x\|^2 = |t|^2\|x\|^2 \tag{1.13}$$

which proves N2.

Let $x, y \in H$ be arbitrary. If we put $t = 1$ in (1.9) and apply the Cauchy–Schwartz' inequality we get

$$\begin{aligned} \|x+y\|^2 = (x+y, x+y) &= \|x\|^2 + \|y\|^2 + 2\Re(x,y) \le \\ \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| &= (\|x\| + \|y\|)^2 \end{aligned} \tag{1.14}$$

and hence

$$\|x+y\| \le \|x\| + \|y\|$$

which proves the triangle inequality. $\square$

Corollary 1.9 *$\|\cdot\|_2$ as defined in* (1.1) *is a norm on $\mathbb{C}^n$ and $\mathbb{R}^n$.*

Proof: This follows directly from Theorem 1.8 and Example 1.5. $\square$

Let us finish this section with yet another example of an inner product space.

Example 1.10 *Let $a, b \in \mathbb{R}$ with $a < b$ and let $C([a,b], \mathbb{C})$ denote the set of all continuous functions $f : [a,b] \to \mathbb{C}$. If we define sum and multiplication by scalars as in Example 1.3, $C([a,b], \mathbb{C})$ becomes a complex vector space. If $f, g \in C([a,b], \mathbb{C})$, we put*

$$(f,g) = \int_a^b f(t)\overline{g(t)}dt.$$

It is readily verified that $(\cdot,\cdot)$ is an innner product on $C([a,b],\mathbb{C})$ and that the corresponding norm is given by:

$$\|f\|_2 = (\int_a^b |f(t)|^2 dt)^{\frac{1}{2}} \quad \textit{for alle } f \in C([a,b],\mathbb{C}).$$

Exercise 1.11 *Prove the statements in Example 1.10*

2 Metric spaces and their topologies

In Section 1 we have seen how to define the lenght of a vector in a normed vector space, namely as the norm of the vector. In this section we shall introduce so–called metric spaces where it is possible to define distances between points. We start with:

Definition 2.1 *Let S be a non–empty set. A function $d : S \times S \to \mathbb{R}$ is called a metric (or a distance function) if it satisfies the following conditions:*

M1 $d(x,y) \geq 0$ *for all* $x,y \in S$.

M2 $d(x,y) = d(y,x)$ *for all* $x,y \in S$.

M3 $d(x,y) \leq d(x,z) + d(z,y)$ *for all* $x,y,z \in S$.

M4 $\forall x,y \in S : d(x,y) = 0 \Leftrightarrow x = y$

The pair (S,d) is called a metric space.

If you look on it intuitively M1, M2 and M4 are natural requirements for a distance function; the distance between two points should be non–negative, the distance from x to y should be the same as the distance from y to x, and the distance between two points should be 0 if and only if they are equal. If you think about how you measure distances in the plane or in our physical world (3–dimensional space), you will also see that M3 is a natural requirement. M3 is called *the triangle inequality*.

If it is clear from the contents what the metric d is, we shall often write "the metric space S" instead of "the metric space(S,d)".

Let us now look on some examples.

Example 2.2 *Let S be any non–empty set and define d by $d(x,y) = 1$ if $x,y \in S$ with $x \neq y$ and $d(x,y) = 0$ if $x,y \in S$ with $x = y$. It is clear from the definition that M1, M2 and M4 are satisfied. To prove M3 we let $x,y,z \in S$ be arbitrary. If $x = y$, then clearly*

$$d(x,y) = 0 \leq d(x,z) + d(z,y).$$

If $x \neq y$, then either $x \neq z$ or $z \neq y$ and hence

$$d(x,y) = 1 \leq d(x,z) + d(z,y)$$

which shows that also M3 holds. Hence d is a metric on S. It is called the discrete metric on S.

This example is a bit exoctic, though sometimes useful. The next theorem will provide some more important examples.

Theorem 2.3 *If $X, \|\cdot\|)$ is a normed space and d is defined by*

$$d(x,y) = \|x-y\| \quad \textit{for all } x,y \in X,$$

then d is a metric on X turning it into metric space. d is called the metric induced by the norm on X.

Proof: Let $x,y,z \in X$ be arbitrary. By N1 $d(x,y) = \|x-y\| \geq 0$ so M1 holds. By N2 we get

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

so M2 holds. By N3 we have

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

so that M3 holds. Finally N4 gives that $d(x,y) = \|x-y\| = 0$ if and only if $x-y=0$, if and only if $x=y$. Hence d is a metric on X. $\square$

Combining Theorem 2.3 with Theorem 1.8 we get that every inner product space is a metric space in a natural manner. In particular Corollary 1.9 gives:

Corollary 2.4 *Let $n \in \mathbb{N}$. d_2 defined by*

$$d_2(x,y) = \|x-y\|_2 \quad \textit{for all } x,y \in \mathbb{C}^n \ (x,y \in \mathbb{R}^n)$$

is a metric on $\mathbb{C}^n$, respectively $\mathbb{R}^n$.

d_2 is called the Euclidian metric on $\mathbb{R}^n$ or $\mathbb{C}^n$ and is the one we will use on these spaces unless otherwise stated. The reason is of course that this is the distance concept we normally use in real life on $\mathbb{R}$, $\mathbb{R}^2$, and $\mathbb{R}^3$. If we write it out in coordinates we get for all $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{C}^n$ and all $\mathbf{y} = (y_1, y_2, \cdots, y_n) \in \mathbb{C}^n$:

$$d_2(\mathbf{x},\mathbf{y}) = (\sum_{k=1}^{n} |x_k - y_k|^2)^{\frac{1}{2}}.$$

We could also use the norm $\|\cdot\|_1$ from Example 1.2 to define a metric d_1 on $\mathbb{C}^n$ or $\mathbb{R}^n$. Indeed, if $\mathbf{x}$ and $\mathbf{y}$ are as above, then

$$d_1(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_1 = \sum_{k=1}^{n} |x_k - y_k|.$$

Any carpenter who is going to make a table for you will of course think that you are nuts if you insist that all measurements in $\mathbb{R}^2$ should be made using the metric d_1. However, in some mathematical proofs where metric spaces are applied , e.g. the proof of the existence of solutions to certain differential equations, d_1 is used instead of d_2 of technical reasons: simply, no square roots appear in the definition of d_1. In the exercise hours you will see that d_1 and d_2 are equivalent metrics (see the definition below).

In the rest of this section (S, d) will denote a fixed metric space.

If $A \subseteq S$ is non–empty, we can consider the restriction d_A of d to $A \times A$, that is

$$d_A(x, y) = d(x, y) \quad \text{for all } x, y \in A.$$

Clearly d_A is a metric on A and is called the metric on A *induced by* d. Thus any non–empty subset of a metric space is itself a metric space with the induced metric. We shall not use the notation d_A much in the sequel, but will simply write d and remember that when we consider A as a metric space, we only evaluate d in points in $A \times A$. We shall also need:

Proposition 2.5 *Let (X, d_X) and (Y, d_Y) be metric spaces. If $\mathbf{x}, \mathbf{y} \in X \times Y$, say $\mathbf{x} = (x_1, y_1)$ and $\mathbf{y} = (x_2, y_2)$ where $x_1, x_2 \in X$ and $y_1, y_2 \in Y$. If we put*

$$d(\mathbf{x}, \mathbf{y}) = \max\{d_X(x_1, x_2), d_Y(y_1, y_2)\},$$

then d is a metric on $X \times Y$.

If $x \in S$ and $r > 0$, the set $B(x, r) = \{y \in S \mid d(x, y) < r\}$ is called the ball with center x and radius r. We shall sometimes write the open ball with center x and radius r to emphasize the sharp inequality in a definition. $B(x)$ is called a ball with center x if $B(x) = B(x, r)$ for some $r > 0$. The closed ball with center x and radius r is the set $B_c(x, r) = \{y \in S \mid d(x, y) \leq r\}$.

To illustrate that balls may look differently than anticipated given our intuition from $\mathbb{R}^n$, we give the following example:

Example 2.6 *We consider $\mathbb{Z} \subseteq \mathbb{R}$ with the metric induced by the metric on $\mathbb{R}$. The ball with center 1 and radius 1 is*

$$B(1, 1) = \{n \in \mathbb{Z} \mid |n - 1| < 1\} = \{1\},$$

while the closed ball with center 1 and radius 1 is the set $B_c(x, r) = \{n \in \mathbb{Z} \mid |n - 1| \leq 1\} = \{0, 1, 2\}$.

We now introduce a very important concept.

Definition 2.7 *A subset $G \subseteq S$ is called open if for all $x \in G$ there exists an $r > 0$ so that $B(x, r) \subseteq G$. A subset $F \subseteq S$ is called closed if $S \setminus F$ is open.*

As a quick example we see that an interval of the form $]a, b[$ with $a < b$ is an open subset of $\mathbb{R}$. If we in $\mathbb{R}^2$ let F be the line segment between the points $(1, 2)$ and $(3, 4)$ including the endpoints, F is a closed subset of $\mathbb{R}^2$. Make a drawing of the situation! If we let A be the same line segment without the endpoints, A is neither open nor closed.

We need the following lemma which is a consequence of the triangle inequality.

Lemma 2.8 *Let $x \in S$ and $r > 0$. The open ball $B(x, r)$ is an open set. The closed ball $B_c(x, r)$ is a closed set.*

Proof: Let $y \in B(x, r)$ be arbitrary. Since $d(x, y) < r$, $s = r - d(x, y) > 0$ and we wish to show that $B(y, s) \subseteq B(x, r)$. To do this we let $z \in B(y, s)$ and get by the triangle inequality:

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

which shows that $z \in B(x, r)$. Hence we have proved that $B(y, s) \subseteq B(x, r)$. This shows that $B(x, r)$ is open.

To prove that $B_c(x, r)$ is closed we let $y \in S \setminus B_c(x, r)$ be arbitrary and observe that $s = d(x, y) - r > 0$. We wish to show that $B(y, s) \subseteq S \setminus B_c(x, r)$. Hence let $z \in B(y, s)$ and use the triangle inequaly to get

$$d(x, y) \leq d(x, z) + d(y, z) < d(x, z) + s$$

so that $d(x, z) > d(x, y) - s = r$. This shows the inclusion. Thus we have proved that $S \setminus B_c(x, r)$ is open and hence $B_c(x, r)$ is closed. $\square$

Corollary 2.9 *Open intervals in $\mathbb{R}$ are open subsets of $\mathbb{R}$. Closed intervals in $\mathbb{R}$ are closed subsets of $\mathbb{R}$.*

The following proposition is very useful.

Proposition 2.10 *(S, d) has the following properties:*

(i) If $x, y \in S$ with $x \neq y$, there exist a ball $B(x)$ with center x and a ball $B(y)$ with center y so that $B(x) \cap B(y) = \emptyset$

(ii) Any singleton $\{x\} \subseteq S$ is closed.

Proof: (i): Let $x, y \in S$ with $x \neq y$ be arbitrary and put $r = \frac{1}{2}d(x, y) > 0$. We claim that $B(x, r) \cap B(y, r) = \emptyset$. Assume namely that this is not the case and choose a point $z \in B(x, r) \cap B(y, r)$. The triangle inequality then gives

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

which is a contradiction. This proves (i).

To prove (ii) we let $x \in S$ be arbitrary and have to prove that $S \setminus \{x\}$ is open. If $y \in S \setminus \{x\}$, (i) gives that we can find balls $B(x)$ and $B(y)$ with centers x and y respectively so that $B(x) \cap B(y) = \emptyset$; in particular $x \notin B(y)$ which means that $B(y) \subseteq S \setminus \{x\}$. This proves (ii). $\square$

The next exercise shows that whether a given set is open or closed or neither depends on which metric space we consider it a subset of.

Exercise 2.11 *Let $\mathbb{Z}$ be equipped with the metric induced by $\mathbb{R}$ as in Example 2.6. Prove that $\{1\}$ considered as a subset of $\mathbb{Z}$ is both open and closed. Prove next that if $\{1\}$ is considered as a subset of $\mathbb{R}$, then it is closed, but not open.*

In higher dimensions it can often be difficult to decide whether a given set is open just by using the definition. In the next section we shall get some strong tools to decide that. Here we can prove:

Theorem 2.12 *The following statements are satisfied.*

(i) An arbitrary union of open subsets of S is open.

(ii) A finite intersection of open subsets of S is open.

(iii) An arbitrary intersection of closed subsets of S is closed.

(iv) A finite union of closed subsets of S is closed.

(v) S and $\emptyset$ are both open and closed.

Proof: To prove (i) we let I be an arbitrary index set and let $\{G_j \mid j \in I\}$ be a family of open subsets of S. We have to prove that $G = \cup_{j\in I} G_j$ is open. If $x \in G$ is given, there exists a $j_0 \in I$ so that $x \in G_{j_0}$ and since G_{j_0} is open, there exists an $r > 0$ with $B(x,r) \subseteq G_{j_0}$. This gives us immediately that

$$B(x,r) \subseteq G_{j_0} \subseteq \cup_{j\in I} G_j = G$$

and hence G is open.

(ii): let $n \in \mathbb{N}$ and finitely many open subsets $H_k \subseteq S$, $1 \leq k \leq n$ be given and put $H = \cap_{k=1}^n H_k$. We have to show that H is open so choose $x \in H$ arbitrarily and observe that this implies that $x \in H_k$ for all $1 \leq k \leq n$. For each $1 \leq k \leq n$ H_k is open, and hence we can find an $r_k > 0$ so that $B(x, r_k) \subseteq H_k$. Put now $r = \min\{r_k \mid 1 \leq k \leq n\}$ and note that since there are only finitely many r_k's and r is the smallest of them, also $r > 0$. We claim that $B(x,r) \subseteq H$. If we fix a $1 \leq k \leq n$, we get that $B(x,r) \subseteq B(x,r_k) \subseteq H_k$. Since this holds for all $1 \leq k \leq n$ we get that actually $B(x,r) \subseteq \cap_{k=1}^n H_k = H$. Thus we have proved that H is open.

(iv): Let I be an arbitrary index set and let $\{F_j \mid j \in I\}$ be a family of closed subsets of S and put $F = \cap_{k\in I} F_j$. Since for each $j \in I$ F_j is a closed set, $S \setminus F_j$ is open. Standard set manipulations give us $S \setminus F = \cup_{j\in I}(S \setminus F_j)$ which is open by (i). Hence F is closed.

In a similar manner (iv) follows by taking complements and use (ii).

(v): Any ball with center in a point in S is by definition contained in S which shows that S is open and hence $\emptyset$ is closed. $\emptyset$ is open since any statement that starts with "for all $x \in \emptyset$" is a true statement. Therefore also S is closed. $\square$

If we take infinite intersections of open subsets of S, the conclusion in (ii) need not be true and similarly if we take infinite unions of closed subsets of S, the conclusion in (iv) need not be true. Here is an example:

Example 2.13 *For every $n \in \mathbb{N}$ we consider the interval $]-\frac{1}{n}, \frac{1}{n}[\subseteq \mathbb{R}$ which is clearly open. We claim that $\{0\} = \cap_{n=1}^{\infty}]-\frac{1}{n}, \frac{1}{n}[$. Clearly "$\subseteq$" holds and if $x \in \cap_{n=1}^{\infty}]-\frac{1}{n}, \frac{1}{n}[$, then $-\frac{1}{n} < x < \frac{1}{n}$ for all $n \in \mathbb{N}$ so letting $n \to \infty$, we obtain that $x = 0$. Hence the set equality above holds. But $\{0\}$ is clearly not an open subset of $\mathbb{R}$.*

The example shows that even in the most simple cases the conclusion of (ii) in Theorem 2.12 is untrue if we take infinite intersections. This means that one has to be very careful.

Let us define:

$$\mathcal{G}_d = \{G \subseteq S \mid G \text{ is open in } S\}. \tag{2.1}$$

Concepts that can be characterized solely using $\mathcal{G}_d$ are called *topological concepts*. We shall see below that many concepts are topological concepts though it apriori does not seem so. We need the following definition:

Definition 2.14 *Two metrics d and σ on S are called equivalent if they determine the same class of open sets, that is $\mathcal{G}_d = \mathcal{G}_\sigma$.*

In the next result we denote balls in (S, d) by B^d and balls in (S, σ) by B^σ

Proposition 2.15 *Two metrics d and σ are equivalent, if and only if*

(i) $\forall x \in S \quad \forall B^d(x) \quad \exists B^\sigma(x) : B^\sigma(x) \subseteq B^d(x)$,

(ii) $\forall x \in S \quad \forall B_1^\sigma(x) \quad \exists B_1^d(x) : B_1^d(x) \subseteq B_1^\sigma(x)$.

Proof: Assume first that d and σ are equivalent. Let $x \in S$ and a ball $B^d(x)$ be given. Since the two metrics are equivalent, $B^d(x)$ is open in (S, σ) and therefore there exists a ball $B^\sigma(x)$ so that $B^\sigma(x) \subseteq B^d(x)$. Hence (i) holds. By interchanging the roles of d and σ in this argument we get (ii).

Assume next that (i) and (ii) are satisfied and let $G \in \mathcal{G}_d$ be arbitrary. If $x \in G$ we can by definition find a ball $B^d(x) \subseteq G$. By (i) we can find a ball $B^\sigma(x) \subseteq B^d(x) \subseteq G$. Since $x \in G$ was arbitrary, we get that G is open in (S, σ), i.e $G \in \mathcal{G}_\sigma$. Hence we have proved that $\mathcal{G}_d \subseteq \mathcal{G}_\sigma$. By interchanging the roles of d and σ and using (ii) we get the other inclusion. □

As mentioned earlier we shall in the exercise hours see that on $\mathbb{C}^n$ and $\mathbb{R}^n$ the two metrics d_1 and d_2 are equivalent. Here we will give another example, but first we need a useful lemma.

Lemma 2.16 *The following two statements are equivalent for a metric space (S, d):*

(i) Every singleton is an open subset of S.

(ii) Every subset of S is open.

Proof: Assume first that (i) holds, that is $\{x\}$ is open for all $x \in S$. If $A \subseteq S$ is an arbitrary subset, then clearly $A = \cup_{x \in A}\{x\}$. Theorem 2.12 (i) now gives that A is open.

If (ii) holds, then every subset of S is open so in particular $\{x\}$ is open for every $x \in S$. □

Example 2.17 *Let d denote the metric on $\mathbb{Z}$ induced by the usual metric on $\mathbb{R}$ as in Example 2.6. In Exercise 2.11 we saw that $\{1\}$ was open in $(\mathbb{Z}, d)$ and in the same way we can of course get that $\{n\}$ is open for every $n \in \mathbb{Z}$. Using Lemma 2.16 we then get that every subset of $\mathbb{Z}$ is open in $(\mathbb{Z}, d)$. Let σ be the discrete metric on $\mathbb{Z}$ as defined in Example 2.11. If $n \in \mathbb{Z}$, then $B^\sigma(n, 1) = \{n\}$ and hence $\{n\}$ is open $(\mathbb{Z}, \sigma)$. Again Lemma 2.16 gives that every subset of $\mathbb{Z}$ is open in $(\mathbb{Z}, \sigma)$. This shows that d and σ are equivalent metrics on $\mathbb{Z}$.*

For sequences in metric spaces we have the following definition.

Definition 2.18 *A sequence $(x_n) \subseteq S$ is said to converge to a point $x \in S$ (notation: $x_n \to x$ for $n \to \infty$) if*

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \in \mathbb{N} : n \geq n_0 \Rightarrow d(x, x_n) < \varepsilon. \tag{2.2}$$

x is called the limit of (x_n) or the limit point of (x_n).

If $x_n \to x$ for $n \to \infty$, we shall often write $x = \lim_{n\to\infty} x_n$ which in particular is useful when a limit point occurs in some calculations. Instead of saying "(x_n) converges to x" we shall often say "(x_n) tends to x".

Note also that since (2.2) must hold for all $\varepsilon > 0$, it does not matter whether we write "$< \varepsilon$" or "$\leq \varepsilon$". Finally we observe that $x_n \to x$ in S if and only if $d(x, x_n) \to 0$ for $n \to \infty$.

Example 2.19 *It is assumed well known to the reader that $\frac{1}{n} \to 0$ in $\mathbb{R}$ for $n \to \infty$ and this was actually used in Example 2.13. If the reader is uncomfortable with this, here is a quick proof. Let $\varepsilon > 0$ be arbitrarily given. By Theorem 0.13 there exists an $n_0 \in \mathbb{N}$ so that $n_0 > \frac{1}{\varepsilon}$ and hence if $n \geq n_0$, $\frac{1}{n} \leq \frac{1}{n_0} < \varepsilon$, which proves the statement.*

We really need the following proposition.

Proposition 2.20 *Any sequence $(x_n) \subseteq S$ has at most one limit point.*

Assume that $x_n \to x \in S$ and $x_n \to y \in S$ for $n \to \infty$ and let $\varepsilon > 0$ be arbitrary. Since $x_n \to x$, there exists an $n_1 \in \mathbb{N}$ so that $d(x, x_n) < \varepsilon$ for all $n \geq n_1$. Similarly, since $x_n \to y$, there exists an $n_2 \in \mathbb{N}$ so that $d(y, x_n) < \varepsilon$ for all $n \geq n_2$. If we put $m = \max(n_1, n_2)$, then $d(x, x_m) < \varepsilon$ since $m \geq n_1$ and $d(y, x_m) < \varepsilon$ since $m \geq n_2$. The triangle inequality gives that

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

Since ε was arbitrary, we have proved that $d(x, y) < 2\varepsilon$ for all $\varepsilon > 0$. Since the left hand side of this inequality does not depend on ε, we can let $\varepsilon \to 0$ to obtain that $d(x, y) = 0$, hence $x = y$. $\square$

We have the following proposition:

Proposition 2.21 *Let $(x_n) \subseteq S$ be a sequence and let $x \in S$. $x_n \to x$ if and only if:*

$$\forall U \in \mathcal{G}_d \quad \text{with } x \in U \quad \exists n_0 \in \mathbb{N} : n \geq n_0 \Rightarrow x_n \in U.$$

Consequently convergence of sequences in S is a topological property.

The last example of this section is based on Example 1.3:

Example 2.22 *We let $B([0,1],\mathbb{C})$ denote the space of all bounded functions $f : [0,1] \to \mathbb{C}$ equipped with the norm $\|\cdot\|_\infty$ as defined in Example 1.3. The corresponding metric d_∞ is given by $d_\infty(f,g) = \|f-g\|_\infty$ for all $f,g \in B([0,1],\mathbb{C})$. If we for every $n \in \mathbb{N}$ define the function $f_n(t) = \frac{1}{n}t$ for all $t \in [0,1]$, then clearly $0 \le f_n(t) \le \frac{1}{n}$ for all $t \in [0,1]$ so that $f_n \in B([0,1],\mathbb{C})$ with $\|f_n\|_\infty \le \frac{1}{n}$. Therefore $d_\infty(f_n,0) = \|f_n\|_\infty \le \frac{1}{n} \to 0$ for $n \to \infty$ and $f_n \to 0$ in $B([0,1],\mathbb{C})$. This is actually not surprising.*

3 Continuous functions on metric spaces

In this section we shall define continuous functions between metric spaces and investigate their properties.

We tacitly assume that the reader is familar with the continuity properties of the classical functions defined on the $\mathbb{R}$ or intervals of $\mathbb{R}$. However, let us recall that if $I \subseteq \mathbb{R}$ is an interval and $f : I \to \mathbb{R}$ is a function, then f is said to be continuous in a point $x_0 \in I$, if

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in I : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon. \tag{3.1}$$

Intuitively speaking, this means that when x gets close to x_0, then $f(x)$ gets close to $f(x_0)$. In a metric space it has meaning to say that two points are close, namely if their distance is small and therefore it has meaning to generalize (3.1) to metric spaces in order to define continuity of functions between metric spaces. One should of course not generalize things for their own sake, but the reader will soon find that the concept of continuity of functions between metric spaces is extremely useful.

In the rest of this section we let (X, d_X) and (Y, d_Y) denote metric spaces

Definition 3.1 *A function $f : X \to Y$ is said to be continuous in the point $x_0 \in X$, if the following condition holds:*

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in X : d_X(x, x_0) < \delta \Rightarrow d_Y(f(x), f(x_0)) < \varepsilon. \tag{3.2}$$

f is said to be continuous if it is continuous in all points of X.

The reader is invited to compare equations (3.1) and (3.2) to check that (3.2) is really an extension of (3.1).

Remark: If we let B_X denote balls in X and B_Y denote balls in Y we get that (3.2) is equivalent to

$$\forall B_Y(f(x_0)) \quad \exists B_X(x_0) : f(B_X(x_0)) \subseteq B_Y(f(x_0)). \tag{3.3}$$

Like in the case of functions from $\mathbb{R}$ to $\mathbb{R}$ one can check continuity by working with sequences. We have:

Theorem 3.2 *Let $f : X \to Y$ be a function and $x_0 \in X$. The following two statements are equivalent:*

(i) f is continuous in x_0.

(ii) For every sequence $(x_n) \subseteq X$ we have that if $x_n \to x_0$, then $f(x_n) \to f(x_0)$.

Proof: (i)⇒(ii): Assume (i) holds, let $(x_n) \subseteq X$ be an arbitrary sequence with $x_n \to x_0$, and let $\varepsilon > 0$ be arbitrary. Since f is continuous in x_0, we can find a $\delta > 0$ so that if $x \in X$ with $d_X(x, x_0) < \delta$, then $d_Y(f(x), f(x_0)) < \varepsilon$. Further, since $x_n \to x$, there exists an $n_0 \in \mathbb{N}$ so that $d_X(x, x_n) < \delta$ for all $n \geq n_0$. If we now fix an arbitrary $n \geq n_0$, then $d_X(x_0, x_n) < \delta$ and therefore by the above $d_Y(f(x_0), f(x_n)) < \varepsilon$. Hence we have found an n_0 so that whenever $n \geq n_0$, then $d_Y(f(x_0, f(x_n)) < \varepsilon$. This is exactly that $f(x_n) \to f(x_0)$.

(ii)⇒(i): We assume that f is not continuous in x_0 and have to find a sequence $(x_n) \in X$ so that (x_n) converges to x_0, but $(f(x_n))$ does not converge to $f(x_0)$. The discontinuity of f in x_0 means that:

$(*)$ There exists an $\varepsilon > 0$ so that for all $\delta > 0$ there exists an $x_\delta \in X$ so that $d_X(x_0, x_\delta) < \delta$ and $d_Y(f(x_0, x_\delta) \geq \varepsilon$.

Think this over!!

Since $(*)$ holds for all $\delta > 0$, we can for an arbitrary $n \in \mathbb{N}$, put $\delta = \frac{1}{n}$ and $(*)$ will provide us with an $x_n \in X$, so that $d_X(x_0, x_n) < \frac{1}{n}$ and $d_Y(f(x_0), f(x_n)) \geq \varepsilon$. Think that over too!!. This gives us on one hand that $d_X(x_0, x_n) \to 0$ and hence $x_n \to x_0$. One the other hand $d_Y(f(x_0), f(x_n)) \geq \varepsilon$ for all $n \in \mathbb{N}$ so $(f(x_n))$ does not converge to $f(x_0)$. Hence we have proved what we wanted. $\square$

The next theorem is well known for classical continuous functions.

Theorem 3.3 *Let (X, d_X), (Y, d_Y), and (Z, d_Z) be metric spaces If $f : X \to Y$ is continuous in $x_0 \in X$ and $g : Y \to Z$ is continuous in $f(x_0) \in Y$, then the composition $g \circ f$ is continuous in x_0.*

In particular, if f and g are continuous, then $g \circ f$ is continuous.

Proof: To prove it we shall apply Theorem 3.2. Let $(x_n) \in X$ be an sequence so that $x_n \to x_0$. Since f is continuous in x_0, $f(x_n) \to f(x_0)$, and since g is continuous in $f(x_0)$, $g(f(x_n)) \to g(f(x_0))$. According to Theorem 3.2 this implies that $g \circ f$ is continuous in x_0. $\square$

We now wish to prove a theorem which relates continuity to open sets. In fact, it turns out that the question of continuity of a given function $f : X \to Y$ is a topological property.

Theorem 3.4 *Let $f : X \to Y$ be a function. The following three statements are equivalent:*

(i) f is continuous.

(ii) For every open set $G \subseteq Y$ $f^{-1}(G)$ is an open subset of X.

(iii) For every closed set $F \subseteq Y$ $f^{-1}(F)$ is a closed subset of X.

Proof: (i)⇒(ii): We assume that f is continuous and let $G \subseteq Y$ be an arbitrary open subset. If $f^{-1}(G) = \emptyset$, there is nothing to prove so assume that $f^{-1}(G) \neq \emptyset$ and let $x \in f^{-1}(G)$ be arbitrary. We have to find a ball $B_X(x)$ with center x so that $B_X(x) \subseteq f^{-1}(G)$.

Since $f(x) \in G$ and G is open, we can find a ball $B_Y(f(x))$ with center $f(x)$ so that $B_Y(f(x)) \subseteq G$. Since f is continuous in x, we can according (3.3) in the remark just after Definition 3.1 find a ball $B_X(x)$ so that $f(B_X(x)) \subseteq B_Y(f(x)) \subseteq G$. Taking f^{-1} on both sides of this inequality we get that $B_X(x) \subseteq f^{-1}(G)$ and we have proved what we wanted.

(ii)⇒(i): We assume (ii), let $x \in X$ be arbitrary, and have to prove that f is continuous in x. If $B_Y(f(x))$ is an arbitrary open ball with center $f(x)$, then by (ii) $f^{-1}(B_Y(f(x))$ is an open subset of X which clearly contains x. Therefore there exists a ball $B_X(x)$ with center x so that $B_X(x) \subseteq f^{-1}(B_Y(f(x))$, or equivalently $f(B_X(x)) \subseteq B_Y(f(x))$. (3.3) now gives that f is continuous in x.

To see that (ii)⇔(iii) we just observe that $F \subseteq Y$ is closed in Y if and only if $Y \setminus F$ is open and that $f^{-1}(F) = X \setminus f^{-1}(Y \setminus F)$ □

Theorem 3.4 can alternatively be used to prove the second part of Theorem 3.3. Indeed, let $f : X \to Y$ and $g : Y \to Z$ be continuous. We wish to show that $g \circ f$ is continuous. Let therefore $G \subseteq Z$ be an arbitrary open set. Since g is continuous, $g^{-1}(G)$ is open in Y, and since f is continuous, $(g \circ f)^{-1}(G) = f^{-1}(g^{-1}(G))$ is open in X, which by Theorem 3.4 implies that $g \circ f$ is continuous.

Theorem 3.4 can also be used to prove that certain sets are open. Here is an example:

Example 3.5 *$A \subseteq \mathbb{R}^2$ defined by:*

$$A = \{(x,y) \in \mathbb{R}^2 \mid \arctan\left(x^7y^3 \sin\left(x^{10}y^{17}+5\right)\right) + \cos\left(x+5y\right) < \frac{1}{100}\}$$

is an open subset of $\mathbb{R}^2$. To see this we define the function $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x,y) = \arctan\left(x^7y^3 \sin\left(x^{10}y^{17}+5\right)\right) + \cos\left(x+5y\right)$$

for all $(x,y) \in \mathbb{R}^2$. Since f is a composition of classical continuous functions, it is continuous and since the interval $]-\infty, \frac{1}{100}[$ is open, we get that $A = f^{-1}(]-\infty, \frac{1}{100}[$ is an open subset of $\mathbb{R}^2$.

If $f : X \to Y$ is a continuous function, it is by definition continuous in all points of X and this can be written with quantifiers as:

$$\forall \varepsilon > 0 \quad \forall x \in X \quad \exists \delta > 0 \quad \forall y \in X : d_X(x,y) < \delta \Rightarrow d_Y(f(x), f(y)) < \varepsilon. \tag{3.4}$$

It is always a bit dangerous to interchange quantifiers in a logical statement because the statement changes radically. Let us anyway look on the following statement:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in X \quad \forall y \in X : d_X(x,y) < \delta \Rightarrow d_Y(f(x), f(y)) < \varepsilon. \tag{3.5}$$

If we do a little text analysis of the two statements we see that in (3.4) the δ depends on ε and x while in (3.5) the δ only depends on ε and thus works for all $x, y \in X$. The statement (3.5) makes perfectly sence and gives rise to the following definition:

Definition 3.6 *A function $f : X \to Y$ is called uniformly continuous if it satisfies* (3.5).

The word "uniformly" is used because given $\varepsilon > 0$, one can use the same δ for all $x, y \in X$. The next statement is really an example, but we formulate it as a proposition.

Proposition 3.7 *Let $f : [1, \infty[\to \mathbb{R}$ be defined by $f(x) = \sqrt{x}$ for all $1 \leq x < \infty$. Then f is uniformly continuous.*

Proof: Let $x, y \geq 1$ be arbitrary. Since f is differentiable, we can by the mean value theorem find a ξ between x and y so that

$$f(x) - f(y) = f'(\xi)(x - y).$$

Since $\xi \geq 1$ and $f'(\xi) = \frac{1}{2\sqrt{\xi}}$, we get that $|f'(\xi)| \leq \frac{1}{2}$ and hence

$$|f(x) - f(y)| \leq \frac{1}{2}|x - y|$$

which holds for all $x, y \geq 1$. If now $\varepsilon > 0$ is arbitrary, we can choose a $0 < \delta < 2\varepsilon$ and if $|x - y| < \delta$, then by the above:

$$|f(x) - f(y)| \leq \varepsilon.$$

This shows that f is uniformly continuous. $\square$

We shall later prove that any continuous function defined on a closed and bounded interval of $\mathbb{R}$ is uniformly continuous. Combining this with Proposition 3.7 we get that the square root function is in fact uniformly continuous on $[0, \infty[$. The next example shows that even very nice continuous functions need not be uniformly continuous.

Example 3.8 *let $g : \mathbb{R} \to \mathbb{R}$ be defind by $g(x) = x^2$ for all $x \in \mathbb{R}$. We claim that g is not uniformly continuous. It is clearly enough to prove that g is not uniformly continuous on* $[0, \infty]$. *To see this we put $\varepsilon = 1$ and let $0 < \delta \leq 1$ be arbitrary. If $x \geq 0$, we get*

$$0 \leq g(x + \delta) - g(x) = (x + \delta)^2 - x^2 = (2x + \delta)\delta.$$

For all $x > \frac{1}{2}(\delta^{-1} - \delta)$ we get that

$$g(x + \delta) - g(x) > 1$$

which shows that g is not uniformly continuous.

4 Topological spaces

In many applications of mathematics one studies some objects which can be considered as elements of a set X together with a system τ of subsets of X. We would then like to find a metric on X so that the open sets are precisely the sets belonging to τ. This is not always possible, but if the sets belonging to τ satisfies the same conditions as open sets do, then we are still able to do something, though we have to leave the metric situation. This is the background of this section.

If X is a set, we let $\mathcal{P}(X)$ denote the set of all subsets of X; that is

$$\mathcal{P}(X) = \{A \mid A \subseteq X\}.$$

We make the following definition.

Definition 4.1 *Let X be a set. A subset $\tau \subseteq \mathcal{P}(X)$ is called a topology on X if the following conditions are satisfied:*

T1 If $\{G_i \mid i \in I\}$ is an arbitrary system of subsets from τ, then

$$\cup_{i \in I} G_i \in \tau.$$

T2 If $n \in \mathbb{N}$ and $G_k \in \tau$ for $1 \leq k \leq n$, then

$$\cap_{k=1}^{n} G_k \in \tau.$$

T3 $\emptyset \in \tau$ and $X \in \tau$.

Let us look on some examples:

Example 4.2 *We shall here consider three examples.*

- *If we put $\tau = \mathcal{P}(X)$, then obviously T1–T3 are satisfied, so $\mathcal{P}(X)$ is a topology on X. It is called the discrete topology.*
- *If we put $\tau = \{\emptyset, X\}$, then we also get a topology on X, called the indiscrete topology.*
- *This example makes the connection to the previous sections. If (X, d) is a metric space and we put $\tau = \mathcal{G}_d$, then it follows from Theorem 2.12 that T1–T3 are satisfied. Hence $\mathcal{G}_d$ is a topology on X. It is called the topology generated by the metric d.*

Definition 4.3 *If τ is a topology on a set X, the pair (X, τ) is called a topological space. The subsets of X which are elements of τ are called the open subsets of the topological space (X, τ). A subset $F \subseteq X$ is called closed if $X \setminus F$ is open.*

If there is no doubt about the topology on X we shall often in the sequel just write the "topological space X".

Using the properties of taking complements the following proposition follows immediately from Definition 4.1 og Definition 4.3.

Proposition 4.4 *If (X, τ) is a topological space, the following statements hold:*

(i) Every intersection of closed sets is again closed.

(ii) Every finite union of closed sets is again closed.

(iii) X and $\emptyset$ are both open and closed.

It follows from Example 4.2 that every metric space is a topological space. A topological space (X, τ) is called *metrizable* if there exists a metric d on X, so that $\mathcal{G}_d = \tau$.

Proposition 4.5 *Let $X, \tau)$ be a topological space and $A \subseteq X$. If we define $\tau_A \subseteq \mathcal{P}(A)$ by*

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

then τ_A is a topology on A. It is called the induced topology on A.

The proof is left to the reader.

Hence every subset A of a topological space is itself a topological space when we equip it with the induced topology τ_A. We shall always use that topology when we consider a subset as a topological space. Note that if (X, d) is a metric space and $A \subseteq X$, then $\tau_A = \mathcal{G}_{d_A}$ where d_A is the induced metric on A as defined in Section 2.

We also need to be able to turn a product of two topological spaces into a topological space. This is the contents of the next proposition.

Proposition 4.6 *Let (X, τ) and $Y, \mathcal{S})$ be topological spaces. We define $\tau_{X \times Y}$ be be the subset of $\mathcal{P}(X \times Y)$ consisting af all sets of the form*

$$\cup_{i \in I}(U_i \times V_i),$$

where I is an arbitrary index set and $U_i \in \tau$ and $V_i \in \mathcal{S}$ are arbitrary for all $i \in I$. Then $\tau_{X \times Y}$ is a topology on $X \times Y$, called the product topology on $X \times Y$.

Exercise 4.7 *Prove Proposition 4.6.*

If (X, d_X) and (Y, d_Y) are metric spaces, then it is readily verified that $\tau_{X \times Y} = \mathcal{G}_d$ where d is the metric on $X \times Y$ defined in Proposition 2.5.

In the rest of this section we shall study topological spaces in greater detail and this investigation will also give new results on metric spaces. In the rest of this section we let (X, τ) denote a topological space.

Definition 4.8 *Let $x \in X$. A subset $M \subseteq X$ is called a neighbourhood of x if there exist an open subset $G \subseteq X$ with $x \in G \subseteq M$. An open set G with $x \in G$ is called an open neighbourhood of x.*

Note that according to the definition every non–empty open subset of X is a neighbourhood of any of its points.

If $x \in X$, we shall in the sequel put

$$\mathcal{U}_x = \{U \subseteq X \mid U \text{ is a neighbourhood of } x\}$$

If (X, d) is a metric space we see that a subset $M \subseteq X$ is a neighbourhood of a point $x \in X$ if and only if there exists a ball $B(x)$ with center x so that $B(x) \subseteq M$. Due to this the only interesting neighbourhoods of a point x in a metric space are the balls with center x. Our first result should be compared to Definition 2.7.

Theorem 4.9 *A subset $W \subseteq X$ is open if and only if every point $x \in W$ has a neighbourhood M with $M \subseteq W$.*

Proof: If W is open in X, then it is a neigbourhood of any of its points x, so in this case we can put $M = W$ for all $x \in W$.

Assume next that every $x \in W$ has a neigbourhood $M_x \subseteq W$.By definition we can to every $x \in W$ find an open set U_x with $x \in U_x \subseteq M_x$. If we can prove

$$W = \cup_{x \in X} U_x, \tag{4.1}$$

we can conclude that W is open, since by T1 the right hand side of (4.1) is an open set. To prove (4.1) we proceed af follows: By definition we have that $U_x \subseteq W$ for all $x \in X$ and hence $\cup_{x \in X} U_x \subseteq W$. If $z \in W$, then $z \in U_z \subseteq \cup_{x \in X} U_x$ which shows the other inclusion. $\square$

Definition 4.10 *Let $W \subseteq X$. A point $x \in W$ is called an inner point of W, if there exists a neighbourhood U_x of x so that $U_x \subseteq W$. The subset of W consisting of all innner points of W is called the interior of W and denoted $int(W)$.*

It follows immediately that if (X, d) is a metric space and $W \subseteq X$, then a point $x \in W$ is an inner point of W, if and only if there exists a ball $B(x)$ with center x so that $B(x) \subseteq W$.

It can easily happen that $W \neq \emptyset$ while $int(W) = \emptyset$. Let for example L be a line in the plane $\mathbb{R}^2$. If $x \in L$ is any point, it is readily seen that every ball with center x will intersect $\mathbb{R}^2 \setminus L$ so that x is not an inner point. Hence $int(L) = \emptyset$.

We have the following result on the interior of a set.

Theorem 4.11 *Let $W \subseteq X$. The following statements hold for $int(W)$:*

(i) *$int(W)$ is an open subset of X.*

(ii) *If U is an open subset of X with $U \subseteq W$, then $U \subseteq int(W)$.*

(iii) *$int(W) = \cup\{U \subseteq W \mid U \text{ is open in } X\}$.*

Proof: (i): If $int(W) = \emptyset$, there is nothing to prove so assume that $int(W) \neq \emptyset$. If $x \in int(W)$, we can find an open neighbourhood U of x so that $U \subseteq W$. We wish to show that actually $U \subseteq int(W)$. To see this let $y \in U$ be arbitrary. Since U is open, it is a neighbourhood of y and since $U \subseteq W$, we get that y is an interior point of W. Hence $U \subseteq int(W)$.

Let us summarize what we have done. We started with a point $x \in int(W)$ and found an open neighbourhood U of x so that $U \subseteq int(W)$. An application of Theorem 4.9 now gives that $int(W)$ is open. This proves (i).

(ii): Let U be an open subset of X so that $U \subseteq W$. If $x \in U$ is arbitrary, then U is a neighbourhood of x which by the above is fully contained in W. This shows that $x \in int(W)$. Hence we have proved that $U \subseteq int(W)$.

(iii): Let V be equal to the union on the right hand side of (iii). Since V is a union of open sets, V is open and clearly $V \subseteq W$. By (ii) $V \subseteq int(W)$. Since $int(W)$ is open by (i), it is one of the U's that appear in the union on the right hand side of (iii) and therefore $int(W) \subseteq V$. This shows the equality in (iii). $\square$

Together (i) and (ii) shows that $int(W)$ is the largest open subset of W.

We have the following corollary:

Corollary 4.12 *A subset $W \subseteq X$ is open if and only if $W = int(W)$.*

Proof: If W is open, then by Theorem 4.11 (ii) $W \subseteq int(W)$, so $W = int(W)$. If we start with $W = int(W)$, then W is open since by Theorem 4.11 (i) $int(W)$ is open. $\square$

We now need the following definition:

Definition 4.13 *Let $A \subseteq X$.*

(i) A point $x \in X$ is called a boundary point of A if every $U \in \mathcal{U}_x$ contains points from A and from $X \setminus A$, that is $U \cap A \neq \emptyset$ and $U \cap (X \setminus A) \neq \emptyset$. The set of boundary points is denoted by ∂A

(ii) A point $x \in X$ is called a contact point of A if every $U \in \mathcal{U}_x$ intersects A, that is $U \cap A \neq \emptyset$. The set of contact points of A is denoted $\overline{A}$ and is called the closure of A.

(iii) A point $x \in X$ is called an accumulation point of A if for every $U \in \mathcal{U}_x$ $(U \setminus \{x\}) \cap A \neq \emptyset$.

It follows immediately from the definition that every boundary point and every accumulation point belongs to $\overline{A}$. So $\partial A \subseteq \overline{A}$ and $A \subseteq \overline{A}$. Actually $\overline{A} = int(A) \cup \partial A$. This can be seen as follows: If $x \in \overline{A} \setminus int(A)$, then every $U \in \mathcal{U}_x$ intersects A. On the other hand, since $x \notin int(A)$, we also have that every $U \in \mathcal{U}_x$ intersects $X \setminus A$ and hence $x \in \partial A$.

The definition contains many concepts so therefore it is adequate to come with an example.

Example 4.14 *Let $A \subseteq \mathbb{R}^2$ be defined by:*

$$A = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\} \cup \{(1, 0)\} \cup \{(2, 7)\}.$$

A consists of the open disc with radius 1 and the points $(1, 0)$ and $(2, 7)$. If $\mathbf{x} = (x, y)$ with $x^2 + y^2 \leq 1$ and $B(\mathbf{x})$ is an arbitrary ball with center $\mathbf{x}$, then $B(\mathbf{x}) \cap A$ is an infinite set and

therefore in particular $\mathbf{x}$ is an accumulation point. The point $(1,0)$ is a boundary point which actually belongs to A. If we take a ball with center $(2,7)$ and radius 1, it only intersects A in the point $(2,7)$. Hence $(2,7)$ is a contact point for A, but not an accumulation point. If $\mathbf{x}=(x,y)$ with $x^2+y^2=1$, then clearly any ball with center $\mathbf{x}$ will intersect both A and $\mathbb{R}^2 \setminus A$, so $\mathbf{x}$ is a boundary point. Since $(2,7) \in A$ every ball with center $(2,7)$ of course intersects A, but it is also clear that any ball with center $(2,7)$ intersects $\mathbb{R}^2 \setminus A$ so $(2,7)$ is a boundary point of A which actually belongs to A. If $\mathbf{x} \in \mathbb{R}^2$ with $x^2+y^2>1$ and $\mathbf{x} \neq (2,7)$, then it is also clear that there exists a ball with center $\mathbf{x}$ which does not intersect A. The conclusion is then that

$$\overline{A} = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 \leq 1\} \cup \{(2,7)\},$$

$$\partial A = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 = 1\} \cup \{(2,7)\}$$

while the set of accumulation points of A is equal to

$$\{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 \leq 1\}.$$

We have the following result on the closure $\overline{A}$ of a subset A of the toplogical space X:

Theorem 4.15 *If $A \subseteq X$, the following statements hold:*

(i) *$\overline{A}$ is a closed set.*

(ii) *If $F \subseteq X$ is a closed set with $A \subseteq F$, then $\overline{A} \subseteq F$.*

(iii) *$\overline{A} = \cap\{F \subset X \mid A \subseteq F \quad F \text{ closed}\}$.*

Proof: (i): We have to prove that $X \setminus \overline{A}$ is open so let $x \in X \setminus \overline{A}$ be arbitrary. Since $x \notin \overline{A}$, x is not a contact point for A and therefore there is an open $U \in \mathcal{U}_x$ so that $U \cap A = \emptyset$. We claim that actually $U \cap \overline{A} = \emptyset$. If namely there exists a $z \in U \cap \overline{A}$, then z is a contact point for A and since U is open, it is a neigbourhood of z which implies that $U \cap A \neq \emptyset$. This is a contradiction. Hence $U \cap \overline{A} = \emptyset$ which means that $U \subseteq X \setminus \overline{A}$. Hence $X \setminus \overline{A}$ is open.

(ii): Let F be a closed subset of X with $A \subseteq F$. We shall prove (ii) by contradiction so assume that $\overline{A} \setminus F \neq \emptyset$ and choose a point z in that set. Since $z \in X \setminus F$, which is open, there exists a neigbourhood U of z so that $U \subseteq X \setminus F$. But z is a contact point for A, so $U \cap A \neq \emptyset$. This is a contradiction, since $A \subseteq F$ and $U \cap F = \emptyset$.

(iii): Since $\overline{A}$ is closed by (i), it is one of the sets in the intersection on the right hand side of (iii) and hence $\overline{A}$ contains that intersection. On the other hand, if F is a closed subset of X with $A \subseteq F$, then by (ii) $\overline{A} \subseteq F$. Therefore $\overline{A}$ is contained in the intersection on the right hand side of (iii). This proves (iii). $\square$

As a corollary we obtain:

Corollary 4.16 *If $A \subseteq X$, then A is closed if and only if $A = \overline{A}$.*

Proof: If $A = \overline{A}$, then A is closed because $\overline{A}$ is closed by Theorem 4.15 (i). If on the other hand A is closed, (ii) of that theorem shows that $\overline{A} \subseteq A$ and hence $A = \overline{A}$. $\square$

The use of Corollary 4.16 is the most common way to prove that a given set A is closed.

We also need:

Theorem 4.17 *If $A \subseteq X$ and $B \subseteq X$, then we have:*

(i) $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

(ii) $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$.

Proof: (i): Since $A \cup B \subseteq \overline{A} \cup \overline{B}$ and the latter set is closed, we get from Theorem 4.15(i) that $\overline{A \cup B} \subseteq \overline{A} \cup \overline{B}$. The theorem also clearly gives that $\overline{A} \subseteq \overline{A \cup B}$ and $\overline{B} \subseteq \overline{A \cup B}$ hence $\overline{A} \cup \overline{B} \subseteq \overline{A \cup B}$. This proves (i).

(ii): Since $A \cap B \subseteq \overline{A} \cap \overline{B}$ and the latter set is closed we get again from Theorem 4.15 that $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$. $\square$

One cannot expect equality in Theorem 4.17 (ii). Indeed, let us consider the subsets $A = [0, 1[$ and $B =]1, 2]$ of $\mathbb{R}$. Since $A \cap B = \emptyset$, also $\overline{A \cap B} = \emptyset$. On the other hand $\overline{A} = [0, 1]$ and $\overline{B} = [1, 2]$ so that $\overline{A} \cap \overline{B} = \{1\}$.

We now wish to define convergence of sequences in topological spaces in the same way as we did for metric spaces.

Definition 4.18 *Let $(x_n) \subseteq X$ and let $x \in X$. We say that (x_n) converges (or tends to) x (notation: $x_n \to x$), if*

$$\forall U \in \mathcal{U}_x \quad \exists n_0 \in \mathbb{N} : n \geq n_0 \Rightarrow x_n \in U.$$

This definition is clearly a generalization of Definition 2.18 for metric spaces.

Theorem 4.19 *Let $A \subseteq X$ and let $x \in X$. Consider the statements:*

(i) There exists a sequence $(x_n) \subseteq A$ so that $x_n \to x$.

(ii) $x \in \overline{A}$.

If (X, τ) is a general topological space, then (i) implies (ii). If (X, d) is a metric space, then (i) and (ii) are equivalent.

Proof: Let us first consider the case of a general topological space. If $(x_n) \subseteq A$ with $x_n \to x$, we have to show that $x \in \overline{A}$. Hence let $U \in \mathcal{U}_x$ be arbitrary. Since $x_n \to x$, there is an $n_0 \in \mathbb{N}$ so that $x_n \in U$ for all $n \geq n_0$. But then $\{x_n \mid n \geq n_0\} \subseteq U \cap A$ so that $U \cap A \neq \emptyset$ which shows that $x \in \overline{A}$

Let now (X, d) be a metric space. We have to prove that (ii) implies (i) so let $x \in \overline{A}$ be arbitrary. Since x is a contact point for A, $B(x, \frac{1}{n}) \cap A \neq \emptyset$ for all $n \in \mathbb{N}$ and therefore we can choose a sequence (x_n) so that $x_n \in B(x, \frac{1}{n}) \cap A$ for all $n \in \mathbb{N}$. This sequence is a sequence from A and since $d(x, x_n) \leq \frac{1}{n}$, we get that $d(x, x_n) \to 0$ which means that $x_n \to x$. $\square$

In general topological spaces (ii) does not imply (i).

As a corollary we get

Corollary 4.20 *If $A \subseteq \mathbb{R}$ is an upper bounded and closed subset, then $\sup A \in A$. Similarly, if $B \subseteq \mathbb{R}$ is a lower bounded and closed set, then $\inf B \in B$.*

Proof: If $n \in \mathbb{N}$ is arbitrary, $\sup A - \frac{1}{n}$ is not an upper bound for A and therefore there exists an $x_n \in A$ so that $\sup A - \frac{1}{n} < x_n \leq \sup A$. Clearly $x_n \to \sup A$ and therefore $\sup A \in \overline{A}$ by Theorem 4.19, but since A is closed $A = \overline{A}$. Similarly for the set B. $\square$

The next result on metric spaces explains the name accumulation point.

Theorem 4.21 *If A is a subset of a metric space (X, d) and x is an accumulation point of A, then every ball with center x (or for that matter every neighbourhood of x) contains infinitely many points from A.*

Proof: Let $R > 0$ be arbitrary. By induction we will construct a sequence $(x_n) \subseteq B(x, R) \cap A$ so that $x_n \neq x_m$ for all $n, m \in \mathbb{R}$ with $n \neq m$ and $x_n \neq x$ for all n. Since x is an accumulation point of A, there is an $x_1 \in B(x, R) \cap A$ with $x_1 \neq x$. Assume next that we have constructed $\{x_k \mid 1 \leq k \leq n\} \subseteq B(x, R) \cap A$ so that they are all different and different from x. Since $d(x, x_k) > 0$ for all $1 \leq k \leq n$, we can find an $0 < r < R$ with $r < d(x, x_k)$ for all $1 \leq k \leq n$. Again since x is an accumulation point, there is an $x_{n+1} \in B(x, r) \cap A$ with $x_{n+1} \neq x$. Clearly $d(x, x_{n+1}) < d(x, x_k)$ when $1 \leq k \leq n$ so $x_{n+1} \neq x_k$ and since $r < R$ we also get that $x_{n+1} \in B(x, R) \cap A$. This finishes the induction. Since all the x_n's are different we have found infinitly many points in $B(x, R) \cap A$. $\square$

The reason for that the above proof works is of course that in metric spaces we can separate points with balls. There is nothing in the definition of a topological space which can give us a similar separation of points by neigbourhoods. This is the background of the next definition.

Definition 4.22 *X is called a Hausdorff space, if there for all $x, y \in X$ with $x \neq y$ exists a $U \in \mathcal{U}_x$ and a $V \in \mathcal{U}_y$ with $U \cap V = \emptyset$.*

It follows from Propossition 2.10 that all metric spaces are Hausdorff spaces. Most of the spaces you will meet will be Hausdorff spaces. Similar to Theorem 4.21 we get:

Theorem 4.23 *If A is a subset of a Hausdorff space X and x is an accumulation point of A, then every neighbourhood of x contaions infinitely many points from A.*

Proof: We will try to mimic the proof of Theorem 4.21 so let $U \in \mathcal{U}_x$ be an arbitrary open neighbourhood (it is enough to consider open neighbourhoods). By induction we shall construct a sequence $(x_n) \subseteq U \cap A$, so that $x_n \neq x$ and $x_n \neq x_m$ for all $n, m \in \mathbb{N}$, $n \neq m$.

Since x is an accumulation point for A, there exists an $x_1 \in U \cap A$ with $x \neq x_1$

Assume next that we have constructed $\{x_k \mid 1 \leq k \leq n\}$, fulfilling our conditions. By the Hausdorff property we can for every $1 \leq k \leq n$ find an open neigborhood V_k of x so that $x_k \notin V_k$. If we put $V = U \cap \cap_{k=1}^{n} V_k$, then V is an open neighbourhood of x, not containing any of the x_k's and with $V \subseteq U$. Again since x is an accumulation point for A, we can find an $x_{n+1} \in V \cap A$ which is different from x. This finishes the induction.

Since all the x_n's are different, we have found infinitely many points in $U \cap A$. $\square$

Notice that in the proof we have only used the Hausdorff property to conclude that X has a somewhat weaker property; namely that if two points in X are different, then each of the points has a neighborhood not containing the other point. This property is called T_1, but we shall not use it in these notes. The stronger Hausdorff property is more important to us.

It is practical to know the next result:

Proposition 4.24 *A sequence in a Hausdorff space has at most one limit point.*

Proof: Let (x_n) be a sequence in a Hausdorff space (X, τ) and assume that there are $x, y \in X$ so that $x_n \to x$ and $x_n \to y$ and $x \neq y$. By the Hausdorff property we can find a $U \in \mathcal{U}_x$ and a $V \in \mathcal{U}_y$ with $U \cap V = \emptyset$. Since $x_n \to x$, there is an $n_1 \in \mathbb{N}$ so that $x_n \in U$ for all $n \geq n_1$ and since $x_n \to y$, there is an $n_2 \in \mathbb{N}$ so that $x_n \in V$ for all $n \geq n_2$. If we put $n = \max\{n_1, n_2\}$, then $x_n \in U \cap V$ which is a contradiction. $\square$

Let us end this section by a few results on so–called dense subsets of topological spaces. We start with the definition.

Definition 4.25 *A subset A of the topological space X is said to be dense in X if $\overline{A} = X$.*

The definition immediately implies the next proposition.

Proposition 4.26 *A subset A of X is dense if and only if every non–empty open subset of X contains at least one point of A.*

Proof: Left to the reader. $\square$

Definition 4.27 *X is called* separable *if it contains a countable dense subset.*

Let us end this section with two important result on very concrete topogical spaces.

Theorem 4.28 *The set $\mathbb{Q}$ of rational numbers is dense in the set $\mathbb{R}$ of real numbers. In particular $\mathbb{R}$ is separable.*

Proof: It is clearly enough to show that any bounded open interval $]a, b[$ contains points from $\mathbb{Q}$ and since $0 \in \mathbb{Q}$ we need only consider the cases $a \geq 0$ or $b \leq 0$ (it is actually enough to consider intervals of length less than one. Why?).

Assume that $a \geq 0$ and choose $n \in \mathbb{N}$ so that $n > (b-a)^{-1}$. Since the set $B = \{k \in \mathbb{N} \mid an < k\}$ is non–empty, it contains a first element m. We claim that $a < \frac{m}{n} < b$. To see this we observe that by the definition of B we get $a < \frac{m}{n}$ and since m is the first element in B, then either $m-1 = 0$ or $m-1 \in \mathbb{N} \setminus B$. In both cases we get that $\frac{m-1}{n} \leq a$ and hence $\frac{m}{n} = \frac{m-1}{n} + \frac{1}{n} < a + (b-a) = b$. Hence we have proved that $\frac{m}{n} \in]a, b[$.

If $b \leq 0$, we can use the map $x \to -x$ to move the interval $]a, b[$ to the non–negative part of $\mathbb{R}$, apply the first case, and then go back.

Hence we have proved that $\overline{\mathbb{Q}} = \mathbb{R}$ and since $\mathbb{Q}$ is countable by [4, Exercise 12], we conclude that $\mathbb{R}$ is separable. □

You can think about the proof of Theorem 4.28 as follows: After having chosen $n \in \mathbb{N}$ with $\frac{1}{n} < b - a$ you walk along the real line starting in 0 and take steps of length $\frac{1}{n}$. Since your step length is strictly smaller than the interval length, you have to hit the interval $]a, b[$ after a finite number of steps.

Theorem 4.28 can be generalized to $\mathbb{R}^n$. Indeed, if one does the above proof on each coordinate we get the following:

Theorem 4.29 *If $n \in \mathbb{N}$, then $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$. In particular $\mathbb{R}^n$ is separable.*

The fact that $\mathbb{Q}^n$ is countable can be established by applying induction on the Exercises 11 and 12 in [4].

The importance of the notion of density comes in the situation where you have to check whether a given topological space has a certain property; in many cases it is enough to check that a dense subset has the property in question.

5 Continuous functions on topological spaces

In this section we will generalize the concept of continuity to functions between topological spaces and obtain some additional results on continuity which also cover the case of metric spaces. We start with the following definition which is a generalization of Definition 3.1, see also equation (3.3) in the remark just after.

Definition 5.1 *Let X and Y be topological spaces, and $x_0 \in X$. A function $f : X \to Y$ is said to be continuous in x_0, if the following condition holds:*

$$\forall V \in \mathcal{U}_{f(x_0)} \quad \exists U \in \mathcal{U}_{x_0} : f(U) \subseteq V. \tag{5.1}$$

f is said to be continuous if it is continuous in every $x_0 \in X$.

Our first theorem should be compared to Theorem 3.2:

Theorem 5.2 *Let X and Y be topological spaces, $x_0 \in X$, and $f : X \to Y$ a function. Consider the statements:*

(i) f is continuous in x_0.

(ii) For all sequences $(x_n) \subseteq X$ we have that if $x_n \to x_0$, then $f(x_n) \to f(x_0)$.

In general (i) implies (ii). If X is a metric space, (i) and (ii) are equivalent.

Proof: The proof looks much as that of Theorem 3.2. Assume first that (i) holds, let $(x_n) \subseteq X$ be a sequence with $x_n \to x_0$, and let $V \in \mathcal{U}_{f(x_0)}$ arbitrary. The continuity of f gives us a $U \in \mathcal{U}_{x_0}$ so that $f(U) \subseteq V$. Since $x_n \to x_0$, there is an $n_0 \in \mathbb{N}$ so that $x_n \in U$ for all $n \geq n_0$, but then also $f(x_n) \in V$ for all $n \geq n_0$. This shows that $f(x_n) \to f(x_0)$.

In the case where (X, d) is a metric space, we assume that f is discontinuous in x_0 and have to prove that (ii) does not hold. The discontinuity of f in x_0 means:

$(*)$ there exists a $V \in \mathcal{U}_{f(x_0)}$ so that for all $\delta > 0$ there exists an $x_\delta \in X$ with $d(x_0, x_\delta) < \delta$ and $f(x_\delta) \notin V$.

Since $(*)$ holds for all $\delta > 0$, we can for every $n \in \mathbb{N}$ put $\delta = \frac{1}{n}$ and $(*)$ will provide us with an $x_n \in X$ so that $d(x_0, x_n) < \frac{1}{n}$ and $f(x_n) \notin V$. Since $d(x_0, x_n) \to 0$, $x_n \to x_0$, but since $f(x_n) \notin V$ for any n, $(f(x_n))$ cannot converge to $f(x_0)$. This shows that (ii) does not hold. □

The next theorem is a generalization of Theorem 3.3.

Theorem 5.3 *Let X, Y, and Z be topological spaces and $x_0 \in X$. If $f : X \to Y$ $g : Y \to Z$ so that f is continuous in x_0 and g is continuous in $f(x_0)$, then $g \circ f$ is continuous in x_0.*

In paticular, if f and g are continuous, then $g \circ f$ is continuous.

Proof: Let $W \in \mathcal{U}_{g(f(x_0))}$ be arbitrary. Since g is continuous in $f(x_0)$, there is a $V \in \mathcal{U}_{f(x_0)}$ with $g(V) \subseteq W$ and since f is continuous in x_0, there is a $U \in \mathcal{U}_{x_0}$ with $f(U) \subseteq V$. If now $x \in U$, we get that $f(x) \in V$ and that $g(f(x)) \in W$. Hence we have proved that $(g \circ f)(U) \subseteq W$ and the continuity of $g \circ f$ in x_0 is established. □

Now we generalize Theorem 3.4.

Theorem 5.4 *Let X and Y be topological spaces and $f : X \to Y$ a function. The following statements are equivalent.*

(i) f is continuous.

(ii) For every open subset $G \subseteq Y$ $f^{-1}(G)$ is an open subset X.

(iii) For every closed subspace $F \subseteq Y$ $f^{-1}(F)$ is a closed subspace of X.

Proof: $(i) \Rightarrow (ii)$: We assume that f is continuous and let $G \subseteq Y$ be an arbitrary open subset of Y. If $f^{-1}(G) = \emptyset$, there is nothing to prove so assume $f^{-1}(G) \neq \emptyset$ and let $x \in f^{-1}(G)$ be arbitrary. We have to find a $U \in \mathcal{U}_x$ with $U \subseteq f^{-1}(G)$.

Since $f(x) \in G$ and G is open, G is a neighbourhood of $f(x)$ and the continuity of f provides us with a $U \in \mathcal{U}_x$ so that $f(U) \subseteq G$ which is equivalent to $U \subseteq f^{-1}(G)$. This shows that (i) implies (ii).

$(ii) \Rightarrow (i)$: We assume (ii), let $x \in X$ be arbitrary, and have to show that f is continuous in x. Hence let V be an arbitrary open neighbourhood of $f(x)$ and put $U = f^{-1}(V)$. Since $x \in U$ and U is open by (ii), U is a neighbourhood of x. Clearly $f(U) \subseteq V$. This shows (i).

The equivalence of (ii) and (iii) can be proved exactly as in Theorem 3.4. □

We end this section with a useful application.

Theorem 5.5 *Let X and Y be topological spaces and assume that Y is a Hausdorff space. If $f, g : X \to Y$ are continuous, the set $\{x \in X \mid f(x) \neq g(x)\}$ is open.*

Proof: Put $A = \{x \in X \mid f(x) \neq g(x)\}$and let $x \in A$ be arbitrary. Since $f(x) \neq g(x)$ and Y is a Hausdorff space, we can find open neighbourhoods $V \in \mathcal{U}_{f(x)}$ and $W \in \mathcal{U}_{g(x)}$ with $V \cap W = \emptyset$ and since f and g are continuous, $f^{-1}(V)$ and $g^{-1}(W)$ are open. Clearly $x \in f^{-1}(V) \cap g^{-1}(W)$ so that this set is an open neighbourhood of x. If $z \in f^{-1}(V) \cap g^{-1}(W)$, then $f(z) \in V$ and $g(z) \in W$ so that $f(z) \neq g(z)$ This shows that $f^{-1}(V) \cap g^{-1}(W) \subseteq A$ and hence that A is open. □

Corollary 5.6 *Let X and Y be as in Theorem 5.5 and let B be a dense subset of X.
If $f, g : X \to Y$ are continuous and $f(x) = g(x)$ for all $x \in B$, then $f = g$.*

Proof: Since f and g are continuous, the set $\{x \in X \mid f(x) = g(x)\}$ is closed by Theorem 5.5 and contains B by assumption. Therefore

$$Y = \overline{B} \subseteq \{x \in X \mid f(x) = g(x)\}$$

which shows that $f(x) = g(x)$ for all $x \in X$, i.e. $f = g$. □

6 Compact sets

This section is probably the most important in these notes because we shall here define the notion of compactness which is applied vastly in all areas of mathematics. Let us discuss the arrangement of this section in greater detail. First we prove two classical results concerning closed and bounded intervals of $\mathbb{R}$ and from these we draw some consequences about real–valued continuous functions defined on closed and bounded intervals of $\mathbb{R}$. Inspired by the classical results we define compact subsets of topological spaces and prove many deep results on the structure of such sets and how continuous functions behave on them. After this we return to the classical case and use the general theory to give a complete description of the compact subsets of $\mathbb{R}^n$ for all $n \in \mathbb{N}$.

We need some new notions in this section. This is included in the following definition.

Definition 6.1 *Let X be a topological space and $A \subseteq X$. A family $\mathcal{C}$ of subsets of X is said to be a cover of A if*

$$A \subseteq \bigcup_{G \in \mathcal{C}} G.$$

If every $G \in \mathcal{C}$ is open, $\mathcal{C}$ is called an open cover of A.

A subcover of $\mathcal{G}$ is a subfamily $\mathcal{D} \subseteq \mathcal{C}$ so that $\mathcal{D}$ is still a cover of A.

In the connection with covers we shall also use terms like "$\mathcal{C}$ covers A" or "the G's from $\mathcal{C}$ cover A".

6.1 Two classical results

We start with a lemma which sharpens Theorem 0.14.

Lemma 6.2 *If $([a_n, b_n])$ is a sequence of closed and bounded intervals so that $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$ for all $n \in \mathbb{N}$, we have:*

(i) There exists $a, b \in \mathbb{R}$ with $a \leq b$ so that

$$\bigcap_{n=1}^{\infty} [a_n, b_n] = [a, b]. \tag{6.1}$$

(ii) If in addition the interval length $b_n - a_n \to 0$ for $n \to \infty$, then $a = b$ and hence

$$\bigcap_{n=1}^{\infty} [a_n, b_n] = \{a\}$$

Proof: (i) is clearly the same as Theorem 0.14 so we only have to prove (ii). Hence assume that $b_n - a_n \to 0$. Since the sequence of intervals is decreasing, we get as in the proof of Theorem 0.14 that $a_n \leq a_m < b_m \leq b_n$ for all $n \leq m$ so that (a_n) is an increasing sequence and (b_n) is decreasing. We also saw that

$$a = \sup\{a_n \mid n \in \mathbb{N}\}$$
$$b = \inf\{b_n \mid n \in \mathbb{N}\}.$$

Since (a_n) is increasing, it follows from [5, Exercise 13] that $a_n \to a$ and the same exercise also gives that $b_n \to b$. Hence $b_n - a_n \to b - a$ so that $a = b$. □

We are now able to prove:

Theorem 6.3 *(Weierstrass' accumulation theorem)*
Let $a, b \in \mathbb{R}$ with $a < b$. Every infinite subset $A \subseteq [a, b]$ has at least one accumulation point in $[a, b]$.

Proof: Let $A \subseteq [a, b]$ be an infinite set. By successive division of an interval in the middle we shall by induction construct a sequence $([a_n, b_n])$ of closed and bounded intervals so that

(i) $A \cap [a_n, b_n]$ is an infinite set for all $n \in \mathbb{N}$.

(ii) $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n] \subseteq [a, b]$ for all $n \in \mathbb{N}$.

(iii) $b_{n+1} - a_{n+1} = \frac{1}{2}(b_n - a_n)$ for all $n \in \mathbb{N}$.

For $n = 1$ we put $a_1 = a$ and $b_1 = b$ and it follows from our assumption that $[a_1, b_1]$ satisfies our conditions (only (i) is relevant in the first step of the induction).

Let now $n \geq 1$ and assume that we have constructed the intervals $[a_k, b_k]$ for all $1 \leq k \leq n$ so that (i)–(iii) are satisfied. We then have to construct the $(n+1)$th interval. This is done as follows:

We divide the interval $[a_n, b_n]$ in the middle and get the two intervals $[a_n, \frac{a_n+b_n}{2}]$ and $[\frac{a_n+b_n}{2}, b_n]$. By intersecting with A we get

$$A \cap [a_n, b_n] = (A \cap [a_n, \frac{a_n + b_n}{2}]) \cup (A \cap [\frac{a_n + b_n}{2}, b_n]).$$

Since by the induction hypothesis the set on the left side of this equality is infinite, at least one of the sets forming the union on the right side must also be infinite. If $A \cap [a_n, \frac{a_n+b_n}{2}]$ is infinite, we put $a_{n+1} = a_n$ and $b_{n+1} = \frac{a_n+b_n}{2}$ and if it is finite, we put $a_{n+1} = \frac{a_n+b_n}{2}$ and $b_{n+1} = b_n$. In both cases the interval $[a_{n+1}, b_{n+1}]$ will satisfy (i)–(iii). Hence we have constructed the sequence of intervals.

Iterating (iii) we obtain that $b_n - a_n = 2^{-n+1}(b - a)$ so that $b_n - a_n \to 0$ and hence by Lemma 6.2 we get an $x \in [a, b]$ with

$$\bigcap_{n=1}^{\infty} [a_n, b_n] = \{x\}.$$

We wish to show that x is an accumulation point for A. Let therefore $r > 0$ be arbitrary. We saw in Lemma 6.2 that both sequences (a_n) and (b_n) tend to x for $n \to \infty$ and hence we can find n_1, n_2 so that $x - r < a_n$ for all $n \geq n_1$ and $b_n < x + r$ for all $n \geq n_2$. If we now put $n = \max\{n_1, n_2\}$, we obtain that $[a_n, b_n] \subseteq]x - r, x + r[$ and hence also $A \cap [a_n, b_n] \subseteq A \cap]x - r, x + r[$. Since by construction $A \cap [a_n, b_n]$ is an infinite set, it follows that also $A \cap]x - r, x + r[$ is infinite. Since this holds for all $r > 0$, x is accumulation point for A. $\square$

The next important classsical result we shall prove is the following:

Theorem 6.4 *(***Heine–Borel***)*
Let $a, b \in \mathbb{R}$ with $a < b$. Every open cover of $[a, b]$ has a finite subcover.

Proof: We shall prove the theorem by contradiction so assume that $\mathcal{C}$ is an open cover of $[a, b]$ without a finite subcover. We will reach a contradiction using the same technique as in the proof of Theorem 6.3. By induction we shall construct a sequence $([a_n, b_n])$ of bounded and closed intervals so that

(i) For no $n \in \mathbb{N}$ can $[a_n, b_n]$ be covered by finitely many G's from $\mathcal{C}$.

(ii) $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n] \subseteq [a, b]$ for all $n \in \mathbb{N}$.

(iii) $b_{n+1} - a_{n+1} = \frac{1}{2}(b_n - a_n)$ for all $n \in \mathbb{N}$.

For $n = 1$ we put $a_1 = a$ and $b_1 = b$.

Let now $n \geq 1$ and assume that we have constructed the intervals $[a_k, b_k]$ for all $1 \leq k \leq n$ so that (i)–(iii) are satisfied. To construct the interval $[a_{n+1}, b_{n+1}]$ we proceed as follows:

Since by assumption $[a_n, b_n]$ cannot be covered by finitely many G's from $\mathcal{C}$ at least one of the intervals $[a_n \frac{a_n+b_n}{2}]$ and $[\frac{a_n+b_n}{2}, b_n]$ cannot be covered by finitely many G's from $\mathcal{C}$. If this happens to the first mentioned interval, we put $a_{n+1} = a_n$ and $b_{n+1} = \frac{a_n+b_n}{2}$, else we choose the other interval and get that $[a_n, b_n]$ satisfies (i)–(iii). This finishes the induction.

As in the proof of Theorem 6.3 we get that $b_n - a_n \to 0$ and hence there is an $x \in [a, b]$ so that

$$\bigcap_{n=1}^{\infty} [a_n, b_n] = \{x\}.$$

Since $\mathcal{C}$ is a cover of $[a, b]$, there is a $G \in \mathcal{C}$ so that $x \in G$ and since G is open, G is a neighbourhood of x and therefore there is an $r > 0$ with $]x - r, x + r[\subseteq G$. Since as before both (a_n) and (b_n) tend to x, there is an $n_0 \in \mathbb{N}$ so that $a_n, b_n \in]x - r, x + r[$ for all $n \geq n_0$; in particular $[a_{n_0}, b_{n_0}] \subseteq]x - r, x + r[\subseteq G$.

Hence we have obtained that a single G from $\mathcal{G}$ covers $[a_{n_0}, b_{n_o}]$ which contradicts (i). $\square$

Some of the readers might know the next theorem, but have probably never seen a proof of it. It is a consequence of the results we have obtained so far and the proof illustrates quite well the techniques to be used in this section.

Theorem 6.5 *Let $a, b \in \mathbb{R}$ with $a < b$. If $f : [a, b] \to \mathbb{R}$ is continuous, then f is bounded; i.e. there exists a $K > 0$ so that $|f(x)| \leq K$ for all $x \in [a, b]$.*

Proof: Since f is continuous we can for every $x \in [a, b]$ find a $\delta_x > 0$ so that $|f(y) - f(x)| \leq 1$ for all $y \in [a, b]$ with $|y - x| < \delta_x$ (we use here the continuity condition with $\varepsilon = 1$). Hence also $|f(y)| \leq |f(y) - f(x)| + |f(x)| \leq 1 + |f(x)|$ for all $y \in [a, b]$ with $|y - x| < \delta_x$. Clearly the system

$$\mathcal{C} = \{]x - \delta_x, x + \delta_x[\quad | \; x \in X\}$$

is an open cover of $[a, b]$ and therefore it follows from Theorem 6.4 that there exist finitely many points $x_1, x_2, \cdots, x_n \in [a, b]$ so that

$$[a, b] \subseteq \bigcup_{j=1}^{n}]x_j - \delta_{x_j}, x_j + \delta_{x_j}[. \tag{6.2}$$

Put

$$K = \max\{|f(x_j)| \quad | \; 1 \leq j \leq n\} + 1.$$

We claim that $|f(x)| \leq K$ for all $x \in [a, b]$. If namely $x \in [a, b]$ is arbitrary, then according to (6.2) there exists a $1 \leq j \leq n$ so that $x \in]x_j - \delta_{x_j}, x_j + \delta_{x_j}[$ and hence

$$|f(x)| \leq |f(x) - f(x_j)| + |f(x_j)| \leq 1 + |f(x_j)| \leq K.$$

□

We could also have used Theorem 6.3 to prove Theorem 6.5. We shall a little later strengthen Theorem 6.5 considerably.

The last theorem of this subsection marks a small detour from our theme, but we have included it here because it is important and because its proof resembles those of the Theorem 6.3 and 6.4.

Theorem 6.6 *The set $\mathbb{R}$ of real numbers is uncountable.*

Proof: Let $(x_n) \subseteq \mathbb{R}$ be an arbitrary sequence with $x_n \neq x_m$ for all $n, m \in \mathbb{N}$ with $n \neq m$. If we can prove that

$$\mathbb{R} \neq \{x_n \mid n \in \mathbb{N}\}, \tag{6.3}$$

it will follow that $\mathbb{R}$ is uncountable. To achieve this we will by induction construct a sequence $([a_n, b_n])$ of closed and bounded intervals so that

(i) $x_n \notin [a_n, b_n]$ for all $n \in \mathbb{N}$,

(ii) $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$ for all $n \in \mathbb{N}$,

(iii) $b_{n+1} - a_{n+1} = \frac{1}{3}(b_n - a_n)$.

For $n = 1$ we choose a bounded interval $[a_1, b_1]$ so that $x_1 \notin [a_1, b_1]$. Let next $n \geq 1$ and assume that we have constructed the intervals $[a_k, b_k]$ for all $1 \leq k \leq n$ so that (i)–(iii) are satisfied. If we divide $[a_n, b_n]$ in three intervals of equal length, at least one of these will not contain x_{n+1} so choose this and call it $[a_{n+1}, b_{n+1}]$. Clearly our conditions are satisfied. This finishes the induction.

Since $b_n - a_n \to 0$, we get from Lemma 6.2 that there exists an $x \in \mathbb{R}$ so that

$$\{x\} = \bigcap_{n=1}^{\infty} [a_n, b_n].$$

If $n \in \mathbb{N}$ is arbitrary, then $x \in [a_n, b_n]$ while $x_n \notin [a_n, b_n]$ which shows that $x \neq x_n$. This proves (6.3). □

6.2 Compactness and sequentially compactness

In this subsection we let X denote a topological space.

Inspired by the two classical results above we give the following two definitions:

Definition 6.7 *A subset $K \subseteq X$ is called compact if every open cover of K has a finite subcover.*

Definition 6.8 *A subset $K \subseteq X$ is called sequentially compact if every infinite subset of K has an accumulation point in K.*

In the defintion it is important that the accumulation point belongs to K.

From Theorem 6.3 and Theorem 6.4 we get immediately:

Corollary 6.9 *Every closed and bounded interval of $\mathbb{R}$ is both compact and sequentially compact.*

In general we have:

Theorem 6.10 *Every compact subset of X is sequentially compact.*

Proof: Let $K \subseteq X$ be compact. It is clearly equivalent to prove that any subset of K without an accumulation point in K is finite so let $A \subseteq K$ be a subset without any accumulation point in K. Hence if $x \in K$, there exists an open $U_x \in \mathcal{U}_x$ so that $U_x \cap A = \{x\}$ if $x \in A$ and $U_x \cap A = \emptyset$ if $x \notin A$. We have clearly that

$$K \subseteq \bigcup_{x \in K} U_x$$

so that $\{U_x \mid x \in K\}$ is an open cover of K. Since K is compact, there exists a finite set of points, say $\{x_j \mid 1 \leq j \leq n \subseteq K\}$ so that

$$K \subseteq \bigcup_{j=1}^{n} U_{x_j}$$

and if we intersect with A, we get

$$A = A \cap K \subseteq \bigcup_{j=1}^{n} (U_{x_j} \cap A).$$

Each of the sets in the union above contains at most one element and therefore the union is a finite set and hence also A is finite. $\square$

We shall a bit later prove that in metric spaces compactness and sequentially compactness are equivalent notions but let us begin by investigating compactness a little. The first result is easy, but quite useful.

Theorem 6.11 *If X is compact, every closed subset of X is also compact.*

Proof: Let $K \subseteq X$ be closed and let $\mathcal{C}$ be an open cover of K, that is

$$K \subseteq \bigcup_{G \in \mathcal{C}} G,$$

from which we get that

$$X = \bigcup_{G \in \mathcal{C}} G \cup (X \setminus K). \tag{6.4}$$

Since K is closed, $X \setminus K$ is open and therefore (6.4) shows that $\mathcal{C} \cup \{X \setminus K\}$ is an open cover of X. Since X is compact, we can find finitely many $G_1, G_2, \cdots, G_n \in \mathcal{C}$ so that

$$X = \bigcup_{j=1}^{n} G_j \cup (X \setminus K),$$

but taking intersection with K we get from this that

$$K \subseteq \bigcup_{j=1}^{n} G_j.$$

Hence we have found a finite subcover of $\mathcal{C}$ which shows that K is compact. $\square$

The next theorem is somewhat in the other direction.

Theorem 6.12 *Let X be a Hausdorff space and let $K \subseteq X$ be compact. If $x \in (X \setminus K)$, then there exist open sets U and V so that $x \in U$, $K \subseteq V$, and $U \cap V = \emptyset$.*

In particular K is closed.

Proof: Since X is a Hausdorff space, we can to every $y \in K$ find open sets U_y and V_y so that $y \in V_y$, $x \in U_y$, and $U_y \cap V_y = \emptyset$. Clearly

$$K \subseteq \bigcup_{y \in K} V_y$$

so that $\{V_y \mid y \in K\}$ is an open cover of K which by the compactness of K has a finite subcover, say $\{V_j \mid 1 \leq j \leq n\}$. If we put $U = \cap_{j=1}^{n} U_j$ and $V = \cup_{j=1} V_j$, then U and V are both open with $x \in U$ and $K \subseteq K$. We claim that $U \cap V = \emptyset$. Indeed, for every $1 \leq j \leq n$ we have $U \cap V_j \subseteq U_j \cap V_j = \emptyset$ and therefore $U \cap V = \cup_{j=1}^{n} (U \cap V_j) = \emptyset$. Hence we have proved the first part of the theorem.

To prove the second part we observe that if $x \in (X \setminus K)$ and U is constructed as above, then U is a neighbourhood of x with $U \subseteq (X \setminus K)$. Since this works for all $x \in (X \setminus K)$, we have proved that $X \setminus K$ is open and hence K is closed. $\square$

If X is not a Hausdorff space, there exist points $x, y \in X$ which cannot be separated by open sets and since singletons clearly are compact sets, the first part of Theorem 6.12 is false in this case. If X is not a T_1–space, there are singletons which are not closed so also the second part of the theorem is false.

The following definition and proposition are often handy:

Definition 6.13 *A subset $K \subseteq X$ is said to have the finite intersection property if the following condition is satisfied:*

For every family $\mathcal{F}$ of closed subset of X with $K \cap \cap_{F \in \mathcal{F}} F = \emptyset$ there exist an $n \in \mathbb{N}$ and $F_1, F_2, \cdots, F_n \in \mathcal{F}$ with $K \cap \cap_{j=1}^{n} F_j = \emptyset$.

For a proof of the next propostion see [4, Exercise 20].

Proposition 6.14 *If $K \subseteq X$, the following two statements are equivalent.*

(i) K is compact.

(ii) K has the finite intersection property.

As a corollary we get the following result which is often used in existence proofs.

Corollary 6.15 *If $K \subseteq X$ is compact and $\mathcal{F}$ is a family of closed subsets of X so that $K \cap \cap_{F \in \mathcal{F}_f} F \neq \emptyset$ for all finite subfamilies $\mathcal{F}_f$ of $\mathcal{F}$, then $K \cap \cap_{F \in \mathcal{F}} F \neq \emptyset$.*

Let us now look a little on the metric situation. We need the following definition:

Definition 6.16 *A subset K of a metric space (X, d) is called bounded if it is a subset of an open ball in X.*

If we for $K \subseteq X$ define $diam(K) = \sup\{d(x, y) \mid x, y \in K\}$, then it easily follows from the triangle inequality that K is bounded if and only if $diam(K) < \infty$.

Note also that boundedness of a set in a metric space is not a topological property, but depends on the specific metric d. Indeed, if (X, d) is a metric space, we can according to [4, Exercise 4] find an equivalent metric d' on X so that $d'(x, y) \leq 1$ for all $x, y \in X$ and hence every subset of X will be bounded with respect to d'.

Theorem 6.17 *In a metric space (X, d) every compact subspace is closed and bounded.*

Proof: Let $K \subseteq X$ be compact and let $x_0 \in X$. Since

$$K \subseteq X = \bigcup_{k=1}^{\infty} B(x_0, k),$$

the family $\{B(x, k) \mid k \in \mathbb{N}\}$ is an open cover of K and therefore there exists an $n \in \mathbb{N}$ so that

$$K \subseteq \bigcup_{k=1}^{n} B(x_0, k) = B(x_0, n)$$

which shows that K is bounded.

It follows from Theorem 6.12 that K is closed. $\square$

In a general metric space a closed and bounded subset need not be compact. Indeed, if we define the metric d' from [4, Exercise 4], then $\mathbb{R}$ is closed and bounded in that metric, but $\mathbb{R}$ is clearly not compact.

However, if we equip $\mathbb{R}^n$ with the usual metric, then a subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded. This result is called the Theorem of Heine–Borel and is one of the most important theorems of these notes. The theorem will be proved in the next subsection.

We shall now look on sequential compactness, in particular in metric spaces. We need the following:

Definition 6.18 *Let $(x_n) \subseteq X$ be a sequence. If $(n_k) \subseteq \mathbb{N}$ with $n_k < n_{k+1}$ for all $n \in \mathbb{N}$, then (x_{n_k}) is called a subsequence of (x_n).*

Here is a trivial lemma which readers often forget to take into account.

Lemma 6.19 *If $(x_n) \subseteq X$ and the set $\{x_n \mid n \in \mathbb{N}\}$ is finite, then (x_n) has a constant subsequence (x_{n_k}). In particular (x_{n_k}) is convergent.*

Proof: Since the finite set $A = \{x_n \mid n \in \mathbb{N}\}$ is indexed by the infinite set $\mathbb{N}$ there must be an $a \in A$ so that $N_1 = \{n \in \mathbb{N} \mid x_n = a\}$ is an infinite set. If we number N_1 by a sequence (n_k) with $n_k < n_{k+1}$ for all $k \in \mathbb{N}$, we get that $x_{n_k} = a$ for all k. A constant sequence is clearly convergent. $\square$

We are now able to give the following characterization of sequentially compact subsets of metric spaces.

Theorem 6.20 *Let K be a subset of the metric space (X, d). The following statements are equivalent.*

(i) K is sequentially compact.

(ii) Every sequence $(x_n) \subseteq K$ has a convergent subsequence with limit point in K.

Proof: $(i) \Rightarrow (ii)$: Let $K \subseteq X$ be sequentially compact and let $(x_n) \subseteq K$ be a sequence. Because of Lemma 6.19 it is no loss of generality to assume that the set $A = \{x_n \mid n \in \mathbb{N}\}$ is infinite and therefore A has an accumulation point $x \in K$. By induction we shall construct a subsequence (x_{n_k}) of (x_n) so that

$$x_{n_k} \in B(x, \frac{1}{k}) \quad \text{for all } k \in \mathbb{N}. \tag{6.5}$$

Since x is an accumulation point for A we can choose a point $x_{n_1} \in B(x, 1)$.

Assume next that $k \geq 1$ and that we have constructed $\{x_{n_j} \mid 1 \leq j \leq k\}$ so that (6.5) holds. Since by Theorem 4.21 the set $B(x, \frac{1}{k+1}) \cap A$ is infinite, we can find a $n_{k+1} > n_k$ so that $x_{n_{k+1}} \in B(x, \frac{1}{k+1})$. This finishes the induction.

It follows from (6.5) that $d(x, x_{n_k}) \to 0$ so that $x_{n_k} \to x$. This proves $(i) \Rightarrow (ii)$.

$(ii) \Rightarrow (i)$: Assume that (ii) holds and let D be an infinite subset of K. We have to prove that D has an accumulation point in K. Since D is infinite, we can find a sequence $(x_n) \subseteq D$ with $x_n \neq x_m$ for all $n \neq m$ which then by assumption has a convergent subsequence (x_{n_k}) with $x = \lim_k x_{n_k} \in K$. If $B(x)$ is any ball with center x, there is a $k_0 \in \mathbb{N}$ so that $x_{n_k} \in B(x)$ for all $k \geq k_0$. Since the x_{n_k}'s are mutually different, this means that $B(x)$ contains infinitely many points from D so that x is an accumulation point of D. $\square$

Note that we used the Axiom of Choice in $(i) \Rightarrow (ii)$.

The next theorem which states that compactness and sequentially compactness are equivalent notions in metric spaces is one of our main theorems.

Theorem 6.21 *If K is a subset of the metric space (X, d), the following statements are equivalent.*

(i) *K is compact.*

(ii) *K is sequentially compact*

For general topological spaces (ii) need not imply (i).

The proof of Theorem 6.21 is technically somewhat difficult and we need two lemmas before we can give the proof.

Lemma 6.22 *If (x_n) is a sequence in a metric space (X, d) and there is an $r > 0$ so that $d(x_n, x_m) \geq r$ for all $n \neq m$, then (x_n) has no convergent subsequences.*

Proof: Assume that (x_n) has a convergent subsequence (x_{n_k}) with limit x. Hence we can find a $k_0 \in \mathbb{N}$ so that $d(x, x_{n_k}) < \frac{r}{2}$ for all $k \geq k_0$. If now $k, j \geq k_0$ with $k \neq j$, then by the triangle inequality

$$d(x_{n_k}, x_{n_j}) \leq d(x_{n_k}, x) + d(x, x_{n_j}) < r$$

which contradicts our assumptions on the sequence (x_n). $\square$

The next lemma is more involved:

Lemma 6.23 *Let K be a sequentially compact subset of a metric space (X, d). If $\mathcal{C}$ is an open cover of K, there exists an $r > 0$ so that there for all $x \in K$ exists a $U \in \mathcal{C}$ with $B(x, r) \subseteq U$.*

Proof: For every $n \in \mathbb{N}$ we define

$$A_n = \{x \in K \mid B(x, \frac{1}{n}) \setminus U \neq \emptyset \quad \text{for all } U \in \mathcal{C}\} \tag{6.6}$$

and claim that there is a $m \in \mathbb{N}$ so that $A_m = \emptyset$.

To prove this claim by contradiction we assume that $A_n \neq \emptyset$ for all $n \in \mathbb{N}$ and choose an $x_n \in A_n$ for all n. Since K is sequentially compact, there exist by Theorem 6.20 a subsequence (x_{n_k}) and an $x_0 \in K$ so that $x_{n_k} \to x_0$. Since $\mathcal{C}$ is a cover of K, there is a $U_0 \in \mathcal{C}$ with $x_0 \in U_0$ and since U_0 is open, there is an $\varepsilon > 0$ with $B(x_0, \varepsilon) \subseteq U_0$. Let us choose k_0 so that $n_{k_0}^{-1} < \frac{\varepsilon}{2}$. Since $x_{n_k} \to x_0$, we can find a k_1 so that

$$d(x_{n_k}, x_0) < \frac{\varepsilon}{2} \quad \text{for all } k \geq k_1.$$

We now put $k = \max\{k_0, k_1\}$ and claim that $B(x_{n_k}, n_k^{-1}) \subseteq B(x_0, \varepsilon)$. Indeed, if $y \in B(x_{n_k}, n_k^{-1})$, then

$$d(y, x_0) \leq d(y, x_{n_k}) + d(x_{n_k}, x_0) < n_k^{-1} + \frac{\varepsilon}{2} \leq \varepsilon$$

which shows that $y \in B(x_0, \varepsilon)$. Summarizing we have proved that

$$B(x_{n_k}, n_k^{-1}) \subseteq B(x_0, \varepsilon) \subseteq U_0$$

which clearly contradicts that $x_{n_k} \in A_{n_k}$.

Hence we have proved that there is an $m \in \mathbb{N}$ so that $A_m = \emptyset$. We claim that $r = m^{-1}$ is the number we are looking for. Indeed, if $x \in K$ is arbitrary, then of course $x \notin A_m$ and by the definition of A_m there is a $U \in \mathcal{C}$ so that $B(x, m^{-1}) \subseteq U$. $\square$

Note that the important information in the lemma of course is that the number r does not depend on the points $x \in K$ but only on the covering $\mathcal{C}$ of K.

The reader will also observe that the Axiom of Choice was used in the proof.

We are finally ready for:

Proof of Theorem 6.21: It follows from Theorem 6.10 that (i) implies (ii) for all topological spaces so we only need to prove that (ii) implies (i) which we want to do by contradiction. If K is sequentially compact and not compact, we can find an open cover $\mathcal{C}$ of K which does not have a finite subcover and let r be the number from Lemma 6.23. By induction we shall construct a sequence $(x_n) \subseteq K$ so that for all $n \in \mathbb{N}$ we have:

$$x_{n+1} \notin \bigcup_{k=1}^{n} B(x_k, r) \tag{6.7}$$

For $n = 1$ we just pick some $x_1 \in K$. Let next $n \geq 1$ and assume that we have found $\{x_k \mid 1 \leq k \leq n\}$ so that (6.7) holds. By Lemma 6.23 we can to every $1 \leq k \leq n$ find $U_k \in \mathcal{C}$ so that $B(x_k, r) \subseteq U_k$ and hence also $\cup_{k=1}^{n} B(x_k, r) \subseteq \cup_{k=1}^{n} U_k \neq K$ since $\mathcal{C}$ does not have a finite subcover. Hence we can find an $x_{n+1} \subseteq K \setminus \cup_{k=1}^{n} B(x_k, r)$ which finishes the induction so we have constructed our sequence (x_n).

It follows immediately from (6.7) that $d(x_n, x_m) \geq r$ for all $n \neq m$ and therefore (x_n) does not have any convergent subsequence according to Lemma 6.22. This contradicts that K is sequentially compact. $\square$

6.3 Heine–Borels Theorem: Characterization of the compact subsets of $\mathbb{R}^n$

The main result of this section is The Heine Borel Theorem, but before we can prove it we need a little notation and a general result on products of compact metric spaces.

If $n \in \mathbb{N}$ and $(X_1, d_1), (X_2, d_2), \cdots, (X_n, d_n)$ are metric spaces, the product topology on $\prod_{j=1}^n X_j$ is metrizable. Indeed, if $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in \prod_{j=1}^n$ and $\mathbf{y} = (y_1, y_2, \cdots, y_n) \in \prod_{j=1}^n X_j$, the following equivalent metrics induce the product topology:

$$\mathbf{d_1}\mathbf{x}, \mathbf{y}) = \sum_{j=1}^n d_j(x_j.y_j), \tag{6.8}$$

$$\mathbf{d_2}(\mathbf{x}, \mathbf{y}) = (\sum_{j=1}^n d_j(x_j, y_j))^{\frac{1}{2}} \tag{6.9}$$

$$\mathbf{d}_\infty(\mathbf{x}, \mathbf{y}) = \max\{d_j(x_j, y_j) \mid 1 \leq j \leq n\}. \tag{6.10}$$

That (6.8), (6.9) and (6.10) define metrics which are equivalent can be proved as in [4, Exercises 1, 2]. Note that in case $X_j = \mathbb{R}$ for all $1 \leq j \leq n$, it is the metric $\mathbf{d_2}$ we normally use on $\mathbb{R}^n$; compare with Corollary 2.4.

We shall write a sequence in $\prod_{j=1}^n X_j$ as $(\mathbf{x}(k)) \subseteq \prod_{j=1}^n X_j$ because we want to use subscripts to indicate coordinates.

Our first lemma is left to the reader:

Lemma 6.24 *Let $(\mathbf{x}(k)) \subseteq \prod_{j=1}^n X_j$ be a sequence and let us for each $1 \leq j \leq n$ denote the j'th cordinate of $\mathbf{x}(k)$ by $x_j(k)$. If $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in \prod_{j=1}^n X_j$, then $\mathbf{x}(k) \to \mathbf{x}$ if and only if $x_j(k) \to x_j$ for all $1 \leq j \leq n$.*

We are now ready to prove the following important result.

Theorem 6.25 *If X_j $1 \leq k \leq n$ are compact metric spaces, then $\prod_{j=1}^n X_j$ is compact.*

Proof: We shall prove the theorem by induction on n but we will first prove it for $n = 2$ since the argument for this is the same as that for the induction step.

The case $n = 2$**:** Since $X_1 \times X_2$ is a metric space, it follows from Theorem 6.21 that it is enough to prove that $X_1 \times X_2$ is sequentially compact. Hence let $(\mathbf{x}(k)) \subseteq X_1 \times X_2$ be an arbitrary sequence from which we want to extract a convergent subsequence. Let us for each k write $\mathbf{x}(k) = (x_1(k), x_2(k))$. We proceed as follows:

Since X_1 is sequentially compact, we can find a convergent subsequence $(x_1(k_i(1))$ of $(x_1(k))$ and let us put $x_1 = \lim x_1(k_i(1))$. Since also X_2 is sequentially compact, we can find a convergent subsequence $(x_2(k_i(2))$ of $(x_2(k_i(1))$ and put $x_2 = \lim x_2(k_i(2))$.

We observe that since $(x_1(k_i(2)))$ is a subsequence of $(x_1(k_i(1)))$, we still have $x_1 = \lim x_1(k_i(2))$ and therefore it follows from Lemma 6.24 that if we put $\mathbf{x} = (x_1, x_2)$, then $\mathbf{x}(k_i(2)) \to \mathbf{x}$ in $X_1 \times X_2$. This proves that $X_1 \times X_2$ is sequentially compact.

The induction: For $n = 1$ it is just the assumption on X_1 so let us assume that we have proved the theorem for some $n \geq 1$. Hence $\prod_{j=1}^{n} X_j$ is compact, but since

$$\prod_{j=1}^{n+1} X_j = (\prod_{j=1}^{n} X_j) \times X_{n+1},$$

it follows from the case $n = 2$ that $\prod_{j=1}^{n+1} X_j$ is compact. This finishes the induction and hence also the proof of the theorem. $\square$

A subset $Q \subseteq \mathbb{R}^n$ is called a box if it of the form $Q = \prod_{j=1}^{n} I_j$ where I_j is an interval for all $1 \leq j \leq n$. If all the intervals I_j are bounded, we shall call Q a bounded box and if they are all closed, we shall call Q a closed box. Hence Q is a closed and bounded box if all the intervals I_j are closed and bounded.

We can now prove:

Theorem 6.26 *(***Heine– Borel***)*
Let $K \subseteq \mathbb{R}^n$. The following two statements are equivalent.

(i) K is closed and bounded.

(ii) K is compact.

Proof: Since the implication $(ii) \Rightarrow (i)$ is true in all metric spaces, we only have to prove that (i) implies (ii) and we divide the proof into two parts.

Part 1: Let $Q \subseteq \mathbb{R}^n$ be a closed and bounded box, say $Q = \prod_{j=1}^{n} [a_j, b_j]$ where $a_j, b_j \in \mathbb{R}$ with $a_j < b_j$ for all $1 \leq j \leq n$. By Theorem 6.4 $[a_j, b_j]$ is compact for all $1 \leq j \leq n$ and therefore we get from Theorem 6.25 that Q is compact.

Part 2: Let now $K \subseteq \mathbb{R}^n$ be closed and bounded. Since K is bounded, we can find a closed and bounded box Q with $K \subseteq Q$ and since Q is compact by Part 1 and K is closed, it follows from Theorem 6.11 that K is compact. $\square$

We will end this section with two results on infinite products of topological spaces, but we shall not go into details.

If $(X_n, d_n))$ is a sequence of metric spaces, then the infinite product $\prod_{n=1}^{\infty} X_n$ is metrizable in its product topology. Indeed, we may by [4, Exercise 4] assume that all the d_n's are bounded by

1 and the metric we need is e.g.

$$\mathbf{d}(\mathbf{x},\mathbf{y}) = \sum_{n=1}^{\infty} 2^{-n} d_n(x_n, y_n).$$

Theorem 6.27 *If (X_n) is a sequence of compact metric spaces, then $\prod_{n=1}^{\infty} X_n$ is compact in its product topology.*

The proof of Theorem 6.27 resembles that of Theorem 6.25. Indeed we take a sequence in the product space and continue to take subsequences of the coordinate sequences. To get a subsequence at the end which will work for all coordinates we use the diagonal technique from [4, Exercise 37]. The Axiom of Choice is used.

The most general result in this direction, the proof of which is outside the scope of these notes, is the following.

Theorem 6.28 *(***Tychonov's Theorem***)*
Let I be any index set. If $\{X_i \mid i \in I\}$ is a family of compact topological spaces, then $\prod_{i\in I} X_i$ is compact in its product topology.

It can be proved that Tychonov's Theorem is equivalent to The Axiom of Choice.

6.4 Continuous functions on compact spaces

In this subsection we shall investigate the behaviour of continuous functions defined on compact topological spaces and will prove some fundamental results in Mathematical Analysis. These will also lead to new results in the classical situation of real–valued continuous functions defined on a closed and bounded interval of $\mathbb{R}$. We start with:

Theorem 6.29 *Let X and Y be topological spaces and let $f : X \to Y$ be a continuous function. If $K \subseteq X$ is compact, then $f(K)$ is compact.*

Proof: Let $\mathcal{C}$ be an arbitrary open cover of $f(K)$, that is

$$f(K) \subseteq \bigcup_{G\in\mathcal{C}} G,$$

which implices that

$$K \subseteq \bigcup_{G\in\mathcal{C}} f^{-1}(G). \tag{6.11}$$

Since f is continuous and every $G \in \mathcal{C}$ is open, $f^{-1}(G)$ is open for all $G \in \mathcal{C}$ and therefore $\{f^{-1}(G) \mid G \in \mathcal{C}\}$ is an open cover of K. By the compactness of K we can find finitely many $G_1, G_2, \cdots, G_n \in \mathcal{C}$ so that

$$K \subseteq \bigcup_{k=1}^{n} f^{-1}(G_k)$$

and if we now use f on both sides of this, we get:

$$f(K) \subseteq \bigcup_{k=1}^{n} G_k.$$

Hence we have found a finite subcover of $\mathcal{C}$ and can conclude that $f(K)$ is compact. $\square$

If we combine Theorem 6.29 with the Heine–Borel Theorem we get the following fundamental results.

Theorem 6.30 *If X is a compact topological space and $f : X \to \mathbb{R}$ is continuous, then $f(X)$ is bounded and f attains its maximal and minimal values.*

Proof: By Theorem 6.29 $f(X)$ is compact and therefore closed and bounded and Corollary 4.20 gives then that $\sup f(X) \in f(X)$ and $\inf f(X) \in f(X)$. Hence we can find an $x \in X$ and a $y \in X$ with $f(x) = \sup f(X)$ and $f(y) = \inf f(X)$. $\square$

An immediate consequence is:

Corollary 6.31 *Let $a, b \in \mathbb{R}$ with $a < b$. If $f : [a, b] \to \mathbb{R}$ is continuous, then it is bounded and attains its maximal and minimal values.*

Proof: The Heine–Borel Theorem gives that $[a, b]$ is compact and the conclusion then follows from Theorem 6.30. $\square$

Some readers without any prior knowledge of topology might feel that they know Corollary 6.31, but since compactness is needed, they have certainly never seen a proof of it. Many of you probably know the corollary with the extra assumption that f is differentiable in $]a, b[$.

We end this subsection with an important result on uniform continuity which was already mentioned earlier. Before we go on, the reader should recall the defintion of uniform continuity in Definition 3.6.

Theorem 6.32 *Let (X, d_X) and (Y, d_Y) be metric spaces so that X is compact. If $f : X \to Y$ is continuous, then it is uniformly continuous.*

Proof: Let $\varepsilon > 0$ be given arbitrarily. Since f is continuous, we can to every $x \in X$ find a $\delta_x > 0$ so that

$$\forall y \in X : d_X(x, y) < \delta_x \Rightarrow d_Y(f(x), f(y)) < \frac{\varepsilon}{2}. \tag{6.12}$$

Clearly

$$X = \bigcup_{x \in X} B(x, \frac{\delta_x}{2}) \tag{6.13}$$

so that $\{B(x, \frac{\delta_x}{2}) \mid x \in X\}$ constitutes an open cover of X. Since X is compact, we can find finitely many points $x_1, x_2, \cdots, x_n \in X$ so that already

$$X = \bigcup_{k=1}^{n} B(x_k, \frac{\delta_{x_k}}{2}). \tag{6.14}$$

We now put

$$\delta = \min\{\frac{\delta_{x_k}}{2} \mid 1 \leq k \leq n\} \tag{6.15}$$

and note that since we have only taken minimum of finitely many positive numbers in (6.15), $\delta > 0$. We claim that this is the δ we need to finish the proof and let therefore $x, y \in X$ be arbitrary with $d_X(x,y) < \delta$ and have to estimate the distance between $f(x)$ and $f(y)$.

By (6.14) there is a j so that $x \in B(x_j, \frac{\delta_{x_j}}{2})$, that is $d_X(x, x_j) < \frac{\delta_{x_j}}{2}$, and therefore by (6.12) we get that

$$d_Y(f(x), f(x_j)) < \frac{\varepsilon}{2}. \tag{6.16}$$

We now wish to estimate the distance from y to x_j and get by the triangle inequality:

$$d_X(y, x_j) \leq d_X(y, x) + d(x, x_j) < \delta + \frac{\delta_{x_j}}{2} \leq \delta_{x_j} \tag{6.17}$$

so that again (6.12) gives us that

$$d_Y(f(y), f(x_j)) < \frac{\varepsilon}{2}. \tag{6.18}$$

If we now compare (6.16) and (6.18) and use the triangle inequality once more, we get:

$$d_Y(f(x), f(y)) \leq d_Y(f(x), f(x_j)) + d(f(x_j), f(y)) < \varepsilon.$$

This proves that the δ was the one we needed. $\square$

We can now finish the investigation of the square root function which we started in Proposition 3.7:

Example 6.33 *Let $f : [0, \infty[\to \mathbb{R}$ be defined by $f(x) = \sqrt{x}$ for all $x \in [0, \infty[$. Then f is uniformly continuous. Indeed, by the classical theory f is a continuous function, and since the interval $[0,1]$ is compact, f is uniformly continuous on $[0,1]$. By Proposition 3.7 f is also uniformly continuous on $[1, \infty[$ so together this gives that f is uniformly continuous on $[0, \infty[$.*

The function x^2 is uniformly continuous on every closed and bounded interval of $\mathbb{R}$ but for this we do not need Theorem 6.32. We just differentiate and use the mean value theorem and [5, Exercise 8].

7 Connected sets

In this section we shall study connected subsets of topological spaces and how continuous functions behave on them. Loosely speaking, a topological spaces is connected if we cannot split it up in disjoint open subsets. Here is the formal definition:

Definition 7.1 *A topological space X is called connected if X and $\emptyset$ are the only subsets of X which are both open and closed.*

A subset $A \subseteq X$ is called connected if it is connected in the topology induced by the topology on X.

A subset of X is called disconnected if it is not connected.

Before we give any examples it is better to prove:

Theorem 7.2 *Let X be a topological space. The following conditions are equivalent:*

(i) X *is connected.*

(ii) If A and B are disjoint open subsets of X with $X = A \cup B$, then either $A = \emptyset$ (hence $B = X$) or $B = \emptyset$ (hence $A = X$).

(iii) If C and D are disjoint closed subsets of X with $X = C \cup D$, then either $C = \emptyset$ (hence $D = X$) or $D = \emptyset$ (hence $C = X$).

Proof: $(i) \Rightarrow (ii)$: Let X be connected and let A and B be open subsets of X with $X = A \cup B$ and $A \cap B = \emptyset$. Since $A = X \setminus B$, A is also closed, so either $A = \emptyset$ or $A = X$ from which (ii) follows.

$(ii) \Rightarrow (i)$: Assume (ii) and let G be a subset of X which is both open and closed and hence also $X \setminus G$ is also both open and closed. Since $X = G \cup (X \setminus G)$, (ii) gives that either $G = \emptyset$ or $G = X$.

$(ii) \Leftrightarrow (iii)$: This follows from the fact that if A and B are disjoint open sets with $X = A \cup B$, then A and B are also closed. Similarly if A and B are disjoint closed sets with $X = A \cup B$. $\square$

Before we go on, let us focus a little on the induced topology on an $A \subseteq X$. If A is open, then it follows easily from the definition that a set $G \subseteq A$ is open in the induced topology if and only if it is open considered as a subset of X. Similarly, if A is closed, then a set $F \subseteq A$ is closed in the induced topology if and only if F is closed as a subset of X. If A is neither open nor closed, the open sets and the closed sets can be pretty ugly. It is therefore appropriate to formulate the connectedness of an $A \subseteq X$ in terms of the open subsets of X:

Proposition 7.3 *Let X be a topological space and $A \subseteq X$. The following statements are equivalent:*

(i) A is connected.

(ii) If U and V are open subsets of X with $A \subseteq U \cup V$ and $U \cap V \cap A = \emptyset$, then either $A \subseteq U$ or $A \subset V$.

Proof: Left to the reader as [4, Exercise 28]. $\square$

Some theorems are needed to determine whether even "nice" sets like intervals are connected or not, but let us here give a simple example.

Example 7.4 *Let $x, y \in \mathbb{R}^2$ with $x \neq y$, let $r, s > 0$ with $r+s < d_2(x, y)$ and put $A = B(x, r) \cup B_c(y, s)$. Since $r+s < d_2(x, y)$, it follows from the triangle inequality that $B(x, r) \cap B_c(y, s) = \emptyset$ and since $A \cap B(x, r)$, $B(x, r)$ is open in the induced topology on A. $B_c(y, s)$ is also open in A; indeed, if we choose a $u > s$ so that still $r + u < d_2(x, y)$, then $B(y, u) \cap A = B_c(y, s)$ This shows that A is disconnected.*

Using Proposition 7.3 we easily get:

Proposition 7.5 *If A is a connected subset of a topological space X, then also $\overline{A}$ is connected.*

Proof: Let U and V be open subsets of X so that $U \cap V \cap \overline{A} = \emptyset$ and $\overline{A} \subseteq U \cup V$. Since A is connected, Proposition 7.3 gives that either $A \subseteq U$ or $A \subseteq V$ and we can without loss of generality assume that $A \subseteq U$. We want so show that also $\overline{A} \subseteq U$. Indeed, if $x \in \overline{A}$, then $x \in U \cup V$, but since $V \cap A = \emptyset$, V cannot be a neighbourhood of the contact point x and hence $x \in U$.

From Proposition 7.3 we conclude that $\overline{A}$ is connected. $\square$

To get some confidence we have:

Lemma 7.6 *If $a, b \in \mathbb{R}$ with $a < b$, then the interval $[a, b]$ is connected.*

Proof: We note that it follows from the remarks prior to Proposition 7.3 that a subset of $[a, b]$ is closed in the induced topology if and only if it is closed as a subset of $\mathbb{R}$.

Assume that $[a, b]$ is not connected. Then we can find closed non–empty subsets $A, B \subseteq \mathbb{R}$ so that $[a, b] = A \cup B$ and $A \cap B = \emptyset$ which we have named so that $b \in B$. Since A is bounded upwards by b, the supremum $\sup A = c$ exists and since A is closed, $c \in A$ by Corollary 4.20. Further since $b \notin A$, we must have $c < b$ so that $]c, b] \subseteq B$. This shows that every open interval with center c contains points from B so that c is a contact point of B and therefore $c \in B$ since B is closed. We have obtained that $c \in A \cap B$ which is a contradiction. $\square$

Like compactness, connectedness is stable under continuous functions as the next theorem shows.

Theorem 7.7 *Let X and Y be topological spaces and $f : X \to Y$ a continuous function. If X is connected, then also $f(X)$ is connected.*

Proof: Let $G \subseteq f(X)$ be both open and closed in $f(X)$. Since f is also continuous considered as a function from X to $f(X)$, we get that $f^{-1}(G)$ is both open and closed in X and therefore

$f^{-1}(G) = \emptyset$ or $f^{-1}(G) = X$. In the first case we get that $G = \emptyset$ and in the second case we get that $G = f(X)$. Hence $f(X)$ is connected. $\square$

We now wish to define a more intuitve concept of connectedness.

Definition 7.8 *If X is a topological space, then a continuous function $\gamma : [0, 1] \to X$ is called a path. We put $\gamma^* = \gamma([0, 1])$ and call γ^* the curve parametrized by γ.*

Definition 7.9 *A topological space X is called path connected if there for all $x, y \in X$ exists a path $\gamma : [0, 1] \to X$ so that $\gamma(0) = x$ and $\gamma(1) = y$.*

X being path connected means geometrically that any two points in X can be joined by a curve which stays inside X. It is clear that the set in Example 7.4 is not path connected.

Luckily we have the following theorem.

Theorem 7.10 *A path connected topological space X is connected.*

Proof: We assume that X is path connected and let $A \subseteq X$ be a non-empty subset of X which is both open and closed. We have to prove that $A = X$. Since $A \neq \emptyset$, we can choose a fixed point $x \in A$. If $y \in X$ is arbitrary, there is a path $\gamma : [0, 1] \to X$ so that $\gamma(0) = x$ and $\gamma(1) = y$ and since γ is continuous, $\gamma^{-1}(A)$ is both open and closed in $[0, 1]$ and because $x \in A$, we have $0 \in \gamma^{-1}(A)$. Since by Lemma 7.6 $[0, 1]$ is connected, we have that $\gamma^{-1}(A) = [0, 1]$ and in particular $1 \in \gamma^{-1}(A)$ so that $y = \gamma(1) \in A$. This shows that $A = X$. $\square$

The two notions of connectedness are unfortunately not equivalent as the following example shows.

Example 7.11 *Consider the following two sets in $\mathbb{R}^2$:*

$$L = \{(x, y) \in \mathbb{R}^2 \mid x = 0 \quad \textit{and} -1 \leq y \leq 1\}$$

$$E = \{(x, y) \in \mathbb{R}^2 \mid y = \cos(\frac{1}{x}) \quad \textit{and } 0 < x \leq 1\}$$

and put $S = E \cup L$. The set looks like this:

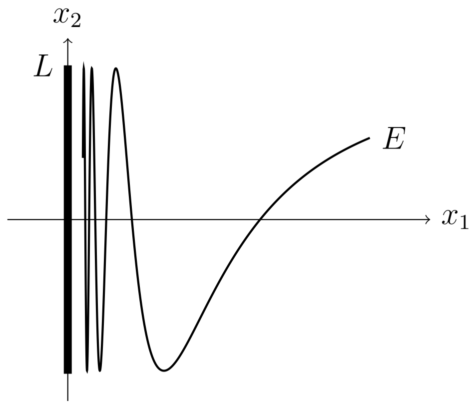

Figure 1: Test

We want to show that S is connected but not path connected.

Let us first prove that $\overline{E} = S$ and for this it suffices to prove that $L \subseteq \overline{E}$. Hence let $(0, y) \in L$ and let $\varepsilon > 0$ be arbitrary. Consider the open box Q with center $(0, y)$ defined by:

$$Q =]-\varepsilon, \varepsilon[\times]y - \varepsilon, y + \varepsilon[$$

(actually only the right half of Q is important here). Since cosine takes all values between –1 and 1 on the interval $]\varepsilon^{-1}, \infty[$, we can find an $x \in]0, \varepsilon[$ (actually infinitely many) so that $\cos(\frac{1}{x}) \in]y - \varepsilon, y + \varepsilon[$. Equivalently $(x, \cos(\frac{1}{x})) \in Q$ so that $Q \cap E \neq \emptyset$ which means that $(0, y) \in \overline{E}$

Since E is a curve itself, it is of course path connected and therefore connected by Theorem 7.10. Proposition 7.5 now gives that also $S = \overline{E}$ is connected.

We assume now that S is path connected and will prove that this leads to a contradiction.

By assumption we can find a path $\gamma : [0, 1] \to E$ so that $\gamma(0) = (0, 1)$ and $\gamma(1) = (\frac{2}{\pi}, 0)$. Since γ is continuous and L is closed in S, $\gamma^{-1}(L)$ is closed in $[0, 1]$ and hence if

$$t_0 = \sup \gamma^{-1}(L),$$

then $t_0 \in \gamma^{-1}(L)$ so that there is a $y_0 \in [-1, 1]$ with $\gamma(t_0) = (0, y_0)$. Note that $\gamma(t) \in E$ for all $t_0 < t \leq \frac{2}{\pi}$. Since γ is continuous, there is a $\delta > 0$ so that $\|\gamma(t) - \gamma(t_0)\|_2 < 1$ for all $t_0 \leq t < t_0 + \delta$. Let x_0 denote the first coordinate of $\gamma(t_0 + \delta)$.

The curve $\gamma([t_0, t_0 + \delta])$ is of course path connected and therefore we can to every $x \in]0, x_0]$ find a $t \in]t_0, t_0 + \delta]$ so that $\gamma(t) = (x, \cos(\frac{1}{x}))$. Since cosine attains all values between –1 and 1 on the interval $]0, x_0]$ we can in particular find $t_1, t_2 \in]t_0, t_0 + \delta[$ so that the second coordinate of $\gamma(t_1)$ is equal to 1 and the second coordinate of $\gamma(t_2)$ is equal to –1. But then $1 - y_0 < 1$ and $1 + y_0 < 1$ which is a contradiction.

For certain classes of sets it turns out that connectedness and path connectedness are actually equivalent notions and this we will go into now.

We recall that a subset K of a vector space V is called *convex* if for all points $x, y \in K$ and all $\lambda \in [0,1]$ we have that $\lambda x + (1-\lambda)y \in K$. Geometrically this means that the line segment connecting x with y is a subset of K. We have:

Proposition 7.12 *Every convex set of a normed space* $(X, \|\cdot\|)$ *is path connected. In particular every ball (open or closed) in* X *is path connected.*

Proof: Let $K \subseteq X$ be convex and let $x, y \in K$. The map γ defined by

$$\gamma(t) = tx + (1-t)y \quad \text{for all } t \in [0,1]$$

is clearly continuous and by definition $\gamma([0,1]) \subseteq K$.

Let us for $x_0 \in X$, and $r > 0$ consider the open ball

$$B(x_0, r) = \{x \in X \mid \|x - x_0\| < r\}$$

which we want to prove is convex. To this end we let $x, y \in B(x_0, r)$ and $t \in [0,1]$ and get

$$\begin{aligned}&\|tx + (1-t)y - x_0\| = \|t(x - x_0) + (1-t)(y - x_0)\| \leq \\ &t\|x - x_0\| + (1-t)\|y - x_0\| < r.\end{aligned}$$

This shows that $B(x_0, r)$ is convex and therefore path connected. A similar calculation for a closed ball. $\square$

We can now prove the following:

Theorem 7.13 *An open subset* G *of as normed space is connected if and only if it is path connected.*

Proof: Let $G \subseteq X$ be an open, connected set. If $G = \emptyset$, there is nothing to prove, so assume that $G \neq \emptyset$ and choose a point $x \in G$ which we keep fixed in the following. Define:

$$A = \{y \in G \mid \exists \gamma^* \subseteq G \quad \text{connecting } x \text{ with } y\}.$$

We want to prove that $A = G$ and do that by showing that A is both open and closed as a subset of G. Note that $x \in A$ so that $A \neq \emptyset$.

A **open:** Let $y \in A$ be arbitrary. Since G is open, we can find a ball $B(y)$ with center y so that $B(y) \subseteq G$. We wish to show that actually $B(y) \subseteq A$. To see this we pick a $z \in B(y)$ and note that by Proposition 7.12 the line segment L between y and z is contained in $B(y)$ and therefore also in G. Since there is a curve γ^* in G which connects x with y, the composition of γ^* and L wil give us a curve contained in G and connecting x with z. Hence $B(y) \subseteq A$. Since this works for all $y \in A$ we have proved that A is open.

A **closed in** G**:** We shall do this by showing that $G \setminus A$ is an open subset of G and the argument is quite similar to the one above. For an arbitrary $y \in (G \setminus A)$ we can find a ball $B(y) \subseteq G$ and claim that $B(y) \subseteq (G \setminus A)$. If this is not the case, we can find a point $z \in B(y) \cap A$. z can then be connected to x by a curve contained in G, but z can also be connected to y by a line segment contained in $B(y) \subseteq G$. Putting this together we can connect y with x by a curve contained in G and this contradicts $y \notin A$. Hence $B(y) \subseteq (G \setminus A)$ and we have proved that $G \setminus A$ is open.

Since G is connected and A is both open and closed and $x \in A$, we get that $A = G$. If now $y, z \in G$ are arbitrary, each of them can be connected to x by curves contained in G. Adding this together we can connect z with y by a curve contained in G. This shows that G is path connected. □

Corollary 7.14 *An open subset of* $\mathbb{R}^n$ *is connected if and only if it is path connected.*

In $\mathbb{R}^n$ we can actually do better so we introduce:

Definition 7.15 *A path* $\gamma : [0,1] \to \mathbb{R}^n$ *is called a step–path if there exists points* $0 = t_0 < t_1 < t_2 < \cdots < t_m = 1$ *so that* $\gamma([t_{j-1}, t_j])$ *is a line seqment parallel to one of the coordinate axes in* $\mathbb{R}^n$ *for all* $1 \leq j \leq m$. *We shall call* $\gamma([0,1])$ *a step–line.*

We need a result similar to Proposition 7.12.

Proposition 7.16 *If* B *is a ball in* $\mathbb{R}^n$ *and* $x, y \in B$, *then* x *and* y *can be joined by a step–line contained in* B.

Proof: Since every ball is a translate of a ball with center 0, it is enough to prove the statement for a ball

$$B(0,r) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 < r\}$$

where $r > 0$. It is clearly enough to show that any point in $B(0,r)$ can be joined with 0 by a step–line contained in $B(0,r)$. Hence let $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in B(0,r)$. Since $\mathbf{x_1} = (x_1, 0, 0, \cdots, 0) \in B(0,r)$, the line segment from 0 to $\mathbf{x_1}$ is contained in $B(0,r)$ and parallel to the first coordinate axis. Since also $\mathbf{x_2} = (x_1, x_2, 0, 0, \cdots, 0)$ belongs to the ball the line segment from $\mathbf{x_1}$ to $\mathbf{x_2}$ is also containd in the ball and parallel to the second coordinate axis. We now continue this way until we have reached $\mathbf{x}$. Clearly we get a step–line. □

The next result is very useful, in particular in Complex Function Theory.

Theorem 7.17 *If* $G \neq \emptyset$ *is an open connected subset of* $\mathbb{R}^n$, *then any two points in* G *can be joined by a step–line contained in* G.

Proof: In the proof of Theorem 7.13 we just write ”step–line” instead of ”curve” and use Proposition 7.16 instead of Proposition 7.12. □

Let us now characterize the connected subsets of $\mathbb{R}$.

Theorem 7.18 *A subset $S \subset \mathbb{R}$ is connected if and only if it is an interval.*

Proof: Assume first that $S \subseteq \mathbb{R}$ is not an interval. We wish to show that then S is disconnected. By assumption we can find points $x, y, z \in \mathbb{R}$ with $x < y < z$ and $x, z \in S$, but $y \notin S$. The sets

$U =]-\infty, y] \cap S$ and $V =]y, \infty[\cap S$

are both open in the induced topology from $\mathbb{R}$ and clearly $S = U \cup V$ and $U \cap V = \emptyset$. Since $x \in U$ and $z \in V$, both sets are non–empty and this proves that S is disconnected.

If S is an interval and $x, y \in S$, then $tx + (1-t)y \in S$ for all and this shows that S is path–connected and therefore connected. $\square$

As a consequence we get the following classical result:

Theorem 7.19 *Let $I \subset \mathbb{R}$ is an interval. If $f : I \to \mathbb{R}$ is a continuous function, then $f(I)$ is an interval.*

In particular, if f attains two values a and b, then it also attains any value between a and b.

Proof: Since I is connected and f is continuous, $f(I)$ is also connected and therefore an interval. The last statement in the clear. $\square$

Theorem 7.20 *Let X be a connected, compact topological sxpace. If $f : X \to \mathbb{R}$ is a continuous function, then $f(X) = [\min f(X), \max f(X)]$.*

Proof: Since X is connected and f is continuous, $f(X)$ is an interval and since X is compact, f attains its maximal and minimal value by Theorem 6.30. This proves the statement. $\square$

8 Homeomorfier

If X and Y are topological spaces and $f : X \to Y$ is a continuous bijection, then $f^{-1} : Y \to X$ need not be continuous as the following example shows:

Example 8.1 *Let $f : [0,1] \cup]2,3] \to [0,2]$ be defined by $f(x) = x$ for all $x \in [0,1]$ and $f(x) = x - 1$ for all $x \in]2,3]$ (make a sketch of the graph of f). Clearly f is a continuous bijection and it is also pretty clear that $f^{-1} : [0,2] \to [0,1] \cup]2,3]$ is not continuous in 1. The last statement can however also be seen as follows: Since f^{-1} maps a connected set onto a disconnect set, it follows from Theorem 7.7 that it cannot be continuous.*

We introduce the following definition:

Definition 8.2 *Let X and Y be topological spaces. A bijection $f : X \to Y$ is called a homeomorphism if both f and its inverse f^{-1} are continuous.*

If there exists a homeomorphism from X to Y, then X and Y are called topological equivalent.

We note that if $f : X \to Y$ is a homeomorphism, then a subset $U \subseteq X$ is open if and only if $f(U)$ is open, U is compact if and only if $f(U)$ is compact and U is connected if and only if $f(U)$ is connected. Also X is Hausdorff if and only if Y is Hausdorff. In other words X and Y have the same topological structures.

There is a very useful case where a continuous bijection automatically is a homeomorphism:

Theorem 8.3 *Let X be a compact space, Y a Hausdorff space and $f : X \to Y$ a continuous bijection. Then f is a homeomorphism.*

Proof: We have to show that f^{-1} is continuous and this will follow if we prove that for every closed $A \subseteq X$ $(f^{-1})^{-1}(A) = f(A)$ is closed in Y. Hence let $A \subseteq X$ be closed. Since X is compact, A is compact, and since f is continuous, $f(A)$ is compact. Since Y is a Hausdorff space, it follows that $f(A)$ is closed in Y. $\square$

Let us now discuss intervals in $\mathbb{R}$. We have:

Theorem 8.4 *If $a, b \in \mathbb{R}$ with $a < b$, then the following statements hold:*

(i) $[0,1]$ *is homeomorphic to* $[a,b]$

(ii) $]0,1[$ *is homeomorphic to* $]a,b[$.

(iii) $[0,1[$ *is homeomorphic to* $[a,b[$

(iv) *(*$]0,1]$ *is homeomorphic to* $]a,b]$

Proof: Let I be any of the four intervals with endpoints 0 and 1. The required homeomorphism is in all four cases given by $f(t) = ta + (1-t)b$ for all $t \in I$. $\square$

We have also:

Theorem 8.5 $\arctan$ *is a homeomorphism of* $\mathbb{R}$ *onto* $]-\frac{\pi}{2}, \frac{\pi}{2}[$.

Proof: $\arctan$ is continuous and it inverse $\tan$ is also continuous. $\square$

If we combine the Theorems 8.4 and 8.5 we obtain:

Corollary 8.6 *Every open interval in $\mathbb{R}$ is homeomorphis to $\mathbb{R}$*

We can also mention a negative result:

Proposition 8.7 *If $a \in \mathbb{R}$ and $a < b \leq \infty$, then there is no continuous bijection of $[a,b[$ onto an open interval.*

Proof: Let $I \subseteq \mathbb{R}$ be an open interval and assume that $f : [a,b[\to I$ is bijection. Clearly $f(]a,b[) = I \setminus \{f(a)\}$, and since $]a,b[$ is connected and $I \setminus \{f(a)\}$ is disconnected, f cannot be continuous. □

References

[1] C. Berg, *Metriske rum*, University of Copenhagen 1997.

[2] V.L. Hansen, *Grundbegreber i den moderne analyse*, The Tecnical University of Denmark, 1992.

[3] W.A. Sutherland, *Introduction to Metric and Topolgical Spaces*, Oxford Science Publications, 2006.

[4] *Opgaver til MM508 og MM509*, IMADA, 2008.

[5] *Additional Exercises for MM508*, IMADA, 2008.

[6] *Supperende noter til MM509*, IMADA, 2007.