METRIC & TOPOLOGICAL SPACES

JUHAN ARU

[1] This course is about topology. The word "topology" comes from Greek, where "topos" means place or space, and "logos" means study or science. [2] Poincaré has said that geometry is often called the art of reasoning about badly drawn figures. In topology this is even more true, as for a topologist all triangles are equal, a triangle and an octagon are equal, even a triangle and a circle are the same.

The most basic mathematical object is a set - just a list of distinct elements. A set has basically no geometry, no relation between its elements - its only characteristic property is its size. Topology is the weakest structure that one can attach to the set in order to give it some geometric content, to turn it into something that feels like a space. In a topological space one can already give a meaning to intuitive concepts like continuity and connectedness. Our basic questions will be:

- What can I say about a given topological space?
- How are the different properties of a topological spaces related between each other?
- Given two topological spaces, how can we decide if they are the same or different? (Recall that we cannot draw them!)

It might be helpful to think of the following table, where the geometric structure is increasingly detailed from left to right. Each next column is a specific case of the more general notion (e.g. metric spaces form a specific subclass of topological spaces) and thus each next column also contains all the structure and notions from the previous column:

	Set	Topological space	Metric space	Normed vector space
Geometric structure:	none	nearness	distance function	norm, linear structure
Geometric notions:	size	continuity, connectedness, compactness...	balls, boundedness, completeness...	straight line, convexity, origin...

Each class (column) comes with a notion of equivalence that says when two spaces are equal: roughly this is the case when all intrinsic properties of the spaces are the same. This is usually expressed by saying that two objects A and B are equivalent, if there is a function $f : A \to B$ (often called a map) with special properties, guaranteeing that the structure of A carries over unchanged to B.

In the case of sets, two sets are called equivalent if there is a bijection between them. For normed vector spaces equivalence means the existence of a bijective linear map that

[1] All kinds of feedback, including smaller or bigger typos, is appreciated - juhan.aru@epfl.ch

[2] Historically, topology carried a second name, analysis situs, but this got lost over time. Interestingly, this is also the name of an article by Poincaré in 1895 where he introduced several fundamental concepts of algebraic topology, like the fundamental group.

preserves the norm. For topological spaces such a map will be called a homeomorphism, for metric spaces an isometry. We will see for example that the usual $\mathbb{R}$ and $(0, 1)$ are equivalent as sets or topological spaces, but different as metric spaces. We will also see that $\mathbb{R}$ and $\mathbb{R}^2$ differ as topological spaces.

In fact, the description we have given fits into a more general theory of mathematical structures, called category theory. In this language such objects - sets with some structure together with structure-preserving maps - are called concrete categories. This notion is more general than just geometric structure, also groups with group homomorphism form a concrete category, or measure spaces with measurable maps, that we will meet in the probability theory course.

The aim of this course is to on the one hand see how the notion of topology gives the right context for basic concepts of analysis like continuity and connectedness. In fact, several notions of topology were introduced late 19th century to better understand analysis. We will for example revisit:

- Bolzano Weierstrass theorem on existence of convergent subsequences in bounded domains;
- Heine-Borel theorem on closed and bounded subsets of $\mathbb{R}^n$;
- Intermediate value theorem.
- Extreme value theorem.

On the other hand we will also look at a topological space as an interesting object on its own - by now topology has become a branch independent of analysis, has given rise to important subbranches like algebraic topology and has many applications in physics, biology, data science.

The course will consist of four chapters:

(1) Topology and continuity - here we introduce the basic formalism of topology in terms of open sets, and see how to formulate some basic notions like convergence and continuity.
(2) Connectedness - we discuss the abstract way to determine whether a space is connected or not, we also see a way to distinguish between $\mathbb{R}$ and $\mathbb{R}^2$ as topological spaces, and between $\mathbb{R}^2$ and $\mathbb{R}^3$.
(3) Compactness - compactness is the generalization of closed and bounded sets in $\mathbb{R}^n$, and is one of the most important and powerful concepts in topology. On compact intervals continuous functions are bounded and attain extrema.
(4) Metric spaces - metric spaces are a special subclass of topological spaces, where a distance function is defined between points. We will see that in this context, some notions have easier, more intuitive definitions. We study in some detail the space of continuous functions on a compact metric space, but also notions of separability. We touch upon the question - when can one construct a metric on a topological space?

We will begin though with a short reminder of basic set theory.

Section 0

Basics on set theory and functions

Axiomatic set theory, the so called Zermelo-Fraenkel (ZF) axioms form the foundation of mathematics. In this course we will not study this foundation, but take intuitive statements about sets for granted; only at some rare times will stop to ponder upon them. An axiomatic treatment can be found for example in the book of Munkres.

The following basic notions and results, that you have met in school / in the first year of EPFL will be used throughout the course:

(1) A set A is a collection of elements $a \in A$.
(2) For two sets A, B:
   - we say that A is a subset of B, denoted $A \subseteq B$, iff $\forall a \in A$, it holds that $a \in B$;
   - we say that A and B are equal, denoted $A = B$, iff $A \subseteq B$ and $B \subseteq A$:
   - we say that A is a proper subset of B if $A \subseteq B$, but $A \neq B$;
(3) Some useful set identities:
   - $(A \cup B) \cap (C \cup D) = (A \cap C) \cup (A \cap D) \cup (B \cap C) \cup (B \cap D)$;
   - $(A \cap B) \cup (A \cap C) = A \cap (B \cup C)$;
(4) De Morgan's laws on taking complements. Suppose $A, B \subseteq D$ and let $A^c = D \backslash A$ denote the complement. Then:
   - $(A \cap B)^c = A^c \cup B^c$;
   - $(A \cup B)^c = A^c \cap B^c$.
(5) The (Cartesian) product of sets $A \times B$, is defined as the set of ordered pairs (a, b). The square of a set A^2 denotes the product of A with itself.
(6) A subset R of $A \times A$ is called a relation. If $(a, b) \in R$, we say that $a \sim b$. A relation is symmetric if $a \sim b$ iff $b \sim a$, reflexive if $\forall a \in A, a \sim a$ and transitive if $a \sim b$ and $b \sim c$ imply that $a \sim c$. If a relation satisfies all these three properties, then it is an equivalence relation. An equivalence class of an element a, usually denoted $[a]$ contains all elements $b \in A$ such that $a \sim b$. Naturally, equivalence classes partition the set A. Similar definition of an equivalence relation can be given even if A is too big to be a set, e.g. if A correspond to all possible sets.
(7) A function $f : A \to B$
   - is called injective if $f(a) = f(a')$ implies $a = a'$;
   - is called surjective if $\forall b \in B$, there is some $a \in A$ with $f(a) = b$;
   - is called bijective if it is injective and surjective. A function is bijective if and only if f^{-1} can be defined everywhere. In this case both f and f^{-1} are bijections and $f^{-1} \circ f(a) = a$. An injection $f : A \to B$, always gives rise to a bijection $f : A \to f(A)$.
(8) For function $f : A \to B$, and some sets $A_0 \subseteq A$, $B_0 \subseteq B$ we define $f(A_0) = \{f(a) : a \in A_0\}$, and $f^{-1}(B_0) = \{a : f(a) \in B_0\}$. Notice that the latter is defined even if f^{-1} is not defined as a function.
(9) If $f : A \to B$, I_A, I_B are some sets of indexes, $(A_i)_{i \in I_A}$ is any collection of subsets of A, and $(B_i)_{i \in I_B}$ is any collection of subsets of B, then:
   - $f(\cup_{i \in I_A} A_i) = \cup_{i \in I_A} f(A_i)$;
   - $f(\cap_{i \in I_A} A_i) \subseteq \cap_{i \in I_A} f(A_i)$;

- $f^{-1}(\cup_{i \in I_B} B_i) = \cup_{i \in I_B} f^{-1}(B_i)$;
- $f^{-1}(\cap_{i \in I_B} B_i) = \cap_{i \in I_B} f^{-1}(B_i)$.

Exercise 0.1. *Prove the above identities for sets and functions. Prove that being in bijection induces an equivalence relation between sets.*

From the point of view of set theory all sets that are in bijection are the same. For example, for set theory sets $\{1, 2, 3\}$, $\{A, B, C\}$ and $\{John, Jack, Tom\}$ are all the same. If A and B are in bijection, we sometimes say that they are equivalent and write that $A \cong B$ - that this is really an equivalence relation was the content of the exercise just above.

A natural question to ask is when exactly are two sets equivalent, i.e. when is there a bijection between them? Is there some easy criteria for that? This would help us also somehow classify all possible sets. Let us start from finite sets.

0.1 Finite sets

Definition 0.1. *If a set A is in bijection with $\{1, \dots, n\}$ for some $n \in \mathbb{N}$, then we say that A is finite and define its size $|A|$ by $|A| := n$. By definition an empty set is also finite and has size 0. All other sets are called infinite.*

But why is the size well-defined? Why couldn't there be two different sets $\{1, \dots, n\}$ and $\{1, \dots, m\}$ in bijection with a finite set A?

One would like to argue as follows: Suppose that there are bijections $f : A \to \{1, \dots, n\}$ and $g : A \to \{1, \dots, m\}$. Then $g \circ f^{-1}$ is a bijection from $\{1, \dots, n\} \to \{1, \dots, m\}$, but this is possible only if $n = m$.

But why is this final step true? Why can $\{1, \dots, n\}$ not be in bijection with $\{1, \dots, m\}$ for some $m \neq n$? It would be very odd if a finite set is in bijection with its subset, so this statement is very intuitive, but is it mathematically true? We will for example see that in the case of infinite sets this is well possible!

The axiomatic setting of mathematics is exactly there to stop such discussions and to establish a firm answer to these questions, starting from the axioms. In other words, one just agrees on some set of axioms, on rules of logic and sees whether this is true or not in this system.

For example, in our case, let's agree (or if you wish, take as an "axiom") that an empty set cannot be in bijection with a non-empty set.[3] Agreed on this, the above statement can be and has to be proved:

Lemma 0.2. *For any $n \in \mathbb{N}$, the set $\{1, \dots, n\}$ cannot be in bijection with any of its proper subsets.*

I may to skip this proof in the first read, as it is not that interesting nor important, but I still wanted to give it for completeness...and of course I hope your curiosity still brings you to read it!

[$\star$ This proof is non-examinable $\star$]

[3] In fact, this "axiom" follows directly from the ZF axioms as in this axiomatic set-up, by definitions there is no function from the 1-element set to the empty set.

Proof. The case $m = 1$ follows from our "axiom": the only proper subset of $\{1\}$ is the empty set, and the empty set is not in bijection with the non-empty set $\{1\}$

Next, we use mathematical induction. So suppose the statement of the claim holds for any set $\{1, \dots, k\}$ for $k \leq n$. Now, consider the set $A := \{1, \dots, n+1\}$. Suppose for contradiction that $B \subseteq A$ is some proper subset of A and that there is a bijection $f : A \to B$. As B is a proper subset of A, there exists some $a \in A$ but $a \notin B$. Consider the sets $A' = A \backslash \{a\}$ and $B' = B \backslash \{f(a)\}$. Then $f : A' \to B'$ is still a bijection and $B' \subseteq A'$ is still a proper inclusion.

We will now draw a contradiction from this, by looking at different cases.

- If $n+1 \notin B$, we can take $a = n+1$. In this case $A' = \{1, \dots, n\}$ and $B' \subseteq \{1, \dots, n\}$ is a proper subset of A'. But this is in contradiction with the induction hypothesis.
- So suppose $n+1 \in B$. Pick some $a \notin B$ to obtain the sets A', B' and a bijection $f : A' \to B'$. We will have two sub-cases here.
  1. If $f(a) = n+1$, then B' is a proper subset of $\{1, \dots, n\}$. On the other hand, it is also in bijection with $\{1, \dots, n\}$: indeed, consider the function $h : \{1, \dots, n\} \to A'$ given by $h(b) = b$ if $b \neq a$ and $h(a) = n+1$. The function h is both injective and surjective, and thus bijective. Hence $f \circ h : \{1, \dots, n\} \to B'$ is a bijection as a composition of bijections and we get a contradiction with the induction hypothesis.
  2. So suppose $f(a) \neq n+1$ and consider $B'' := (B' \backslash \{n+1\}) \cup \{a\}$. Then by definition B'' is a proper subset of $\{1, \dots, n\}$. To draw a contradiction, let us show that $\{1, \dots, n\}$ is in bijection with B''. Indeed, let $h : \{1, \dots, n\} \to A'$ be as in the point above. Similarly, consider $g : B' \to B''$ given $g(b) = b$ if $b \neq n+1$ and $g(n+1) = a$, which is also a bijection. Then $g \circ f \circ h : \{1, \dots, n\} \to B''$ is a bijection as a composition of bijections. But B'' is a proper subset of $\{1, \dots, n\}$, and thus we again obtain a contradiction with the induction hypothesis.

□

[⋆ End of the non-examinable proof ⋆]

From this we can easily establish the corollary that the size of a finite set is well-defined, we can take a good breath and move on:

Corollary 0.3. *Each finite set A has a well-defined size.*

Proof. Suppose that there are bijections $f : A \to \{1, \dots, n\}$ and $g : A \to \{1, \dots, m\}$. Then $g \circ f^{-1}$ is a bijection from $\{1, \dots, n\} \to \{1, \dots, m\}$ and thus we deduce from the claim above that $n = m$. □

A very similar argument now shows that for finite sets size is really the only characterizing property:

Lemma 0.4. *Two finite sets are in bijection if and only if they have the same size.*

Proof. Denote these two sets by A, B.

By definition, if A, B have finite and equal size, then for some $n \in \mathbb{N}$, there exists bijections $f : A \to \{1, \dots, n\}$, $g : B \to \{1, \dots, n\}$. Then by composition law for bijections, $f^{-1} \circ g$ is a bijection from $B \to A$.

In the other direction, by assumption there is a bijection $\phi : A \to B$. But as A is finite, there is some $n \in N$ and a bijection $f : A \to \{1, \ldots, n\}$. Thus, by composition law for bijections, $\phi \circ f^{-1}$ is a bijection from $\{1, \ldots, n\}$ to B, showing that B has also size n. □

0.2 Infinite sets

But you already know that not all sets are finite, you have already met many "infinite" sets like $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{R}^n$. We should still check that they are infinite according to the definition above: for example,

Lemma 0.5. $\mathbb{N}$ *is infinite.*

Proof. Suppose for a contradiction that $\mathbb{N}$ is finite. Then for some $n \in N$, there should be a bijection $f : \mathbb{N} \to \{1, \ldots, n\}$. But then there would in particular exist an injection from $\{1, \ldots n+1\} \to \{1, \ldots, n\}$, and hence (as an injection is in bijection with its image) a bijection of $\{1, \ldots, n+1\}$ to its proper subset, which we know is impossible by Lemma 0.2. □

A natural question is how do the sizes of these infinite sets compare between each other?

Definition 0.6 (Countable/uncountable)**.** *If a set A is in bijection with $\mathbb{N}$ then it is called countably infinite. If a set is finite or countably infinite, then it is called countable. All other sets are called uncountable.*

One can verify that another way to say this is the following:

Exercise 0.2. *Prove that a non-empty set A is countable if and only if there is an injection $f : A \to \mathbb{N}$.*

In particular this means that any subset of $\mathbb{N}$ is countable, and thus either in bijection with $\mathbb{N}$ or with some $\{1, \ldots, n\}$ for some $n \in \mathbb{N}$. In plain words, any such set can be counted as $1, 2, \ldots, ...$ (possibly stopping at a finite number), thus the name countable. The whole numbers, the pairs of whole numbers and rationals are all countable:

Lemma 0.7. *The set $\mathbb{Z}$ is countably infinite.*

Proof. Consider the function $f : \mathbb{Z} \to N$ defined by $f(m) = 2m + 1$, if $m \geq 0$, and $f(m) = -2m$, if $m < 0$. It is then easy to check that this is a bijection. □

Exercise 0.3. *Prove that $\mathbb{Z}^2$ and $\mathbb{Q}$ are countably infinite. What about $\mathbb{Q}^{100}$?*

The real numbers, however, are not countable:

Proposition 0.8 (V)**.** *The set $\mathbb{R}$ is uncountable.*

Proof. This proof is called Cantor's diagonal argument. $\mathbb{R}$ is not finite by a similar argument that we gave for $\mathbb{N}$. We suppose by contradiction that there is a bijection from $\mathbb{N} \to \mathbb{R}$, and then show that we can construct some element of $\mathbb{R}$ without a preimage.

Indeed, suppose for contradiction that there is a bijection $f : \mathbb{N} \to \mathbb{R}$. Write the fractional part of $f(n)$ in its decimal expansion $0.x_1(n)x_2(n)\ldots$.

Our aim is now to define define a number $y \in \mathbb{R}$ that is not in the list $f(1), f(2), \ldots$. We will define this number using its decimal expansion: $y = 0.y_1y_2y_3\ldots$ as follows: for each $i \in \mathbb{N}$ just pick the digit y_i different from $x_i(i)$, say concretely $y_i := x_i(i) + 1 \mod 10$. Moreover, to guarantee that we are not producing an alternative decimal representation of a number (which, recall, is possible only if the decimal expansion ends with $99999\ldots$ or with $000000\ldots$) we further choose y_i different from y_{i-1}, for example in such a case set $y_i = x_i(i) + 2 \mod 10$. Then we know that there is no n such that $f(n) = x$, and thus f cannot be a bijection. □

A very similar argument gives the following result:

Exercise 0.4. *Prove that a set A cannot be in bijection with the set of its subsets. [Hint: It might be a good idea to start from taking $A = \mathbb{N}$ and think of the Cantor argument.] Deduce that there are sets that are bigger than $\mathbb{R}$, i.e. that there exist sets A such that there is an injection $\mathbb{R} \to A$, but no bijection.*

We have now shown that $\mathbb{R}$ is not in bijection with $\mathbb{Q}$ or $\mathbb{Z}$, it is in some sense "bigger". But what about for example $\mathbb{R}^2$. Or what about comparing $(0,1)$ and $[0,1]$? Or $\mathbb{R}$ and the set of continuous functions on $[0,1]$? Are they in bijection? In other words, are there uncountable sets that are not in bijection with $\mathbb{R}$? For example, are there sets of size between $\mathbb{N}$ and $\mathbb{R}$? [4]

0.3 Schröder-Bernstein theorem

To answer these questions, we need to know the existence of bijections. But how to construct a bijection between $(0,1)$ and $[0,1]$. The next theorem is very helpful in saying that the existence of a bijection can be deduced from just the existence of injections. It some sense it is similar to deducing that $a = b$ from knowing that $a \leq b$ and $b \leq a$.

Theorem 0.9 (Schroder-Bernstein). *If there is injection $f : A \to B$ and an injection $g : B \to A$, then there is a bijection between A and B.*

Corollary 0.10. *There is a bijection between $[0,1]$ and $(0,1)$.*

Proof. We want to apply Schroder-Bernstein theorem, so we need to prove the existence of an injection from $(0,1)$ to $[0,1]$ and vice-versa. Now, the identity map $f(x) = x$ gives the injection from $(0,1) \to [0,1]$. On the other hand the map $x \to x/2 + 1/3$ gives the injection from $[0,1]$ to $(0,1)$ and thus we are done. □

Exercise 0.5. *Prove that $\mathbb{R}$ and $\mathbb{R}^2$ are in bijection. What about $\mathbb{R}$ and the set of irrational numbers?*

The proof of Schroder-Bernstein is short, but a clever and tricky, and thus for the curious! [$\star$ The proof of Schroder-Bernstein theorem is for fun and non-examinable $\star$]

Proof. Notice that if $b \in f(A)$, then $f^{-1}(b)$ is well-defined as f is injective. Similarly, if $a \in g(B)$, then $g^{-1}(a)$ is well defined.

Thus, for any $a \in A$, we can consider the sequence of $a, g^{-1}(a), f^{-1} \circ g^{-1}(a), \ldots\ldots$, where we apply in turns g^{-1} and f^{-1} as long as it is possible, i.e. the functions are defined. For example, the sequence would stop at the first step if $a \notin g(B)$, as then $g^{-1}(a)$ is not defined. Notice that for any $a \in A$, there are three options:

[4] It's good to think about it, but not for too long - the Continuum hypothesis says that there is no such set, but in fact it cannot be proved from ZF axioms! It's independent of these axioms.

(1) Either this sequence continues infinitely;
(2) this sequence ends with an element $\overline{a}$ in A;
(3) or the sequence ends with an element $\overline{b}$ in B;

This trichotomy decomposes A into a disjoint union $A = A_1 \cup A_2 \cup A_3$, each corresponding to one of the options. Using this decomposition, we define a map $\phi : A \to B$ as follows:

(1) on A_1 we set $\phi(a) = f(a)$;
(2) on A_2 we also set $\phi(a) = f(a)$;
(3) on A_3 we, however, set $\phi(a) = g^{-1}(a)$.

To prove the theorem, it now suffices to argue that ϕ is a bijection.

To do this, first notice that for any $b \in B$, we could similarly consider the sequence $b, f^{-1}(b), g^{-1} \circ f^{-1}(b), \ldots$. For this sequence there are the same three options, and again they divide the set B into three disjoint subsets B_1, B_2, B_3 based on which of the three scenarios happens.

Now, observe the following:

- If $a \in A_1$, then $f(a) \in B_1$; and conversely, if $b \in B_1$, then $f^{-1}(b)$ is well-defined and $f^{-1}(b) \in A_1$. Thus ϕ restricted to A_1 is a bijection between A_1 and B_1.
- Similarly, if $a \in A_2$, then $f(a) \in B_2$; and conversely, if $b \in B_2$, then $f^{-1}(b)$ is well-defined and $f^{-1}(b) \in A_2$. Thus ϕ restricted to A_2 is a bijection between A_2 and B_2.
- Finally, if $a \in A_3$, then $g^{-1}(a) \in B_3$ and conversely for $b \in B_3$, we have that $g(b) \in A_3$; Thus ϕ restricted to A_3 gives a bijection between A_3 and B_3.

We conclude that $\phi : A \to B$ is a bijection.

$\square$

[$\star$ End of the non-examinable proof $\star$]

Section 1

Topology and continuous functions

1.1 What is a topology?

Let us start by the definition of a metric space - a set together with a distance function. This definition first appeared in Fréchet's PhD thesis in 1906 - his motivation was to understand better spaces of continuous functions and their properties, e.g. under which conditions do continuous functions converge to continuous functions.

Definition 1.1 (Metric space)**.** *Let X be any set. Then a metric is a function $d : X\times X \to \mathbb{R}$, satisfying the following conditions:*

- *reflexivity: $d(x, y) = 0$ iff $x = y$;*
- *symmetry: $d(x, y) = d(y, x)$;*
- *triangle inequality: $d(x, y) + d(y, z) \geq d(x, z)$ for any $x, y, z \in X$.*

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

It follows from these properties that in fact

Claim 1.2. *If d is a metric then $d(x, y) > 0$ if $x \neq y$.*

Proof. Indeed, by symmetry we have that $d(x, y) = d(y, x)$, by the triangle inequality we have that $2d(x, y) = d(x, y) + d(y, x) \geq d(x, x)$ and by reflexivity this equals 0. Thus we conclude that $d(x, y) \geq 0$. But we know that $d(x, y) = 0$ iff $x = y$, giving the conclusion. □

Some examples of metric spaces:

- Take X to be any set and define $d(x, y) = 1$ if $x \neq y$ and $d(x, x) = 0$. It is simple to check the three conditions for being a metric. This is called the discrete metric.
- Take $X = \mathbb{R}^n$ and let $d(x, y) = ||x - y||_2$. This way the Euclidean norm gives rise to a metric (check!). We call it the standard metric on $\mathbb{R}^n$. Any other norm would similarly give rise to a metric.
- One can also put metrics on more complicated objects: e.g. the next exercise asks you to how that the set $C([0, 1])$ of continuous functions on $[0, 1]$ with $d(f, g) = \sup_{x\in[0,1]} |f(x) - g(x)|$ gives rise to a metric space.
- Or, on more real-world objects: consider the set to be all towns in Switzerland and the distances the cycling times in hours. Why does this give rise to a metric space?

Exercise 1.1 (Space of continuous functions)**.** *Consider the set $C([0, 1])$ of continuous functions on $[0, 1]$ with the metric $d(f, g) = \sup_{x\in[0,1]} |f(x)-g(x)|$. Prove that this defines a metric space. What happens if you instead consider continuous functions on $(0, 1)$?*

We now come to the definition of topology that is usually traced back to a textbook by Hausdorff on set theory, published in 1914. This definition might look scary to begin with, but it will hopefully become a friend as the course progresses:

Definition 1.3. *Let X be any non-empty set. Then a topology on the set X is a collection τ of subsets $U \subseteq X$ satisfying the following properties:*

(1) Both $\emptyset \in \tau$ and $X \in \tau$;

(2) the intersection of any finite number of sets in τ, is again in τ;

(3) the union of any (possibly uncountable!) collection of sets in τ, is again in τ.

The sets $U \in \tau$ are called open sets and the pair (X, τ) is called a topological space.

Remark 1.4. *Notice that the second condition could have been also worded as:*

(2') The intersection of any two open sets is again an open set - the case of all finite sets would then follow by induction.

You should not worry if you don't see the intuition behind this definition straight away - at the moment this is just like looking at the architectural plan of the house, but we will also visit and see the house.

Historically it took some 40 years before people agreed on this very definition of topology: the notion of open sets started entering analysis in the very end of 19th century, and was first named in the PhD thesis of Baire in 1899 (open domains in $\mathbb{R}^n$) and Lebesgue in 1902. In the beginning, open sets were just used in the context of analysis on $\mathbb{R}^n$ to clarify the existing definitions and results. The idea of studying an abstract space came roughly at the same time, and was really first put down by Hausdorff: in his introductory book on set theory in 1914, he gave a definition of an abstract topological space, taking as a primitive idea "the neighbourhood of a point". Whereas he didn't use exactly open sets to define the topology [5], and had worked with a special type of topological space (that we will call a Hausdorff space later), this was the real starting point of topology. Interestingly, in the next edition of his book, he replaced general topological spaces again with metric spaces (i.e. spaces with a distance function), but now the idea had started spreading.

The current definition became most likely standard due to the Bourbaki group in France - this was a secret group of high-level French mathematicians who met, argued loudly and drank wine with the aim to write down proper foundations of mathematics, and develop all existing mathematics from these foundations. Whereas they didn't manage to fulfil their complete aim - mathematics was developing quicker than they could write it down - they did set the golden standard for many definitions and notions.

Here are some very basic examples of topologies.

- On the two element set $X = \{0, 1\}$ we could for example take $\tau = \{\emptyset, \{0, 1\}$ or also $\tau = \{\emptyset, \{0\}, \{0, 1\}\}$. It is easy to check that both collections of subsets satisfy the three properties for being a topology.
- On the set of three elements $X = \{0, 1, 2\}$ we could take several different topologies. For example $\tau = \{\emptyset, \{0\}, \{1, 2\}, \{0, 1, 2\}\}$ gives a topology, but for example $\tau = \{\emptyset, \{0, 1\}, \{1, 2\}, \{0, 1, 2\}\}$ does not - the intersection of the sets $\{0, 1\}$ and $\{1, 2\}$ is not contained in τ.

As the above shows, we can put many different topologies on a given set, giving them a different geometric structure. Sometimes two topologies on a set can be comparable: if τ^1 and τ^2 are two topologies on X and $\tau^1 \subseteq \tau^2$, then τ^2 is called finer and τ^1 coarser. Not all topologies are comparable (find an example!).

The two extremes topologies on any set are the discrete and the indiscrete topology:

[5] he used something called "neighbourhoods", but they also don't quite correspond to the intuitive idea of a neighbourhood in a town or city.

- The discrete topology τ_D on X contains all possible subsets and is thus the finest topology on any set. (Check that it is a topology!).
- The indiscrete topology τ_I contains only the empty set and the whole space and is thus the coarsest topology on any set X. (Check that it is a topology!).

It should, however, be mentioned that a finer topology doesn't necessarily mean more "interesting structure". We will make this more precise soon.

Exercise 1.2. *Find a chain of* 5 *topologies* $\tau_1, \ldots, \tau_5$ *on* $\{0, 1, 2\}$ *such that for each* $i = 1 \ldots 4$, τ_i *is coarser than* τ_{i+1}. *Write down* 3 *topologies on* $\{0, 1, 2\}$ *so that no pair is comparable.*

1.1.1 The metric topology

To gain some intuition about the definition of topology, let's go back to what we understand better. Namely, a central example of a topology is the topology induced by a metric space.

To start, we introduce the notion of a metric ball around some point $x \in X$ - the set of points closer than a certain distance to this point.

Definition 1.5 (Metric ball)**.** *Let* (X, d) *be a metric space. Then for any* $\delta > 0$, *we define the open ball to be* $B(x, \delta) := \{y \in X : d(x, y) < \delta\}$.

One way to define open sets of a metric space, i.e. to define the metric topology is to then just declare a set U open if around any point $x \in U$, we can find an open ball $B(x, \epsilon) \subseteq U$:

Lemma 1.6 (Metric topology)**.** *Let* (X, d) *be a metric space. Define a set of subsets* τ_d *as follows:*

- *We declare* $U \subseteq X$ *to be open (this is, we set* U *to be in* τ_d*), if for every* $x \in U$, *we can find some* $\delta > 0$ *such that* $B(x, \delta) \subseteq U$.

Then τ_d *is a topology and is called the metric topology.*

As a quick sanity check, notice that open metric balls are themselves open sets.

Proof. We need to verify the conditions of the topology:

(1) $X \in \tau_d$ as any metric ball is by definition a subset of X, $\emptyset \in \tau_d$ as the condition is void.
(2) Let $U_1 \in \tau_d, U_2 \in \tau_d$. Consider now $x \in U_1 \cap U_2$. Then there exits $\delta_1, \delta_2 > 0$ such that $B(x, \delta_1) \subseteq U_1$ and $B(x, \delta_2) \subseteq U_2$. In particular, if we set $\delta = \min(\delta_1, \delta_2)$ then $B(x, \delta) \subseteq U_1 \cap U_2$.
(3) Let I be any index set and $U_i \in \tau_d$ for all $i \in I$. Consider $x \in \bigcup_{i \in I} U_i$. Then there is some $i \in I$ such that $x \in U_i$ and thus by definition of τ_d, some $\delta > 0$ with $B(x, \delta) \subseteq U_i$. But then $B(x, \delta) \subseteq \bigcup_{i \in I} U_i$ giving the claim.

□

There is another equivalent way to define the metric topology:

Lemma 1.7. *Let* (X, d) *be a metric space and let* τ_d *be the metric topology. Then a set* U *is open if and only if it can be written as a union of open metric balls.*

Proof. As open metric balls are themselves open sets, all possible unions of them are open sets.

In the other direction, by definition of τ_d, for any $U \in \tau_d$ and any $x \in U$, we can find a ball $B(x, \delta_x) \subseteq U$. We now notice that $\bigcup_{x \in U} B(x, \delta_x) = U$. Indeed, as each $B(x, \delta_x) \subseteq U$, then also $\bigcup_{x \in U} B(x, \delta_x) \subseteq U$ and clearly $U \subseteq \bigcup_{x \in U} B(x, \delta_x)$. $\square$

Thus any metric space gives naturally rise to a topological space. In particular,

- The Euclidean metric in $\mathbb{R}^n$ gives rise to a topology on $\mathbb{R}^n$ called either the Euclidean or the standard topology.
- As should be expected, the discrete metric on any set X gives rise to the discrete topology τ_D on X. To prove this, it suffices to show that all singleton sets $\{x\}$ belong to τ_d - all other subsets are given by unions of those. But now, notice that we can write $\{x\} = B(x, 0.5)$. Moreover, by definition the open ball $B(x, 0.5)$ is an open set of the metric topology. Thus $\{x\}$ also is an open set and hence by what we said above indeed $\tau_d = \tau_D$.

1.1.2 Basis of a topology

In the metric topology there was a nice way to write any open set as a union of metric balls. Something similar is possible in all topological spaces, via the notion of a basis. A basis in topology is also a bit analogous to a basis of a vector space, we have some simpler bricks out of which we can build up everything else.

Definition 1.8 (Basis of a topology). *Let X be a set. Then a collection τ^B of subsets of X is called a basis for a topology, if it satisfies two conditions:*

- *The union of all $V \in \tau^B$ covers X (i.e. $\bigcup_{V \in \tau^B} V = X$);*
- *and for any $V_1 \in \tau^B, V_2 \in \tau^B$ and $x \in V_1 \cap V_2$, there is some $V_3 \in \tau^B$ with $x \in V_3 \subseteq V_1 \cap V_2$.*

Notice that a priori we have not yet defined a topology on X, just a certain collection of subsets. However, any such collection naturally generates a topology via the following proposition:

Proposition 1.9. *Given some set X, and any subset τ^B satisfying the conditions of Definition 1.8, consider τ to be the set of all possible unions of $V \in \tau^B$, together with the empty set. Then τ defines a topology on X. We also then say that τ_B is a basis for the topology τ.*

As many proofs in this section, this proof also contains no essential new idea - one hast to just directly verify the conditions for a topology. In fact, the mathematical content in basic topology hides rather in notions, definitions and structures than the proofs, which are more sort of verifications that things have been set up correctly.

Proof. We need to carefully check that the three conditions for a topology hold:

- The empty set belongs to τ by the condition of the proposition, and the full space by the conditions of Definition 1.8.
- Intersection property: $U_1 \in \tau$ and $U_2 \in \tau$, then by the condition of the proposition, we can write $U_1 = \bigcup_{i \in I} V_i$ and $V_2 = \bigcup_{j \in J} V_j$ for some index sets I, J and for some sets $V_i \in \tau^B$, $V_j \in \tau^B$ for all $i \in I$, $j \in J$. Then

$$U_1 \cap U_2 = (\bigcup_{i \in I} V_i) \cap (\bigcup_{j \in J} V_j) = \bigcup_{i \in I, j \in J} (V_i \cap V_j).$$

Now, consider $x \in U_1 \cap U_2$, then $\exists i_x \in I, j_x \in J$ such that $x \in V_{i_x} \cap V_{j_x}$. But by the second condition for being a basis, there is thus an element $V_x \in \tau^B$ such that $x \in V_x$ and $V_x \subseteq V_{i_x} \cap V_{j_x}$. It is now easy to see that $U_1 \cap U_2 = \bigcup_{x \in U_1 \cap U_2} V_x$, and thus $U_1 \cap U_2 \in \tau$.

- Union property: if $(V_i)_{i \in I}$ are some sets in τ, then they are unions of $U \in \tau^B$. But then also any union of them will be an union of $U \in \tau^B$, and hence an element in τ.

□

Notice that any topology containing the subset τ^B has to contain the union of all elements in τ^B. Thus a way to rephrase the previous lemma is to say that:

- The topology τ is the smallest topology containing the basis τ^B.

Conversely, we see that:

Exercise 1.3. *Let (X, τ) be a topological space. Let $\widetilde{\tau}^B \subseteq \tau$ be a collection of subsets such that any set in τ is an union of sets from $\widetilde{\tau}^B$. Prove that $\widetilde{\tau}^B$ is a basis for a topology, and that the topology induced by this basis is τ.*

It is good to verify that we didn't lie before - open balls do form a basis for the metric topology:

Lemma 1.10. *Let (X, d) be a metric space. Then $\tau^B := \{B(x, \delta) : x \in X, \delta > 0\}$ is a basis for the metric topology τ_d.*

Proof. By Exercise 1.3, we need to verify that any $U \in \tau_d$ can be written as a union of elements from τ^B. But we know this from Lemma 1.7. □

In the case of Euclidean topology, things are even nicer, all open sets can be described as the unions of countably many sets:

Exercise 1.4. *Prove that the balls $B(x, \delta)$ where $x \in \mathbb{Q}^n$ and $\delta \in \mathbb{Q} \cap (0, \infty)$ are a basis for the Euclidean topology on $\mathbb{R}^n$. Moreover, prove that there are countably many elements in this basis. Find a metric space which does not have a countable basis.*

One can now also see the difference between a basis in a topology and in a finite-dimensional vector space: in the latter, given a basis, each element has a unique writing as a linear combination of the basis elements, whereas here open sets could be written as many different unions using the sets in a topological basis.

1.1.3 Some non-metric topologies and closed sets

We have seen that metric topologies are quite nice. However, not all topologies stem from metric spaces. To give two quick examples:

- For any set X, the indiscrete topology $\tau_I = \{X, \emptyset\}$ does not stem from a metric. We will prove this shortly, but it might be a good idea to think a bit about it before.
- A slightly more interesting example is that of a co-finite topology. For any set X we take τ to contain all subsets, whose complement w.r.t. X is finite together with the empty set. It is an exercise to prove that this defines a topology. We will also see that this topology can also not stem from an underlying metric, but we need to develop some more concepts to prove this.

Exercise 1.5. *Prove that the co-finite topology defined above is indeed a topology.*

In defining co-finite topology, we saw how complements of open sets naturally enter the scene.

Definition 1.11 (Closed sets)**.** *Let (X, τ) be a topological space. A set C is called closed if it is a complement of some open set, i.e. if we can write $C = X \backslash U$ with $U \in \tau$.*

In fact, we could have defined the topology of a set also using closed sets. The set of closed sets has exactly the same information as the set of open sets. The closed sets would have had to satisfy then the following properties:

Exercise 1.6. *Let (X, τ) be a topological space. Then both $\emptyset$ and X are closed sets. Prove that any finite union of closed sets is again closed, and an arbitrary intersection of closed sets is closed.*

It is important to remark that there are many sets that are neither open nor closed: for example, if X is any set that contains at least two points, then for the indiscrete topology (X, τ_I) all the singletons $\{x_0\}$ for $x_0 \in X$ are neither open nor closed.

On the contrary, in the discrete topology every singleton is both open and closed: indeed, by definition of the discrete topology $\{x_0\}$ is open for any $x_0 \in X$; but as $X \backslash \{x_0\} = \bigcup_{x \in X; x \neq x_0} \{x\}$, we see that $\{x_0\}$ is also closed.

Claim 1.12. *In a metric space (X, d) all singletons $\{x\}$ are closed sets of the metric topology. Also, all closed balls $\overline{B(z, \delta)} = \{x \in X : d(x, z) \leq \delta\}$ are closed sets.*

Proof. Let $x \in X$. Then for any $y \in X$ the ball $B(y, 0.5d(x, y))$ does not contain x. Thus we can write $X \backslash \{x\}$ as $X \backslash \{x\} = \bigcup_{y \in X} B(y, 0.5d(x, y))$, which is open. Hence $\{x\}$ is closed.

The second part is on the example sheet.

□

In particular, as for the indiscrete topology on the space of at least two points the singletons are nor open, nor closed we have that:

Corollary 1.13. *The indiscrete topology on a space X with at least two points cannot be a metric topology for some metric space (X, d).*

1.1.4 Interior/closure/boundary

By definition closed and open sets are complements of each other. However, there are other natural relations between open and closed sets, defined in terms of certain operations.

Definition 1.14 (Interior and closure)**.** *Let (X, τ) be a topological space. Then for any subset $A \subseteq X$, the interior of A, denoted $int(A)$ is defined as the largest open set contained in A, and the closure of A, denoted $cl(A)$ is defined as the smallest closed set containing A.*

One should wonder why is this largest or smallest well-defined? For example $(0, 1)$ has no smallest element, so the question is well-justified. However, in our case things are very nice, because open sets are stable under union, and closed sets stable under intersection [6]:

Lemma 1.15. *Let (X, τ) be a topological space and A some subset. Then the interior and closure of A are well defined. The interior is equal to the union of all open sets contained in A, and the closure is equal to the intersection of all closed sets containing A.*

[6] And in particular, for those who are interested - because of this stability there is no need for Zorn's lemma here...

Proof. Let us consider the case of the interior. Then if $(U_i)_{i\in I}$ is the collection of open sets contained in A, we have that $V = \bigcup_{i\in I} U_i$ is open. But by definition of V, any open set $U_i \subseteq A$ satisfies $U_i \subseteq V$. Thus $V = int(U)$. □

It follows from the definition, for any open set U, $int(U) = U$ and for any closed set C, we have that $cl(C) = C$. Interestingly, it was in terms of closures that one first described an abstract topological space equivalent to our definition:

Exercise 1.7 (Kuratowski's formulation of a topological space)**.** *In 1922 Kuratowski considered an arbitrary set X with an operation named "closure" $\overline{cl}$ from the set of subsets of X to itself, satisfying the following axioms:*

(1) For any two subsets A, B of X, we have $\overline{cl}(A \cup B) = \overline{cl}(A) \cup \overline{cl}(B)$;
(2) $A \subseteq \overline{cl}(A)$;
(3) $\overline{cl}(\emptyset) = \emptyset$ and $\overline{cl}(X) = X$.

Show that if we start from a topological space (X, τ), then the closure defined by Definition 1.14 satisfies all these conditions.

In the opposite direction, suppose we have a set X and the operation $\overline{cl}$. Define a set C to be closed iff it equals its closure. Show that the sets $X\backslash C$ form a topology on X.

The notions of interior and closure motivate us to think of open sets as of an interior of a garden. A closed set then corresponds to a garden together with its fence, and closure to adding a fence to your piece of land. But what about the notion of a fence itself?

Definition 1.16 (Boundary of a set)**.** *Let (X, τ) be a topological space and $A \subseteq X$. Then the boundary of A, denoted ∂A, is defined as $cl(A)\backslash int(A)$.*

Notice that the boundary could also be empty, this happens iff the set is both closed and open. For example there are no boundaries in the discrete topology.

One would intuitively want to say that $\partial A = \partial(X\backslash A)$. And indeed, this is the case. To prove this we first find a maybe slightly less appealing definition of the boundary:

Lemma 1.17. *Let (X, τ) be a topological space and $A \subseteq X$. Then $x \in \partial A$ if and only if for any open set U containing x, we have that both $U \cap A$ and $U \cap (X\backslash A)$ are non-empty.*

Proof. If $x \notin \partial A$, then either $x \in int(A)$ or $x \in X\backslash cl(A)$. If $x \in int(A)$, then $int(A)$ is an open set surrounding x that doesn't intersect $X\backslash A$. Similarly, if $x \in X\backslash cl(A)$, then $X\backslash cl(A)$ is an open set containing x that doesn't intersect A.

In the opposite direction, suppose there is some open set U containing the point x, such that either $U \cap A$ is empty or $U \cap (X\backslash A)$ is empty. In the first case, notice that the closed set $X\backslash U$ must contain A. In particular, by definition of closure, we have that $cl(A) \subseteq X\backslash U$ and hence $x \notin cl(A)$. In the second case $U \subseteq A$, and as U is open it means that $U \subseteq int(A)$. But then again $x \notin cl(A)\backslash int(A)$ and the lemma follows. □

As a corollary we can now affirm our intuition:

Corollary 1.18. *Let (X, τ) be a topological space and $A \subseteq X$. Then $\partial A = \partial(X\backslash A)$.*

Proof. Indeed, the equivalent definition of the boundary in Lemma 1.17 is symmetric w.r.t A and $X\backslash A$. □

1.2 Convergence and continuous maps

As mentioned in the introduction, a topology endows the set X with a notion of nearness. As the topology is defined by a collection of open sets, it is then tempting to think that the open sets give some sort of neighbourhoods around points. There is some truth to it:

- If we deal with a metric topology, then open balls are also open sets, and open balls by definition contain points close to the centre of the ball.
- Moreover, sequences and continuity can be defined using open sets.

Indeed, the notion of convergence is a direct generalization of the convergence in $\mathbb{R}^n$, when we replace open balls with open sets.

Definition 1.19 (Convergence)**.** *Let (X, τ) be a topological space and $(x_n)_{n\geq 1}$ a sequence of points in X. We say that x_n converges to x, denoted $x_n \to x$ if for any open set U containing x there is some $n_U \in \mathbb{N}$ such that $\forall n \geq n_U : x_n \in U$.*

But one should also keep in mind that:

- There are many open sets that one would not intuitively think of as of neighbourhoods: even in a metric topology also unions of distant disjoint balls are open sets.
- Statements like "y is closer to x than z because it is in more open sets" usually don't make literal sense. The notion of nearness present in topology does not usually give a generic meaning to "is closer to". In fact, we will see that if you take $\mathbb{R}^2$ with the standard topology then you can map any distinct m points to any distinct m points without changing the topology.
- It is possible that all open sets are unimaginably big - e.g. think of open sets in the indiscrete or co-finite topology for $\mathbb{R}$.

Moreover, concepts like convergence might not always behave like we expect. Firstly, limits are not necessarily unique. Indeed, for example in a topological space with the indiscrete topology any sequence will converge to all points. A more interesting example is the following:

Exercise 1.8. *Consider $\mathbb{N}$ with its cofinite topology.*

- *Find a sequence that converges simultaneously to all $n \in \mathbb{N}$.*
- *Can you find a sequence that converges exactly to the points 1 and 2?*

Secondly, sequences may not suffice to describe the structure of the topological space. For example, from the Euclidean topology on $\mathbb{R}^n$ we are used to the fact that the boundary of a set can be described using sequences of points inside the set: i.e. if for some set A we have that $x \in \partial A$, then there is some sequence of points $(x_n)_{n\geq 1}$ in A such that $x_n \to x$. This is not necessarily the case in topological spaces - we will not be always able to define the boundary using sequences.

Exercise 1.9 ($\star$ McMullen p. 19-20)**.** *Consider $\mathbb{R}$ with the co-countable topology. Show that the closure of $(0, 1)$ is the whole space. On the other hand show that there is no sequence in $(0, 1)$ converging to 2.*

It might be worth mentioning that these two counter-intuitive things happen for different reasons. The first one has more to do with the fact that the topology might be too sparse - there are not enough open sets to separate points. The second one happens in fact when the topology is too big - there are just too many open sets to describe everything using sequences. We will come back to these two points at the end of the chapter.

1.2.1 Continuity at a point

Let us now get to continuity and start from looking at continuity at some fixed point. This way, the definition of continuity can be also seen as a direct generalization from the usual $\epsilon - \delta$ definition in real analysis. Let us see how this generalization happens step by step:

- A function $f : \mathbb{R} \to \mathbb{R}$ is called continuous at x if $\forall \epsilon > 0$, there is some $\delta > 0$ such that whenever $|z - x| < \delta$, then $|f(x) - f(x)| < \epsilon$.
- Now, rewrite the latter conditions using metric balls: A function $f : \mathbb{R} \to \mathbb{R}$ is called continuous at x if $\forall \epsilon > 0$, there is some $\delta > 0$ such that $f(B(x, \delta)) \subseteq B(f(x), \epsilon)$.
- We know that each open set around $f(x)$ contains an open ball around $f(x)$, so it is equivalent to require that: A function $f : \mathbb{R} \to \mathbb{R}$ is called continuous at x if for all open sets U around $f(x)$, there is some $\delta > 0$ such that $f(B(x, \delta)) \subseteq U$.
- Similarly, it is also equivalent to replace the second open ball by a general open set: A function $f : \mathbb{R} \to \mathbb{R}$ is called continuous at x if for all open sets U around $f(x)$, there is some open set V around x, such that $f(V) \subseteq U$.

This final definition generalizes nicely to topological spaces:

Definition 1.20 (Continuity at a point). *A map $f : X \to Y$ from a topological space (X, τ_X) to a topological space (Y, τ_Y) is continuous at a point $x \in X$, if for any open set U containing $f(x)$, there is some open set V_U containing x, such that $f(V_U) \subseteq U$.*

Exercise 1.10. *Verify carefully that this definition is equal to the usual definition for real functions.*

Remark 1.21. *From now onwards we will always, when working with continuous maps also keep track of the topologies: i.e. although the function is defined on the points of X, we will denote $f : (X, \tau_X) \to (Y, \tau_Y)$ to keep track of the topologies too. Different choices of topologies might make the same function continuous or discontinuous.*

It comes out that even in an abstract topological space, this definition implies that:

Lemma 1.22. *Suppose $(X, \tau_X), (Y, \tau_Y)$ are topological spaces and suppose that the map $f : (X, \tau_X) \to (Y, \tau_Y)$ is continuous at $x \in X$. Then for any sequence $(x_n)_{n \geq 1} \to x$, we have that $(f(x_n))_{n \geq 1} \to f(x)$.*

Proof. Consider a sequence $(x_n)_{n \geq 1}$ and for each $n \in \mathbb{N}$, let $y_n = f(x_n)$. Let further $y = f(x)$. We need to prove that for any open set $U \in \tau_Y$ containing y, there exists some $n_U \in \mathbb{N}$ such that $\forall n_U \geq n$, we have that $y_n \in U$.

So, consider an open set U containing y. By continuity of f at the point x (Definition 1.20), there is some open set V_U containing x such that $f(V_U) \subseteq U$. By the definition of convergence of $(x_n)_{n \geq} \to x$, there exists some $n_V \in \mathbb{N}$ such that $\forall n \geq n_V$ we have that $x_n \in V_U$, and thus by the definition of V_U, for all $n \geq n_V$ we have that $y_n = f(x_n) \in U$. $\square$

Remark 1.23. *Interestingly, the converse of this lemma does not hold in general topological spaces - i.e. one find topological spaces $(X, \tau_X), (Y, \tau_Y)$ and a function $f : X \to Y$ such that for some point x we know that for any sequence $(x_n)_{n \geq 1} \to x$, we have that $(f(x_n))_{n \geq 1} \to f(x)$; and yet f is not continuous at x. We will later see that the converse does hold for the metric topology.*

1.2.2 Continuity of a map

The global continuity of a map can be formulated in even a slicker way:

Proposition 1.24 (Continuous map)**.** *A map $f : X \to Y$ from a topological space (X, τ_X) to a topological space (Y, τ_Y) is continuous at every point $x \in X$ iff the pre-image of any open set is an open set, i.e. iff for any open set U of (Y, τ_Y) we have that $f^{-1}(U)$ is open in (X, τ_X).*

Often when we check continuity of a map, we will in fact use this condition. An equivalent condition would be to require that the pre-image of any closed set is a closed set. It is key to remember here is that continuity has to do with preimages:

Exercise 1.11 ($\star$)**.** *Find a function $f : \mathbb{R} \to \mathbb{R}$ such that $f(U)$ is open for all open sets U, but f is not continuous. Find a continuous function $f : \mathbb{R} \to \mathbb{R}$ that doesn't map open sets to open sets.*

Proof of Proposition 1.24. Suppose first that for any $U \in \tau_Y$ we have that $f^{-1}(U) \in \tau_X$. We want to show that f is continuous at all $x \in X$. So pick some $x \in X$ and an open set $U \in \tau_Y$ with $f(x) \in U$ (at least the whole space is such an open set). Then we know that $V_U := f^{-1}(U)$ is open and moreover $x \in V_U$, thus as $f(V_U) \subseteq U$, we see that f is continuous at x, according to Definition 1.20.

In the opposite direction, suppose that f is continuous at each $x \in X$ according to Definition 1.20. Take some open set U in Y. If $f^{-1}(U)$ is empty, then we are done as the empty set is an open set. Otherwise, consider any $x \in V := f^{-1}(U)$. Then by the continuity of f at x we know that there is some open set V_x with $f(V_x) \subseteq U$ and thus $V_x \subseteq V$. Now define $\hat{V} = \bigcup_{x \in V} V_x$. Then $\hat{V}$ is open as an union of open sets. Moreover, we have that $\hat{V} \subseteq V$ as for every $x \in V$ we have that $V_x \subseteq V$. On the other hand $V \subseteq \hat{V}$ as for any $x \in V$, we know that $x \in V_x \subseteq \hat{V}$. Thus $f^{-1}(U) = V = \hat{V}$ is open and hence indeed the preimage of every open set is open. $\square$

Here are some basic examples:

- The identity map from (X, τ_X) to itself is always continuous.
- The constant map from any space to any space is continuous.
- Consider $X = \{0, 1\}$, $\tau_1 = \{\emptyset, \{0\}, \{0, 1\}\}$ and $\tau_2 = \{\emptyset, \{1\}, \{0, 1\}\}$. Then the identity map from (X, τ_1) to (X, τ_2) is not continuous as the preimage of $\{1\}$ is $\{1\}$ and it is not an open set in τ_1. Yet, the map f that sends $1 \to 0$ and $0 \to 1$ is continuous.
- Any map from a topological space (X, τ_X) to a topological space (Y, τ_I) with indiscrete topology is continuous: indeed the only open sets in the indiscrete topology are the empty set and the whole space. Their preimages are again the empty set and the whole space. Indeed, it is always the case that $f^{-1}(Y) = X$ as every point of X by definition maps somewhere.
- Any map from a topological space with discrete topology to any topological space is continuous: this is just because all sets in the discrete topology are open.

In general for a map $f : X \to Y$, it is easier to be continuous if X has many open sets, and harder if Y has many. You should make sure to understand why this makes sense in an intuitive way.

There are three nice facts about continuous maps. First, there is a neat fact - continuous maps are sufficient to describe the topology (X, τ_X):

Exercise 1.12. *Consider $\{0,1\}$ with the topology $\tau = \{\emptyset, \{1\}, \{0,1\}\}$. For any subset $A \subseteq X$ define the function $f_A : X \to \{0,1\}$ to be equal to 1 if $x \in A$ and 0 otherwise. Show that f_A is continuous exactly when A is open.*

Second, there is an important fact - the composition of continuous maps remains continuous:

Lemma 1.25. *Suppose $(X, \tau_X), (Y, \tau_Y), (Z, \tau_Z)$ are topological spaces and $f : (X, \tau_X) \to (Y, \tau_Y)$, $g : (Y, \tau_Y) \to (Z, \tau_Z)$ are continuous maps. Then $g \circ f : (X, \tau_X) \to (Z, \tau_Z)$ is also continuous.*

Proof. Consider any open set $U \subseteq Z$. Then by continuity of g, we have that $V = g^{-1}(U)$ is open in Y. Thus, by continuity of f, $f^{-1}(V) = f^{-1} \circ g^{-1}(U) = (g \circ f)^{-1}(U)$ is open in X. Hence the lemma follows from Proposition 1.24. $\square$

And finally, an useful fact is that it suffices to check the continuity only for a basis:

Exercise 1.13. *Let (X, τ_X) and (Y, τ_Y) be two topological spaces and $f : X \to Y$ a map. Let further τ_Y^B be a basis for τ_Y and suppose that for any $U \in \tau_Y^B$, we know that $f^{-1}(U)$ is open in X. Prove that f is continuous.*

1.3 Equivalence of topological spaces - homeomorphisms

In the beginning of the course we mentioned that each class of spaces - sets, topological spaces, metric spaces, normed vector spaces - comes with a notion of equivalence, i.e. a notion of when two spaces are equal from the point of view of this class. For sets this notion was given by a bijection. For topological spaces, this notion is called a homeomorphism:

Definition 1.26 (Homeomorphism). *Let (X, τ_X), (Y, τ_Y) be two topological spaces. Then $f : X \to Y$ is called a homeomorphism if f is bijective and both f and f^{-1} are continuous.*

It is important to notice that a continuous bijective map might not yet be a homeomorphism: indeed, consider a set X with its indiscrete topology (X, τ_I), and with its discrete topology (X, τ_D). Then we know that the identity map from (X, τ_D) to (X, τ_I) is continuous, as a map to the indiscrete space is always continuous. However, as soon as X is larger than a point, the inverse which is a identity map from (X, τ_I) to (X, τ_D) is not continuous.

Definition 1.27 (Homeomorphic spaces). *Two topological spaces (X, τ_X) and (Y, τ_Y) are called homeomorphic or equivalent as topological spaces if there exists a homeomorphism $f : X \to Y$. We denote this by $(X, \tau_X) \cong (Y, \tau_Y)$.*

Let's try to put this into words: as long as we deform a topological space continuously in a way that can be continuously undone, the space actually remains (from the point of view of topology) unchanged. This is why topology is sometimes called "rubber geometry"- you can deform and mould a piece of rubber as you wish, as long as you are not pinching holes, tearing the rubber into several pieces nor gluing it together at some points, you are not changing the topology of your piece of rubber. Notice that this means much more freedom, then if you had to keep distance between different rubber points fixed!

Exercise 1.14. *Verify that being homeomorphic induces an equivalence relation on topological spaces.*

Here are some examples of homeomorphic spaces:

- We saw already that (X, τ_I) and (X, τ_D) are not homeomorphic. In particular different topologies on the same space can induce topological spaces that are inequivalent.
- On the other hand, consider $X = \{0, 1\}$ and two topologies given by $\tau_1 = \{\emptyset, \{0\}, \{0, 1\}\}$ and $\tau_2 = \{\emptyset, \{1\}, \{0, 1\}\}$. Then the identity map between (X, τ_1) and (X, τ_2) is not continuous, yet the two spaces are homeomorphic via f that swaps 0 and 1.
- The open interval $(0, 1)$ with the Euclidean topology is homeomorphic to $(0, M)$ with the Euclidean topology for any $M > 0$ using $x \to Mx$. Similarly, $(0, 1)$ is homeomorphic to $(0, \infty)$ using $x \to x^{-1} - 1$.

In fact, there are many important examples and non-examples from the realm of Euclidean spaces, some of which will be proved only much later in the course:

- The spaces $(0, 1)$ and $\mathbb{R}$ with Euclidean topology are homeomorphic;
- The spaces $(0, 1)$ and $[0, 1]$ with Euclidean topology are not homeomorphic;
- The spaces $\mathbb{R}$ and $\mathbb{R}^2$ with Euclidean topology are not homeomorphic.
- More generally, the spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ with Euclidean topology are homeomorphic iff $m = n$.

To show that two topological spaces are homeomorphic we usually have to find an explicit homeomorphism:

Exercise 1.15. *Prove that* $(0, 1)$ *and* $\mathbb{R}$ *are homeomorphic.*

Only sometimes we are lucky, and some easy property is equivalent to being homeomorphic:

Lemma 1.28. *Two topological spaces* (X, τ_D), (Y, σ_D) *with discrete topology are homeomorphic if and only if there is a bijection between the sets* X *and* Y.

Proof. One direction is clear, as by definition a homeomorphism is in particular a bijection on the underlying sets. In the other direction, let $f : X \to Y$ be a bijection. We know that any map from a space with the discrete topology is continuous. Thus, both f and f^{-1} are continuous and hence f is a homeomorphism between (X, τ_D) and (Y, σ_D). $\square$

A similar claim holds also when you consider two topological spaces with indiscrete topology, and maybe more surprisingly, when you consider the co-finite topology (Why?).

To show that two topological spaces (X, τ_X) and (Y, τ_Y) cannot be homeomorphic, it suffices to find some property that holds for (X, τ_X) and holds for any space that is homeomorphic to (X, τ_X), but doesn't hold for (Y, τ_Y). Such a property is called a topological invariant and we will see several important topological invariants throughout the course. Here are some examples:

(1) The property 'all singletons are open' is a topological invariant: if there is a homeomorphism $f : X \to Y$ between (X, τ_X) and (Y, τ_Y) then as f^{-1} is bijective, each $f^{-1}(\{y\})$ is a singleton and as f^{-1} is continuous each $f^{-1}(\{y\})$ is open iff $\{y\} \in \tau_Y$.
(2) It is easy to check that the property that each converging sequence has a unique limit is a topological invariant.
(3) The property that each sequence has a convergent subsequence is also a topological invariant (we will return to this later in the course).

1.3.1 Hausdorff spaces

Let us look at one important example of a topological invariant in more detail. When Hausdorff first introduced topological spaces in his 1914 book he included an extra condition - he axiomatically asked open sets to separate individual points. Nowadays these spaces are called Hausdorff spaces:

Definition 1.29 (Hausdorff space)**.** *A topological space (X, τ_X) is called Hausdorff if for any two distinct points x, y we can find two disjoint open sets U_x, U_y such that $x \in U_x$ and $y \in U_y$.*

This is a reasonable thing to ask, as then automatically sequences do have unique limits:

Lemma 1.30. *If (X, τ_X) is Hausdorff, then any convergent sequence has a unique limit.*

Proof. Suppose that a sequence $(x_n)_{n\geq 1}$ converges to a point x and let y be some other point. As the space is Hausdorff, then there are some disjoint open sets U_x, U_y containing respectively x and y. But if $(x_n)_{n\geq 1} \to x$, then from some n_0 onwards all $(x_n)_{n\geq n_0}$ belong to U_x and thus not to U_y. Hence by the definition of convergence,$(x_n)_{n\geq 1}$ cannot converge to y. $\square$

The discrete topology is always Hausdorff, and indiscrete topology is not Hausdorff as soon as we have a space with at least two points. A slightly more interesting example is the metric topology that is also Hausdorff: indeed, for any $x \neq y$ we can take $U_x := B(x, d(x, y)/3)$ and $U_y := B(y, d(x, y)/3)$. We saw that singletons are closed in metric spaces. This holds more generally for Hausdorff spaces:

Exercise 1.16. *Prove that in a Hausdorff space all singletons are closed.*

As promised we should check that being Hausdorff is a topological invariant:

Lemma 1.31. *Suppose that (X, τ_X) is Hausdorff and $f : (X, \tau_X) \to (Y, \tau_Y)$ a homeomorphism. Then (Y, τ_Y) is also Hausdorff.*

Proof. Pick two distinct points $y_1 = f(x_1)$ and $y_2 = f(x_2)$ in Y. Then as (X, τ_X) is Hausdorff, we can find U_1, U_2 two disjoint open sets such that $x_1 \in U_1$ and $x_2 \in U_2$. But then, as f^{-1} is continuous, $f(U_1)$ and $f(U_2)$ are both open, and as f is injective, they are disjoint. Thus $V_1 = f(U_1)$ and $V_2 = f(U_2)$ are the desired disjoint open sets separating y_1 and y_2 and hence (Y, τ_Y) is Hausdorff. $\square$

Finally, we should mention that we have also already met a more interesting non-Hausdorff topology:

Exercise 1.17. *Prove that an infinite set with the co-finite topology is not Hausdorff and deduce that cofinite topology on infinite sets cannot be induced by a metric.*

1.4 New from old: unions, subspaces and product spaces

Let us now look at the several ways to form new topological spaces from already existing ones. Indeed, there are several ways to construct new sets from old sets: we can either take unions, subsets or products of sets. In all cases, if the initial set is endowed with a topology, there is a natural way to induce a topology also on those sets. We will here treat the case subspaces and product spaces. The case of unions is easier and will be left to the exercise sheet; the case of quotient topology is harder and is left to the starred section of the exercise sheet, and to the next semester.

1.4.1 Subset topology

We can define a topology on $[0,1]$ by using the metric topology induced by $d(x,y) = |x-y|$. However, $[0,1]$ is also naturally a subspace of $\mathbb{R}$ and thus one would expect that there is a way to restrict the topology to a subset. And indeed, there is a natural way to do this:

Proposition 1.32 (Subspace topology)**.** *Let (X, τ_X) be a topological space and A a subset of X. Then define $\tau_{X,A}$ to be the collection of sets of the form $A \cap U$, where $U \in \tau_X$. Then $\tau_{X,A}$ defines a topology on A that is called the subspace topology.*

Proof. One can directly check that the three conditions for being a topology hold:

- As $\emptyset \cap A = \emptyset$ and $X \cap A = A$, we see that the empty set and the whole space are in τ_A^S.
- The intersection property: Let V_1, V_2 be two sets in $\tau_{X,A}$. Then there exist some U_1, U_2 in τ_X such that $V_1 = A \cap U_1$ and $V_2 = A \cap U_2$. Thus $V_1 \cap V_2 = (A \cap U_1) \cap (A \cap U_2) = A \cap (U_1 \cap U_2)$. But $U_1 \cap U_2 \in \tau_X$ and hence $V_1 \cap V_2 \in \tau_{X,a}$.
- The union property follows similarly: if $V_i = A \cap U_i$, then $\bigcup_{i \in I} V_i = A \cap (\bigcup_{i \in I} U_i)$.

□

It is important to check that the topology defined this way indeed coincides with what we expect:

Exercise 1.18. *Let (X, d) be a metric space, then it induces a topological space (X, τ_X) via the metric topology. Now consider $A \subseteq X$. If we restrict d to $A \times A$, we obtain a metric space (A, d) and this induces a topological space (A, τ_A). Prove that τ_A coincides with the subspace topology $\tau_{X,A}$.*

Now consider the case where $X = \mathbb{R}$ (with the standard topology) and $A = [0,1]$ and find some open subsets in A that are not open in X.

Still, doubts might remain - is this the most natural topology? It could also feel natural to define a topology $\widetilde{\tau}_{X,A}$ by

$$\widetilde{\tau}_{X,A} := \{U : U \in \tau_X, U \subseteq A\},$$

together with the set A. One can check that this also defines a topology. Or one could think of many other topologies. How do we know that the one we defined is the 'right' one?

There are two possible ways to answer this question:

(1) based on well-chosen examples we can rule out some choices straight away
(2) we can try to find a natural criteria that distinguishes one specific topology among the others.

So let us first consider an example: consider $\mathbb{R}$ with its standard topology and $\mathbb{Z} \subseteq \mathbb{R}$. Then the subspace topology induces a discrete topology on $\mathbb{Z}$, as we have that $\{n\} = \{n\} \cap B(n, 1/3)$. The topology $\widetilde{\tau}_{X,A}$ is however the indiscrete topology. Indeed, any non-empty open set of $\mathbb{R}$ with the standard topology contains at least some open interval (as open intervals from a basis and any open set can be written as a union of basis elements), and no subset of $\mathbb{Z}$ contains an interval. Thus, at least in this comparison the subspace topology wins, as clearly the standard topology of $\mathbb{R}$, restricted to $\mathbb{Z}$ should at least separate all points.

Second, there is an abstract way to pick out $\tau_{X,A}$ as the natural candidate among all choices:

Lemma 1.33. *Let (X, τ_X) be a topological space and $(A, \tau_{X,A})$ a subspace with the subspace topology. Then $\tau_{X,A}$ is the smallest topology $\widetilde{\tau}$ for which the inclusion map $i : (A, \widetilde{\tau}) \to (X, \tau_X)$ defined on A by identity is continuous.*

Proof. First, let us check that the inclusion map is continuous for $(A, \tau_{X,A})$. Indeed, as i is identity on A, then for any open set $U \in \tau_X$, we have that $i^{-1}(U) = A \cap U$. Thus i is continuous.

Now, if $i : (A, \widetilde{\tau}) \to (X, \tau_X)$ is continuous, then for any open set $U \in \tau_X$, we have that $i^{-1}(U) = A \cap U$ is open in $\widetilde{\tau}$. But this means that $\tau_{X,A} \subseteq \widetilde{\tau}$ giving the claim. $\square$

This implies the following characterisation of the subspace topology:

Exercise 1.19 ($\star$)**.** *Consider a topological space (X, τ_X) and let $A \subseteq X$. Then the subspace topology $\tau_{X,A}$ is the only topology on A with the following property: for any topological space (Y, τ_Y), and any map $g : (Y, \tau_Y) \to (A, \tau_{X,A})$, the map g is continuous if and only if $i \circ g : (Y, \tau_Y) \to (X, \tau_X)$ is continuous, where i is the inclusion map as before.*

Moreover, this definition of a topology on a subspace satisfies several other natural properties:

Exercise 1.20. *Let (X, τ_X) be a topological space and $(A, \tau_{X,A})$ a subspace with the subspace topology.*

- *Prove that if (Y, τ_Y) is another topological space and $f : (X, \tau_X) \to (Y, \tau_Y)$ is continuous, then also f restricted to A is a continuous map from $(A, \tau_{X,A}) \to (Y, \tau_Y)$.*
- *In particular, prove that if $f : (X, \tau_X) \to (Y, \tau_Y)$ is a homeomorphism and $f(A) = B$ for some $B \subseteq Y$, then the restriction of f to A induces a homeomorphism between A and B with their respective subspace topologies.*

Finally, let us check that the Hausdorff property descends nicely to subspaces:

Lemma 1.34. *Let (X, τ_X) be a Hausdorff topological space. Then $(A, \tau_{X,A})$ is also Hausdorff.*

Proof. Consider some $x_1 \neq x_2 \in A$. Then as X is Hausdorff, we can find disjoint open sets $U_1, U_2 \in \tau_X$ such that $x_1 \in U_1$ and $x_2 \in U_2$. Now define $V_1 = U_1 \cap A$, and define $V_2 = U_2 \cap A$. Then $V_1, V_2 \in \tau_{X,A}$, they are disjoint and contain x_1, x_2 respectively. Thus $(A, \tau_{X,A})$ is Hausdorff. $\square$

1.4.2 Product topology on finite products

Recall that the Cartesian product $X \times Y$ of two sets is given by the set of ordered pairs: $X \times Y = \{(x, y) : x \in X, y \in Y\}$. Now, if both X and Y are endowed with a topology, it is natural to ask what would be a natural topology on $X \times Y$.

What should be the open sets for $X \times Y$? Recall, that in metric topologies, being open was characterised by the following: around each point x in this open set, you can find a small ball that is contained in this set, i.e. you have some freedom to move around. Now consider a set $U \times V$ with U open for X, and V open for Y. Then if $(x, y) \in U \times V$, there is "space" to move in both the first and the second coordinate - so it makes sense to declare $U \times V$ to be open for a natural topology on $X \times Y$:

Proposition 1.35 (Product topology on $X \times Y$)**.** *Consider two topological spaces (X, τ_X) and (Y, τ_Y). Define $\tau^B_{X\times Y}$ to be the collection of the subsets of $X \times Y$ of the form $U \times V$, where U is open in X and V is open in Y. Then $\tau^B_{X\times Y}$ is a basis for a topology, and the topology $\tau_{X\times Y}$ induced by it is called the product topology on $X \times Y$.*

Proof. We need to verify that the two conditions for being a basis for a topology are met. Then we are done as any basis induces a topology by taking unions of its elements.

First, by taking $U = X$ and $V = Y$, we see that the collection of subsets in $\tau^B_{X\times Y}$ covers $X \times Y$.

Second, let $W_1, W_2 \in \tau^B_{X\times Y}$. Then we can find U_1, U_2, be open in (X, τ_X) and V_1, V_2 open in (Y, τ_Y) such that $W_1 = U_1 \times V_1$, $W_2 = U_2 \times V_2$. We have that

$$W_1 \cap W_2 = (U_1 \times V_1) \cap (U_2 \times V_2) = (U_1 \cap U_2) \times (V_1 \cap V_2).$$

But $(U_1 \cap U_2) \times (V_1 \cap V_2)$ is also by definition in $\tau^B_{X\times Y}$. Thus the second condition for being a basis is also satisfied and thus $\tau^B_{X\times Y}$ is a basis for a topology. $\square$

In fact the proof can be strengthened to show that one can consider a smaller basis to define the product topology :

Exercise 1.21. *Let (X, τ_X), (Y, τ_Y) be two topological spaces. Suppose τ^B_X is a basis for (X, τ_X) and τ^B_Y is a basis for (Y, τ_Y). Show that the set*

$$\widetilde{\tau}^B_{X\times Y} := \{\widehat{U} \times \widehat{V} : \widehat{U} \in \tau^B_X, \widehat{V} \in \tau^B_Y\}$$

is a basis for the product topology.

Exercise 1.22. *Consider $(\mathbb{R}, \tau_E)$, i.e. $\mathbb{R}$ with Euclidean topology and the product $\mathbb{R} \times \mathbb{R}$ with the product topology. Show that the resulting space is homeomorphic to $\mathbb{R}^2$ with its Euclidean topology.*

It is important to remember that not all open sets of the product topology are given by products of open sets: for example in $\mathbb{R}^2$ with its standard topology, the set $(0, 1)^2 \cup (1, 2)^2$ is open, but not given by the product. The same also holds for the ball $B(1, 1)$, for example: indeed if the latter was given by $U \times V$, then $(0, 1) \subseteq U$ and $(0, 1) \subseteq V$, and thus $(0, 1)^2 \subseteq B(1, 1)$ which we know is not true.

Again, although the definition of the product topology felt again very reasonable, there could be possibly other natural ways to define a topology on the product. So how can we mathematically pin down why this is the natural choice? Similarly to the case of the subspace topology, the following lemma gives one such way:

Lemma 1.36. *Let (X, τ_X) and (Y, τ_Y) be topological spaces and consider $X \times Y$ with the product topology. Then the product topology $\tau_{X\times Y}$ is the smallest topology $\widetilde{\tau}$ on $X \times Y$ such that the projection maps $p_X : (X \times Y, \widetilde{\tau}) \to (X, \tau_X)$ given by $p_X(x, y) := x$ and $p_Y : (X \times Y, \widetilde{\tau}) \to (Y, \tau_Y)$, given by $p_Y(x, y) := y$ are both continuous.*

Proof. Again, let us first check that p_X, p_Y are continuous for the product topology. For any open set U of (X, τ_X) we have that $p_X^{-1}(U) = U \times Y$, and this belongs to $\tau^B_{X\times Y}$. Similarly, for any open set V of (Y, τ_Y) we have that $p_Y^{-1}(V) = X \times V \in \tau^B_{X\times Y}$, and thus the continuity follows.

Now, suppose $p_X : (X \times Y, \widetilde{\tau}) \to X$ and $p_Y : (X \times Y, \widetilde{\tau}) \to Y$ are continuous. Then, by above all sets of the form $U \times Y$ with $U \in \tau_X$ and $X \times V$ with $V \in \tau_Y$ have to belong to $\widetilde{\tau}$. But then also $(U \times Y) \cap (X \times Y) \in \widetilde{\tau}$ and thus in particular $\widetilde{\tau}$ contains the basis $\tau^B_{X\times V}$. But then, as $\widetilde{\tau}$ is a topology, it has to contain the topology induced by this basis, i.e. $\tau_{X\times Y}$, giving the claim. ☐

Let us now quickly consider the interplay between the product topology and continuous functions. Firstly, a function f from some space Z to the product space $X \times Y$, written $f(z) = (f_1(z), f_2(z))$ is continuous if and only if both f_1 and f_2 are continuous:

Exercise 1.23. *Let (X, τ_X) and (Y, τ_Y) be topological spaces and consider $X \times Y$ with the product topology. Let further (Z, τ_Z) be another topological space and $f : (Z, \tau_Z) \to (X \times Y, \tau_{X\times Y})$. Prove that f is continuous if and only if both $f_1 := p_X \circ f : (Z, \tau_Z) \to (X, \tau_X)$ and $f_2 : p_Y \circ f : (Z, \tau_Z) \to (Y, \tau_Y)$ are continuous.*

Now, suppose that instead of a map $f : Z \to X \times Y$, you consider a map $f : X \times Y \to Z$. Moreover, suppose that it is continuous separately in both coordinates, i.e. $f(x, \cdot) : (Y, \tau_Y) \to (Z, \tau_Z)$ is continuous for each $x \in X$ and $f(\cdot, y) : (X, \tau_X) \to (Z, \tau_Z)$ is continuous for each $y \in Y$. One might hope that this also means that $f : (X \times Y, \tau_{X\times Y}) \to (Z, \tau_Z)$ would be continuous. However, as in real analysis, this is not the case. In fact the very same counter-example will work here too: we can for example take $f(x, y) : \mathbb{R}^2 \to \mathbb{R}$ to be equal to $\frac{2xy}{x^2+y^2}$ if one of the points is non-zero, and equal to 0 at $(0, 0)$. [7]

Let now (X_i, τ_{X_i}) with $i = 1 \dots n$ be topological spaces. How should we define the product topology on $X_1 \times \cdots \times X_n$? One way would be to go about inductively: having defined a topology $\widetilde{\tau}_{n-1}$ on $X_1 \times \cdots \times X_{n-1}$, we can define a topology $\widetilde{\tau}_n$ on $X_1 \times \cdots \times X_n$ by taking the product topology for $(X_1 \times \cdots \times X_{n-1}, \widetilde{\tau}_{n-1})$ and (X_n, τ_{X_n}). But in doing so, we are making some arbitrary choices here - for example, why does the resulting topology not depend on the fact that we went step by step from X_1 to X_n, and not vice versa?

This question is nicely answered by finding a canonical definition for the product topology for arbitrary finite products:

Proposition 1.37 (The product topology on a finite product space)**.** *Let us consider $(X_1, \tau_{X_1}), \dots, (X_n, \tau_{X_n})$, a finite collection of topological spaces. Let*

$$\tau^B_{X_1\times\cdots\times X_n} := \{U_1 \times \cdots \times U_n : U_i \in \tau_{X_i} \forall i = 1 \dots n\}.$$

Then $\tau^B_{X_1\times\cdots\times X_n}$ is a basis for a topology, and this topology is called the product topology on $X_1 \times \cdots \times X_n$.

Proof. The proof is exactly the same as for two spaces. ☐

[7] Geometrically this can be just thought of as follows: consider the vector (x, y) and then reflect it w.r.t the line $x = y$ to get the vector (y, x). Then f is just equal to the $\cos(\theta)$ where θ is the angle between the vectors (x, y) and (y, x). Clearly this angle changes continuously when you fix a non-zero x or a non-zero y. When one of them is equal to 0, then the angle is 90 degrees and thus the function is constantly 0. However when $x = y$, then the angle is 0 degrees independently of the size of $x = y$, and thus the function is not continuous at 0.

It now just remains to verify that the inductive definition induces the same basis for its topology:

Exercise 1.24. *Show that for all $n \geq 2$, the sets of the form $U_1 \times \cdots \times U_n$ with $U_i \in \tau_{X_i}$ for all $i = 1 \ldots n$ form a basis for $\widetilde{\tau}_n$ as defined above, and thus that $\widetilde{\tau}_n$ is the product topology on $X_1 \times \cdots \times X_n$.*

1.4.3 Product topology on infinite products

With an infinite number of spaces one has to be more careful - it comes out that there are several natural ways to define a topology!

Let us however, start by reminding how to think about infinite products. Let I be some infinite index set and $(X_i)_{i \in I}$ a collection of sets. If I is countably infinite, then it is in bijection with $\mathbb{N}$ and we can think of the product $\Pi_{i \in I} X_i$ as of sequences $(x_1, \ldots, x_n, \ldots)$. However, a better way, which generalizes to any index set is to think of $\Pi_{i \in I} X_i$ as of functions $x : I \to \cup X_i$ such that $x(i) \in X_i$. One way to picture this is to think of I as the horizontal axis and $\cup X_i$ as the vertical axis - for each $i \in I$, we then mark an element in $\cup X_i$, like we usually do when drawing a graph.

Remark 1.38. *Somewhat surprisingly, it does not follow from the usual axioms of the set theory, the so-called Zermelo-Fraenkel (ZF) axioms, that an arbitrary product of non-empty spaces is non-empty. In fact, such a statement - that any product of non-empty sets is non-empty is (equivalent to) the axiom of choice, stated just below. This axiom is independent of ZF theory, in the sense that either this axiom or its negation can be added to the ZF axioms to get a consistent mathematical theory.*

We will come back to this axiom later in the course. For now we just mention that in concrete settings there is sometimes no need for the Axiom of choice. For example, in the case when we have the product of one identical space X, then we know that the product space is non-empty, as we can always pick some $x \in X$ and then the element defined by $x_i = x$ for all $i \in I$ is certainly in $\Pi_{i \in I} X_i$. The axiom of choice becomes important when the spaces X_i are really arbitrary and we have no additional information about them.

- **Axiom of choice**: if $(X_i)_{i \in I}$ is any collection of non-empty sets, then also $\Pi_{i \in I} X_i$ is non-empty.

Now, having determined that $\Pi_{i \in I} X_i$ is non-empty, if all spaces are non-empty, let us put a topology on it. It comes out that the right way to put a topology on the infinite product is as follows:

Proposition 1.39 (The product topology on an infinite product space)**.** *Let now I be some infinite index set and $((X_i, \tau_{X_i}))_{i \in I}$ a collection of topological spaces. Let $\tau^B_{\Pi_{i \in I} X_i}$ be the collection of subsets of $\Pi_{i \in I} X_i$ of the form $\Pi_{i \in I} U_i$, where each $U_i \subseteq X_i$ is open in X_i and $U_i \neq X_i$ only for finitely many $i \in I$. Then τ^B is a basis for a topology, and this topology is called the product topology on $\Pi_{i \in I} X_i$.*

The proof is really the same as in the case of two spaces and thus omitted.

Remark 1.40 (Box topology)**.** *In fact, the choice above might not have been your first guess. Maybe you would have preferred to define the basis using the collection of subsets of the form $\Pi_{i \in I} U_i$, where each $U_i \subseteq X_i$ is open in X_i? This indeed defines a topology on*

$\Pi_{i\in I}X_i$, however, it comes out that this topology is too large to be of much use. If interested, see the exercise in the starred section.

One way to motivate the choice of the product topology is again to look for a natural property. And indeed, if we come back to the projection maps, we obtain the following:

Lemma 1.41. *Let now I be some infinite index set and $((X_i, \tau_{X_i}))_{i\in I}$ a collection of topological spaces. Then the product topology is the smallest topology on $\Pi_{i\in I}X_i$ such that all co-ordinate maps are continuous.*

Again, we will not repeat the proof here, as it is the same as for two spaces. However, if you do revisit the proof try to notice where the fact that only finitely many U_i differ from X_i in the basis elements enters in the proof.

Maybe another good reason for the product topology is the following description of convergence in product spaces with the product topology:

Lemma 1.42. *Let now I be some index set and $((X_i, \tau_{X_i}))_{i\in I}$ a collection of topological spaces. Then a sequence $(x_n)_{n\geq 1}$ converges to x in $\Pi_{i\in I}X_i$ with the product topology if and only if it converges pointwise, i.e. iff for all $i \in I$, $(x_n(i))_{n\geq 1}$ converges to $x(i)$ in (X_i, τ_{X_i}).*

Proof. Let us first assume that $(x_n)_{n\geq 1}$ converges to x in $\Pi_{i\in I}X_i$ with the product topology. Fix some $i_0 \in I$ and some open set $U_{i_0} \in \tau_{X_{i_0}}$ containing $x(i)$. Consider the open set $W_{i_0} := \Pi_{i\in I}V_i$, where $V_{i_0} = U_{i_0}$ and $V_i = X_i$ otherwise. Then by the definition of convergence, there exists some $n_{i_0} \in \mathbb{N}$ such that for all $n \geq n_{i_0}$, we have that $x_n \in W_{i_0}$. In particular $x_n(i_0) \in U_{i_0}$ and thus $(x_n(i_0))_{n\geq 1} \to x(i_0)$ in $(X_{i_0}, \tau_{X_{i_0}})$.

Conversely, suppose that for all $I \in I$, we have that $(x_n(i))_{n\geq 1} \to x(i)$ in (X_i, τ_{X_i}). As any open set is a union of the basis elements, and any $x \in \Pi_{i\in I}X_i$ is contained in some basis element, it suffices to show that for any open set U in the basis $\tau^B_{\Pi_{i\in I}X_i}$ there is some $n_U \in \mathbb{N}$ such that for all $n \geq n_U$, we have that $x_n \in U$ [8].

Now, any $U \in \tau^B_{\Pi_{i\in I}X_i}$ is of the form $\Pi_{i\in I}U_i$, where each $U_i \subseteq X_i$ is open in X_i and $U_i \neq X_i$ for only finitely many $i \in I$. Denote this finite set by I_0. As $(x_n(i_0))_{n\geq 1} \to x(i_0)$ for all $i_0 \in I$, there exists for any $i_0 \in I$ some $n_{U_{i_0}} \in \mathbb{N}$ such that for all $n \geq n_{U_{i_0}}$ we have that $x_n(i_0) \in U_{i_0}$. But then if we take $n_U = \max_{i_0 \in I_0} n_{U_{i_0}}$, we have that $x_n \in U$ as soon as $n \geq n_U$. $\square$

An interesting case is the space $\{0, 1\}^{\mathbb{N}}$ with the product topology, when each $\{0, 1\}$ is given the discrete topology. Can you find some other set that it is homeomorphic to?

Exercise 1.25 ($\star$). *Consider a infinite dyadic tree, i.e. an infinite connected graph, where exactly one vertex v has degree 2, every other vertex has degree 3 and which contains no cycles. Consider the set X of all branches starting from v. Any two branches b_i and b_j start from v, then stay together for a bit, and then separate once and for all. Let $n_{i,j}$ be the number of common vertices for two branches and define a distance $d(b_i, b_j) := n_{i,j}^{-1}$. Check that d defines a metric on X, and thus also induces a topology τ_X. Find a bijection between X and $\{0, 1\}^{\mathbb{N}}$. Moreover, prove that (X, τ_X) is in fact homeomorphic to $\{0, 1\}^{\mathbb{N}}$, where we consider $\{0, 1\}$ with the discrete topology and then take the product topology.*

[8] In fact, we proved a nice general lemma here: to check convergence to some $x \in X$, it suffices to check that for all basis elements U containing x, the sequence is eventually contained in U

Exercise 1.26 ($\star$, McMullen p. 20-21). *Recall the standard Cantor set obtained as follows: we start from the unit interval, and then remove the middle third* $(1/3, 2/3)$. *In the next step we remove the middle third of both* $[0, 1/3]$ *and* $[2/3, 1]$. *We continue infinitely. The resulting space endowed with the subspace topology induced from* $\mathbb{R}$ *is called the Cantor space. Prove that the Cantor space is homeomorphic to the space* (X, τ_X) *of the previous exercise, and thus also to* $\{0,1\}^{\mathbb{N}}$ *with the product topology as above.*

1.4.4 Disjoint unions

Maybe the simplest is the case is actually that of disjoint union. It's a good way to test your understanding of the chapter:

Exercise 1.27 (Disjoint unions). *Let* (X, τ_X) *and* (Y, τ_Y) *be topological spaces such that the sets* X, Y *are disjoint.*

- *Define a reasonable topology on* $X \cup Y$.
- *Explain why your choice is a reasonable topology via examples: e.g. show that the disjoint union of two discrete spaces is still discrete.*
- *Is the Euclidean topology on* $[0, 2]$ *induced by taking the disjoint union of* $([0, 1], \tau_E)$ *and* $((1, 2], \tau_E)$*?*
- *Find a way to single out your topology: e.g. by showing that it is the finest topology making both inclusion maps* $X \to X \cup Y$ *with* $x \to x$ *and* $Y \to X \cup Y$ *with* $y \to y$ *continuous.*

You might wonder what to do in case the sets are not disjoint. Can you still take some union?

Recall that actually sets are only defined up to bijection, and topological spaces up to homeomorphism. So if you are given any sets (or top. spaces), you can always just take some representatives of their equivalence class so that the underlying sets are disjoint.

For example, if you have a collection of sets $(X_i)_{i \in I}$, then the sets $\hat{X}_i := \{(x, i) : x \in X_i\}$ are disjoint so that $X_i \cong \hat{X}_i$. If in addition each space comes with a topology τ_{X_i}, you can further use the bijections $f_i : X_i \to \hat{X}_i$ to induce a topology $\tau_{\hat{X}_i}$ on each $\hat{X}_i$ so that in fact $(X_i, \tau_{X_i}) \cong (\hat{X}_i, \tau_{\hat{X}_i})$: you just say that $U \in \tau_{\hat{X}_i}$ iff $f^{-1}(U) \in \tau_{X_i}$.

Section 2

Connectedness

We will now discuss connectedness of a topological space - do we have single piece of rubber or does it consist of several disjoint pieces? Pulling or cutting a piece of rubber into several pieces feels like highly discontinuous - some points that were close together, will now be far apart. Thus it feels intuitive that being connected is something that is preserved by continuous transformations.

Again the first question is how to even define connectedness in the formalism of topology. Let us start with some examples that could guide our definitions:

- It seems reasonable to say that $\mathbb{R}^n$ and $[0,1]$ are connected, as we can always find a continuous path joining two points, and to say that $[-2,-1]\cup[1,2]$ is not connected.
- But do you think for example $\mathbb{Q}$ with the induced topology is a connected space or not, or $\mathbb{R}\backslash\mathbb{Q}$? Intuitively, we would still probably agree that there are gaps between the points and thus the spaces should be disconnected.
- And what about the following example in $\mathbb{R}^2$ (called topologist's sine curve): consider the graph of the curve $y = \sin(1/x)$ on $x \in (0,1]$, together with the line segment $y \in [-1,1]$ on the $y-$axis. Is this plane figure connected or not? (See Proposition 2.18.)

In fact, it comes out that there are (at least) two reasonable ways to define connectedness, following two different intuitive ideas:

(1) We could say that a space is connected if for all points x and y we can find a continuous path going from one point to the other.
(2) Alternatively, we could try to formalize the existence of a "gap" and say that a space is disconnected, if we can decompose it into two pieces A and B, so that for all points in A, sufficiently small neighbourhoods around them are disjoint from B (i.e. they cannot cross the gap);

It comes out that indeed both of these notions can be formalized, and whereas they agree on the first two examples we gave, they disagree on the third one - on the topologist sine curve.

2.0.1 Connectedness in topological spaces

Let us start from the definition formalizing the notion of "gaps":

Definition 2.1 (Connectedness). *A topological space (X, τ_X) is called connected if for any open partition of X, i.e. for any writing of $X = U \cup V$ with $U, V \in \tau_X$ and $U \cap V = \emptyset$, either U or V has to be the empty set (or equivalently the whole space). A subset of X is called connected if it is connected in the subspace topology.*

Notice that this corresponds at least to some extent to our intuition about gaps: if a space can be decomposed into two disjoint open sets U and V, then both of these sets are also closed as complements of open sets. In particular their boundaries are empty. Thus there are no points so that all their open neighbourhoods intersect both U and V, at least hinting that there is a separation between the sets - an element is either really in the interior of U or in the interior of V and there is emptiness in-between.

Based on this, an equivalent way of stating connectedness is apparent:

- A topological space is connected if the only subsets of X that are both open and closed are $\emptyset$ and X.

Indeed, for any set $U \subseteq X$ we have a partition $X = U \cup (X \backslash U)$. Now if U is both open an closed, this gives an open partition and thus X being connected means that any such U is either empty or the whole space.

Let us also consider some examples:

- firstly, one-point space is clearly connected by this definition;
- secondly, any space with the indiscrete topology is connected, as the only open sets are the whole set and the empty set;
- finally, the discrete topology on more than one point is not connected, as for any point x in the space both $\{x\}$ and $X \backslash \{x\}$ are open and non-empty.

Exercise 2.1. *Consider a set X with the co-finite topology. When are the resulting topological spaces connected, when are they disconnected?*

Showing from the definitions that any interval of the real line with its standard topology is connected, is already a bit trickier:

Proposition 2.2. *The interval $[a, b]$ with its standard topology is connected.*

Proof. Consider any decomposition $[a, b] = U \cup V$ of the interval into two disjoint open sets. WLOG suppose that $a \in U$. Consider now $s := \sup\{z \in [a, b] : [a, z] \subseteq U\}$. To show that $[a, b]$ is connected, we need to prove that $s = b$.

First, we check that $s > a$: as U is open and $a \in U$, we have that for some $[a, a + \delta) \subseteq U$, hence $[a, a + \delta/2] \subseteq U$ and thus $s > a$.

Now we check that $s \in U$: indeed, suppose by contradiction that $s \in V$. Then by assumption V is open and hence if $s \in V$, then also $(s - \delta', s + \delta') \subseteq V$ for some $\delta'_s > 0$. As V and U are disjoint, this means that $s - \delta'/2 \notin U$, contradicting the definition of s.

Finally, let us argue that $s = b$. Suppose for contradiction that $s < b$. As U is open, there is again some $\delta > 0$ such that $(-\delta + s, s + \delta) \subseteq U$ and thus $s + \delta/2 \in U$ giving again a contradiction with the definition of s.

Hence $s = b$ and the proposition follows. □

Remark 2.3. *Basically the proof would also give connectedness for open or half-open intervals, also all types of half-line and $\mathbb{R}$ itself.*

In fact, in $\mathbb{R}$ one can characterise all possible connected subsets:

Proposition 2.4. *Let $A \subseteq \mathbb{R}$ be any subset of $\mathbb{R}$. Then A is connected if and only if A is an interval, a half-line or $\mathbb{R}$ itself.*

Proof. The proof is on the exercise sheet. □

2.0.2 Path-connectedness in topological spaces

The idea of connectedness as any two points being connected by a path is called path-connectedness. Let us first define

Definition 2.5 (Path)**.** *A continuous map $\gamma : ([0, 1], \tau_E) \to (X, \tau_X)$ is called a path. If $\gamma(0) = x$ and $\gamma(1) = y$, then γ is called a path from x to y (sometimes denoted $x \to y$).*

Definition 2.6 (Path-connectedness). *A topological space (X, τ_X) is called path-connected if for any two points $x, y \in X$, there is a path γ from x to y.*

Again let us start from some examples:

- Any set X with the indiscrete topology is path-connected, as any function to a indiscrete space, in particular any function from $([0,1], \tau_E)$ to (X, τ_I) is continuous and thus given $x, y \in X$ we can for example just define $\gamma(t) = x$ for $t \in [0,1)$ and $\gamma(t) = y$ for $t = 1$.
- Any interval I in $(\mathbb{R}, \tau_E)$ is path-connected as for any $x < y \in I$ we can map $[0,1]$ to $[x, y]$ continuously by scaling and translation.

One can concatenate paths - if we have a continuous path from $x \to y$ and from $y \to z$, then we can concatenate the paths to get a path from $x \to z$.

Lemma 2.7 (Concatenation of paths). *Let (X, τ_X) be a topological space. Suppose that $x, y, z \in X$ and γ_1 is a path from $x \to y$ and γ_2 is a path from $y \to z$. Define $\tilde{\gamma} : [0,1] \to X$ by $\tilde{\gamma}(t) = \gamma_1(2t)$ for $t \in [0, 1/2]$ and $\tilde{\gamma}(t) = \gamma_2(2t-1)$ for $t \in [1/2, 1]$. Then $\tilde{\gamma}$ is a path from $x \to z$.*

This Lemma follows directly from an important general lemma, sometimes called pasting or gluing lemma:

Lemma 2.8 (Pasting / gluing lemma). *Let (X, τ_X), (Y, τ_Y) be topological spaces and A, B closed subsets of X such that $X = A \cup B$. Suppose that f is a map from (X, τ_X) to (Y, τ_Y) such that the restriction of f to A is continuous as a map from $(A, \tau_{X,A})$ to (Y, τ_Y), and the restriction of f to B is continuous as a map from $(B, \tau_{X,B})$ to (Y, τ_Y). Then f is continuous as a map from (X, τ_X) to (Y, τ_Y).*

Proof. The proof is on the exercise sheet. □

2.0.3 Connectedness vs path-connectedness in topological spaces

Path-connectedness is a stronger notion than connectedness:

Theorem 2.9. *Any topological space that is path-connected is also connected.*

We start from a useful lemma, that helps to encode connectedness using continuous functions:

Lemma 2.10. *A topological space (X, τ_X) is connected if and only if any continuous map from (X, τ_X) to $\{0, 1\}$ with discrete topology is constant.*

Proof. First suppose there exists some continuous non-constant function $f : X \to \{0, 1\}$. Then by continuity both $f^{-1}(\{0\})$ and $f^{-1}(\{1\})$ are open, disjoint and by assumption both are different from the empty set. Moreover, also $X = f^{-1}(\{0\}) \cup f^{-1}(\{1\})$ and thus X is not connected.

In the other direction, suppose that X is not connected. Then there exists disjoint open $U, V \subseteq X$ that are both non-empty an open such that $X = U \cup V$. Now set $f = 0$ on U and $f = 1$ on V. Then by construction f is both non-constant and continuous. □

The proof of the Theorem 2.9 now follows:

Proof of Theorem 2.9. Let (X, τ_X) be a path-connected topological space. We aim to use Lemma 2.10 as a criterion.

Consider any continuous function $f : X \to \{0, 1\}$, where the two-point space is taken with the discrete topology. We want to show that f is constant. Fix some point $x \in X$ and let $y \in X$ be any other point. Then as (X, τ_X) is path-connected, there exists a continuous path $\gamma : [0, 1] \to X$ with $\gamma(0) = x$ and $\gamma(1) = y$. Now, by the composition rule for continuous functions (Lemma 1.25) we have that $f \circ \gamma : ([0, 1], \tau_E) \to (\{0, 1\}, \tau_D)$ is also continuous. But by Proposition 2.2 the interval $[0, 1]$ is connected in the Euclidean topology. Thus we conclude from Lemma 2.10 that $f \circ \gamma$ has to be constant and in particular $f(x) = f(y)$. But this holds for any $y \neq x$, implying that f is constant. But f was an arbitrary continuous function, and hence Lemma 2.10 implies that (X, τ_X) is connected. $\square$

On the other hand, the connectedness is a strictly weaker notion - there are spaces that are connected, yet are not path-connected. The most famous example is the topologist's sine curve. This planar subset is defined as the graph of the curve $y = \sin(1/x)$ on $x \in (0, 1]$, together with the line segment $y \in [-1, 1]$ on the $y-$axis, i.e. as the union $\{(x, \sin(\frac{1}{x}) : x \in (0, 1]\} \cup \{(0, y) : y \in [-1, 1]\}$.

Proposition 2.11. *The topologist sine-curve is connected, but not path-connected as a subset of $\mathbb{R}^2$.*

We will postpone the proof for now, but will come back to it after having developed some convenient methods for working with connectedness. There is another example of a connected but path-connected space on the example sheet. Also, even if connectedness and path-connectedness are not equivalent in full generality, it comes out that in certain special settings the two notions are equivalent. Notice the word "open" in the following statement:

Theorem 2.12. *Consider $\mathbb{R}^n$ with the Euclidean topology. Then an open set is connected if and only if it is path-connected.*

As is often the case in basic topology, theorems about "concrete" spaces are a bit trickier to prove than the very abstract general results - some special properties of the concrete setting have to enter, but on the other hand you have to be able to use these properties in the general formalism.

Proof. We already showed that path-connected implies connected, so it remains to prove the inverse. The key observation, resulting from special properties of $\mathbb{R}^n$ is very simple:

Claim 2.13. *The open balls $B(x, \delta)$ for any $x \in \mathbb{R}^n$ and any $\delta > 0$ are path-connected.*

Proof. This can be directly verified by writing down a straight-line path between any pair of points. $\square$

Now suppose U is some connected open set in $\mathbb{R}^n$ and fix some point $x \in U$. Define further U_x as the set of points $y \in U$ that are path-connected to x. Notice that U_x is non-empty as $x \in U_x$.

We claim that U_x is open: indeed if $y \in U_x$, there is some path γ in U from $x \to y$. Moreover, as U is open and the open balls generate the Euclidean topology, there is some $\delta_y > 0$ such that $B(y, \delta_y) \subseteq U$. But by the claim above, each ball is path-connected, thus for any $z \in B(y, \delta_y)$ there exists some path $\gamma_z : y \to z$. Using the concatenation lemma for the paths γ and γ_z we obtain a path $\tilde{\gamma}$ from $x \to z$. Thus $z \in U_x$, hence $B(y, \delta_y) \subseteq U_x$ and thus U_x is open.

On the other hand, define V_x to be the set of points $z \in U$ that are not path-connected to x. Then $V_x = U \backslash U_x$. We claim that V_x is also open. Indeed, if it is empty, then it is open by definition. Otherwise, consider any $z \in V_x$. As U is open, there exists again $\delta_z > 0$, such that $B(z, \delta_z) \subseteq U$. Consider some $w \in B(z, \delta_z)$. Then by the claim above, there is some path $\gamma_w : w \to z$. But now if $w \in U_x$, then there is also a path γ from $x \to w$, and again we could concatenate γ and γ_w to get a path from $x \to z$, contradicting $z \in V_x$. Thus $w \in V_x$, hence $B(z, \delta_z) \subseteq V_x$ and we conclude that V_x is also open.

In conclusion, we can write $U = U_x \cup V_x$ with both U_x, V_x open and disjoint. Thus, as U is connected, one of U_x and V_x has to be empty. But we know that U_x is non-empty and hence V_x is empty, giving the claim. □

Again, let us stress the importance of considering open subsets: the topologist sine-curve is also a subset of $\mathbb{R}^2$ for which this theorem does not hold.

2.1 More on connectedness

Let us now study how connectedness behaves under operations on topological spaces

2.1.1 Connectedness and continuity

Firstly, connectedness behaves nicely under continuous maps.

Proposition 2.14. *Let (X, τ_X) be a connected topological space and f a continuous map to some other topological space (Y, τ_Y). Then $f(X)$ is connected in the subspace topology.*

Similarly, to the proof of Theorem 2.9, we aim to use the criterion of Lemma 2.10:

Proof. Consider $f(X) \subseteq Y$ with the subspace topology inherited from (Y, τ_Y), and consider a continuous map g from $(f(X), \tau_{Y,f(X)})$ to $\{0,1\}$ with the discrete topology. Now look at the map $g \circ f : X \to \{0,1\}$. Then as the composition of continuous maps is continuous by Lemma 1.25, we deduce that $g \circ f$ is continuous as a map from (X, τ_X) to $(\{0,1\}, \tau_D)$. Thus from connectivity of X and Lemma 2.10 it follows that $g \circ f$ is constant. But this implies that g is constant on $f(X)$ and thus $f(X)$ is connected, by applying again Lemma 2.10. □

In particular this implies that connectedness is a topological invariant:

Corollary 2.15. *Suppose that (X, τ_X) is connected and homeomorphic to (Y, τ_Y). Then (Y, τ_Y) is also connected.*

This allows to argue easily that S^1 is not homeomorphic to $[0,1]$. Indeed, by Exercise 3 on sheet 4 we know that if f is a homeomorphism from $[0,1]$ to S^1 (say $S^1 \subseteq \mathbb{R}^2$, $S^1 = \{(x,y) : x^2+y^2 = 1\}$), then f restricted to $[0,1]\backslash\{1/2\}$ is also a homeomorphism between $[0,1]\backslash\{1/2\}$ and $S^1\backslash\{f(1/2)\}$. Now the former is not connected, as all the connected subspaces of $\mathbb{R}$ are intervals, half-lines or $\mathbb{R}$. On the other hand, the latter is connected: by rotating the sphere we can assume $f(1/2) = (1,0)$, then $g(x) = (\cos(2\pi x), \sin(2\pi x))$ is a homeomorphism from $(0,1)$ to $S^1\backslash\{(1,0)\}$, and as $(0,1)$ is connected, so is $S^1\backslash\{(1,0)\}$. A similar question on the exercise sheet shows that $\mathbb{R}$ and $\mathbb{R}^n$ cannot be homeomorphic for $n \geq 2$.

Another direct consequence of Proposition 2.14 is a generalization of the Intermediate value theorem:

Theorem 2.16 (Intermediate value theorem)**.** *Let (X, τ_X) be a connected topological space and $f : X \to \mathbb{R}$ a continuous map. Suppose that for some $x < y$, we have that $x \in f(X)$ and $y \in f(X)$. Then in fact $[x, y] \subseteq f(X)$.*

Proof. By Proposition 2.14 we know that $f(X)$ is connected. But we proved before that all connected subsets A of $\mathbb{R}$ are either intervals, half-intervals or the whole of $\mathbb{R}$ and for each of them it holds that if $x \in A$ and $y \in A$ with $x < y$, then also $[x, y] \subset A$. □

On the exercise sheet you will see that path-connectedness is also a topological property.

Exercise 2.2. *Prove that being path-connected is also a topological invariant. Start by proving that if (X, τ_X) is path-connected and f a continuous map to some other topological space Y, then $f(X)$ is path-connected.*

2.1.2 Connectedness does not imply path-connectedness

We saw that both path-connectedness and connectedness behave well under continuous maps. However, this is not the case when taking closure - connectedness is stable under taking closure and path-connectedness not. First,

Lemma 2.17. *If A is a connected subset of (X, τ_X), then $cl(A)$ is also connected.*

Proof. This will be on the exercise sheet. □

Exercise 2.3. *Show by example that even if A is connected $int(A)$ might be disconnected.*

In the other direction, the example of topologist sine-curve will show that the closure of a path-connected space is not necessarily path-connected. To do show this we start from the main proposition showing that connectedness does not imply path-connectedness:

Proposition 2.18. *The topologist sine-curve $S := \{(x, \sin(\frac{1}{x}) : x \in (0, 1]\} \cup \{(0, y) : y \in [-1, 1]\}$ is connected, but not path-connected as a subset of $\mathbb{R}^2$.*

Noticing that $S = cl(\{(x, \sin(\frac{1}{x}) : x \in (0, 1]\}$ and that $\{(x, \sin(\frac{1}{x}) : x \in (0, 1]\}$ is path-connected, we deduce:

Corollary 2.19. *Path-connectedness is not stable under closure.*

Proof of Proposition 2.18. Let us denote the topologist sine-curve by $S = I \cup C$, with $C = \{(x, \sin(\frac{1}{x}) : x \in (0, 1]\}$ for 'curve' and $I = \{(0, y) : y \in [-1, 1]\}$ for 'interval'.

Part 1: S is connected.
As C is the image of $(0, 1]$ under the continuous map $f(x) = \sin(\frac{1}{x})$, it is connected by Proposition 2.14. Moreover, notice that $S = cl(C)$: indeed, one can verify that all of I is the boundary of C using Lemma 1.17. Thus by Lemma 2.17, the set S is connected.

Part 2: S is not path-connected.
We will prove this by contradiction. Indeed, we suppose for contradiction that S is path-connected. In particular this means that would mean that there exists a continuous function $\gamma : ([0, 1], \tau_E) \to (\mathbb{R}^2, \tau_E)$ going from $\gamma(0) = (1, \sin(1))$, $\gamma(1) = (0, 0)$ to $\gamma([0, 1]) \subseteq S$. Our aim is to show that in fact such a γ cannot be continuous and we do it again in two steps. The intuition is that at some point the path has to cross from C to L, however due to the

oscillations it cannot do it in a continuous way.

Step 2(a): locating the point of discontinuity. The point of discontinuity should be where the path first enters I. Thus we define $s = \sup\{t \in [0,1] : \gamma([0,t]) \in C\}$. We claim that $\gamma(s) \in I$.

Write $\gamma(t) = (x_\gamma(t), y_\gamma(t))$. Then from Lemma 1.36 it follows that both $x_\gamma(t)$ and $y_\gamma(t)$ are continuous on $[0,1]$. By definition of s and by continuity of γ, we have that $s > 0$. Now, if $s = 1$, then by definition of γ, we have $\gamma(s) = (0,0) \in I$. If, however, $s < 1$, then by definition of s we can find $\forall \epsilon > 0$ some $s_\epsilon \in (s, s+\epsilon)$ such that $x_\gamma(s_\epsilon) = 0$. But by assumption $x(t)$ is continuous at $t = s$, and we conclude that $x_\gamma(s) = 0$ and hence $\gamma(s) \in I$.

Step 2(b): proving discontinuity: The idea (that is easy to graphically verify) is to pick two sequences $(a)_{n\geq 1}$ and $(b)_{n\geq 1}$ such that

- $\forall n \geq 1$, both $a_n \leq s$ and $b_n \leq s$ and both $(a_n)_{n\geq 1} \to s$ and $(b_n)_{n\geq 1} \to s$;
- it holds that $y_\gamma(a_n) = 0$ and thus $y_\gamma(a_n) \to 0$ as $a_n \to s$;
- it also holds that $y_\gamma(b_n) = 1$ thus $y_\gamma(b_n) \to 1$ as $a_n \to s$.

Indeed, doing this proves discontinuity of the function $\gamma(t)$ at $t = s$: if it was continuous, then $y_\gamma(t)$ would be continuous at $t = s$ and $y_\gamma(a_n)$ and $y_\gamma(b_n)$ would converge to the same value.

It thus remains to justify that we can pick such sequences. Graphically/intuitively this is rather clear (why?), but it requires a bit of justification to do it rigorously (why?).

For example, define $a_n = \sup_{t<s}\{x_\gamma(t) = \frac{1}{\pi n}\}$ and $b_n = \sup_{t<s}\{x_\gamma(t) = \frac{2}{\pi(4n+1)}\}$. It remains to verify that these sequences satisfy the properties above. Let us do it here for a_n. By the Intermediate Value Theorem the sets we take the infimum / maximum over are non-empty, thus the value a_n exists. By continuity $x_\gamma(a_n) = \frac{1}{\pi m}$ and thus $y_\gamma(a_n) = 0$. By definition $a_n \leq s$ and finally we claim that a_n converges to s. If this was not the case, then as a_n is a bounded sequence, by the usual Bolzano-Weierstrass Theorem there would be a subsequence $(a_{n_k})_{k\geq 1}$ converging to some $s' < s$. But we have $x_\gamma(s') = \lim_{k\to infty} \frac{1}{\pi n_k} = 0$ and thus $s' \in I$, contradicting the definition of s.

□

2.1.3 Products and connectedness

Both connectedness and path-connectedness behave nicely under taking products. For path-connectedness this is direct:

Exercise 2.4. *Suppose that (X_i, τ_{X_i}) are path-connected topological spaces for all $i \in I$. Prove that the product $\Pi_{i\in I} X_i$ with its product topology is path-connected. [Hint: first, generalize Exercise 5 from the Exercise sheet 4 to infinite product spaces.]*

For connectedness, we will satisfy ourselves with only finite product spaces for now, although the result is also true for arbitrary products.

Proposition 2.20. *Let $(X_1, \tau_{X_1}), \dots, (X_n, \tau_{X_n})$ be connected topological spaces. Then $(X_1 \times \dots \times X_n, \tau_{X_1\times \dots X_n})$ is also connected.*

There are many ways of proving this proposition. We will depart from the following useful lemma:

Lemma 2.21. *Let (X,τ_X), (Y,τ_Y), (Z,τ_Z) be topological spaces and suppose that $f:(X\times Y,\tau_{X\times Y})\to(Z,\tau_Z)$ is continuous. Then $\forall x\in X$ the function $f_x(y):=f(x,y):(Y,\tau_Y)\to(Z,\tau_Z)$ is continuous, as is $\forall y\in Y$ the function $f_y(x):=f(x,y):(X,\tau_X)\to(Z,\tau_Z)$.*

We saw before that continuity along coordinates does not imply joint continuity, and this lemma says that in the other direction things do go well: join-continuity does imply continuity along each coordinate.

Proof. The proof of this lemma is left as an exercise. □

We are now ready to prove the proposition:

Proof of Proposition 2.20. Let us first prove the case $n=2$. To do this consider any continuous $f:(X_1\times X_2,\tau_{X_1\times X_2})\to(\{0,1\},\tau_D)$. By Lemma 2.10 it suffices to show that f is constant. Now fix some $(x_1,x_2)\in X_1\times X_2$ and consider any $(y_1,y_2)\in X_1\times X_2$. By Lemma 2.21 we have that $f_{x_1}(x):(X_2,\tau_{X_2})\to(\{0,1\},\tau_D)$ is continuous, and as X_2 is connected, Lemma 2.10 implies that f_{x_1} is constant. Thus $f(x_1,x_2)=f(x_1,y_2)$. Similarly we obtain that $f_{y_2}(x)$ is constant and hence $f(x_1,y_2)=f(y_1,y_2)$. But then $f(x_1,x_2)=f(y_1,y_2)$, hence f is constant and the case $n=2$ follows.

The case for arbitrary n now follows by induction. Indeed, suppose we know that $(X_1\times\cdots\times X_{n-1},\tau_{X_1\times\ldots X_{n-1}})$ is connected. Then by the case $n=2$ we know that the product of $X_1\times\cdots\times X_{n-1}$ and X_n with its product topology is also connected. As we know that this topology is exactly the product topology on $X_1\times\ldots X_n$, the proposition follows. □

2.1.4 Connected components

For now we only looked at the topological invariant of being connected, but this doesn't allow us to tell the difference between $[0,1]\cup[2,3]$ and $[0,1]\cup[2,3]\cup[4,5]$. So it is useful to introduce the notion of connected components. It is intuitive then that the number of connected components should also be a topological invariant.

Definition 2.22 (Connected component)**.** *Let (X,τ_X) be a topological space. The connected component C_x of $x\in X$, is the union of all connected subsets containing x.*

Exercise 2.5. *Prove that connected components are indeed connected, and that any connected component is necessarily closed. Show by example that connected components are not necessarily open. Can you find a criteria for all components to be open?*

Proposition 2.23. *Let (X,τ_X) be a topological space. Any two connected components of X are either disjoint, or coincide. Connected components $(C_i)_{i\in I}$ form a partition of X, i.e. $X=\bigcup_{i\in I}C_i$. Moreover, any homeomorphism between topological spaces (X,τ_X) and (Y,τ_Y) induces a bijection between the sets of connected components of X and Y.*

Proof. The proof is left as an exercise □

In particular, it follows that the number of connected components is a topological invariant, i.e. if (X,τ_X) has $n\in\mathbb{N}$ connected components then so does a homeomorphic space (Y,τ_Y). The same holds if we replace $n\in\mathbb{N}$ by "countably many" or "uncountably many".

A space where every point is a separate connected component are called totally disconnected. For example any topological space with the discrete topology is a totally disconnected space.

Exercise 2.6 ($\star$). *Prove that the Cantor space (see e.g. Exercise Sheet 4, exo 12) is a totally disconnected space.*

One can similarly define path-connected components of a point $x \in X$: the path-component P_x of x is the set of all points y such that there is path γ from x to y. Both in case of path-components and connected components, there is also a nice way to think about these components. Namely, we can define an equivalence relation on X such that $x \simeq y$ if there is a path from $x \to y$, then path-components are exactly the equivalence classes under this relation.

2.2 Simple-connectedness - can we identify the holes?

We saw that one way to think of connectedness, or rather disconnectedness, was to think about the existence of "gaps" between different components. In the rubber example - if you take a piece of rubber and tear it into two, you create a gap and clearly there is something discontinuous in this operation as some points that were very close are now distant. Another violent operation (in terms of continuity) would be to push a hole through rubber. Indeed, in this case also some points that were very close together would become far apart, even though the whole piece might remain connected.

Hence the existence and the number of such holes should help us differentiate between different topological spaces. In that sense a croissant and a doughnut should be topologically different sweets, and "u" and "o" topologically different letters. Continuing with everyday examples, should a T-shirt and jeans topologically the same?

Our aim is to develop some notions that will help to detect and describe the existence of such holes [9]. As this is already in the realm of algebraic topology, we will remain very brief.

In this section we will only consider path-connected spaces X. Recall that path-connectedness means that for any pair of points x and y, the point x can be continuously transported to y in the space X. The definition of path-homotopy gives a notion of connectedness for paths - two paths γ_1, γ_2 from x to y are "connected" (called path-homotopic) if one can be deformed or transported continuously to the other one. If this is possible for all pairs of paths, the space will be called simply-connected.

We will soon make this definition precise, and later see that there are other equivalent ways of defining simply-connectedness (like e.g. given in your course of complex analysis). It comes out that simply-connectedness is a very useful concept. It is the first topological property (we haven't yet seen that it is a topological property, but we will see), that really puts strong constraints on the space and helps us classify different spaces:

- You have seen or will see in complex analysis the Riemann mapping theorem: for any simply-connected open subset $D \subseteq \mathbb{C}$ that is not equal to $\mathbb{C}$, there is a holomorphic, bijective map from the unit disk $\mathbb{D}$ to D.
- In further topology courses you would see that any orientable, compact 2D surface without boundary that is simply-connected is homeomorphic to the sphere.
- Poincaré conjectured that the previous statement is true in all dimensions. Only in 2006 it was proved if a orientable, compact 3D manifold without boundary is simply connected, then it is homeomorphic to the sphere.

[9] There are also other types of holes - for example if you consider the sphere in 3D, then it also in some sense has a hole inside. This is not accounted for in this chapter

2.2.1 Homotopy of paths and simply-connectedness

Let us start now formalizing the ideas above.

Definition 2.24 (Homotopy of paths)**.** *Let (X, τ_X) be a topological space and $\gamma_1 : [0,1] \to X$, $\gamma_2 : [0,1] \to X$ two continuous paths in X with the same starting point x_0, and the same end-point x_1. We say that γ_0 is (path-)homotopic to γ_1 if there exists a continuous map $F : ([0,1]^2, \tau_E) \to (X, \tau_X)$ such that $\forall s \in [0,1]$ we have that $F(s,0) = \gamma_0(s)$ and $F(s,1) = \gamma_1(s)$, and moreover $F(0,t) = x_0$ and $F(1,t) = x_1$ for any $t \in [0,1]$. We say that F is a path-homotopy between γ_0 and γ_1*

Remark 2.25. *Notice that from Lemma 2.21 we know that for any fixed $t \in [0,1]$, the function $\gamma_t(s) := F(s,t)$ is a continuous map from $([0,1], \tau_E)$ to (X, τ_X), and in particular as $F(0,t) = x_0$ and $F(1,t) = x_1$, $\gamma_t(s)$ is a path from x_0 to x_1. In other words, the homotopy map F does give an interpolation from the path γ_0 to γ_1 via continuous paths with fixed endpoints.*

Remark 2.26. *Sometimes you will also find definitions of homotopy where the roles of s, t are exchanged. This is just a matter of convention and you should be ready to meet both versions.*

Definition 2.27 (Simply-connected I)**.** *A path-connected space (X, τ_X) such that for any $x, y \in X$, any two paths γ_0, γ_1 from x to y are path-homotopic is called simply-connected.*

You can now prove directly from this definition that $\mathbb{R}^n$ is simply-connected - you just have to come up a path-homotopy, i.e. a natural way to interpolate between any two paths with the same endpoints.

Exercise 2.7. *Prove that in $\mathbb{R}^n$ with the Euclidean topology any two paths with the same starting point and the same end-point are path-homotopic and thus that $(\mathbb{R}^n, \tau_E)$ is simply-connected.*

In any topological space one can equivalence relation between points, so that two points are in the same equivalence class if and only if they are path-connected. Equivalence classes then corresponded to path-connected components. Similarly, the notion of path-homotopy also provides an equivalence relation between continuous paths and partitions them into separate equivalence classes:

Lemma 2.28 (Path-homotopy is an eq. relation)**.** *Let (X, τ_X) be a topological space. Path-homotopy induces an equivalence relation $\simeq$ on continuous paths $\gamma : [0,1] \to X$ with a fixed starting point $x_0 \in X$ and a fixed endpoint $x_1 \in X$. We denote the equivalence class of a path γ by $[\gamma]$ and the set of equivalence classes by Γ_{x_0,x_1}*

Proof. The proof is left as an exercise. □

An intuitive and useful fact is that time-reparametrization does not change the homotopy-equivalence class of a path:

Lemma 2.29. *Let γ be a path in (X, τ_X) from x to y and $\psi : ([0,1], \tau_E) \to ([0,1], \tau_E)$ be continuous with $\psi(0) = 0$ and $\psi(1) = 1$. Then $\gamma \circ \psi$ is also a path from x to y and moreover $\gamma \simeq \gamma \circ \psi$.*

Proof. First, notice that $\gamma \circ \psi$ is indeed a path from x to y: $\gamma \circ \psi$ is a continuous function from $([0,1], \tau_E)$ to (X, τ_X) as a composition of continuous functions. Moreover, $\gamma \circ \psi(0) = \gamma(0) = x$ and $\gamma \circ \psi(1) = \gamma(1) = y$.

Let us next define the homotopy map. To do this, we will just interpolate between the two time-parametrizations linearly. In other words, define

$$F(s,t) := \gamma\left((1-t)s + t\psi(s)\right).$$

We claim that F is a homotopy between γ and $\gamma \simeq \gamma \circ \psi$. Indeed, $F(s,0) = \gamma(s)$, $F(s,1) = \gamma \circ \psi$, $F(0,t) = \gamma(0)$ as $\psi(0) = 0$ and similarly $F(1,t) = \gamma(1)$. It remains to argue that F is continuous. One can directly check that $g : ([0,1]^2, \tau_E) \to ([0,1], \tau_E)$ given by $(s,t) \to (1-t)s + t\psi(s)$ is continuous, and then $F = \gamma \circ g$ is a composition of continuous functions and thus continuous. $\square$

2.2.2 Algebraic operations on homotopy eq. classes

An interesting fact is that whereas the equivalence classes of path-connectedness, i.e. path-components just form a set, the homotopy equivalence classes can be endowed with algebraic operations:

- Firstly, for any two paths γ_1 from x to y and γ_2 from y to z, we already defined the concatenated path $\gamma_3 := \gamma_1 * \gamma_2$ in Lemma 2.7:
$$\gamma_3(t) = \begin{cases} \gamma_1(2t) & \text{if } t \leq 1/2 \\ \gamma_2(2t-1) & \text{if } t \geq 1/2 \end{cases}$$
- Secondly, for any path γ we can define its time-reverse by $\overleftarrow{\gamma}(t) = \gamma(1-t)$.

It comes out that these operations induce operations not only on individual loops, but actually on the space of path-homotopy equivalence classes of loops. In other words, we claim that

- if $\gamma_1, \simeq \widetilde{\gamma}_1$ are homotopic paths from x to y and $\gamma_2 \simeq \widetilde{\gamma}_2$ are homotopic paths from y to z, then $\gamma_1 * \gamma_2 \simeq \widetilde{\gamma}_1 * \widetilde{\gamma}_2$ as paths from x to z
- similarly, if $\gamma_1, \simeq \widetilde{\gamma}_1$, then $\overleftarrow{\gamma}_1 \simeq \overleftarrow{\widetilde{\gamma}_1}$.

Let us state this as a lemma:

Lemma 2.30. *Concatenation and time-reversal induce well-defined operations*

- $* : \Gamma_{x,y} \times \Gamma_{y,z} \to \Gamma_{x,z}$, *defined by* $[\gamma_1] * [\gamma_2] := [\gamma_1 * \gamma_2]$ *and*
- $\overleftarrow{[\gamma]} : \Gamma_{x,y} \to \Gamma_{y,x}$, *defined by* $\overleftarrow{[\gamma]} := [\overleftarrow{\gamma}]$,

on the sets of homotopy equivalence classes.

For this proof it is very helpful to draw a picture of the different homotopies:

Proof. We will explain here how to do time-reversal and leave the more interesting case of concatenation of paths to the exercise sheet. In both cases one has to define suitable homotopy maps F_1.

Consider $\gamma_1, \simeq \gamma_2$. We want to show that $\overleftarrow{\gamma}_1 \simeq \overleftarrow{\gamma_2}$. Now, let F be the path-homotopy between γ_1 and γ_2. Consider

$$F_1(s,t) = F_1(1-s,t)$$

. Let us check that F_1 is a homotopy between $\overleftarrow{\gamma}_1$ and $\overleftarrow{\gamma_2}$. First, F_1 is continuous from $([0,1]^2, \tau_E)$ as a composition of continuous maps. Second, by definition for all $(s,t) \in [0,1]^2$, we have that $F_1(s,0) = \overleftarrow{\gamma}_1(s)$, $F_1(s,1) = \overleftarrow{\gamma_2}(s)$, $F_1(0,t) = F(1,t) = y$ and $F_1(1,t) = F(0,t) = x$ and thus indeed F_1 is a path-homotopy.

□

As we see in the next chapter, these operations behave even more nicely on a set of closed loops rooted at some point x, and give rise to the group structure.

2.2.3 The fundamental group

We are now ready to prove the main theorem of this section, stating that the set of equivalence classes $\Gamma_x := \Gamma_{x,x}$ of closed paths starting and ending at x forms a group under the operation $*$. This is the basic result of algebraic topology proved by H. Poincaré in the end of 19th century that brings into light a wonderful interplay between geometry and algebra:

Theorem 2.31. *Let (X, τ_X) be a topological space and $x \in X$. Then the set of homotopy equivalence classes Γ_x of closed paths rooted at x (i.e. paths from x to x), equipped with*

- *the operation $* : \Gamma_x \times \Gamma_x \to \Gamma_x$,*
- *and the identity given by eq. class of the constant path $[e_x(t)] := x$,*
- *the inverse given by the time-reversal $[\overleftarrow{\gamma}]$,*

forms a group, called the fundamental group of X at the point x, that is denoted by $\pi_1(X, x)$.

To prove the theorem we have to just verify that the axioms of a group are satisfied for the operation $*$. Firstly, the set Γ_x is closed under the operation $*$, thus it remains to prove associativity, the existence of identity elements and of inverses. We state all this in a slightly more general form, that is useful for us later on:

Lemma 2.32. *Let γ_1 be a path from x to y, γ_2 a path from y to z, and γ_3 a path from z to w. The concatenation operation $*$ on equivalence classes of paths satisfies the following properties:*

(1) (Associativity) We have that $[\gamma_1] * ([\gamma_2] * [\gamma_3]) = ([\gamma_1] * [\gamma_2]) * [\gamma_3]$;
(2) (Identity elements) If we denote by e_x the path that stays constantly at the point x, then $[e_x] * [\gamma_1] = [\gamma_1] = [\gamma_1] * [e_y]$;
(3) (Inverses) We have that $[\gamma_1] * [\overleftarrow{\gamma}_1] = [e_x]$ and $[\overleftarrow{\gamma}_1] * [\gamma_1] = [e_y]$.

Again, I encourage you to draw pictures of the different homotopies needed in the proof:

Proof. Take some representatives $\gamma_i \in [\gamma_i]$. Then a path γ in $[\gamma_1] * ([\gamma_2] * [\gamma_3])$ can be written as

$$\gamma(t) = \begin{cases} \gamma_1(2t) & \text{if } t \leq 1/2 \\ \gamma_2(4t-2) & \text{if } 1/2 \leq t \leq 3/4 \\ \gamma_3(4t-3) & \text{if } t \geq 3/4 \end{cases}$$

Similarly a path $\widetilde{\gamma}$ in $([\gamma_1] * [\gamma_2]) * [\gamma_3]$ can be written as

$$\widetilde{\gamma}(t) = \begin{cases} \gamma_1(4t) & \text{if } t \leq 1/4 \\ \gamma_2(4t-1) & \text{if } 1/4 \leq t \leq 1/2 \\ \gamma_3(2t) & \text{if } t \geq 1/2 \end{cases}$$

Thus we can see that $\widetilde{\gamma}(t) = \gamma \circ \psi(t)$ with $\psi(t)$ a piece-wise linear function:

$$\psi(t) = \begin{cases} 2t & \text{if } t \leq 1/4 \\ t + 1/4 & \text{if } 1/4 \leq t \leq 1/2 \\ t/2 + 1/2 & \text{if } t \geq 1/2 \end{cases}$$

In particular $\psi(t)$ is continuous and thus it follows from Lemma 2.29 that γ and $\widetilde{\gamma}$ are path-homotopic.

An analogous argument also shows the second point. Let us here only consider the first half, i.e. let us show that $[e_x] * [\gamma_1] = [\gamma_1]$, the other half following similarly. Let $\gamma_1 \in [\gamma_1]$. Then $e_x * \gamma_1$ can be written as $\gamma_1 \circ \psi(t)$ with $\psi(t) = 0$ when $t \leq 1/2$ and $\psi(t) = 2t - 1$ for $t \geq 1/2$. As $\psi(t)$ is continuous, it again follows from Lemma 2.29 that $e_x * \gamma_1$ and γ_1 are path-homotopic.

Finally, let us prove the third part of the lemma. Again we prove the first half of the statement, the other half following similarly. Let $\gamma_1 \in [\gamma_1]$ and write $\gamma(t) := \gamma_1 * \overleftarrow{\gamma}_1(t)$. Then $\gamma(t) = \gamma_1(2t)$ for $t \leq 1/2$ and $\gamma(t) = \gamma_1(2 - 2t)$ for $t \geq 1/2$. Now define $F(s,t)$ by

$$F(s,t) = \begin{cases} \gamma(s) & \text{if } s \leq 1/2 - t/2 \\ \gamma(s) & \text{if } s \geq 1/2 + t/2 \\ \gamma(1/2 - t/2) & \text{if } 1/2 - t/2 \leq s \leq 1/2 + t/2 \end{cases}$$

We claim that F is a homotopy between γ and e_x. First, notice that $\gamma(1/2 - t/2) = \gamma(1/2 + t/2) = \gamma_1(1 - t)$. Thus, for any fixed t the path $F(s,t)$ just goes up to the point $\gamma_1(1-t)$ along γ_1 with a certain speed, stays there for some time and then returns (Again do draw a picture!). Let us now check the conditions for being a homotopy. We have that $F(s,0) = \gamma$ and $F(s,1) = x = e_x$. Also, $F(0,t) = F(1,t) = x$. It remains to check that F is continuous. But this follows from the Pasting Lemma (Lemma 2.8) as $F(s,t)$ can be checked to be continuous on the closed sets $S_1 := [0,1]^2 \cap (\{(s,t) : s \leq 1/2 - t/2\} \cup \{(s,t) : s \geq 1/2 + t/2\})$ and $S_2 := [0,1]^2 \cap \{(s,t) : 1/2 - t/2 \leq s \leq 1/2 + t/2\}$. Indeed, on S_1 we just have that $F(s,t) = \gamma(s)$ and on S_2, $F(s,t) = \gamma(1/2 - t/2)$ so in both cases the continuity follows from the continuity of γ. $\square$

A priori we have a different group at every point of the space. However, the nice thing is that in a path-connected space all these groups are isomorphic, meaning that there is a bijection between the groups that preserves the group structure.

Proposition 2.33. *Let (X, τ_X) be a path-connected topological space. Then for any $x, y \in X$ the groups $\pi_1(X,x)$ and $\pi_1(X,y)$ are group-isomorphic.*

Proof. Let η be a path from x to y. We claim that the map $G : \Gamma_y \to \Gamma_x$ defined for $\gamma_y \in \Gamma_y$ by $G([\gamma_y]) = [\eta] * [\gamma_y] * [\overleftarrow{\eta}]$ induces a group-isomorphism between $\pi_1(X,x)$ and $\pi_1(X,y)$. In other words, we need to prove that G is a bijection and that it preserves the group structure: for any $[\gamma_1], [\gamma_2] \in \Gamma_y$ we should have that $G([\gamma_1]) * G([\gamma_2]) = G([\gamma_1] * [\gamma_2])$.

Let us start from the latter point, i.e from the fact that the group structure is preserved by G. By definition of G, we have that

$$G([\gamma_1]) * G([\gamma_2]) = ([\eta] * [\gamma_1] * [\overleftarrow{\eta}]) * ([\eta] * [\gamma_2] * [\overleftarrow{\eta}])$$

By associativity of $*$, i.e. Lemma 2.32 we can write this as

$$[\eta] * ([\gamma_1] * ([\overleftarrow{\eta}] * [\eta]) * [\gamma_2]) * [\overleftarrow{\eta}].$$

But again by Lemma 2.32 we have that $[\overleftarrow{\eta}] * [\eta] = [e_y]$ and moreover that $[\gamma_1] * [e_y] * [\gamma_2] = [\gamma_1] * [\gamma_2]$. Thus

$$G([\gamma_1]) * G([\gamma_2]) = [\eta] * ([\gamma_1] * [\gamma_2]) * [\overleftarrow{\eta}] = G([\gamma_1] * [\gamma_2]).$$

It remains to check that G is bijective. This can be done by just writing out the inverse map, but let us here check step by step:

- G is injective: indeed, let $[\gamma_1] \in \Gamma_y$. Then by associativity of $*$, we have that
$$[\overleftarrow{\eta}]G([\gamma_1])[\eta] = [\eta] * ([\overleftarrow{\eta}] * [\gamma_1] * [\eta]) * [\overleftarrow{\eta}] = \gamma_1,$$
and thus if $G([\gamma_1]) = G([\gamma_2])$ then $[\gamma_1] = [\gamma_2]$.
- G is surjective: let $[\widetilde{\gamma}] \in \Gamma_x$. Then $[\overleftarrow{\eta}] * [\widetilde{\gamma}] * [\eta]$ is in Γ_y. Moreover, by associativity of $*$,
$$G([\overleftarrow{\eta}] * [\widetilde{\gamma}] * [\eta]) = [\eta] * ([\overleftarrow{\eta}] * [\widetilde{\gamma}] * [\eta]) * [\overleftarrow{\eta}] = [\widetilde{\gamma}].$$

$\square$

This proposition justifies the usage of "the fundamental group of space X is ...", although one should bear in mind that each point gives a different representative and there is no canonical isomorphism between the different groups (it may depend on the chosen path).

There is now a very nice way to restate simple-connectedness:

Lemma 2.34 (Simply-connected space II). *A path-connected topological space (X, τ_X) is called simply-connected if and only if its Fundamental group is trivial (one-element) group.*

Often this is condition - that the fundamental group is trivial - is used as the definition for simple-connectedness.

Proof. Let x be any point in X. Then as (X, τ_X) is simply connected, any path in Γ_x is path-homotopic to the constant path e_x. Thus we see that $\Pi_1(X, x)$ is trivial. In the other direction consider any two paths γ_1 and γ_2 from x to y. Then $[\gamma_1] * [\overleftarrow{\gamma}_2] \in \Gamma_x$. But now $\Pi_1(X, x)$ is trivial and thus $[\gamma_1] * [\overleftarrow{\gamma}_2] = [e_x]$. Acting now by $*[\gamma_2]$ from the left, we obtain that $[\gamma_1] = [\gamma_2]$, as claimed. $\square$

Continuous maps induce homomorphism on fundamental groups, and based on this one can see that the fundamental groups of homeomorphic spaces are isomorphic:

Lemma 2.35. *Let $(X, \tau_X), (Y, \tau_Y)$ be a path-connected topological spaces and $f : X \to Y$ a continuous map with $f(x) = y$. Then $f_* : \pi_1(X, x) \to \pi_1(Y, y)$ defined by $f_*([\gamma]) = [f \circ \gamma]$ is a (group) homomorphism. Moreover, if $(X, \tau_X) \cong (Y, \tau_Y)$, then the fundamental groups of (X, τ_X) and (Y, τ_Y) are (group) isomorphic.*

In particular this means that the fundamental group is a topological invariant.

Proof. The proof is left as an exercise. $\square$

By showing that all paths in $\mathbb{R}^n$ were path-homotopic, we already showed that $\mathbb{R}^n$ is simply-connected. Our next aim is to quickly discuss cases where the fundamental group is something more exciting. In particular, we look at the case of the circle S^1 and the punctured plane $\mathbb{R}^2 \backslash \{0\}$ - their fundamental groups are isomorphic to the additive group $\mathbb{Z}$. In general fundamental groups are not necessarily commutative, and indeed might turn out to be complicated and not easy to even compute. You will see some exciting cases in further topology courses. Here, however, we will content ourselves with stating and discussing, but not proving the following theorem:

Theorem 2.36. *The fundamental group of the circle S^1 is isomorphic to the additive group $\mathbb{Z}$.*

The idea of the proof is very nice and works in a much more general setting: we set up a map (a covering map) between the circle S^1 and $\mathbb{R}$ by heuristically wrapping $\mathbb{R}$ around S^1 infinitely - in this case it is just the map $f(x) = e^{2\pi ix}$. Notice that via this map for each point $p \in S^1$, we have that $f^{-1}(p)$ is given by a shifted copy of Z. Even more, if you consider the set of homeomorphisms $\tau : \mathbb{R} \to \mathbb{R}$ such that $f \circ \tau = f$, then this set of homeomorphisms with the operation of composition is also isomorphic to $\mathbb{Z}$. In fact this is the key insight in the slightly technical proof, given in the handout - there is a correspondence between the group of closed loops in S^1 and the group of automorphisms of $\mathbb{R}$, that are compatible with f. As the elegant theory of covering spaces is out of the scope of this course, the complete proof (using only hands-on methods accessible to us) is non-examinable and sketched on an independent exercise sheet.

Exercise 2.8. *Now we can prove in a few steps that $\mathbb{R}^2$ is not homeomorphic to any $\mathbb{R}^n$, $n > 2$ (we already know the case $n = 1$):*

- *Show that the fundamental group of $\mathbb{R}^2\backslash\{0\}$ is also isomorphic to $\mathbb{Z}$.*
- *Show that $\mathbb{R}^n\backslash\{0\}$ is simply-connected.*

$\star$ (non-examinable) The fundamental group of S^1

We parametrize $S^1 = \{e^{i2\pi\theta} : \theta \in [0,1)\}$ and set $p_0 = (1,0)$. A natural distance d_S in S^1 is given by the 'angular' difference $d_S(e^{i2\pi\theta_1}, e^{i2\pi\theta_2}) := \min(|\theta_1 - \theta_2|, 1 - |\theta_1 - \theta_2|)$ (check it's a distance!). A natural distance on paths on S^1 is then given by $\delta_S(\gamma_1, \gamma_2) := \sup_{s\in[0,1]} d_S(\gamma_1(s), \gamma_2(s))$. On $\mathbb{R}$ we use the usual distance d and the path distance $\delta(\gamma_1, \gamma_2) := \sup_{s\in[0,1]} d(\gamma_1(s), \gamma_2(s))$.

The following steps should help you prove that the fundamental group of S^1 is isomorphic to $\mathbb{Z}$.

(1) Defining the covering map:
   - Consider the map $f(x) = e^{i2\pi x}$ from $\mathbb{R}$ to S^1. Observe that for any $p \in S$ we have that $f^{-1}(p) = a_p + \mathbb{Z}$ for some $a_p \in [0,1)$. For each $p \in S^1$, denote by U_p the open interval of d_S-length $1/4$ centred at p in S^1. Show that $f^{-1}(U_p) = \cup_{n\in\mathbb{Z}} V_{a_p+n}$, where a_p is as above and V_x denotes an open interval of Euclidean-length $1/4$ around $x \in \mathbb{R}$.
   - Argue that f is locally a homeomorphism: for any $x \in \mathbb{R}$, prove that $h_x : V_x \to U_{f(x)}$ given by restricting f to V_x is a homeomorphism.

(2) Lifting the paths: we aim to show that for any path γ in S^1 from p_0 to p_0 there exists a unique path $\widetilde{\gamma}$ in $\mathbb{R}$ starting from 0 such that $f \circ \widetilde{\gamma}_n = \gamma$. Observe that $\widetilde{\gamma}(1) \in \mathbb{Z}$.
   - Define s as the supremum of $t \in [0,1]$ such that there is unique $\widetilde{\gamma}$ for the path defined on $[0,t]$. Prove that $s > 0$.
   - Prove by contradiction that s is not smaller than 1 and conclude the existence and uniqueness of $\widetilde{\gamma}$.
   - Argue that this lifting maps close-by paths to close-by paths: using the fact that f is locally a homeomorphism and the specific expression for f, show that if two

paths γ_1, γ_2 from p_0 to p_0 satisfy $\delta_S(\gamma_1, \gamma_2) < \epsilon$ for $\epsilon > 0$ small enough, then in fact $\delta(\widetilde{\gamma}_1, \widetilde{\gamma}_2) = \delta_S(\gamma_1, \gamma_2)$.

(3) Lifting the homotopy: we now show that two paths γ_1 and γ_2 from p_0 to p_0 are path-homotopic in S^1 if and only if the paths $\widetilde{\gamma}^1$ and $\widetilde{\gamma}^2$ are path-homotopic in $\mathbb{R}$. Denote by F the homotopy between γ_1 and γ_2.
   - Define $\widetilde{F} : [0,1]^2 \to \mathbb{R}$ by lifting for each fixed $t \in [0,1]$ the path $\gamma_s(t) = F(s,t)$, i.e. set $\widetilde{F}(s,t) = \widetilde{\gamma}_s(t)$.
   - Argue that $\widetilde{F}$ is continuous by using continuity of F and the fact that close-by paths are mapped to close-by paths. For example, define u as the supremum of $t \in [0,1]$ so that $\widetilde{F}$ is continuous on $[0,1] \times [0,t]$. Prove as above that $t > 0$, and prove further that $t = 1$.

(4) Show that if two paths are homotopic in $\mathbb{R}$, then they are homotopic in S^1.

(5) We now define a group isomorphism $j : \pi_1(S^1, p_0) \to \mathbb{Z}$:
   - For any path γ in S^1 from p_0 to p_0, we can look at the endpoint of the lift, i.e. assign $m_\gamma := \widetilde{\gamma}(1)$. Using the previous points (and simple-connectedness of $\mathbb{R}$), argue that m_γ only depends on the equivalence class of γ. Thus define $j([\gamma]) := m_\gamma$.
   - Show that j is an injection from $\pi_1(S^1, p_0)$ to $\mathbb{Z}$ (by using simple-connectedness of $\mathbb{R}$).
   - By considering a sufficiently large class of paths in S^1, show that j is a surjection from $\pi_1(S^1, p_0)$ to $\mathbb{Z}$.
   - Finally, by arguing that the lift of $\gamma_1 * \gamma_2$ is given by $\widetilde{\gamma}_1 * (m_{\gamma_1} + \widetilde{\gamma}_2)$. show that j is a homomorphism, i.e. that $j([\gamma_1 * \gamma_2]) = j([\gamma_1]) + j([\gamma_2])$.

[$\star$ End of an non-examinable section $\star$]

Section 3

Compactness

Oxford dictionary of English tells us what compact means - closely and neatly packed together. And this is indeed very close to its meaning in mathematics. In Analysis II you saw three equivalent definitions for compactness in $\mathbb{R}^n$:

(1) $K \subseteq \mathbb{R}^n$ is compact iff it is closed and bounded;
(2) K is compact iff every sequence in K has a convergent subsequence with a limit in K;
(3) K is compact iff every covering of K with open balls admits a finite subcover - i.e. if $(B_i)_{i\in I}$ are open balls with $K \subseteq \bigcup_{i\in I} B_i$, then $K \subseteq \bigcup_{i\in I_0} B_i$ for some finite set $I_0 \subseteq I$.

We will start a fresh, try to generalize these definitions to arbitrary topological spaces and to then study the resulting notions. In fact, the first of these definitions does not naturally generalize to arbitrary topological spaces as there is no notion of boundedness. The other two do generalize, but actually give rise to slightly different concepts, as we will see soon. As a by-product we will reprove the equivalences for $\mathbb{R}^n$ as well.

To start, let us fix some helpful vocabulary: we say that a collection of sets $(U_i)_{i\in I}$ covers another set K if $K \subseteq \bigcup_{i\in I} U_i$ [10]. We will use the term open cover, when all the sets in this collection are open in the underlying topological space.

Definition 3.1 (Compactness)**.** *A topological space (X, τ_X) is called compact if any open cover of X admits a finite subcover, i.e. if I is any index set, U_i are open for all $i \in I$ and $\bigcup_{i\in I} U_i = X$, then there exists a finite subset $I_0 \subseteq I$ such that $\bigcup_{i\in I_0} U_i = X$.*

To get well-acquainted with this definition, let us start from some basic examples:

- A space (X, τ_I) with the indiscrete topology is compact as every cover has to contain the whole space and we can choose just this single set to be our subcover.
- A space (X, τ_D) with the discrete topology is compact iff it is finite.
  Indeed, if X is finite and $(U_i)_{i\in I}$ is an open cover, then for any $x \in X$ we can pick some U_{i_x} containing x. Then the sets U_{i_x} with $x \in X$ form a finite subcover.
  If X is infinite, we can choose the collection of sets $\{x\}$ - they form an open cover without a finite subcover.
- $(\mathbb{R}, \tau_E)$ is not compact: we can take $U_i = (i-1, i+1)$, then $(U_i)_{i\in\mathbb{Z}}$ form an open cover without a finite subcover.
- $(0,1)$ is not compact: we can take $U_i = (i^{-1}, 1)$ - you can again check that they form an open cover without a finite subcover.

More examples are on the example sheet, for now let us verify that the canonical example of a neatly packed set is indeed compact:

Proposition 3.2. *The interval $[0,1]$ with its standard topology is compact.*

Proof. It suffices to prove that for any covering of $[0,1]$ with open sets admits a finite subcover. So consider such a covering $(U_i)_{i\in I}$. Mimicking the proof of connectedness of $[0,1]$, let s be the supremum of $x \in [0,1]$ such that $[0,x]$ can be covered with a finite subcover of

[10] Notice that a priori the same set could appear several times.

$(U_i)_{i\in I}$. Notice that $s > 0$: the point 0 can be covered with a single open set among U_i and there is some $\delta > 0$ so that $[0, \delta) \subseteq U_i$.

We aim to first show that $s = 1$. Suppose for contradiction that $s < 1$. Consider the open interval U_{i_s} covering s. For some $\delta > 0$ we have that $(s - \delta, s + \delta) \subseteq U_{i_s}$. But then by adding U_{i_s} to the existing finite subcover of $[0, s - \delta/2]$ we obtain a finite subcover of $[0, s + \delta]$, contradicting the choice of s.

Now, 1 is also covered by some set U_j and thus again there is some $\delta > 0$ such that $(1 - \delta, 1] \subseteq U_j$. As $s = 1$, we have that $[0, 1 - \delta/2]$ has a finite subcover; adding U_j to this cover gives a finite subcover of $[0, 1]$. □

Remark 3.3. *A quick way to check whether you understand the proof is to identify where the argument fails for the half-open interval* $[0, 1)$*.*

Let us now introduce the other notion of compactness:

Definition 3.4 (Sequentially compact)**.** *A topological space* (X, τ_X) *is called sequentially compact, if any sequence* $(x_n)_{n\geq 1}$ *in* X *admits a convergent subsequence.*

In fact, it comes out that there are many different definitions of compactness, often with no implications between the definitions. For example compactness and sequentially compactness are in general just different:

- there are spaces that are compact, but not sequentially compact, and also spaces that are sequentially compact but not compact.

We will meet such examples in a bit, but we will also see later on that in the realm of metric spaces these definitions are equivalent.

So which is the 'good' definition or notion? It depends on the context, but if the criteria is neatness, then compactness often wins. Let us illustrate this by the proof of the Boundedness theorem.

Theorem 3.5 (Boundedness theorem)**.** *Let* (X, τ_X) *be a compact topological space and* $f : X \to \mathbb{R}$ *a real-valued continuous function. Then* f *is bounded on* (X, τ_X)*, i.e. there exist* $i \in \mathbb{R}, s \in \mathbb{R}$ *such that* $i \leq f(x) \leq s$ *for all* $x \in X$*.*

Proof. Cover $\mathbb{R}$ with open intervals $U_i = (i-1, i+1)$. Define $V_i := f^{-1}(U_i)$. Then $\bigcup_i V_i = X$. As f is continuous, all V_i are also open in X. Thus $(V_i)_{i\in\mathbb{Z}}$ form an open cover of X. Now X is compact, and thus there exist some $i_1, \ldots, i_m$ such that $X = \bigcup_{k=1\ldots m} V_{i_k}$. But then $f(X) \subseteq \bigcup_{k=1\ldots m} U_{i_k}$ and thus for all $x \in X$ we have that $\min_{k=1\ldots m} i_k - 1 \leq f(x) \leq \max_{k=1\ldots m} i_k + 1$ □

The similar result holds for sequentially compact spaces, but the proof argues by contradiction and is not half as neat. Thus we will now mainly work with compactness.

3.1 Some main properties of compactness

We will often consider subsets of a compact space, so here is an useful criteria in verifying that they are compact (with their subspace topology), by just using open sets of the ambient space. This lemma is so intuitive that we will often even forget mentioning that we use it.

Lemma 3.6. *Let* (X, τ_X) *be a topological space and consider* $K \subseteq X$*. Then* $(K, \tau_{X,K})$ *is compact as a topological space if and only if every covering of* K *with open sets of* X *admits a finite subcover.*

The proof is a direct definition-chasing, which is left as an exercise:

Proof. On the exercise sheet. □

Another useful property is that, as with continuity, compactness can be checked only for a more restricted collection of open sets - those forming a basis for the topology:

Lemma 3.7. *Let (X, τ_X) be a topological space, τ_X^B a basis and A some subset. Suppose that any covering of A with sets from τ_X^B admits a finite cover. Then A is compact.*

Proof. On the exercise sheet. □

A central property is the fact that continuous maps preserve compactness:

Proposition 3.8. *Let (X, τ_X) be a compact topological space and $f : (X, \tau_X) \to (Y, \tau_Y)$ be continuous. Then $f(X)$ is compact.*

The proof reminds us of the proof of the Boundedness theorem (there is a good reason for that!):

Proof. By Lemma 3.6, it suffices to show that any cover $(U_i)_{i \in I}$ of $f(X)$ by open sets of Y admits a finite subcover. So consider such a cover. Then as f is continuous we have that $V_i := f^{-1}(U_i)$ are open in X and as $f^{-1}(Y) = X$, they form an open cover of X. But X is compact, and thus we can find a finite set $I_0 \subseteq I$ such that $X = \bigcup_{i \in I_0} V_i$. But then $f(X) \subseteq \bigcup_{i \in I_0} U_i$ and thus we have exhibited a finite subcover of $f(X)$. □

As a direct corollary we see that compactness is a topological property:

Corollary 3.9. *Suppose that (X, τ_X) is compact and (Y, τ_Y) is homeomorphic to (X, τ_X). Then (Y, τ_Y) is compact.*

Using this corollary we can directly see that $([0, 1], \tau_E)$ and $((0, 1), \tau_E)$ cannot be homeomorphic as the closed interval is compact and the open one not.

Another nice corollary of Proposition 3.8 is a strengthening of the Boundedness theorem - the Extreme value theorem, which says that a continuous real-valued function on a compact space X is not only bounded, but it also attains its bounds:

Theorem 3.10 (Extreme value theorem)**.** *Let (X, τ_X) be a compact topological space and $f : X \to \mathbb{R}$ a real-valued continuous function. Then f is bounded on (X, τ_X) and attains its bounds at some points $x_i, x_s \in X$: i.e. there exist $x_i \in X, x_s \in X$ such that $f(x_i) \leq f(x) \leq f(x_s)$ for all $x \in X$.*

Proof. On the exercise sheet. □

3.2 Compact vs closed and the Hausdorff property

In $\mathbb{R}^n$ compact and closed sets go hand in hand. It comes out that they remain friends in the context of general topological spaces, but sometimes take a bit of distance. As a starter, we remark that compactness can be also stated using closed sets:

Lemma 3.11. *A topological space (X, τ_X) is compact if and only if for any collection $(C_j)_{j \in J}$ of closed subsets of X such that the intersection $\bigcap_{j \in J} C_j$ is empty, there exists some finite subset $J_c \subseteq J$ such that $\bigcap_{j \in J_c} C_j$ is empty.*

This also has a nice corollary:

Corollary 3.12. *Let (X, τ_X) be a compact topological space and $(C_n)_{n\geq 1}$ a sequence of nested closed non-empty subsets of X, i.e $\forall n \in \mathbb{N}$: we have $C_n \supseteq C_{n+1}$. Then $\bigcap_{n\in\mathbb{N}} C_n$ is non-empty.*

The proofs of both of these statements are on the exercise sheet. Moreover, it is important to note that this property does not necessarily hold if we are in a non-compact space.

Exercise 3.1. *Find a topological space (X, τ_X) and a sequence of closed sets $(C_n)_{n\geq 1}$ such that each C_n is non-empty, for each $n \geq 1$ we have that $C_n \supseteq C_{n+1}$, but the intersection $\bigcap_{n\in\mathbb{N}} C_n$ is empty.*

The next proposition can be read as saying that compactness has a hereditary nature: if the whole space is compact, then also are its closed sets.

Proposition 3.13. *Let (X, τ_X) be a compact topological space. Then every closed subset of X is compact.*

Proof. Let C be some closed set in (X, τ_X) and $(U_i)_{i\in I}$ a cover of C by open sets in X. Then as $X\backslash C$ is open, we have that $(U_i)_{i\in I}$ together with $X\backslash C$ form an open cover of X. But X is compact, and thus there exists some finite set I_0 such that $X = (X\backslash C) \cup (\bigcup_{i\in I_0} U_i)$. But then $C \subseteq \bigcup_{i\in I_0} U_i$ and hence $(U_i)_{i\in I_0}$ form an open subcover. Thus (by Lemma 3.6) C is compact. □

Although based on the example of $\mathbb{R}^n$, one would believe that the opposite also holds - that the compact subsets would be closed - this does not hold in full generality. For example think of a set endowed with the indiscrete topology - then all subsets are compact, but the only closed sets are the empty set and the whole space. The opposite does hold, however, in Hausdorff spaces:

Proposition 3.14. *Let (X, τ_X) be a Hausdorff topological space. Then every compact subset of X is closed.*

Proof. Consider some compact subset C. To prove that C is closed, it suffices to show that for any $y \in X\backslash C$ we can find some open set contained in $X\backslash C$ and containing y. Fix such $y \in X\backslash C$. As X is Hausdorff, then for any $x \in C$ we can find disjoint open sets $x \in U_x, U_{y,x}$ with $x \in U_x$ and $y \in U_{y,x}$ separating x from y. Now, observe that $(U_x)_{x\in C}$ form an open cover of C. As C is compact, there exist $x_1, \dots, x_m$ so that $C \subseteq U_{x_1} \cup \dots \cup U_{x_m}$. If we now set $U_y =: U_{y,x_1} \cap \dots \cap U_{y,x_m}$, then U_y is an intersection of finitely many open sets and thus open. We also have that $U_y \cap (U_{x_1} \cup \dots \cup U_{x_m}) = \emptyset$ and hence $U_y \subseteq (X\backslash C)$. Hence U_y is the open set that we were looking for. □

In fact, in some textbooks (mainly French), compactness entails the Hausdorff property, i.e. a compact space is by definition always Hausdorff. The reason behind this is that compact Hausdorff spaces behave more like we would like them to behave - for example we already saw that in this case compact subsets are closed. Compact Hausdorff spaces sometimes even behave better than hoped:

Theorem 3.15. *A continuous bijection between two compact Hausdorff spaces is a homeomorphism.*

Proof. Let $(X, \tau_X), (Y, \tau_Y)$ be compact Hausdorff spaces and consider a continuous bijection $f : (X, \tau_X) \to (Y, \tau_Y)$. It suffices to prove that for any closed set C, we have that $f(C)$ is also closed. As X is compact and C is closed, then by Proposition 3.13 we have that C is compact. Further by Proposition 3.8, as C is compact and f is continuous, we see that $f(C)$ is compact. Now, finally as Y is Hausdorff and $f(C)$ is compact, from Proposition 3.14 we deduce that $f(C)$ is closed. □

Remark 3.16. *Notice that in fact the proof only used the compactness of the domain X and the Hausdorff property of the image-space Y. In other words any continuous bijection from a compact space X to a Hausdorff space Y is a homeomorphism.*

Finally, let us mention another beautiful property of compact Hausdorff spaces. Namely, in Hausdorff spaces one can not only separate distinct points with disjoint open sets, but also any two disjoint closed sets can be separate via disjoint open sets.

Definition 3.17 (Normal space)**.** *A topological space (X, τ_X) is called normal if for any two closed disjoint sets C_1, C_2 we can find open sets U_1, U_2 such that $C_1 \subseteq U_1$, $C_2 \subseteq U_2$ and $U_1 \cap U_2 = \emptyset$.*

Lemma 3.18. *Any compact Hausdorff space (X, τ_X) is also normal.*

Proof. The proof is on the exercise sheet. □

Most often one considers Hausdorff normal spaces. However, there are also normal spaces that are not Hausdorff - for example, if you consider the set $X = \{1, 2, 3, 4\}$ with the topology given by $\tau_X = \{\emptyset, X, \{1, 2\}, \{3, 4\}\}$, then X is normal yet not Hausdorff.

3.3 Compactness for finite product spaces and Heine-Borel in $\mathbb{R}^n$

Let us now see that compactness behaves well w.r.t. taking finite products. The case of infinite products is considerably more difficult, and we come back to this at the end of the section.

Proposition 3.19. *Let $(X_1, \tau_{X_1}), \ldots, (X_n, \tau_{X_n})$ be compact topological spaces. Then also $X_1 \times \cdots \times X_n$ with its product topology is compact.*

Proof. The case of general n follows from the case $n = 2$ by induction, exactly as for connectedness, so this argument is not repeated here and we will only derive the case $n = 2$. For lightness of notation, consider thus two topological spaces (X, τ_X) and (Y, τ_Y) and their product $(X \times Y, \tau_{X \times Y})$. Our aim is to use Lemma 3.7, i.e. to show that any cover of basis elements admits a finite subcover. Recall that any basis element for the product topology is given by $U \times V$ with $U \in \tau_X$ and $V \in \tau_Y$. So consider a cover $(U_i \times V_i)_{i \in I}$.

We will first use compactness of Y: Pick some $x \in X$ and consider the set I^x of $i \in I$ for which $x \in U_i$. We then have that $\{x\} \times Y \subseteq \bigcup_{i \in I^x} U_i \times V_i$. In particular $(V_i)_{i \in I^x}$ is an open cover of Y. But Y is compact, and thus there exists some finite set $I_0^x \subseteq I^x$ such that $(V_i)_{i \in I_0^x}$ is a finite subcover. Then $\{x\} \times Y \subseteq \bigcup_{i \in I_0^x} U_i \times V_i$ and moreover, if we set $W_x := \bigcap_{i \in I_0^x} U_i$ then $W_x \times Y \subseteq \bigcup_{i \in I_0^x} U_i \times V_i$.

But now each W_x is open as a finite intersection of open sets and thus $(W_x)_{x \in X}$ is an open cover of X. As X is compact, there is some finite set $X_0 \subseteq X$ such that $(W_x)_{x \in X_0}$ covers X. But then $(U_i \times V_i)_{i \in I_0^x, x \in X_0}$ is a finite subcover of $(U_i \times V_i)_{i \in I}$ covering $X \times Y$.

□

We are now basically ready to (re)prove the basic characterisation of compactness in $\mathbb{R}^n$:

Theorem 3.20. *[Heine-Borel] Consider $\mathbb{R}^n$ with its standard topology. Then a subset $K \subseteq \mathbb{R}^n$ is compact if and only if it is closed and bounded in the sense that K it is contained in some Euclidean ball $B(0, R)$.*

The only missing part is an almost unnoticeable detail: given spaces $(X_1, \tau_{X_1}), \ldots, (X_n, \tau_{X_n})$ and subsets $K_i \subseteq X_i$, then we have a priori two ways to put a topology on the product $K_1 \times \cdots \times K_n$:

- We take the product topology of the subset topologies: i.e. the product topology for (K_i, τ_{X_i, K_i}) with $i = 1 \ldots n$;
- or we take the subset topology of the product topology: i.e. we consider $K_1 \times \cdots \times K_n$ as a subset of $X_1 \times \cdots \times X_n$ and endow it with the subset topology of the product topology on $X_1 \times \cdots \times X_n$.

It could be a priori possible that these two topologies disagree, but naturally, things are set up so that they agree. Indeed, denote the first topology by τ_1 and the second by τ_2, we want to show that $\tau_1 = \tau_2$. To do this we just observe that for both of them we have a basis given by all sets of the form $(U_1 \cap K_1) \times \cdots \times (U_n \cap K_n)$ with $U_i \in \tau_{X_i}$.

Let us now prove Theorem 3.20. It is maybe interesting to note in the proof, where we use special properties of $\mathbb{R}^n$ and which ones they are. Let us hence already point out some basic results we will use:

(1) The Euclidean topology on $\mathbb{R}^n$ is the same as the product topology on the product of n copies of $\mathbb{R}$. This was on Exercise sheet 4 for $n = 2$, but you can verify that exactly the same proof works for general n, using the explicit basis for the product topology of n spaces.
(2) The Euclidean topology on $\mathbb{R}^n$ is Hausdorff (we saw that all metric spaces are Hausdorff).
(3) The standard Euclidean distance to 0 on $\mathbb{R}^n$ given by $d_E(x, 0) : (\mathbb{R}^n, \tau_E) \to (\mathbb{R}, \tau_E)$ is continuous. (This is on the exercise sheet)

Proof. Suppose K is compact. Then because $\mathbb{R}^n$ is Hausdorff, Proposition 3.14 implies that K is closed. Moreover, as mentioned just above the function $d_E(x, 0) : (\mathbb{R}^n, \tau_E) \to (\mathbb{R}, \tau_E)$ given by the Euclidean distance to the origin is continuous. Thus by the Boundedness Theorem (Theorem 3.5) we know that $d_E(x, 0)$ is bounded on K and hence K is bounded, i.e. contained in some Euclidean ball $B(0, R)$.

In the other direction suppose that K is closed and bounded. Then in particular, we have that $K \subseteq [-m, m]^n$ for some $m > 0$. Now, $[-m, m]$ is homeomorphic to $[0, 1]$ and thus is compact. Moreover, by Proposition 3.19 above, we know that then $[-m, m]^n$ is also compact, and by the small comment just above this also holds when we consider $[-m, m]^n$ with the subset topology of $\mathbb{R}^n$. But then K can be seen as a subset of the compact space $[-m, m]^n$, and thus we can use Proposition 3.13 to deduce that K is compact. $\square$

3.4 Tychonoff's theorem for infinite product spaces and the axiom of choice

Studying compactness for infinite product spaces is considerably harder. Yet - and maybe a bit surprisingly - it comes out that any product space with the product topology is compact:

Theorem 3.21 (Tychonoff's theorem)**.** *Let $((X_i, \tau_{X_i})_{i\in I})$ be any collection of compact topological spaces. Then their product $\Pi_{i\in I}X_i$ with product topology is again compact.*

This is a difficult result and the proof is not part of the course. I'm also not sure there is a clear intuition available for why it should be true. Maybe, rather than seeing Tychonoff's theorem as a statement about the properties of compactness, one should see it as a statement about arbitrary product spaces with the product topology - the product topology tries to keep the product space neatly packed together, i.e. compact.

It comes out that Tychonoff's theorem is mathematically equivalent to the Axiom of Choice, which we briefly already met. This is an axiom of set theory that does not follow from the basic, so called Zermelo-Fraenkel axioms. I believe it was a bit of a controversy whether to accept it or not in the beginning of 20th century, but by today the leaning is heavily towards accepting it. This is mainly because the statement feels just so 'obvious': it roughly says that if you have any collection of non-empty sets, then you can pick an element in each of them.

- **Axiom of choice I**: If $(X_i)_{i\in I}$ is any collection of non-empty sets, then also $\Pi_{i\in I}X_i$ is non-empty.

In fact, this axiom can be reformulated in many ways. For example, here are two reformulations:

- **Axiom of choice II**: Given any collection $(X_i)_{i\in I}$ of non-empty sets, there is a function $x : I \to \bigcup_{i\in I} X_i$ such that $x(i) \in X_i$ (this is called a choice function).
- **Axiom of choice III**: Given any collection of disjoint non-empty sets $(Y_i)_{i\in I}$, there is a set C containing one element of each Y_i.

Here the equivalence of formulations I and II just follows from the definition of the infinite product set. The equivalence with III needs a proof. You will see more on Axiom of Choice in the 3rd year course of logic, if you decide to take it.

A typical situation where it is convenient to use axiom of choice comes up in the proof of Proposition 3.19 - for each x, we are picking a finite cover of $\{x\}\times Y$ and based on this define a set W_x. This can be done without Axiom of choice for any finite number of x, however when we choose W_x for all $x \in X$, we are using Axiom of choice. However, whereas this is a typical example, it is not a very good example, as the proposition can be proved also without the Axiom of choice. But maybe this also illustrates to which extent this axiom is actually accepted.

Exercise 3.2 ($\star$ Compactness of product topology without Axiom of Choice)**.** *Prove Proposition 3.19, i.e. that the product of the compact spaces is compact in the product topology, without using Axiom of choice.*

We will use Axiom of choice in the (non-examinable) proof of Tychonoff's theorem, and it comes out that this is inevitable. In fact Tychonoff's theorem is equivalent to the Axiom of choice:

Theorem 3.22. *Tychonoff's theorem holds if and only if Axiom of Choice holds.*

So if you accept Tychonoff's theorem, you should accept the Axiom of Choice too and vice-versa. This might also be the reason why it's not so easy to put our hand on Tychonoff's theorem - it's already a step beyond the basic Zermelo-Fraenkel axioms for the set theory. Whereas we will not go in more detail with the Axiom of choice, here are some examples to think about.

Suppose you have have a collection $(X_i)_{i\in I}$ of sets X_i.

- Let $X_i = [0,1]$. Then $\Pi_{i\in I} X_i$ is non-empty because you can just define $\bar{x}$ with $\bar{x}(i) = 0$ and that's an element in the product.
- Similarly, suppose that $X_i \subseteq \mathbb{N}$ for all $i \in I$. Prove that $\Pi_{i\in I} X_i$ is non-empty without using the axiom of choice: you can define $\bar{x}(i) = \min\{y \in X_i\}$.
- We were using axiom of choice when showing that a countable union of countable sets is countable. Can you point it down?

The non-examinable proof is given at the end of the chapter.

3.5 Local compactness and compactifications

We saw that $(\mathbb{R}^n, \tau_E)$ is not compact. However, by the Heine-Borel theorem there are many compact subsets of $\mathbb{R}^n$ and they all satisfy the nice properties of compact spaces: for example, continuous real functions defined on them are bounded and take their extremal values. In fact, around each point of $(\mathbb{R}^n, \tau_E)$ we can find a subset that is compact - so at least locally the space is neatly packed. In fact there are many spaces that are locally neatly packed, but globally too large to be compact. This motivates the general definition of local compactness:

Definition 3.23 (Locally compact)**.** *Let (X, τ_X) be a topological space. If for each $x \in X$, we can find an open set U and a compact set K such that $x \in U \subseteq K$, then we say that X is locally compact.*

It is easy to see that $(\mathbb{R}^n, \tau_E)$ is locally compact. Also, any space with the discrete topology is locally compact, as we can take $U = K = \{x\}$. However,

Exercise 3.3. *Prove that $\mathbb{Q}$ with the subspace topology of $(\mathbb{R}, \tau_E)$ is not locally compact.*

So sometimes spaces are really not compact at all. However, there are still several ways to build compact spaces out of them - called compactifications.

Proposition 3.24 (One-point compactification)**.** *Let (X, τ_X) be a topological space. Fix $\infty \notin X$. Let $X' = X \cup \{\infty\}$ and $\tau_{X'}$ be the union of τ_X with the collection of sets of the form $(X \setminus K) \cup \{\infty\}$ where K is any compact and closed subset of (X, τ_X). Then $(X', \tau_{X'})$ is a topological space and that the subspace topology on $X \subseteq X'$ equals τ_X. Moreover, (X', τ') is compact and it is called the one-point compactification of (X, τ_X).*

In the example sheet you will also see that one-point compactifications goes well with homeomorphisms, and deduce from this that the one-point compactifications of $(0,1)$ and $\mathbb{R}$ are both S^1.

Proof. To see that $(X', \tau_{X'})$ is a topological space we need to verify the axioms of a topology for $\tau_{X'}$:

- X' belongs to $\tau_{X'}$ because $\emptyset$ is closed and compact. The empty set belongs to $\tau_{X'}$ as it belongs to τ_X.
- To prove the intersection property notice that as K is closed, any set of the form $(X \setminus K) \cup \{\infty\}$ can be written as $U \cup \{\infty\}$ for some $U \in \tau_X$. Thus if at least one of two open sets $V_1, V_2 \in \tau_{X'}$ is in τ_X, then their intersection is also in τ_X. Otherwise, if both $V_1, V_2 \in \tau_{X'} \setminus \tau_X$, we can write them as $(X \setminus K_1) \cup \{\infty\}$ and $(X \setminus K_2) \cup \{\infty\}$. Thus their intersection is given by $(X \setminus K_1 \cup K_2) \cup \{\infty\}$. But now we know from the exercise sheet that the union of two compact sets is again compact.
- To prove the union property, consider any collection of open sets $(U_i)_{i \in I}$ belonging to $\tau_{X'}$. If all of them are in fact in τ_X, then we are done as τ_X is a topology. Otherwise, we have at least one set U_{i_0} of the form $U_{i_0} = (X \setminus K_0) \cup \{\infty\}$. But notice that then $X \cup \{\infty\} \setminus \bigcup_{i \in I} U_i$ is a closed set that is moreover contained in K_0. As K_0 is compact, Proposition 3.13 implies that $C = X \cup \{\infty\} \setminus \bigcup_{i \in I} U_i$ is also compact. Hence $\bigcup_{i \in I} U_i$ is open in $(X', \tau_{X'})$ as desired.

The second point is clear, as for every set $U \in \tau_{X'}$ we have by definition $U \cap X \in \tau_X$.

Finally, to show that $(X', \tau_{X'})$ is compact, consider any open cover $(U_i)_{i \in I} \in \tau_{X'}$ of X'. This open cover has to also cover the point ∞, and thus there is some set U_{i_0} of the form $U_{i_0} = (X \setminus K) \cup \{\infty\}$. But now $(U_i \cap X)_{i \in I, i \neq i_0}$ is an open cover of K and as K is compact, it admits a finite subcover $(U_i \cap X)_{i \in I_0}$. Then $(U_i)_{i \in I_0}$ together with U_{i_0} is the desired finite subcover of X'. $\square$

3.6 $\star$ [Non-examinable] Proof of Tychonoff's theorem $\star$

In this non-examinable section we will discuss how to prove Tychonoff's Theorem. There are several proofs in the literature, we will provide here one going via the notion of a sub-basis. A key input in the proof is a further simplification for checking compactness.

Definition 3.25 (Subbasis)**.** *Consider a topological space (X, τ_X) and a basis τ_X^B. Then a subbasis S_X^B of τ_X^B is a subset of open sets, i.e. $S_X^B \subseteq \tau_X$ such that any element of τ_X^B can be written as a finite intersection of elements in S_X^B.*

So a sub-basis is a possible even smaller collection of subsets then a basis - whereas a basis generates the topology via unions, for a subbasis we have to take unions of all finite products of its elements. Similarly to basis, one could also just define a subbasis without any reference to a topology and then consider the topology it generates: it would be the smallest topology containing all the unions of finite intersections of sets in the subbasis. In contrary to basis, that had to satisfy some conditions, any collection of subsets (of the power set of X) is a subbasis for some topology:

Exercise 3.4. *Let X be any set and $\tilde{S}_X^B$ any subset of the power set of X. Show that arbitrary unions of finite intersections of sets of $\tilde{S}_X^B$, together with the empty set and X generate a topology τ. Can you write out the basis of τ for which $\tilde{S}_X^B$ would act as a subbasis?*

Let us consider some examples:

- Consider $(\mathbb{R}, \tau_E)$. Then the collection of half-lines $[-\infty, a)$, $(b, \infty]$ over all $a, b \in \mathbb{R}$ forms a subbasis for the basis of open intervals.
- Consider an arbitrary product space $\Pi_{i \in I} X_i$ with the product topology. We have seen that a basis is given by all sets of the form $\Pi_{i \in I} U_i$, where each $U_i \in \tau_{X_i}$ and

$U_i = X_i$ for all but finitely many $i \in I$. A subbasis is given by all sets of the form $\Pi_{i\in I} U_i$, where each $U_i \in \tau_{X_i}$ and $U_i = X_i$ for all but possibly one $i \in I$. It is more concise to write such sets just as $\pi_i^{-1}(U_i)$, where $U_i \in \tau_{X_i}$, i.e. as preimages of the projection maps.

Tychonoff's theorem will follow quite easily from the following criterion, saying that compactness can be checked only using coverings by subbasis elements. Interestingly, whereas the similar statement for just basis elements is a simple exercise, it is quite a bit more demanding in the case of the subbasis.

Theorem 3.26 (Alexander subbasis theorem). *Consider a topological space (X, τ_X) and some subbasis S_X^B. If any covering $(V_i)_{i\in I}$ of X with subbasis elements V_i admits a finite subcover, then X is compact.*

This Theorem also relies on the Axiom of Choice, or rather we will use an equivalent statement to Axiom of Choice called Zorn's lemma. Notice that the similar statement with subbasis replaced by basis did not require any form of the Axiom of Choice. When we deduce Tychonoff's theorem from Alexander subbasis theorem, we use Axiom of Choice once again. In fact, let us start by this - by proving Tychonoff's theorem assuming Alexander subbasis theorem:

Proof of Tychonff's theorem, Theorem 3.21. Consider a covering $(V_m)_{m\in M}$ of X by some subbasis elements, i.e. each V_m is given by $V_m = \pi_{i_m}^{-1}(U_{j_m})$ for some $i_m \in I$ and some $U_{j_m} \in \tau_{X_{i_m}}$. Recall that V_m is then of the form $\Pi_{i\in I} O_i$, where all $O_i \in \tau_{X_i}$, $O_i = X_i$ other than for $i = i_m$, in which case $O_i = U_{j_m}$ is possibly not equal to the full space X_i. In other words, only the coordinate i_m may be non-trivial.

Now pick some coordinate $i_0 \in I$ and consider the set $M^{i_0} \subseteq M$ for which $i_m = i_0$, i.e. the the elements of the cover that are non-trivial exactly in the co-ordinate i_0. Suppose first that the sets $(U_{j_m})_{m\in M^{i_0}}$ cover X_{i_0}. Then as X_{i_0} is compact, there is a finite set $M_0^{i_0}$ such that $X_{i_0} \subseteq \bigcup_{m\in M_0^{i_0}} U_{j_m}$. But then $X \subseteq \bigcup_{m\in M_0^{i_0}} \pi_{i_m}^{-1}(U_{j_m}) = \bigcup_{m\in M_0^{i_0}} V_m$ and thus $(V_m)_{m\in M_0^{i_0}}$ is a finite subcover of X.

Hence, if for some coordinate i_0, we have that $(U_{i_m})_{m\in M^{i_0}}$ covers X_{i_0}, then we are done. So suppose for contradiction that for all $i_0 \in I$, we have that $W_{i_0} := X_{i_0} \backslash (\bigcup_{m\in M^{i_0}} U_{j_m}) \neq \emptyset$. Then by the Axiom of Choice there is some $x \in \Pi_{i_0\in I} W_{i_0}$. But one can verify that this x is not covered by any V_m with $m \in M$, giving a contradiction. Hence our cover of subbasis elements admits a finite subcover, and thus by Alexander subbasis theorem we see that $\Pi_{i\in I} X_i$ with its product topology is compact. □

Let us finally prove Alexander subbasis theorem.

Proof of Alexander subbasis theorem, Theorem 3.26. Let S_X^B be a subbasis for τ_X^B. Suppose that any cover by subbasis elements admits a finite cover. By Lemma 3.7 to show that X is compact, it suffices to show that any cover by basis elements admits a finite subcover.

The countable case. Let us start by the countable case to better present the key ideas: i.e. consider a countable cover $\mathcal{C}_0 = (U_n)_{n\in\mathbb{N}}$ with $U_n \in \tau_X^B$ and let's try to show it has a finite subcover. Suppose for contradiction that this is not the case. We know that U_1 can be written as $U_1 = \bigcap_{j=1\ldots m} V_j^1$ for some subbasis elements $V_j^1 \in S_X^B$.

Claim 3.27. *Suppose that for all $j = 1 \dots m$, the cover given by adding V_j^1 to $\mathcal{C}_0 = (U_n)_{n\in\mathbb{N}}$ admits a finite subcover. Then $(U_n)_{n\in\mathbb{N}}$ admits a finite subcover.*

Proof. From the hypothesis it follows that for all $j = 1 \dots m$, we there is some finite subset $N_j \subseteq \mathbb{N}$ such that $X = V_j^1 \cup (\bigcup_{n\in N_j} U_n)$. Then for each j, we have that $X \backslash V_j^1 \subseteq \bigcup_{n\in N_j} U_n$ and hence by De Morgan's laws $X = (\bigcap_{j=1\dots m} V_j^1) \cup (\bigcup_{j=1\dots m} \bigcup_{n\in N_j} U_n)$. But by definition $U_1 = \bigcap_{j=1\dots m} V_j^1$ and thus in fact $(U_n)_{n\in\mathbb{N}}$ admits a finite subcover. □

Hence, as by assumption $(U_n)_{n\in\mathbb{N}}$ admits no finite subcover, there must be some $j = j(1) \in \{1, \dots, m\}$ such that we can add V_j^1 to the cover $\mathcal{C}_0$, then the resulting cover $\mathcal{C}_1$ still admits no finite subcover. We can now continue recursively - given the cover $\mathcal{C}_{n-1}$ with no finite subcover, we consider the set U_n, which can be written as $U_n = \bigcap_{j=1\dots m_n} V_j^n$. The same argument as before tells us that we can add some V_j^n (with $j = j(n)$) to the cover $\mathcal{C}_{n-1}$ and the resulting cover $\mathcal{C}_n$ has still no finite subcover. The union of all these covers, gives us a cover $\mathcal{C}_\infty$ that contains for each $n \in \mathbb{N}$ some V_j^n [11] and admits no finite subcover (otherwise some finite $\mathcal{C}_n$ would). But we know that $U_n \subseteq V_j^n$ and thus V_j^n is a covering of X by subbasis elements, thus by assumption it should have a finite cover! This gives us a contradiction and hence the initial cover $\mathcal{C}_0$ admits a finite subcover.

The general case. The general case follows a similar philosophy, but the difficulty is that if the index space is uncountable, there is no way to go through all coordinates with the described procedure and add the subbasis elements. Thus, instead we try to add them at once by constructing some sort of maximal cover:

Lemma 3.28. *Given any covering $(U_i)_{i\in I}$ of X by some basis elements $U_i \in \tau_X^B$ that admits no finite subcover, we can find a maximal cover $\mathcal{C} = (\tilde{U}_i)_{i\in\tilde{I}}$ of X by open sets that contains the cover $(U_i)_{i\in I}$, that does not admit a finite subcover and that is maximal in the following sense: if you add any other open set to this cover, then it does admit a finite subcover.*

This lemma follows directly from the Zorn's lemma, stated and explained just below the proof. We will now argue how this lemma proves the theorem. We again start with some cover $(U_i)_{i\in I}$ with $U_i \in \tau_X^B$ and suppose for contradiction that it has no finite subcover. Consider the maximal cover $\mathcal{C}$ obtained in the lemma above.

Now, pick some $i \in I$ and write again $U_i = \bigcap_{j=1\dots m} V_j^i$ with V_j^i elements in the subbasis. Similarly to the countable case, we claim that there must be some $j = j(i) \in \{1 \dots m\}$ such that $V_j^i \in \mathcal{C}$: indeed, if this is not the case then by maximality of the cover $\mathcal{C}$ adding any V_j^i to $\mathcal{C}$, we obtain a finite cover. Then, exactly as in the proof of Claim 3.27, this would imply that $\mathcal{C}$ itself has a finite subcover, giving a contradiction. Hence for each $i \in I$, we have some $V_j^i \in \mathcal{C}$ (with $j = j(i)$). But similarly to the countable case, the collection $(V_j^i)_{i\in I}$ also covers X and by the assumption has a finite subcover. Hence we derive again a contradiction with our starting assumptions - the fact that $(U_i)_{i\in I}$ admits no finite subcover.

□

It remains to describe Zorn's lemma, which we state as a theorem despite its name. To state it properly we need to introduce some vocabulary:

[11] In fact we are using a weaker form of Axiom of Choice here, called the Axiom of Dependent Choice. You may want to think why does the usual mathematical recursion not work here...

(1) A partially ordered set X is a set with a relation $\leq$ that satisfies the following natural conditions:
   - $x \leq x$;
   - if $x \leq y$ and $y \leq x$, then $x = y$;
   - if $x \leq y$ and $y \leq z$, then $x \leq z$.

   Partial here means that there might be some pairs $x, y \in X$ such that there is no relation between x and y, i.e. neither $x \leq y$ nor $y \leq x$. A prime example would be the set of subsets of some set Y with the relation $\leq$ given by set inclusion.

(2) A subset A of X such that any two elements $a, b \in A$ are related, i.e. that we have that either $a \leq b$ or $b \leq a$, is called a chain.

(3) An upper bound for a chain $A \subseteq X$ is an element $u \in X$ such that $a \leq u$ for all $a \in A$.

(4) A maximal element in X is some $x_m \in X$ such that there is no $x \in X$ distinct from x_m with $x_m \leq x$.

Theorem 3.29 (Zorn's lemma)**.** *If in a partially ordered set X any chain has an upper bound, then X has a maximal element.*

In the application to Lemma 3.28 the partially ordered set X would be the set of open coverings containing $(U_i)_{i \in I}$ and admitting no finite subcover. The relation would be the set inclusion: two coverings C_1 and C_2 would be in the relation $C_1 \leq C_2$ if and only if all the open sets in C_1 are also in C_2. I leave it for you to check how Zorn's lemma implies Lemma 3.28.

[$\star$ End of the non-examinable section $\star$]

SECTION 4

Metric spaces

So far the course has been pretty abstract - we have worked in the realm of general topological spaces. Often, however, the spaces used in other domains of mathematics or its applications come with some extra structure. An example of such extra structure is a distance function $d(x, y)$ between each pair of points, giving rise to a metric space.

Given a set X, we recall that a metric $d(x, y) : X \times X \to \mathbb{R}$ has to satisfy three properties:

(1) reflexivity: $d(x, y) = 0$ iff $x = y$;
(2) symmetry: $d(x, y) = d(y, x)$;
(3) triangle inequality: $d(x, y) + d(y, z) \geq d(x, z)$ for any $x, y, z \in X$.

Also, recall that these properties together imply that in fact $d(x, y) \geq 0$ for all $x, y \in X$.

We already saw that metric spaces induce a topology on the underlying space, called the metric topology. Thus all the topological concepts we have introduced also apply to metric spaces. In fact, many of them behave nicer in metric spaces. One of the reasons for this is that in metric spaces sequences are enough to describe many of the topological properties:

- continuity of a function f at a point x is equivalent to the statement that for any sequence $x_n \to x$, we have that $f(x_n) \to f(x)$;
- for any point in the closure of a set, there is some sequence converging to it;
- compactness is equivalent to sequential compactness;
- etc...

We will start by discussing some of the basic topological notions in the realm of metric spaces, and looking into some metric spaces of special interest.

4.1 Basic topology of metric spaces

In the following, we consider a metric space (X, d) with its metric topology τ_d induced by the basis of open balls $B(x, \delta) := \{y \in X : d(y, x) < \delta\}$. This is, we set

$$\tau_d^B := \{B(x, \delta) : x \in X, \delta > 0\}.$$

To start off, let us, ask a very natural question - when do two different metrics on the same set induce the same metric topology? Two metrics that induce the same topology are called topologically-equivalent metrics.

Lemma 4.1. *Consider a set X and two metrics d_1 and d_2. Then the metrics d_1 and d_2 are topologically equivalent if and only if for any $x \in X$ and any $r > 0$, we can find $r_1, r_2 > 0$ such that $B_{d_1}(x, r_1) \subseteq B_{d_2}(x, r)$ and $B_{d_2}(x, r_2) \subseteq B_{d_1}(x, r)$.*

Proof. Suppose that two metrics d_1, d_2 are topologically equivalent and consider some $x \in X$, some $r > 0$. The open ball $B_{d_2}(x, r)$ has to be open in (X, τ_{d_1}). Thus as open balls form a basis for (X, τ_{d_1}), there is some $\delta > 0$ and $y \in X$ such that $x \in B_{d_1}(y, \delta) \subseteq B_{d_2}(x, r)$. But then we can choose $r_1 < \frac{\delta - d(x,y)}{2}$ to obtain $B_{d_1}(x, r_1) \subseteq B_{d_2}(x, r)$. Changing the roles of d_1, d_2 we get the other inequality.

In the other direction, it suffices to prove that for any $x \in X$, $r > 0$, the ball $B_{d_1}(x, r)$ is open for the topology τ_{d_2} and the ball $B_{d_2}(x, r)$ is open for the topology τ_{d_1}. We will prove the first of these two claims.

So consider some $B_{d_1}(x,r)$. Then for any $y \in B_{d_1}(x,r)$, there is some r_y such that $B_{d_1}(y,r_y) \subseteq B_{d_1}(x,r)$. By the hypothesis, for any such y there is some $r_{y,2}$ such that $B_{d_2}(y,r_{y,2}) \subseteq B_{d_1}(y,r_y)$. But this exactly means that $B_{d_1}(x,r) = \bigcup_{y \in B_{d_1}(x,r)} B_{d_2}(y,r_{y,2})$. And thus $B_{d_1}(x,r) \in \tau_{d_2}$. □

Sometimes two metrics d_1 and d_2 defined on some set X are also called Lipschitz-equivalent if there is some $C > 0$ such that for any $x \neq y \in X$ we have that

$$C^{-1} d_2(x,y) < d_1(x,y) < C d_2(x,y).$$

Exercise 4.1. *Show that two Lipschitz-equivalent metrics are topologically equivalent. Are topologically equivalent metrics always also Lipschitz-equivalent?*

4.1.1 Sequences and continuity

It is an easy check that in metric spaces the topological notions of convergence and continuity can be reworded in terms of metric balls.

Lemma 4.2. *Consider two metric spaces (X, d_X) and (Y, d_Y). Then:*

- *A sequence $(x_n)_{n\geq 1}$ converges to x (in the sense of (X, τ_{d_x})-convergence) if and only if for every ball $B(x,\delta)$ there is some $n_\delta \in \mathbb{N}$ such that $x_n \in B(x,\delta)$ for all $n \geq n_\delta$.*
- *A function $f : (X, \tau_{d_X}) \to (Y, \tau_{d_Y})$ is continuous if and only if for each $x \in X$, and each $\epsilon > 0$, there is some $\delta > 0$ such that $f(B(x,\delta)) \subseteq B(f(x),\epsilon)$.*

Proof. The proof of this lemma is on the exercise sheet. □

Let us now verify that (as promised) in metric spaces continuous functions can be defined via sequences:

Proposition 4.3 (Continuity in terms of sequences)**.** *Consider a metric space (X, d) and any topological space (Y, τ_Y). Then a function $f : (X, \tau_d) \to (Y, \tau_Y)$ is continuous at x if and only if for any sequence $(x_n)_{n\geq 1} \to x$, we have that $(f(x_n))_{n\geq 1} \to f(x)$.*

Proof. We have seen that in any topological space the continuity of the function f at a point x implies that for any convergent sequence $(x_n)_{n\geq 1} \to x$, we have that $(f(x_n))_{n\geq 1} \to f(x)$. So it remains to prove the converse.

Suppose for contradiction that for any sequence $(x_n)_{n\geq 1} \to x$, we have that $(f(x_n))_{n\geq 1} \to f(x)$, but f is not continuous at x. Then by assumption and by the definition of continuity at the point x, there is some open set $U \in \tau_Y$ such that there is no open set in X contained in $f^{-1}(U)$.

Now, set $\delta_n := n^{-1}$. Then for any $n \in \mathbb{N}$ we have that $B(x,\delta_n) \backslash f^{-1}(U)$ is non-empty. Hence we can pick some sequence x_n such that $x_n \in B(x,\delta_n)$ and $f(x_n) \notin U$ [12]. But then the sequence $(x_n)_{n\geq 1}$ converges to x, whereas the sequence $(f(x_n))_{n\geq 1}$ cannot converge to $f(x)$, giving a contradiction. □

Remark 4.4. *It is worthwhile to inspect this proof and to ask why it works. Basically, we only used the following property of open sets around x:*

- *there are countably many open sets $(U_n)_{n\in\mathbb{N}}$ containing x, such that for any other open set U containing x, there is some $n \in \mathbb{N}$ such that $U_n \subseteq U$.*

[12] We are using the Axiom of Countable choice here (and we have used it also before...), but the standard is that this axiom is mostly used without mentioning it.

Any topological space satisfying such conditions for any point x is called first-countable - the open sets around each point are encoded in countably many sets. In first-countable spaces sequences are good enough to encode most properties.

Similarly, in metric spaces all points of the set boundary can be described in terms of sequences:

Lemma 4.5. *Consider a subset A of a metric space (X,d). Then for any $x \in \partial A$, there is a sequence $(a_n)_{n\geq 1}$ with $a_n \in A$ for all $n \in \mathbb{N}$ and $(a_n)_{n\geq 1} \to x$.*

This means that in particular in metric spaces the closure of a set A is given by adding all limits of all converging sequences in A.

Proof. On the exercise sheet. □

4.1.2 Products of metric spaces

Consider a finite number of metric spaces $(X_1, d_1), \dots, (X_n, d_n)$. The product $X_1 \times \cdots \times X_n$ has a natural metric given by

$$d_\Pi((x_1, \dots, x_n), (y_1, \dots, y_n)) := d_1(x_1, y_1) + \cdots + d_n(x_n, y_n).$$

It comes out that the underlying topology introduced by this metric is exactly the product topology:

Lemma 4.6. *Let $(X_1, d_1), \dots, (X_n, d_n)$ be metric spaces. Then the metric topology induced by the metric d_Π on $X_1 \times \cdots \times X_n$ defined just above is the product topology.*

Proof. Recall that the product topology was the smallest topology such that the projection maps $\pi_i : X_1 \times \dots X_n$ defined by $\pi_i(x_1, \dots, x_n) = x_i$ are continuous. But using the description of continuity for metric spaces obtained in Lemma 4.2 and the explicit expression of d_Π, we see that all π_i are also continuous when we use the topology induced by d_Π. Thus $\tau_{X_1 \times \cdots \times X_n} \subseteq \tau_{d_\Pi}$.

We now want to show that $\tau_{d_\Pi} \subseteq \tau_{X_1 \times \cdots \times X_n}$. To do this, consider any open ball $B_{d_\Pi}(x, \delta)$ in the τ_{d_Π} topology. It suffices to show that it is also an open set in the product topology. As usual, this follows when we show that around any point in $B_{d_\Pi}(x, \delta)$ we can find some open set of the product topology that is contained in $B_{d_\Pi}(x, \delta)$.

Hence, for any $y \in B_{d_\Pi}(x, \delta)$, define $\delta_y = \delta - d_\Pi(x, y)$. We have that $V_y := B_{d_1}(y_1, \frac{\delta_y}{2n}) \times \cdots \times B_{d_n}(y_n, \frac{\delta_y}{2n})$ is open in the product topology. But for any $z \in V_y$, we have that $d_\Pi(z, y) < \delta_y/2$. Thus by the triangle inequality $V_y \subseteq B_{d_\Pi}(x, \delta)$, and thus $B_{d_\Pi}(x, \delta) = \bigcup_{y \in B_{d_\Pi}(x,\delta)} V_y$ is open in the product topology. □

There is also a nice way to put a metric on a countable product of metric spaces so that obtained topology agrees with the product topology:

Lemma 4.7. *Consider a countable family of metric spaces $(X_n, d_n)_{n\in\mathbb{N}}$. Then the product topology on $X := \Pi_{n\in\mathbb{N}} X_n$ is induced by the following metric: for any x, y in X, writing $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$ with $x_n, y_n \in X_n$ we set*

$$d_\Pi(x, y) = \sum_{n\geq 1} \frac{d_n(x_n, y_n)}{2^n(1 + d_n(x_n, y_n))}.$$

When can we forget about the $1 + d_n(x_n, y_n)$ in the denominator?

Proof. The proof is on the exercise sheet. □

However, notice that the same thing will can not work on uncountable products:

Lemma 4.8. *The product topology on the product of uncountably many metric spaces (that are all bigger than one point) is not metrizable, i.e. there is no metric on the uncountable product space that would induce the product topology.*

The idea of the proof is the following. We noticed that in a metric space X for any point $x \in X$, there are countably many open sets $(U_n)_{n\geq 1}$ containing x such that for any other open set U containing x, we have that $U_n \subseteq U$ for some $n \geq 1$.

Moreover, observe the following: if we can find such countable collection of open sets $(U_n)_{n\geq 1}$, then we can replace U_n with elements of any basis τ_X - indeed, for any U_n there is a basis element $V_n \subseteq U_n$. We will show that in an uncountable product and the usual basis this is not possible.

Proof. The full proof will be on the example sheet as a starred exercise, i.e. the proof is not examinable. □

Remark 4.9. *This could be formulated as saying that the uncountable product of spaces that are bigger then a point with the product topology is not first-countable.*

Let us finish this small section with a little reality check.

Exercise 4.2. *Consider a metric space (X, d_X). Then the function $d(x, y : (X \times X, d_{X\times X}) \to (\mathbb{R}, \tau_E)$ is continuous.*

4.2 Some important and interesting examples of metric spaces

Let us browse through some important metric spaces for keeping in mind, some which we have met and some which we will meet more closely:

Metrics on $\mathbb{R}^n$

- The standard Euclidean metric, given by $d_E(x, y) := \|x - y\|_2$ or indeed the metric induced by any norm $\|x\|$. On exercise sheet 9, we saw that all norms are equivalent, and thus based on above we see that they all induce the same topology - the Euclidean topology.
- There are also less-standard metrics, e.g. the following distance on $\mathbb{R}^2$ is called the Paris metric:
$$d_P(x, y) := \begin{cases} \|x - y\|_2 & \text{if } x \text{ and } y \text{ lie on the same line through the origin} \\ \|x\|_2 + \|y\|_2 & \text{otherwise} \end{cases}$$
It will be an exercise to check that this defines a metric and to study its topology.

Sequence spaces

- Consider the space of all sequences $x := (x_1, x_2, \dots)$ with each $x_i \in \mathbb{R}$. As discussed, you can think of sequences as of points in $\mathbb{R}^{\mathbb{N}}$.
  - Seeing this way, the first metric one would think of is the product metric: $d_\Pi(x, y) = \sum_{n\geq 1} \frac{d_n(x_n, y_n)}{2^n(1+d_n(x_n, y_n))}$, that indeed is a metric on the whole of $\mathbb{R}^{\mathbb{N}}$.

Often, however, this is not the most useful metric, and rather we might want to use:

(1) The sup(remum) metric: $d_\infty(x, y) := \sup_{n\in\mathbb{N}} |x_n - y_n|$.

(2) Or for example the l^2 metric: $d_2(x, y) := \sqrt{\sum_{n\in\mathbb{N}}(x_n - y_n)^2}$.

In both cases, there would be some pairs of elements of $\mathbb{R}^{\mathbb{N}}$ for which the metric is not defined, so in order to really define a metric space with such metrics we need to constrain the set of sequences we need to consider. For example, in the case of d_∞ metric, the distance between $x = (1, 2, 3, \dots)$ and $y = (0, 0, \dots)$ is not well-defined. Instead, one would consider d_∞ on the set of sequence x with $\sup_{i\geq 1} |x_i| < \infty$.

Function spaces

- Similarly to sequence spaces, one could also consider function spaces on, say, $I = [0, 1]$. As mentioned, you could think that this means that instead of $\mathbb{R}^{\mathbb{N}}$ we are now considering the product space $\mathbb{R}^{[0,1]}$. As we saw, in this case there is no natural product metric on this space. So one works directly with smaller subsets:
  (1) For example, we can consider the set of bounded functions, $\mathcal{B}([0, 1])$ with the sup(remum) metric $d_\infty(f, g) = \sup_{x\in[0,1]} d_E(f(x), f(y))$. In fact, we will see that this sup metric behaves even nicer when restricted to only the set of continuous functions $\mathcal{C}([0, 1], \mathbb{R})$.
  (2) The set of continuous functions becomes a metric space also under other distances - for example we could consider the distance $d_1(f, g) = \int_0^1 |f(x)-g(x)|dx$.
  (3) More generally, we could consider for any topological space (X, τ_X) the set of continuous functions $\mathcal{C}(X, \mathbb{R})$ with some metric. As we will see, often the sup metric is the useful one. As a word of caution recall the example of $\mathcal{C}((0, 1), \mathbb{R})$, where we had to tweak the metric d_∞ on $\mathbb{R}$ to obtain a metric space: for example we could define $\widetilde{d}_\infty(f, g) = \min(d_\infty(x, y), 1)$.
  (4) Even more generally, if (Y, d_Y) is any metric space, we could consider the set of continuous functions $\mathcal{C}(X, Y)$ with the uniform or sup-metric defined by $d_\infty(f, g) = \sup_{x\in X} d(f(x), f(y))$. Again, this sometimes defines a metric only after tweaking it a bit. For example, an interesting case would be $X = ([0, 1], \tau_E)$ and Y any metric spaces. Then $\mathcal{C}(X, Y)$ corresponds to the set of paths in Y and it could be useful to look at it as a metric space.

A variety of spaces

- Finally, metric spaces appear also in many different contexts. Let us mention two here:
- The p-adic metric: the whole numbers $\mathbb{Z}$ are naturally a subspace of $(\mathbb{R}, \tau_E)$ and thus can be considered with the Euclidean metric. It comes out that from the point of view of number theory, another useful metric is the $p-$adic metric. This is defined for any prime number p as follows: $d_p(k, n) = m^{-1}$ if $|k - n| = p^{m-1}q$ with q not divisible by p. In words, this metric encodes distance between whole numbers in terms of divisibility by p. We will see that it is also interesting to consider a variant of this metric on the set of rationals.
- The Hausdorff metric on the set of closed subsets of a metric space (X, d). For any closed set C, we can define its $\epsilon-$neighbourhood by $C_\epsilon := \{x \in X : d(x, C) \leq \epsilon\}$.

We can now define the Hausdorff distance between two closed sets C, D, denoted $d_H(C, D)$, as follows:

$$d_H(C,D) := \inf_{\epsilon>0}\{C \subseteq D_\epsilon \text{ and } D \subseteq C_\epsilon\}.$$

It will be an exercise to check that (under certain conditions) this indeed defines a metric.

- In fact also the set of all probability measures on $([0,1], \mathcal{F}_{\mathbb{R}})$ will be a metric space! We will come back to this some time later in the course.

4.3 Compactness in metric spaces

As already mentioned before, in metric spaces compactness and sequential compactness are equivalent:

Theorem 4.10. *A metric space (X, d) is compact if and only if it is sequentially compact.*

Our aim is to now prove Theorem 4.10. We will prove it in two parts. The easier step is:

Proposition 4.11. *Every compact metric space is sequentially compact.*

Here, the proof idea is very natural - we will just have to find a convergent subsequence. We do it by identifying smaller and smaller closed subsets containing an infinite subsequence. The argument is sometimes called Lion-hunting - in some sense we are making the perimeter around the Lion smaller and smaller until we track it down. To make it lighter to write, be record a simple observation:

Lemma 4.12. *In a compact space for every $\epsilon > 0$, one can finitely many points $x_1, \dots, x_n$ such that balls $(B(x_i, \epsilon))_{i\leq n}$ cover the whole space.*

Proof. This follows directly from compactness: the collection $(B(x,\epsilon))_{x\in X}$ is an open cover and by compactness has a finite subcover. ☐

One way to intepret this simple lemma is to say that compact metric spaces satisfy a stronger form of boundedness: for any ϵ we can cover the whole space with a finite number of $\epsilon-$balls.

Proof of Proposition 4.11. Suppose that (X, d) is compact and let $(x_n)_{n\geq 1}$ be any sequence in X. By the lemma above, we find finitely many balls of radius $1/2$ covering X. Now, as there are finitely many balls, some ball $B(z_1, 1/2)$ has to contain infinitely many elements of x_n. We pick the smallest n for which $x_n \in B(z_1, 1/2)$ and define $y_1 := x_n$.

Now, we can also find finitely many balls of radius $1/4$ covering the space, and in particular finitely many of them cover $B(z_1, 1/2)$. For one of them, say $B(z_2, 1/4)$, we have that $B(z_1, 1/2) \cap B(z_2, 1/4)$ must again contain infinitely many elements of the sequence x_n. We again pick such an x_n with the smallest index n and set $y_2 = x_n$. Continue this way to obtain a sequence C_n nested closed sets $C_n := cl(\cap_{m=1\dots n} B(z_m, 2^{-m}))$ and a sequence $(y_n)_{n\geq 1}$ with $y_n \in C_n$.

But we have seen that if (X, d) is compact, then such an intersection of non-empty closed sets $\bigcap_{n\geq 1} C_n$ is non-empty. Pick any element y in this intersection[13]. It is now easy to check that $y_n \to y$ and thus (X, d) is sequentially compact. ☐

[13] In fact there can be only one, but it is not relevant here

Remark 4.13 ($\star$)**.** *It might be interesting to ask yourself, what are the properties we used about metric spaces here?*

The other direction is slightly trickier - how to convert existence of convergent sequences to existence of finite covers? A possible middle-man is provided by the following lemma:

Lemma 4.14 (Lebesgue number lemma)**.** *Let (X,d) be a sequentially compact space. Then for every open cover $(U_i)_{i\in I}$ there exists some $\epsilon > 0$ such that for each $x \in X$, there is some U_i such that $B(x,\epsilon) \subseteq U_i$. The largest such $\epsilon > 0$ for the given cover, is called the Lebesgue number of the cover.*

This is a lemma that clearly directs us towards compactness - firstly, we are dealing already with open covers, and secondly the lemma says that we cannot cover the space without somehow evenly covering each point.

Proof. Suppose that the Lebesgue number is not positive for some cover U_i. Then for this cover, and for any n we can find some sequence $(x_n)_{n\geq 1}$ such that each $B(x_n, 1/n)$ is not contained in any U_i [14].

As X is sequentially compact, then there is some convergent subsequence $(x_{n_k})_{k\geq 1}$ converging to some $x \in X$. Then as $(U_i)_{i\in I}$ covers X, there is some U_i with $x \in U_i$ and in particular as U_i is open and open balls form a basis, there is some $\delta > 0$ such that $B(x,\delta) \subseteq U_i$. Now, we can choose k large enough so that on the on hand $n_k^{-1} \leq \delta/3$ and on the other hand $d(x_{n_k}, x) < \delta/3$. But then by the triangle inequality also $B(x_{n_k}, 1/n_k) \subseteq B(x,\delta) \subseteq U_i$, giving a contradiction with the assumption that $B(x_n, 1/n)$ is not contained in any U_i. ☐

We are now ready to prove the other direction of compactness vs sequential compactness:

Proposition 4.15. *Every sequentially compact metric space is compact.*

Proof. Let (X,d) be sequentially compact space. Consider an open cover $(U_i)_{i\in I}$ of X. By the Lebesgue number lemma, there is some $\epsilon > 0$ such that for each $x \in X$, there is some $i = i(x)$ with $B(x,\epsilon) \subseteq U_i$. In particular, it suffices to prove that X can be covered by finitely many balls $B(x_1,\epsilon), \dots, B(x_n,\epsilon)$ as then the sets $U_{i(x_1)}, \dots, U_{i(x_n)}$ form a finite subcover.

Suppose for contradiction that this is not the case, i.e. that no finite subset of these balls covers the space. Then we can inductively pick a sequence $x_1, x_2, \dots$ such that $d(x_n, x_m) > \epsilon$ for any $m < n$. Indeed, we pick some point x_1 and having picked $x_1, \dots, x_{n-1}$ we pick $x_n \in X \setminus \bigcup_{i=1\dots n-1} B(x_i,\epsilon)$, which is by assumption non-empty. [15]

Now, by assumption (X,d) is sequentially compact. Thus we obtain the contradiction and prove the proposition by showing that

Claim 4.16. *Suppose that $(x_n)_{n\geq 1}$ is a sequence in some metric space (X,d) such that $d(x_n,x_m) > \epsilon$, whenever $n \neq m$. Then $(x_n)_{n\geq 1}$ admits no convergent subsequence.*

Proof of claim. Indeed, suppose for contradiction that some subsequence $(x_{n_k})_{k\geq 1}$ did converge to some x. Then from some point onwards $d(x_{n_k}, x) < \epsilon/2$ and $d(x_{n_{k+1}}, x) < \epsilon/2$, implying by triangle inequality that $d(x_{n_k}, x_{n_{k+1}}) < \epsilon$, contradicting the fact that $d(x_n, x_m) > \epsilon$. Thus $(x_n)_{n\geq 1}$ admits no convergent subsequence. ☐

☐

[14] We are again using countable axiom of choice here, but as already pointed out before - the standard is usually not to even mention it

[15] We are using dependent choice here.

4.4 Cauchy sequences and completeness

You have already met Cauchy sequences in real analysis. Cauchy was an engineer who decided that mathematics is more beautiful than engineering and that kings are better than democracy. Cauchy sequences make sense in general metric spaces.

Definition 4.17 (Cauchy sequence)**.** *Let (X, d) be a metric space. A sequence $(x_n)_{n\geq 1}$ is called Cauchy if for any $\epsilon > 0$, there is some $n_\epsilon \in \mathbb{N}$ such that for any $n_1, n_2 \geq n_\epsilon$ we have that $d(x_{n_1}, x_{n_2}) < \epsilon$.*

It is easy to verify that each convergent sequence is Cauchy:

Lemma 4.18. *Let (X, d) be a metric space. If a sequence $(x_n)_{n\geq 1}$ converges, then it is Cauchy.*

Proof. Let x be the limit of $(x_n)_{n\geq 1}$. Then by definition of convergence, for any $\epsilon > 0$ there exists some n_ϵ such that for all $n \geq n_\epsilon$ we have that $d(x_n, x) < \epsilon/2$. But then by the triangle inequality, for any $n, m \geq n_\epsilon$, we have that $d(x_n, x_m) < \epsilon$ and thus $(x_n)_{n\geq 1}$ is Cauchy. $\square$

As the elements in a Cauchy sequence get closer and closer together, so in some sense the sequence stabilizes It would be reasonable to guess that such a sequence might converge. Whereas this is indeed true in $\mathbb{R}^n$ with its Euclidean metric, it is not true in general. For example,

- consider $\mathbb{Q}$ with its standard metric and look at the sequence x_n given by stopping the decimal expansion of $\sqrt{2}$ at its n-th digit. Then $x_n \in \mathbb{Q}$ and x_n is Cauchy, yet it doesn't converge in $\mathbb{Q}$.
- Or, consider $(0, 1)$ with its standard metric. Then the sequence $x_n = n^{-1}$ is Cauchy, yet does not converge in $(0, 1)$.

However, it is often true for spaces that one wants to work with, so such spaces have earned their own name:

Definition 4.19 (Complete metric space)**.** *A metric space (X, d) is called complete if every Cauchy sequence converges.*

Thus paraphrasing, we just saw that neither $\mathbb{Q}$ nor $(0, 1)$ with their usual metrics are complete. For a simple example of a complete space to keep in mind, consider any set with the discrete metric: this is a complete metric space as any Cauchy sequence is eventually constant (Why?). There are of course also many more interesting spaces:

Exercise 4.3. *Prove that $C([0, 1], \mathbb{R})$ with the sup-norm is complete.*

Let us consider some basic properties of complete spaces (X, d):

- Let K be a closed subset. Then (K, d) is also a complete metric space - indeed, if you take a Cauchy sequence in K, then it is also Cauchy in X and thus converges. But all the limit points of a closed set belong to the set, thus the limit is in K and hence (K, d) is complete.
- Suppose that $(X_1, d_1), \dots, (X_n, d_n)$ are complete - then also $X_1 \times \cdots \times X_n$ are complete with the product metric. (This is on the exercise sheet.)
- But what about continuous images of complete metric spaces? E.g. if $f : (X, d) \to (Y, d_Y)$ is continuous, is $f(X)$ complete? Maybe a somewhat silly way to see that

this is not the case is to consider $X = (0,1)$ with d given by the discrete metric and $(Y, d_Y) = ((0,1), \tau_E)$. Then we know that the former is complete, the latter is not complete and the identity map is continuous as all maps from a discrete space are continuous. You should think of different examples!

You have also already seen in the first year that $\mathbb{R}^n$ with its standard metric is complete. We will deduce it from a more general statement very soon.

4.4.1 Completeness vs compactness

Both completeness and sequential compactness are about convergent sequences. So it is reasonable to expect a relation between the two:

Proposition 4.20. *Every sequentially compact metric space is complete. More generally, if for a metric space (X, d) and some point x_0 we have that all closed balls $\overline{B(x_0, R)}$ are sequentially compact, then (X, d) is also complete.*

Proof. We will directly prove the more general statement. Let x_0 be the point around which we have sequentially compact balls. Let $(x_n)_{n\geq 1}$ be a Cauchy sequence. Then there is some n_0 such that for all $n, m \geq n_0$ we have that $d(x_n, x_m) < 1$. In particular, if we denote $R_0 = d(x_0, x_{n_0})$ then for all $n \geq n_0$ we have that $x_n \in \overline{B(x_0, R_0 + 1)}$. As the space $(\overline{B(x_0, R_0 + 1)}, d)$ is sequentially compact, there is some subsequence $(x_{n_k})_{k\geq 1}$ of $(x_n)_{n\geq n_0}$ that converges to some $x \in \overline{B(x_0, R_0 + 1)}$.

We claim that then the whole sequence converges to the same x:

Claim 4.21. *Let (X, d) be a metric space and $(x_n)_{n\geq 1}$ a Cauchy sequence such that some subsequence $(x_{n_k})_{k\geq 1}$ converges to x. Then in fact $(x_n)_{n\geq 1}$ converges to x.*

Proof of claim. By convergence of $(x_{n_k})_{k\geq 1}$ to x we can pick $m_1 \in \mathbb{N}$ such that for all $n_k \geq m_1$ we have that $d(x_{n_k}, x) < \epsilon/2$. By the Cauchy property of $(x_n)_{n\geq 1}$ we can also pick $m_2 \in \mathbb{N}$ so that for all $n, m \geq m_2$ we have that $d(x_n, x_m) < \epsilon/2$. Now pick some $n_k \geq \max(m_1, m_2)$. Then for all $m \geq \max(m_1, m_2)$ we have by the triangle inequality that

$$d(x, x_m) \leq d(x, x_{n_k}) + d(x_{n_k}, x_m) < \epsilon$$

and thus indeed $(x_n)_{n\geq 1}$ converges to x. □

□

It follows from this result and the fact that compactness and sequential compactness agree in metric spaces, that $\mathbb{R}^n$ is complete. Indeed, we have

Corollary 4.22. *$\mathbb{R}^n$ with its standard metric is complete.*

Proof. We know that for any $R > 0$, the closed ball $\overline{B(0, R)} \subseteq \mathbb{R}^n$ is compact. By the Theorem 4.10 these balls are thus also sequentially compact. But then by Proposition 4.20 we deduce that $\mathbb{R}^n$ is complete. □

Let us now look for a converse of Proposition 4.20. Notice that the converse is not true in a naive sense: for example $\mathbb{R}$ is complete, but not sequentially compact - there are sequences like $1, 2, 3$ where elements can stay far and thus don't have to converge.

Also, just adding say boundedness of the space would not help: any infinite space with the discrete metric is complete and bounded, yet not compact. So what could we say?

Recall that we noticed in Lemma 4.12 that compact metric spaces satisfy a stronger form of boundedness: for any ϵ we can cover the whole space with a finite number of $\epsilon-$balls. Such a property has earned its own name:

Definition 4.23 (Totally bounded metric space)**.** *A metric space (X, d) is called totally bounded if for every $\epsilon > 0$, we can find a finite number of balls of radius ϵ covering X.*

It comes out that this is the missing bit. In fact, again total boundedness on its own, i.e. without completeness does not imply compactness - for example $\mathbb{Q} \cap [0, 1]$ with the usual metric is totally bounded, yet we have seen that it is not compact. However, putting the two - completeness and total boundedness - together, gives a characterisation of compactness for metric spaces:

Theorem 4.24 (Compact = complete + totally bounded)**.** *A metric space is (sequentially) compact if and only if it is complete and totally bounded.*

We have already proved that compactness implies complete and totally bounded. Only the other direction is missing. However, if you inspect the proof of Proposition 4.11, you notice that total boundedness sets us up well for the Lion hunting argument - i.e. to track down a convergent subsequence. Indeed, we can basically re-use the two first paragraphs of that proof, and only modify the last step of identifying the actual limit: instead of the non-empty intersection property for closed sets in compact spaces, we will now use completeness.

Proof. Using the equivalence of compactness and sequential compactness, we see that Proposition 4.20 and Lemma 4.12 provide one direction - they show that compactness implies complete and totally bounded.

So suppose (X, d) is a complete and totally bounded metric space. Let us show that it is sequentially compact, and thus compact.

Suppose that (X, d) is totally bounded and let $(x_n)_{n\geq 1}$ be any sequence in X. Then we can find finitely many balls of radius $1/2$ covering X. Now, as there are finitely many balls, some ball $B(z_1, 1/2)$ has to contain infinitely many elements of x_n, pick the element x_n with smallest n and call it y_1.

Now, we can also find finitely many balls of radius $1/4$ and in particular finitely many of them cover $(B(z_1, 1/2)$. For one of them, say $B(z_2, 1/4)$, we have that $B(z_1, 1/2) \cap B(z_2, 1/4)$ must again contain infinitely many elements of the sequence x_n. We pick one with the smallest index n and call it y_2. Continue this way to obtain a sequence C_n nested closed sets $C_n := cl(\cap_{m=1\dots n} B(z_m, 2^{-m}))$ and a sequence $(y_n)_{n\geq 1}$ with $y_n \in C_n$.

But now, notice that for any $y_n, y_m \in C_{n_0}$ we have that $d(y_n, y_m) \leq 2^{-n_0}$. Thus the sequence y_n is Cauchy! But by completeness it hence converges, and again we have found the convergent subsequence.

□

4.5 Continuous functions with values in metric spaces

We will now concentrate our study on the space of continuous functions defined on some topological space (X, τ_X) and taking values in a metric space (Y, d). The space of such functions is denoted $C(X, Y)$.

We will endow $C(X, Y)$ with the uniform metric (sometimes also called sup metric): for any $f, g \in C(X, Y)$ we want to set $d_\infty(f, g) := \sup_{x\in X} d(f(x), g(x))$. We have seen that this defines a metric space only on a subset - if X is not compact, there might be functions for which the distance becomes infinite. Thus we define $\overline{d}_\infty(f,g) = \min(d_\infty(f,g), 1)$ and call it for clarity sometimes the truncated sup metric. Notice that this metric is well-defined on the larger space Y^X and in fact $(Y^X, \overline{d}_\infty)$ also defines a metric space.

Theorem 4.25. *Let (X, τ_X) be a topological space. If (Y, d) is complete, then so is $C(X,Y)$ with the truncated sup metric $\overline{d}_\infty$. If X is compact, the same conclusion holds for the uniform metric d_∞.*

In particular it says that uniform limits of continuous functions are continuous in quite an abstract setting. We will cut the proof into two independent results. The first says that the product space is complete:

Lemma 4.26. *Let (X, τ_X) be a topological space. If (Y, d) is complete, then so is the space of all functions $f : X \to Y$, i.e. the product space Y^X with the (truncated) uniform metric.*

The second says that any uniform limit of continuous functions is continuous:

Proposition 4.27. *Let (X, τ_X) be a topological space, (Y, d) a metric space and $(Y^X, \overline{d}_\infty)$ the related function space. Consider a sequence $(f_n)_{n\geq 1}$ of continuous functions, i.e. functions in $C(X,Y)$. Suppose that $(f_n)_{n\geq 1}$ converges to $f \in Y^X$ in the $\overline{d}_\infty$ metric. Then in fact $f \in C(X,Y)$.*

The theorem follows directly from these two results:

Proof of Theorem 4.25. Consider a Cauchy sequence $(f_n)_{n\geq 1}$ in $(C(X,Y), \overline{d}_\infty)$. Then $(f_n)_{n\geq 1}$ is also Cauchy in $(Y^X, \overline{d}_\infty)$. As the latter is complete by the lemma above, then $(f_n)_{n\geq 1}$ converges to some $f \in Y^X$. But now by the proposition above in fact $f \in C(X,Y)$ and thus the theorem follows for the $\overline{d}_\infty$ metric. The statement for compact X then follows from the exercise stated just below. □

Exercise 4.4. *Suppose that (X, τ_X) is compact. Recall that then $(C(X,Y), d_\infty)$ is also metric space without truncating the metric. Show that d_∞ and $\overline{d}_\infty$ are topologically equivalent metrics on $C(X,Y)$, but not necessarily Lipschitz-equivalent. Prove that nevertheless $(C(X,Y), d_\infty)$ is complete if and only if $(C(X,Y), \overline{d}_\infty)$ is complete.*

Let us now prove the lemma and the proposition, starting from the lemma:

Proof of Lemma 4.26. Consider a Cauchy sequence $(f_n)_{n\geq 1}$ in $(Y^X, \overline{d}_\infty)$. As then for each $x \in X$, the sequence $(f_n(x))_{n\geq 1}$ is Cauchy and Y is complete, then there is some $y_x \in Y$ such that $(f_n(x))_{n\geq 1}$ converges to y_x. Set $f(x) := y_x$, defining a function $f : X \to Y$, i.e. an element of Y^X. It remains to check that $(f_n)_{n\geq 1}$ converges to f in the $\overline{d}_\infty$ metric. Fix $n \in \mathbb{N}$. Then for any $x \in X$ and any $m \geq n$ we have that

$$d(f_n(x), f(x)) \leq d(f_n(x), f_m(x)) + d(f_m(x), f(x)).$$

As $f_m(x)$ converges to $f(x)$ (with different speed for different x), we have that $d(f_m(x), f(x)) \to 0$ as $m \to \infty$. In particular, for every $x \in X$ and every $\delta > 0$ we can find $m_0 = m_0(x)$ such that $d(f_m(x), f(x)) < \delta$ for all $m \geq m_0$. Thus

$$d(f_n(x), f(x)) \leq d(f_n(x), f_m(x)) + \delta \leq \sup_{m\geq m_0} d(f_n(x), f_m(x)) + \delta.$$

Letting now $\delta \to 0$, we obtain that for any $x \in X$

$$d(f_n(x), f(x)) \leq \limsup_m d(f_n(x), f_m(x)).$$

For n large enough (so that $d_\infty(f_n, f_m) = \overline{d}_\infty(f_n, f_m)$ for all $m \geq n$), the right hand side is further bounded by $\limsup_m \overline{d}_\infty(f_n, f_m)$. Hence also

$$\overline{d}_\infty(f_n, f) \leq \sup_{x \in X} d(f_n(x), f(x)) \leq \limsup_m \overline{d}_\infty(f_n, f_m),$$

which goes to zero as $(f_n)_{n \geq 1}$ is by assumption Cauchy. □

It remains to argue that uniform limits of continuous functions are continuous.

Proof of Proposition 4.27. Consider a sequence $(f_n)_{n \geq 1}$ of functions in $C(X, Y)$ converging in $\overline{d}_\infty$ metric to some $f \in Y^X$. We want to show that $f \in C(X, Y)$. To do this fix some $x_0 \in X$ and consider some open ball $B(f(x_0), \epsilon)$ around $f(x_0)$. Then as $(f_n)_{n \geq 1}$ converges to f in $\overline{d}_\infty$-distance, we can pick n large enough so that for any $x \in X$, we have that $d(f_n(x), f(x)) < \epsilon/3$. But now f_n is continuous, so there is some open set $U_0 \in \tau_X$ containing x_0 such that $f_n(U_0) \subseteq B(f_n(x_0), \epsilon/3)$.

We claim that then $f(U_0) \subseteq B(f(x_0), \epsilon)$ as well: indeed, for any $x \in U_0$, by the choice of U_0 we have $d(f_n(x), f_n(x_0)) < \epsilon/3$. Thus by the triangle inequality

$$d(f(x), f(x_0)) < d(f(x), f_n(x)) + d(f_n(x), f_n(x_0)) + d(f_n(x_0), f(x_0)) < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon,$$

and the claim follows. Hence f is continuous at x_0 and as x_0 was arbitrary, in fact $f \in C(X, Y)$. □

4.6 Compact subsets of $C(X, Y)$

Our next aim is to understand compact subsets of $C(X, Y)$ with the uniform metric d_∞. In other words - given a sequence of functions, when can we extract a convergent subsequence? We have seen that compact subsets of a metric space have to be complete and totally bounded. However, these criteria are a priori not easy to understand in terms of conditions on the functions themselves.

For example, one could ask

- Does the sequence $f_n = \sin(nx)$ on $[0, 1]$ have a convergent sequence?
- Is the subset of differentiable functions on $[0, 1]$ compact? But what if we ask in addition that $|f(x)| \leq 1$? Also that $|f'(x)| \leq 1$?
- Do uniformly bounded holomorphic functions have convergent subsequences?

We will find a pretty useful criteria to answer these questions in the setting where (X, τ_X) is a compact topological space and (Y, d) a metric space such that all closed bounded balls are compact. [16] Recall, that in particular this means that (Y, d) is complete.

Let us start from understanding how functions in $C(X, Y)$ look like in general:

Lemma 4.28. *Let (X, τ_X) be compact and (Y, d) be a metric space. Let $f \in C(X, Y)$. Then*

(1) f is bounded, i.e. there is some $C > 0$ and some $y \in Y$ such that $d(f(x), y) < C$ for all $x \in X$.

[16] One could generalize a bit further, but for us this is quite enough.

(2) f is uniformly continuous, i.e. for all $\epsilon > 0$, we can find a finite open cover $U_1, \ldots, U_n$ of X such that $d((f(x), f(y)) < \epsilon$ whenever $x, y \in U_i$.

Remark 4.29. *It is probably instructive to think why this definition of uniform continuity of general topological spaces is equivalent to the following usual definition for functions in $C([0,1], \mathbb{R})$: $f \in C([0,1], \mathbb{R})$ is uniformly continuous if for all $\epsilon > 0$ we can find $\delta > 0$ such that whenever $x, y \in [0,1]$ are such that $|x - y| < \delta$, then $|f(x) - f(y)| < \epsilon$.*

Proof. Fix some $y \in Y$. As (X, τ_X) is compact and we have seen that $d(f(x), y)$ is a continuous function to $(\mathbb{R}, \tau_E)$, the image of f is compact and thus in particular bounded.

For the second claim consider the open sets $U_y := f^{-1}(B(y, \epsilon/2))$ with $y \in f(X)$. The sets $(U_y)_{y \in f(X)}$ form an open cover of X and thus by compactness of X there are finitely many $y_1, \ldots, y_n$ such that U_{y_i} with $i = 1 \ldots n$ cover X. But by definition, if x_1, x_2 belong to the same U_y, then $d(f(x_1), f(x_2)) < \epsilon$. Thus uniform continuity follows.

□

It comes out that if we now consider a sequence of continuous functions converging to a continuous function, then these two properties are moreover uniform. Let us introduce two definitions to formulate this:

Definition 4.30 (Uniformly bounded functions)**.** *Let $A \subseteq C(X, Y)$ be a subset of continuous functions from (X, τ_X) to (Y, d). Then A is called uniformly bounded, if*

- *there is $C > 0$, and $y \in Y$ such that $d(f(x), y) < C$ for all $x \in X$ and for all $f \in A$.*

Definition 4.31 (Equicontinuous functions)**.** *Let $A \subseteq C(X, Y)$ be a subset of continuous functions from (X, τ_X) to (Y, d). Then A is called equicontinuous, if*

- *for any $\epsilon > 0$, we can find a finite covering of X with open sets $U_1, \ldots, U_m$ such that for all $f \in A$, whenever $x, y \in U_i$ for some $i \in \{1, \ldots, m\}$, we have $d(f(x), f(y)) < \epsilon$.*

Notice that a set consisting of one single continuous function is both uniformly bounded and equicontinuous. Moreover, every finite union of uniformly bounded or equicontinuous sets is again uniformly bounded or equicontinuous, respectively:

Lemma 4.32. *If $A_1, \ldots, A_n$ are uniformly bounded subsets of $C(X, Y)$, then also $A = \bigcup_{i=1\ldots n} A_i$ is uniformly bounded. Similarly, if $A_1, \ldots, A_n$ are equicontinuous subsets of $C(X, Y)$, then also $A = \bigcup_{i=1\ldots n} A_i$ is equicontinuous.*

Proof. By induction it suffices to prove the case for $n = 2$. So let A_1, A_2 be uniformly bounded. Then there exist C_1, C_2 and y_1, y_2 such that $d(f_1(x), y_1) < C_1$ for all $x \in X$ and $f_1 \in A_1$; similarly $d(f_2(x), y_2) < C_2$ for all $x \in X$, $f_2 \in A_2$. Now, consider any $f \in A_1 \cup A_2$. We claim that then $d(f(x), y_1) < d(y_1, y_2) + C_1 + C_2$. This is true if $f \in A_1$ and if $f \in A_2$, then by the triangle inequality

$$d(f(x), y_1) \leq d(f(x), y_2) + d(y_2, y_1) \leq C_2 + d(y_1, y_2),$$

and the claim follows.

Let us now see that $A_1 \cup A_2$ is equicontinuous. Similarly, we can choose $U_1, \ldots, U_m \in \tau_X$ such that whenever $x, y \in U_i$, and $f_1 \in A_1$ then $d(f_1(x), f_1(y)) < \epsilon$; similarly we can choose $V_1, \ldots, V_k \in \tau_X$ such that whenever $x, y \in V_j$ and $f_2 \in A_1$ then $d(f_2(x), f_2(y)) < \epsilon$. But now notice that $W_{i,j} = U_i \cap V_j$ also form a finite open cover, and whenever $x, y \in W_{i,j}$ we have that for every $f \in A_1 \cup A_2$, $d(f(x), f(y)) < \epsilon$. Thus the equicontinuity follows.

□

Moreover, if we have a convergent sequence of functions, then this set of functions has to be both equicontinuous and uniformly bounded:

Proposition 4.33. *Let (X, τ_X) be compact and (Y, d) be a metric space. Let $(f_n)_{n\geq 1}$ be a sequence of functions in $(C(X, Y), d_\infty)$ converging to some $f \in C(X, Y)$. Then the set $A = \{f\} \cup \{f_n : n \in \mathbb{N}\}$ is equicontinuous and uniformly bounded.*

The proofs of the two statements have the following intuition: convergence in the uniform norm means that for n large enough the functions f_n look extremely similar to f and thus the boundedness and uniform continuity of f should imply those of f_n for n large enough, say $n \geq n_0$. Thus we get one equicontinuous and uniformly bounded set with all f_n with $n \geq n_0$. Thereafter, we are left with a finite number of f_n for $n < n_0$, but we already know that finite unions of uniformly bounded and equicontinuous sets are again uniformly bounded and equicontinuous.

Proof. As f is continuous, by Lemma 4.28 we can find $C > 0$, $y \in Y$ such that $d(f(x), y) < C$ for all $x \in X$. Moreover, for every $\epsilon > 0$, we can find $U_1, \ldots, U_m$ of X such that $d((f(x), f(y)) < \epsilon/3$ whenever $x, y \in U_i$.

As $(f_n)_{n\geq 1}$ converges to f in d_∞-metric, there is some $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ we have that $d_\infty(f_n, f) < \epsilon/3$. In particular for all $n \geq n_0$ and all $x, y \in X$ we have that

$$d(f_n(x), f_n(y)) < d(f_n(x), f(x)) + d(f(x), f(y)) + d(f(y), f_n(y) < 2\epsilon/3 + d(f(x), f(y)). \tag{4.1}$$

Let us see how this estimate implies uniform boundedness and equicontinuity.

Firstly, we see that for all $n \geq n_0$ and all $x \in X$ we have that $d(f_n(x), y) < C + 1$. Thus the set $A_0 := \{f\} \cup \{f_n : n > n_0\}$ is uniformly bounded. But then $A = A_0 \cup \{f_1\} \cup \cdots \cup \{f_{n_0}\}$ is a finite union of uniformly bounded sets and thus is uniformly bounded by the lemma just above.

Second, notice that for all $n \geq n_0$ and $x, y \in U_i$ with $i = 1 \ldots m$, we have from Equation (4.1) that $d(f_n(x), f_n(y)) < \epsilon$. Hence the set A_0 is also equicontinuous, and thus similarly by the lemma above A is equicontinuous as a finite union of equicontinuous sets. $\square$

Corollary 4.34. *Let (X, τ_X) be compact and (Y, d) be a metric space such that all bounded closed balls are compact. Then a compact subset A of $(C(X, Y), d_\infty)$ is uniformly bounded and equicontinuous.*

Proof. This is on the exercise sheet. $\square$

We are now ready to state the main result of this section - Arzela-Ascoli theorem. This says that boundedness and equicontinuity are not only necessary but also sufficient to guarantee compactness of a subset of continuous functions:

Theorem 4.35 (Arzela-Ascoli)**.** *Let (X, τ_X) be compact and (Y, d) be a metric space such that all closed bounded balls are compact. Then $cl(A) \subseteq C(X, Y)$ is compact if and only if A is uniformly bounded and equicontinuos .*

We will prove it next week.

We have already basically proved one part of the theorem in Corollary 4.34, up to the following small exercise:

Exercise 4.5. *Show the following statements:*

- *Let X be compact and (Y,d) a complete metric space. Then $A \subseteq (C(X,Y), d_\infty)$ is equicontinuous and uniformly bounded iff $cl(A)$ is.*
- *in any metric space (Y,d) and for any subset $A \subseteq Y$, $cl(A)$ is totally bounded if and only if A is totally bounded.*

Before getting into the proof in the general case, let us sketch the proof in the case $(C([0,1],\mathbb{R}), d_\infty)$. The key step is to prove total boundedness of this set of functions:

- First, because of equicontinuity we can find finitely many points $x_1, \dots, x_m$ such that if we know the value of f at these points, then in fact we know the value of f over whole of $[0,1]$ up to $\epsilon/3-$precision.
- Second, because $f([0,1])$ is bounded, we can also find $y_1 \dots y_k$ such that each $f(x_j)$ is $\epsilon/3-$ close to one of the points y_i.
- This means that if we know for each x_j the point y_i to which $f(x_j)$ is close to (please draw!), then we know f up to $\epsilon-$precision. But now there are only finitely many options assigning to each x_j some y_i and moreover, any two functions giving rise to the same assignment are ϵ.
- This gives us total boundedness, as for every assignment of a point y_i to each x_j, we can pick a function that satisfies such an assignment, if such a function exists. Then ϵ balls around these functions cover the whole set.

Proof. As mentioned, Corollary 4.34 together with the exercise just above proves that if $cl(A)$ is compact, then A has to be uniformly bounded and equicontinuous. So we only need to prove that if A is uniformly bounded and equicontinuous, then $cl(A)$ is compact.

By Theorem 4.25 $(C(X,Y), d_\infty)$ is complete, thus as $cl(A)$ is closed, it is also complete. Hence, by Theorem 4.24 to prove compactness it suffices to show that $cl(A)$ is totally bounded. As the exercise above confirms, it suffices to show that A is totally bounded. Thus for any $\epsilon > 0$, we aim to find functions $f_1, \dots, f_n$ such that the balls $(B(f_i, \epsilon))_{i=1\dots n}$ cover the set A, i.e. that any other function is an ϵ neighbourhood of one of $f_1, \dots, f_n$ in the uniform metric. We will now follow the sketch above.

By equicontinuity of A we can find $U_1, \dots, U_m$ such that for all $f \in A$ and all $x, y \in U_j$ for $j = 1 \dots m$, we have that $d(f(x), f(y)) < \epsilon/3$. By uniform boundedness of A we can also find some $C > 0$ and some $y_0 \in Y$ such that $f(X) \subseteq B(y_0, C)$ for all $f \in A$. By assumption $cl(B(y_0, C))$ is compact and thus in particular totally bounded. Hence we can find $y_1, \dots, y_k$ such that $B(y_0, C) \subseteq \bigcup_{i=1\dots k} B(y_i, \epsilon/6)$.

Pick now points $x_j \in U_j$ for $j = 1 \dots m$. We claim that two functions that are close on all points x_j, then they are close over whole of X:

Claim 4.36. *Suppose that $f_1, f_2 \in A$ are such that for all $j = 1 \dots m$, one can find $i_j \in \{1, \dots, k\}$ with $f_1(x_j) \in B(y_{i_j}, \epsilon/6)$ and $f_2(x_j) \in B(y_{i_j}, \epsilon/6)$. Then in fact $d_\infty(f_1, f_2) < \epsilon$.*

Before proving the claim, let us show how to conclude the proof. Fix some $f_0 \in A$, and define for each $\bar{i} = (i_1, \dots, j_m) \in \{1, \dots, k\}^m$ the function $f_{\bar{i}} \in A$ as follows:

- if there is a function $f \in A$ such that for each $j = 1 \dots m$ we have that $f(x_j) \in B(y_{i_j}, \epsilon/6)$, set $f_{\bar{i}} := f$;
- otherwise set $f_{\bar{i}} := f_0$.

Then by the claim above each $f \in A$ is $\epsilon-$close in the d_∞ metric to $f_{\bar{i}}$ where $\bar{i} = (i_1, \dots, i_m)$ is such that $f(x_j) \in B(y_{i_j}, \epsilon/6)$ for all $j = 1 \dots m$). Thus the theorem follows, once we prove the claim.

Proof of the claim. Fix some $x \in X$. Then there is some U_j with $j \in \{1, \dots, m\}$ such that $x \in U_j$. Then by the triangle inequality,

$$d(f_1(x), f_2(x)) \leq d(f_1(x), f_1(x_j)) + d(f_1(x_j), f_2(x_j)) + d(f_2(x_j), f_2(x)).$$

But now $d(f(x), f(x_j)) < \epsilon/3$ for any $f \in A$ by the definition of U_j and moreover

$$d(f_1(x_j), f_2(x_j)) \leq d(f_1(x_j), y_{i_j}) + d(f_2(x_j), y_{i_j}) < 1\epsilon/3$$

by the conditions of the claim. Thus we conclude that $d(f_1(x), f_2(x)) < \epsilon$. As this holds for all $x \in X$, the claim follows. □

□

4.7 Generic properties in metric and topological spaces

In this subsection we treat a somewhat vague question:

- How does a generic / typical point of a space look like? What properties does it satisfy?

We use typical and generic as synonyms. To give some mathematical content to this question, we will need to make precise the meaning of this generic / typical. It comes out that there are many different ways of doing this.

(1) We could just use the size of a set: for example, it would be reasonable to say that a property of a natural number is generic if it holds for all but a finitely many $n \in \mathbb{N}$. In this vein, being larger than 11 is generic, but being odd is for example not generic. Similarly, one could say that a property of a real number is typical if it holds for all but countably many real numbers - thus for example being irrational would be a typical property. More generally, we could say that a property is typical in a set X if it holds for all $x \in X \backslash A$, where A has strictly smaller size (cardinality).

(2) Generic / typical has also a meaning in the realm of probability theory: for example, there is a notion of a uniform random number on $[0, 1]$, which we denote by X. In this setting, a property of $r \in [0, 1]$ is called typical if it holds with probability 1 for the uniform random number X. For example being irrational has probability 1 and thus it is a typical property, having a digit 5 in the decimal expansion also happens with probability 1 and is thus typical. Notice that in this case the set of $r \in [0, 1]$ that have no digit 5 in their decimal expansion is uncountable, so having a digit 5 would not be typical in the sense of the set theory above.

(3) Finally, there is the notion of generic / typical in the realm of metric and topological spaces. This is what we aim to look at in the current subsection.

We will need a few definitions to get going:

Definition 4.37 (Dense and nowhere dense sets)**.** *Let (X, τ_X) be a topological space. We say that $A \subseteq X$ is dense if $cl(A) = X$. We say that $A \subseteq X$ is nowhere dense if $int(cl(A)) = \emptyset$.*

Lemma 4.38. *Let (X, τ_X) be a topological space.*

- *$A \subseteq X$ is dense if and only if for every non-empty $U \in \tau_X$ we have that $U \cap A \neq \emptyset$.*
- *$A \subseteq X$ is nowhere dense if and only if for every non-empty $U \in \tau_X$, there exists a non-empty $V \in \tau_X$ with $V \subseteq U$ and $V \cap A = \emptyset$.*

Proof. The proof is on the exercise sheet. □

For example, $\mathbb{Q}$ is dense in $(\mathbb{R}, \tau_E)$ as the closure of $\mathbb{Q}$ is equal to $\mathbb{R}$. On the other hand $\mathbb{Q}$ is not dense in $\mathbb{R}$ with the discrete metric - so the definition really does involve also the topology of the underlying space. For more examples - any finite set is nowhere dense in $\mathbb{R}$, any set is dense in a space with the indiscrete metric.

The topological notion of small or negligible is given by the following definition:

Definition 4.39 (Meagre sets)**.** *Let (X, τ_X) be a topological space. A subset $A \subseteq X$ is called meagre if it can be written as a countable union of nowhere dense sets.*

We would like to now say that if a property holds for all $x \in X \backslash A$, and A is meagre, then this property is typical. Notice however that it is very much possible that X itself is small in this topological sense: for example $(\mathbb{Q}, \tau_E)$ can be written as a countable union of $\{q\}$ with $q \in \mathbb{Q}$ and each $\{q\}$ is nowhere dense. Thus this notion is only useful, if the underlying space is not meagre. And luckily, there are many such spaces, called Baire spaces:

Definition 4.40 (Baire space)**.** *A topological space (X, τ_X) is called a Baire space if any $A \subset X$ with non-empty interior is not meagre.*

There is an equivalent definition, which is also frequently used:

Lemma 4.41. *A topological space (X, τ_X) is Baire if and only if for any countable collection of dense open sets $(U_n)_{n \in \mathbb{N}}$ we have that $\bigcap_{n \in \mathbb{N}} U_n$ is dense.*

Proof. This is on the example sheet. $\square$

Baire spaces were introduced in the PhD thesis of René-Louis Baire. He also proved that in fact there are plentiful interesting examples of Baire spaces:

Theorem 4.42 (Baire Category Theorem I)**.** *Every complete metric space is a Baire space.*

Remark 4.43. *The 'I' in the theorem hints that there is a sequel. And indeed, one can also show that every compact Hausdorff space is a Baire space. Whereas this result and its proof are not be examinable, it is on the starred section of the exercise sheet. The proof mimics that of Baire Category Theorem I, but you will need to recall some additional properties of compact Hausdorff spaces...*

The following proof should remind you a bit of the Lion hunting argument: here we are trying to find an element that is not covered by a countable union of nowhere dense sets.

Proof. Let $A \subseteq X$ some subset with non-empty interior. In particular, there is some open set $U \subseteq A$. We aim to show that A cannot be written as a countable union of nowhere dense sets. So let $(C_n)_{n \geq 1}$ be a countable collection of nowhere dense sets.

As C_1 is nowhere dense and U is open, we can find $B(x_1, \delta_1) \subseteq U$ such that $B(x_1, \delta_1) \cap C_1 = \emptyset$. In particular, the closed ball $\overline{B(x_1, \delta_1/2)}$ does not intersect C_1 and is contained in U. Now, C_2 is also nowhere dense and thus we can further find $B(x_2, \delta_2) \subseteq B(x_1, \delta_1/2)$ with $B(x_2, \delta_2) \cap C_2 = \emptyset$. Again, the closed ball $\overline{B(x_2, \delta_2/2)}$ does not intersect C_2. Inductively, having constructed $B(x_1, \delta_1), \dots, B(x_n, \delta_n)$, as C_{n+1} is nowhere dense, we can find $B(x_{n+1}, \delta_{n+1}) \subseteq B(x_n, \delta_n/2)$ such that $B(x_n, \delta_n) \cap C_{n+1} = \emptyset$ and moreover $\overline{B(x_{n+1}, \delta_{n+1}/2)} \cap C_{n+1} = \emptyset$. Observe that we can choose δ_n always as small as we wish, in particular so that $\delta_n < 1/n$ to guarantee that $\delta_n \to 0$.

Hence [17] we obtain a sequence $(x_n)_{n\geq 1}$ that is Cauchy: indeed, for any $n > n_0$ we have that $x_n \in \overline{B(x_{n_0}, \delta_{n_0}/2)}$. In particular, as X is complete, we have that $(x_n)_{n\geq 1}$ converges to some $x \in X$. But now as $\overline{B(x_n, \delta_n/2)}$ is closed, and contains all elements x_m with $m > n$, we see that $x \in \overline{B(x_n, \delta_n/2)}$ for all $n \in \mathbb{N}$. In particular, $x \in A$ and $x \notin C_n$ for all $n \in \mathbb{N}$. Thus A cannot be written as a countable union of nowhere dense sets $(C_n)_{n\geq 1}$ and as this collection was arbitrary the theorem follows.

□

In particular, we see that for example $(\mathbb{R}, d_E)$ is a Baire space and so is $(C([0,1], \mathbb{R}), d_\infty)$. In both cases this has also some interesting consequences. In fact, Baire Category Theorem has surprisingly many interesting consequences. These consequences seem to come in two flavours:

(1) *Existence results:* for example, it can be used to prove the existence of nowhere differentiable continuous functions on $[0,1]$; in fact from Baire Category Theorem it follows that being nowhere differentiable is the generic / typical property of continuous functions. In some cases, the proofs can be much easier than explicit constructions. For example, this seems to be the case when one tries to find a everywhere differentiable but nowhere monotonic functions. In the 19th century the mathematicians really couldn't decide whether they should or should not exist!
(2) *Strengthening of properties:* here most of the examples go out of the scope of the current course, but let us mention a few for the interested: open mapping theorem and uniform boundedness theorem for linear maps between infinite-dimensional vector spaces. Moreover, on the exercise sheet you find the following pretty intriguing fact about holomorphic maps: if a sequence of holomorphic maps on a connected open set U converges pointwise, then the limit is actually holomorphic on a dense open set of U! It is a 'strengthening' in the sense that we extract from a pointwise limit large sets where the function is holomorphic.

Let us now try to illustrate such consequences of the Baire Category Theorem via two examples:

Claim 4.44. *The set of continuous functions on* $[0,1]$ *that is zero on some* $q \in \mathbb{Q}$ *is meagre in* $((C([0,1], \mathbb{R}), d_\infty)$.

In particular this means that typical, or most continuous functions are not zero at any rational point! Here we use the fact that $((C([0,1], \mathbb{R}), d_\infty)$ is complete to see that meagre is negligible, i.e. very small.

Proof. Let A denote the set of functions that are zero at some $q \in \mathbb{Q}$ and let A_q denote the set of functions that are zero at q. Then $A = \bigcup_{q\in\mathbb{Q}} A_q$ and thus to show that A is meagre, it suffices to show that A_q is nowhere dense for every fixed $q \in \mathbb{Q}$.

So fix some $q \in \mathbb{Q}$ and consider some open ball $B(g, \epsilon)$ in $((C([0,1], \mathbb{R}), d_\infty)$. We want to show that there is some smaller ball $B(g_0, \delta) \subseteq B(g, \epsilon)$ such that $B(g_0, \delta) \cap A_q = \emptyset$. We separate two cases:

- First, suppose that $g(q) \neq 0$. Then $|g(q)| > c > 0$ and thus for any $\widetilde{g} \in B(g, c/2)$ we have that $\widetilde{g}(q) > c/2 > 0$. Hence $B(g, c/2) \cap A_q = \emptyset$. As also $B(g, c/2) \subseteq B(g, c)$ we have the desired ball.

[17] Here we are again using a weak version of the Axiom of choice: the Axiom of dependent choice.

- Now, suppose that $g(q) = 0$. Then consider $g_0(x) = g(x) + \epsilon/2$. We claim that $B(g_0, \epsilon/3)$ satisfies the desired conditions: indeed, if $\widetilde{g} \in B(g_0, \epsilon/3)$, then $d_\infty(\widetilde{g}, g) < \epsilon/2 + \epsilon/3 < \epsilon$ by the triangle inequality. Thus $B(g_0, \epsilon/3) \subseteq B(g, \epsilon)$. But also $\widetilde{g}(0) > \epsilon/2 - \epsilon/3 > 0$, thus $B(g_0, \epsilon/3) \cap A_q = \emptyset$.

We conclude that A_q is nowhere dense and thus $A = \bigcup_{q \in \mathbb{Q}} A_q$ is indeed meagre. □

Claim 4.45. *The set of continuous real-valued functions on* $[0, 1]$ *that are differentiable at* $1/2$ *is meagre in* $(C([0, 1], \mathbb{R}), d_\infty)$

In particular this means that typical, or most continuous functions are not differentiable at 1/2! Here we again use the fact that $((C([0, 1], \mathbb{R}), d_\infty)$ is complete to see that meagre is negligible, i.e. very small.

Proof. The proof is detailed on the exercise sheet - it follows a very similar idea as the last proof, it is just a bit more difficult on the technical level (i.e. on the level of choosing δ-s). □

4.8 Separation properties in metric spaces

We end the chapter on metric spaces with a very tiny subsection on separation properties. We already saw that any metric space (X, d) is Hausdorff: for any two distinct points $x, y \in X$, we have that $d(x, y) > 0$ and thus the open ball $B(x, d(x, y)/2)$ is disjoint from y. It comes out that more is true.

Recall, that on Exercise sheet 8 we already met a stronger property of separation - we saw that compact Hausdorff spaces are normal. In a normal space one cannot only separate any two distinct points by open sets, but any two disjoint closed sets can be separated by open sets. It comes out that metric spaces are normal:

Proposition 4.46. *A metric space* (X, d_X) *is a normal space. This is, for any two disjoint closed sets* C_1, C_2 *there are disjoint open sets* U_1, U_2 *such that* $C_1 \subseteq U_1$ *and* $C_2 \subseteq U_2$.

Proof. The proof is detailed on the exercise sheet. □

Thus it follows that any metrizable space has to be normal. In the next chapter we will see a sort of converse - we will see that each topological space (X, τ_X) that is Hausdorff, normal and has a countable basis is metrizable, i.e. there is some metric d on X such that (X, τ_X) is the topology induced by (X, d). That proof will be based on the following important lemma, whose proof is given in the next section, but is not examinable:

Lemma 4.47 (Urysohn's Lemma)**.** *Let* (X, τ_X) *be a normal space. Then for any two disjoint closed sets* K_0, K_1 *we can find a continuous function* $f : (X, \tau_X) \to ([0, 1], \tau_E)$ *such that* $f(x) = 0$ *for* $x \in K_0$ *and* $f(x) = 1$ *for* $x \in K_1$.

I urge you to think about this lemma - how would one go about constructing such a continuous function in an abstract setting? Also, why does the existence of such a continuous function imply that the space is normal?

Section 5

A metrization theorem [$\star$ non-examinable $\star$]

The last non-examinable mini-chapter is a tribute to Pavel Samuilovich Urysohn, who lived short, yet made some fundamental contributions to topology. We will see the metrization theorem, often named after him. But Urysohn also introduced our modern notion of compactness, contributed to the development of the dimension theory of topological spaces and has a metric space named after him - the Urysohn universal space.

Some of this work he did in collaboration with his friend Pavel Sergeevich Aleksandrov, with whom he also went swimming in stormy seas off the coast of Brittany, when 26 years old, to never return. Both Urysohn and Aleksandrov were fluent in German and French and thus were well aware of all the contemporary developments of mathematics. Apparently their papers were succinct, clear and always friendly with attributions of credit. So with no further:

Theorem 5.1 (Urysohn's Metrization Theorem). *Let (X, τ_X) be a topological space that is Hausdorff, normal and has a countable basis. Then (X, τ_X) is metrizable - there is a metric d on X so that the induced topology of (X, d) agrees with τ_X.*

The proof of this theorem is a jewel, with several beautiful ideas and bringing together many ideas of different chapters of the course. There are two key steps:

(1) In the first step, we will show that in a normal space, given two disjoint closed sets one can construct a real-valued continuous function that takes values 0 and 1 on these two sets, respectively. Whereas constructing such continuous functions on metric spaces is relatively easy using the distance function, it is by no means clear in a general topological space.
(2) Thereafter, we see that a countable family of such continuous functions helps to describe the metric structure - indeed, using these functions we construct an embedding (a bi-continuous injection) of our topological space to the product space $[0,1]^{\mathbb{N}}$. As this latter space is endowed by the product metric, it induces a metric on the image of our topological space, and thus on the initial space itself.

The first step is usually separated in a lemma of independent interest, that we already stated in the last section.

Lemma 5.2 (Urysohn's Lemma). *Let (X, τ_X) be a normal space. Then for any two disjoint closed sets K_0, K_1 we can find a continuous function $f : (X, \tau_X) \to ([0,1], \tau_E)$ such that $f(x) = 0$ for $x \in K_0$ and $f(x) = 1$ for $x \in K_1$.*

Before proving the lemma and the theorem, let us quickly revisit the concepts in the statement of the theorem:

- In a normal space, given any two disjoint closed sets K_0 and K_1 we can find disjoint open sets U_0, U_1 containing K_0 and K_1, respectively. Notice that in this case $cl(U_0)$ is also disjoint from U_1 and in particular of K_1. This is just because $X \backslash U_1$ is closed, and hence by definition of the closure $cl(U_0) \subseteq X \backslash U_1$. Recall also that we saw that any metric space is a normal space, so this condition is necessary.

- On the other hand, not every metric spaces has a countable basis. For example $\mathbb{R}$ with its discrete metric has no countable basis. Also, not all basis have to be countable - think of $(\mathbb{R}, \tau_E)$ for example. Existence of a countable basis implies in particular that the underlying space can be approximated by countable sets in the following sense: we say that (X, τ_X) is separable if there is a countable set $A \subseteq X$ with $cl(A) = X$. One can check that if (X, τ_X) has a countable basis $(U_i)_{i \in \mathbb{N}}$, then we can construct such A by picking some $x_i \in U_i$ for all $i \in \mathbb{N}$ and taking their union [18].

Let us now prove Urysohn's lemma:

Proof. Fix two closed disjoint sets K_0, K_1. Let $\mathcal{D}_n = \{\frac{m}{2^n} : m \in \mathbb{N}\}$ and $\mathcal{D} = \bigcup_{n \geq 1} \mathcal{D}_n$. The idea is now to construct a family of open sets $(U_s)_{s \in (0,1) \cap \mathcal{D}}$ so that

- $K_0 \subseteq U_s$ for any $s > 0$;
- for any $s < t$: $cl(U_s) \subseteq U_t$;
- also, $U_t \subseteq X \backslash K_1$ for all $t < 1$.

Before constructing this family, let's see how its existence allows to find the desired continuous function $f : (X, \tau_X) \to ([0,1], \tau_E)$. Indeed, set $f(x) := \inf_{s \in (0,1) \cap \mathcal{D}} \{x \in U_s\}$, if the infimum exists and set $f(x) = 1$ otherwise. Notice that by the conditions on the family U_s, we have that $f(x) = 0$ when $x \in K_0$ and $f(x) = 1$ when $x \in K_1$. So we just need to check the continuity of f. As the sets of the form $[0, a)$ and $(a, 1]$ with $a \in (0,1)$ form a subbasis of open sets for τ_E on $[0,1]$, it suffices to check that their preimages are open in τ_X.

First, by definition if $x \in U_s$, then $f(x) \leq s$. Thus we have that $f^{-1}([0,a)) \supseteq \bigcup_{s \in [0,a) \cap \mathcal{D}} U_s$. On the other hand, if $f(x) < a$, then there is some $a_0 < a$ in $\mathcal{D}$ such that $x \in U_{a_0}$. Thus $f^{-1}([0,a)) \subseteq \bigcup_{s \in [0,a) \cap \mathcal{D}} U_s$. We conclude that $f^{-1}([0,a))$ is open as an union of open sets.

In the other direction, if $x \in X \backslash U_s$, then $f(x) \geq s$. Thus $f^{-1}((a,1]) \supseteq \bigcup_{s \in (a,1] \cap \mathcal{D}} X \backslash cl(U_s)$. But if $f(x) > a$, then we can find $a_0 \in (a, f(x)) \cap \mathcal{D}$ such that $x \notin U_{a_0}$. If we now pick further $a_1 \in (a, a_0) \cap \mathcal{D}$, then as $cl(U_{a_1}) \subseteq U_{a_0}$, we see that $x \in \backslash cl(U_{a_1})$. Thus $f^{-1}((a,1]) \subseteq \bigcup_{s \in (a,1] \cap \mathcal{D}} X \backslash cl(U_s)$. Hence, we see that $f^{-1}((a,1])$ is also open, proving continuity.

It remains to construct a family of open sets $(U_s)_{s \in [0,1] \cap \mathcal{D}}$ with the above conditions. We will do this recursively over the $n \in \mathbb{N}$ of $\mathcal{D}_n$. It is easier to also simultaneously keep track of a family of closed sets $(C_s)_{s \in (0,1) \cap \mathcal{D}}$, such that $C_s := X \backslash U_s$.

The induction basis is the case $n = 1$, i.e. we construct $U_{1/2}, C_{1/2}$. By normality of X, we can find disjoint $S, T \in \tau_X$ such that $K_0 \subseteq S$ and $K_1 \subseteq T$, and we just set $U_{1/2} := S$ and $C_{1/2} := X \backslash S$.

For the induction step: having constructed all U_s, C_s for $s \in \mathcal{D}_n \cap (0,1)$, we construct U_s, C_s with $s = \frac{2m+1}{2^{n+1}}$ as follows:

- if $m \notin \{0, 2^n - 1\}$ consider the disjoint closed sets $C_0 := cl(U_{\frac{m}{2^n}})$ and $C_1 := C_{\frac{m+1}{2^n}}$;
- $m = 0$, set $C_0 := K_0$ and $C_1 := C_{\frac{1}{2^n}}$;
- $m = 2^n - 1$, set $C_0 := cl(U_{\frac{2^n - 1}{2^n}})$ and $C_1 := K_1$.

Again, by normality of X, we can find disjoint open sets S and T such that $C_0 \subseteq S$ and $C_1 \subseteq T$. We set $U_{\frac{2m+1}{2^{n+1}}} := S$ and $C_{\frac{2m+1}{2^{n+1}}} := X \backslash S$ and check that all conditions remain to be satisfied. Thus, we conclude our construction of sets U_1 [19].

□

[18] By using Axiom of countable choice of course...

[19] Yes, we are using Axiom of dependent choice here...

We are now ready to prove the Urysohn's Metrization theorem itself, which would be too beautiful to admit. The idea is as follows: for any two nested open basis sets U_1, U_2 such that $cl(U_1) \subseteq U_2$, we construct the function f_{U_1,U_2} with $K_0 := cl(U_1)$ and $K_1 := X\backslash U_2$ using Urysohn's Lemma. We then show that all these functions f_{U_1,U_2} together actually encode a metric structure, by finding a bi-continuous injection (an embedding) of X into $[0,1]^{\mathbb{N}}$ with its product metric. Here, it is the countability of the basis that allows us to use countably many functions and thus to embed into a countable product space (recall that uncountable product spaces are not metrizable!).

Proof of the Metrization theorem. Consider all pairs of basis sets U_i, V_i such that $cl(U_i) \subseteq V_i$. There are only countable many pairs of the basis sets, and thus there are also only countable many of such nested pairs. Thus we can index all such pairs using $\mathbb{N}$ to obtain $(U_i, V_i)_{i\in\mathbb{N}}$. Notice that this set is indeed non-empty. First of all, as the basis covers X, for any $x \in X$ there is some basis element V_i containing x. Moreover,

Claim 5.3. *For any $x \in X$ and any basis element V_i containing x, there is a further basis element U_i with $x \in U_i$ such that $cl(U_i) \subseteq V_i$.*

Proof of claim. As X is Hausdorff, then $\{x\}$ is closed. As $\{x\}$ and $X\backslash V_i$ are disjoint closed sets and X is a normal space, we can find some disjoint open sets S, T such that $x \in S$ and $X\backslash V_i \subseteq T$. Now, each open set is an union of basis elements, thus there is some further basis element U_i containing x with $U_i \subseteq S$. Moreover as $cl(U_i)$ is disjoint from T, it is in particular contained in V_i and the claim follows. [20] □

For any pair (U_i, V_i) with $i \in I$ define $f_i := f_{U_i,V_i}$ as the continuous function obtained from the Urysohn's lemma with $K_0 = cl(U_i)$ and $K_1 = X\backslash V_i$. Consider now $[0,1]^{\mathbb{N}}$ with its standard product metric d_Π, and the function $f : (X, \tau_X) \to ([0,1]^{\mathbb{N}}, d_\Pi)$ given by $f(x) := (f_1(x), f_2(x), \dots)$. Then f is continuous as each f_i is continuous. We aim to show that f is also injective and f^{-1} is continuous, i.e. that f maps open sets to open sets. Indeed, given this, it would follow that X is homeomorphic to $f(X)$ and as $(f(X), d_\Pi)$ is a metric space, we could conclude that X is metrizable. So let us prove these two properties.

- *f is injective*: pick any disjoint x, y. By the Hausdorff property, there are some disjoint open sets S, T such that $x \in S, y \in T$. Then there must be also some basis element $V_i \subseteq S$ containing x and not y. Thus by the claim above we can also find a basis element U_i with $x \in U_i$ and $cl(U_i) \subseteq V_i$. In particular, by the definition of the function $f_i = f_{U_i,V_i}$ we have that $f(x) = 0$ and $f(y) = 1$.
- *f is an open map, i.e. f maps open sets to open sets*: it suffices to show that $f(V_{i_0})$ is open for any basis set V_{i_0}. So let $x \in V_{i_0}$. It suffices to then argue that there is some open set S of $([0,1]^{\mathbb{N}}, d_\Pi)$ containing $f(x)$ such that $S \cap f(X) \subseteq f(V_{i_0})$. Using the claim above, we can again find a further basis element U_{i_0} containing x with $cl(U_{i_0}) \subseteq V_{i_0}$, and thus also a function $f_{i_0} = f_{U_{i_0},V_{i_0}}$ such that $f_{i_0}(x) = 0$ and $f_{i_0} = 1$ outside V_{i_0}. In particular $f^{-1}([0,1)) \subseteq V_{i_0}$. Thus, if we define $S = \Pi_{i\in\mathbb{N}} S_i$ with $S_i = [0,1]$ for all $i \neq i_0$ and $S_{i_0} = [0,1)$, then S is open and $S \cap f(X) \subseteq f(V_{i_0})$. Thus we conclude that $f(V_{i_0})$ is open.

□

[20]Notice that it might happen that $U_i = V_i$. When?

[$\star$ End of the non-examinable section $\star$]

And that's all there is.