An Introduction to Algebraic Topology

Ulrich Pennig

January 23, 2020

Abstract

These are lecture notes I created for a one semester third year course about (Algebraic) Topology at Cardiff University. The material covered includes a short introduction to continuous maps between metric spaces. We will not assume that the reader is familiar with these notions. However, for the last chapter about the fundamental group we will assume a good understanding of groups and group homomorphisms. Despite some careful proofreading, the document is probably still riddled with typos. If you want to help improve them, please send any comments or corrections to pennigu@cardiff.ac.uk.

Contents

1 What is algebraic topology?

In mathematics we are often faced with *quantitative* questions, like: "What is the value of the integral $\int_0^1 x^2 dx$?", "What is the volume of the 2-dimensional sphere?" or "What is the ratio of a circle's circumference to its diameter?". There are different, but equally important questions that ask for the *qualitative* features of mathematical objects. An example that you have encountered in your Linear Algebra course is: "Which vector spaces are isomorphic to each other?". Questions like this one often lead to interesting mathematical concepts. In the Linear Algebra example we are lead to the definition of the dimension to distinguish vector spaces that are not isomorphic.

Topology deals with the notion of closeness, continuity and continuous deformations in the most general way with the goal to answer qualitative questions like: "Can the surface of a coffee cup be continuously deformed into the surface of a doughnut?". The rough idea of a continuous deformation in this context is that we are allowed to stretch, bend or shrink the object, but not allowed to tear it, cut it or paste something onto it.

Surfaces are not the only objects, where continuous deformations make sense. One of the results of topology is to give a rigorous and very general definition for "things that allow continuous maps between them". These are called topological spaces. Such a space consists of a set of points together with an extra structure – also called a topology – which allows us to talk about points being close together without actually having to talk about distances between them. If X and Y are topological spaces, we can define what it means for a map $f\colon X \to Y$ to be continuous.

A lot of the mathematical objects you might have already dealt with fit this definition: Many geometric structures, like surfaces or in fact all manifolds, are in particular topological spaces. All metric spaces have a natural topology as well. Moreover, there are many operations that create new topological spaces from old ones: We can glue them together using equivalence relations, we can form products and subspaces. Even the set of all continuous maps between nice enough topological spaces can itself be equipped with a useful topology. Therefore the scope of topology is very general and it lies at the root of many modern developments in mathematics.

So - what is algebraic topology? Let us return to the question "Is it possible to continuously deform the surface X into Y?". If the answer to it is "Yes!", we can try to sketch the corresponding deformation. In fact, the surface of a coffee cup in the example above *can* be deformed into that of a doughnut as outlined in Figure 1.

Figure 1: The deformation of a coffee cup into a doughnut.

But what about the cases, where the answer is "No!"? Consider the surface of the doughnut and that of the 2-dimensional sphere (i.e. the surface of a ball) as shown in Figure 2. Can those two be deformed into each other? A few tries reveal that the answer is probably "No.". How do we actually prove that it can not be done? This is the point where algebraic topology enters the scene.

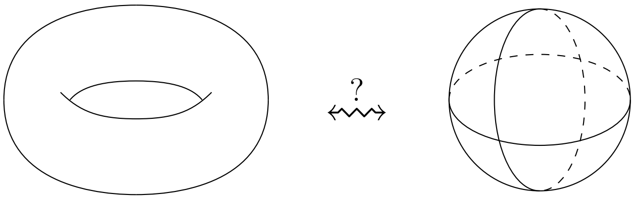


Figure 2: Can we continuously deform the doughnut into the sphere?

Suppose that we could use some construction to attach a number $\chi(X)$ to every surface X in such a way that whenever X can be continuously deformed into Y, we have $\chi(X) = \chi(Y)$. From this we could then deduce that two surfaces can not be deformed into each other if the corresponding numbers are not equal. In other words, $\chi(X)$ would allow us to distinguish such surfaces. This is an example of a ***topological invariant***.

We will later see a more elaborate construction, which associates a group $\pi_1(X, x_0)$ not only to every surface, but to every pair of a topological space X and point $x_0 \in X$. This is done in such a way that any continuous map $f\colon X \to Y$ induces a group homomorphism $\pi_1(X, x_0) \to \pi_1(Y, f(x_0))$. Whenever X can be continuously deformed into Y, the groups will not be equal, but isomorphic. In particular, two topological spaces, which produce non-isomorphic groups, can not be deformed into each other. Such topological invariants are the main object of study in algebraic topology.

1.1 Footballs, chemistry and the Euler characteristic

As explained above, algebraic topology associates algebraic structures, like numbers, groups, rings or modules to topological spaces in such a way that continuous deformations of the underlying space lead to isomorphic algebraic structures, i.e. they give the same number, isomorphic groups, isomorphic rings and so on. In order to develop some intuition for the concepts to come, we will look at a particular algebraic invariant in this section. We will not need the precise definition of a topological space or continuity at this point – just a good intuition about polyhedra.

In 1750 the Swiss mathematician Leonhard Euler made the following surprising observation: Consider a convex polyhedron[1], like a pyramid, a cube or an octahedron. It is built out of vertices, edges and faces, whose numbers we will denote by V, E and F. Then the following formula holds

$$V - E + F = 2 \; . \tag{1}$$

This identity is now known as Euler's polyhedron formula. Some examples are computed explicitly below.

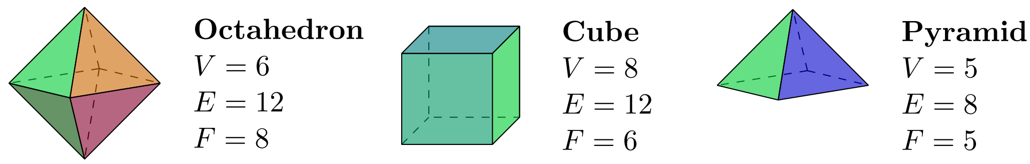


Figure 3: Illustration of Euler's polyhedron formula

For a polyhedron P the left hand side of the equation $\chi(P) = V - E + F$ is called the Euler characteristic of P. Apart from being equal to 2 for convex P, it has another interesting property. It remains constant under the following transformation: We can subdivide a face of P by adding a vertex in its center and adding all edges that connect the new vertex with the vertices of the original face as shown for the cube in Figure 4 on the left. Moving the vertices around does of course not change the values of V, E and F and therefore also not the value of $\chi(P)$. Note that we could have moved the extra point in Figure 4 inside the cube, thereby creating a non-convex polyhedron, which still has $\chi(P) = 2$.

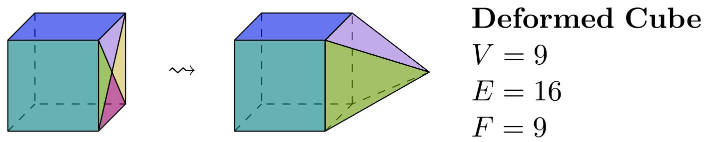


Figure 4: The Euler characteristic remains constant under continuous deformations.

As we see from this, convexity of P is not necessary to have $\chi(P) = 2$. We only need to be able to *continuously deform* P into a convex polyhedron by a sequence of moves

[1]Convexity means that a line connecting any two points on the surface is contained in its interior.

as described in Figure 4. The polyhedra that have this property are also continuously deformable into a sphere. By this we mean the following: Take the cube for example and imagine having a model of it made out of very elastic rubber. Then you could bend the faces outwards and round the edges thereby deforming it into a sphere without tearing it or cutting somewhere. We can deform the Octahedron or the Pyramid in a similar way. It turns out – as we will see much later – that we can extend the definition of the Euler characteristic to surfaces and every surface S that is continuously deformable into a sphere has $\chi(S) = 2$.

This rises the question if there are polyhedra P with $\chi(P) \neq 2$. Of course, such a P can not be deformable into a sphere, in particular it has to be non-convex. In fact, to construct such a P we can start with the cube and cut a hole in it such that the result will be the doughnut shape shown in Figure 5. Note that the four faces shown in darker green are part of it, but there are no interior walls. This particular P is deformable into a torus and has $\chi(P) = 0$.

doughnut shape
$V = 32$
$E = 64$
$F = 32$

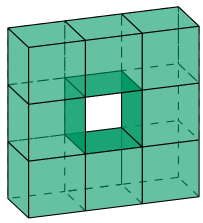


Figure 5: A shape with vanishing Euler characteristic.

We can therefore use the Euler characteristic to distinguish polyhedra that can not be deformed into each other: For example, the calculation above shows that the doughnut in Figure 5 can not be deformed into a cube by successive applications of the transformations sketched in Figure 4. If we were allowed to fill the hole by *cutting* out the four interior faces and the four edges and *gluing* in two new faces, we would end up with the cube again. However, this process changes the Euler characteristic from 0 to 2, since F is reduced by 2 and E by 4. So, cutting and gluing changes $\chi(P)$, while continuous deformations like in Figure 4 do not.

Observe that it is not so easy to reverse the above argument: Just because we know that the Euler characteristics of P_1 and P_2 agree, we can not deduce that one is deformable into the other.

Apart from enabling us to distinguish surfaces, the Euler characteristic also tells us something about which configurations of tiles can be glued together to form convex polyhedra. To illustrate this, consider the “football” depicted in Figure 6. The faces are regular pentagons (with five vertices) and hexagons (with six). How many of each do you need? Are there different possibilities to construct the ball? Let us explore the restrictions we obtain from the Euler characteristic.

Let P be a convex polyhedron composed of pentagons and hexagons and let V, E and F be the number of vertices, edges and faces. Let F_5 be the number of pentagons, F_6

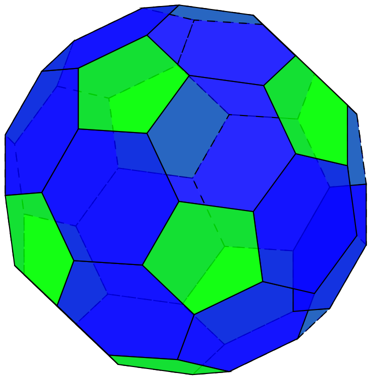


Figure 6: A football is composed of pentagons and hexagons

be the number of hexagons. Then we have

$$F = F_5 + F_6 \ .$$

Each edge is part of two faces and each of the pentagons has five edges, each of the hexagons has six. If we double count the total number of edges we therefore end up with

$$2E = 5F_5 + 6F_6 \ .$$

The angles in a regular pentagon are 108° each. In a regular hexagon they are 120°. To obtain a ball it is necessary that the sum of the angles at each vertex is at most 360°. This means that at most three faces touch at a vertex. Since there are always more than two faces touching at a vertex, we obtain that the number of faces at a vertex has to be three. This is then also the number of edges touching at a vertex. Each vertex belongs to two edges. Hence, we obtain $3V = 2E$, which we can rewrite to

$$V = \frac{2E}{3} = \frac{5F_5 + 6F_6}{3} \ .$$

Now we apply Euler's polyhedron formula, which gives:

$$\begin{aligned} 2 &= V - E + F = \frac{5F_5 + 6F_6}{3} - \frac{5F_5 + 6F_6}{2} + F_5 + F_6 \\ \Leftrightarrow 12 &= 10F_5 + 12F_6 - 15F_5 - 18F_6 + 6F_5 + 6F_6 \\ &= F_5 \end{aligned}$$

Quite miraculously the number F_6 is eliminated in the process. As we can see, the number of pentagons in a football *has* to be exactly 12. This fact is sometimes referred to as the "12 Pentagon Theorem". The restriction for F_6 we obtain from the above is $2F_6 = 20 - V$. In fact, there is a convex polyhedron with $F_6 = 0$ as shown in Figure 7. The football in Figure 6 has $F_6 = 20$.

Although the above considerations look quite abstract, they do have a lot of important applications in other scientific research areas. The 12 Pentagon Theorem is the starting

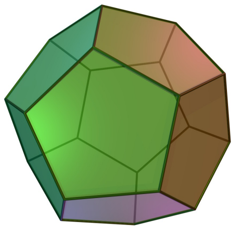


Figure 7: The dodecahedron has $F_5 = 12$ and $F_6 = 0$.

point for studying the topology of fullerenes – highly symmetric spherical molecules, which have the property that their atoms align in pentagons or hexagons. One of them, the buckyball, consists of 60 atoms, that form the vertices of a convex polyhedron. By the above formula, we have $V = \frac{5F_5+6F_6}{3} = 20 + 2F_6$. With $V = 60$ we obtain $F_6 = 20$. Therefore the buckyball has to look roughly like the polyhedron shown in Figure 6. There are other fullerenes, which have a larger number of hexagons, but still only 12 pentagons. An example is shown in Figure 8.

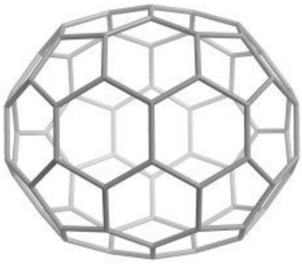


Figure 8: The fullerene C_{70}, known as "the rugby ball"

We have come a far way in this section: Starting off with a surprising observation about convex polyhedra by Euler and ending somewhere in organic chemistry. Along the way we have met our first topological invariant – the Euler characteristic – a number we attached to certain geometric objects, which stays constant under continuous deformations. We have seen some applications of such an invariant as well. However, to proceed we need to be much more precise about the objects we would like to consider and about the notion of continuity. This will lead us to the definition of **topological spaces** and **continuous maps** treated in the next section.

2 Open sets and continuous maps

The intuition behind continuity is the following: A function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous if we can draw its graph without lifting the pen from the paper. The function $f\colon \mathbb{R} \to \mathbb{R}$ with $f(x) = x^3$ is continuous. Its graph is shown on the left hand side in Figure 9. The graph of a non-continuous function contains points where its value jumps upwards or downwards as shown on the right hand side in Figure 9.

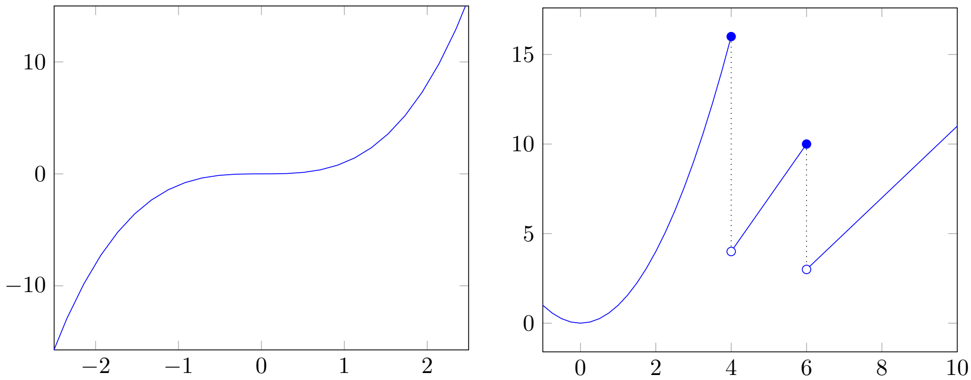


Figure 9: A continuous function (left) vs. a non-continuous one (right)

It is useful to keep this intuitive picture in mind. However, to develop a precise *definition* of continuity, let us work out what distinguishes the points where f is continuous from the ones where it is not. The example function in Figure 9 on the right hand side is given by

$$f\colon \mathbb{R} \to \mathbb{R} \quad ; \quad x \mapsto \begin{cases} x^2 & \text{for } x \leq 4 \\ 3x-8 & \text{for } 4 < x \leq 6 \\ 2x-9 & \text{for } x > 6 \end{cases}$$

By looking at the graph of f we see that the points on the x-axis, where f is *not* continuous, are 4 and 6. The values of f at these points are: $f(4) = 16$ and $f(6) = 10$. Let us take a closer look at the point $x_0 = 6$ with value $(x_0, f(x_0)) = (6, 10)$ in the graph.

The fact that the value of f jumps at x_0 by an amount of at least ϵ can be expressed in the following way: Among all points on the x-axis in an arbitrary small distance from x_0 there are always some points x with $f(x)$ at least ϵ away from $f(x_0)$. To make this precise, fix $\epsilon > 0$ and consider the subset $S_\epsilon(x_0)$ of the graph of f given by

$$S_\epsilon(x_0) = \{(x, f(x)) \in \mathbb{R} \times \mathbb{R} \mid |f(x) - f(x_0)| < \epsilon\}$$

These are all the points, such that the value of f lies between $f(x_0) - \epsilon$ and $f(x_0) + \epsilon$. For $\epsilon = 4$ and $x_0 = 6$ we obtain the blue subset shown on the left hand side in Figure 10. Moreover, we need the subset $T_\delta(x_0)$ of the x-axis given by

$$T_\delta(x_0) = \{x \in \mathbb{R} \mid |x - x_0| < \delta\} \ .$$

It contains all points $x \in \mathbb{R}$ between $x_0 - \delta$ and $x_0 + \delta$. For $\delta = 1$ it is drawn in Figure 10 on the right hand side. It is the union of the red and blue interval around $x_0 = 6$.

We can now rephrase our previous statement: The fact that the value of f *jumps at* x_0 by an amount of at least ϵ can be expressed by saying that for every $\delta > 0$ the subset $T_\delta(x_0)$ contains points $x \in T_\delta(x_0)$, such that $(x, f(x)) \notin S_\epsilon(x_0)$. As we can see all the points on the x-axis in red on the right hand side in Figure 10 correspond to points

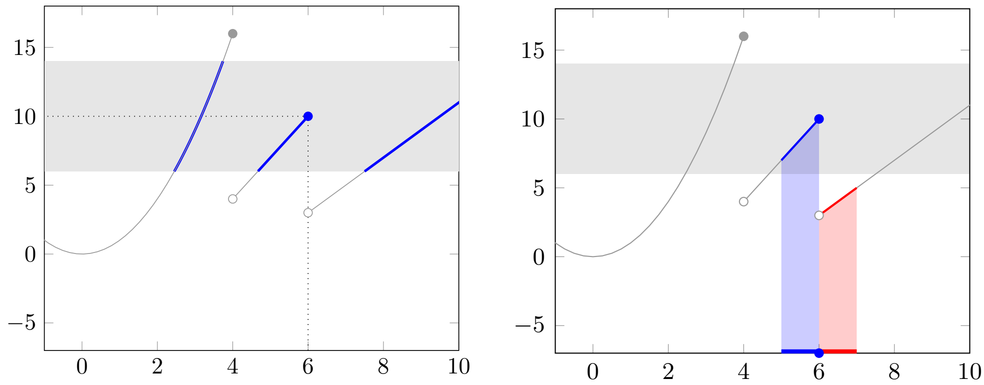


Figure 10: left hand side in blue: all points of the graph with $|f(x) - f(6)| < 4$
right hand side: checking non-continuity via the ϵ-δ definition

$(x, f(x))$ not in $S_4(6)$. Hence, we can conclude that the value of f jumps at $x_0 = 6$ by at least 4.

Observe how this behaviour is different from the one around a point where the function is continuous. Figure 11 shows the function $g\colon \mathbb{R} \to \mathbb{R}$ with $g(x) = x^2$. It is continuous at every point $x_0 \in \mathbb{R}$, in particular at $x_0 = 0$. No matter how small we choose the grey box in the picture around $g(0) = 0$ on the y-axis, there is always an interval around 0 on the x-axis, such that all the points are mapped into the grey region. In other words: No matter how small $\epsilon > 0$ is chosen, there is always a $\delta > 0$, such that for every point $x \in T_\delta(x_0)$ the point $(x, g(x))$ in the graph is in $S_\epsilon(x_0)$. By our observation above this is just rephrasing the fact that the function g does *not* jump at x_0 by *any* amount ϵ.

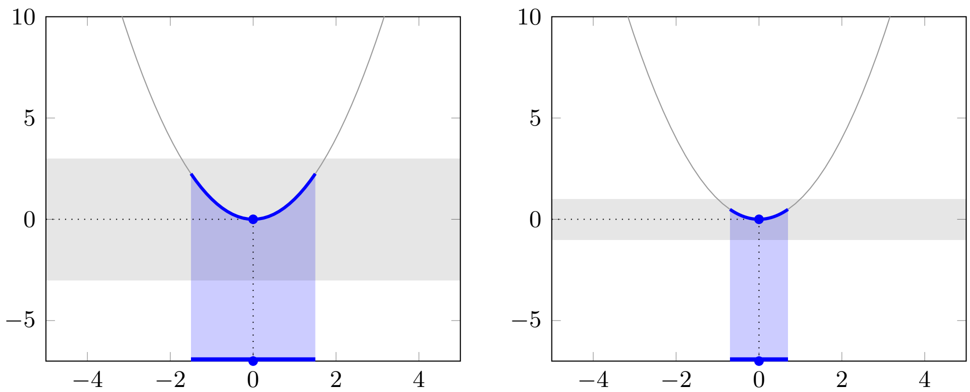


Figure 11: The function $g(x) = x^2$ is continuous at $x_0 = 0$.

Using the definition of $T_\delta(x_0)$ and $S_\epsilon(x_0)$ we can rephrase this to: For every $\epsilon > 0$ there is a $\delta > 0$ such that $|x - x_0| < \delta$ implies that $|g(x) - g(x_0)| < \epsilon$. This leads us to the following definition:

Definition 2.0.1. A function $f\colon \mathbb{R} \to \mathbb{R}$ is ***continuous at*** $x_0 \in \mathbb{R}$ if for all $\epsilon > 0$ there exists a $\delta > 0$, such that for all $x \in \mathbb{R}$ with $|x - x_0| < \delta$ we have that $|f(x) - f(x_0)| < \epsilon$. We will call a function $f\colon \mathbb{R} \to \mathbb{R}$ ***continuous*** if it is continuous at all points $x_0 \in \mathbb{R}$.

Exercise 2.0.2. Let h be the function given by

$$h\colon \mathbb{R} \to \mathbb{R} \quad , \quad x \mapsto \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{else} \end{cases}$$

Prove that it is not continuous.

2.1 Continuity in metric spaces

Definition 2.0.1 only uses the fact that we can measure distances between points in $\mathbb{R}$. The distance function $d\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by $d(x, x') = |x - x'|$ has the following properties:

- It only takes values that are greater or equal to zero.
- The distance between two points is zero if and only if the two points agree.
- It does not matter if the distance is measured from the first point to the second or the other way around.
- The triangle inequality holds: For all $x, y, z \in \mathbb{R}$ we have $|x - z| \leq |x - y| + |y - z|$.

Such a "distance function" is also called a metric and the above are its defining properties. If a set X can be equipped with a metric, we call it a metric space. We summarise this in the following definition:

Definition 2.1.1. A pair (X, d) consisting of a set X together with a map $d\colon X \times X \to \mathbb{R}$ is called a ***metric space*** if it satisfies the following axioms

a) for all $x, y \in X$ we have $d(x, y) \geq 0$,

b) for $x, y \in X$ we have that $d(x, y) = 0$ if and only if $x = y$,

c) for all $x, y \in X$ we have $d(x, y) = d(y, x)$ (this condition is called "symmetry"),

d) for all $x, y, z \in X$ we have the triangle inequality $d(x, z) \leq d(x, y) + d(y, z)$.

The map d is called the ***metric of*** X.

Example 2.1.2. Let $n \in \mathbb{N}$ and define for $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ the norm of v as

$$\|v\| = \left(\sum_{i=1}^{n} v_i^2 \right)^{1/2} .$$

Let $d\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be given by $d(x, y) = \|x - y\|$. This satisfies the axioms of Definition 2.1.1. Therefore $(\mathbb{R}^n, d)$ is a metric space. In particular, we obtain for $n = 1$ the real line $\mathbb{R}$ with the metric given by $d(x, y) = |x - y|$. This is called the ***standard metric*** on $\mathbb{R}^n$.

Example 2.1.3. Let $Y \subset X$ be a subspace of a metric space (X, d) and note that $Y \times Y \subset X \times X$. Define $d_Y : Y \times Y \to \mathbb{R}$ by restriction, i.e. $d_Y = d|_{Y \times Y}$. Then, (Y, d_Y) is again a metric space. In particular, the n-dimensional spheres $S^n \subset \mathbb{R}^{n+1}$ are metric spaces.

Example 2.1.4. Let $C([0,1], \mathbb{R})$ be the set of all real-valued continuous functions on the unit interval $[0,1] \subset \mathbb{R}$ and define

$$d_{\text{sup}}(f, g) = \sup \{ |f(x) - g(x)| \mid x \in [0,1] \} \ .$$

Then $(C([0,1], \mathbb{R}), d_{\text{sup}})$ is a metric space. It is a good exercise to check that d_{sup} in fact satisfies the axioms a) to d) in Definition 2.1.1. Why do we know that $d_{\text{sup}}(f, g)$ is finite for every $f, g \in C([0,1], \mathbb{R})$?

We can now generalise Definition 2.0.1 to the continuity of maps between metric spaces.

Definition 2.1.5. A map $f : X \to Y$ between metric spaces (X, d_X) and (Y, d_Y) is ***continuous at*** $x \in X$ if for all $\epsilon > 0$ there exists a $\delta > 0$, such that for all $x' \in X$ with $d_X(x', x) < \delta$ we have that $d_Y(f(x'), f(x)) < \epsilon$. We will call a function $f : X \to Y$ ***continuous*** if it is continuous at all points $x \in X$.

Exercise 2.1.6. Let $(C([0,1], \mathbb{R}), d_{\text{sup}})$ be the metric space described in Example 2.1.4 and let $x_0 \in [0,1]$. Consider the evaluation map

$$\text{ev}_{x_0} : C([0,1], \mathbb{R}) \to \mathbb{R} \quad ; \quad f \mapsto f(x_0)$$

as a map between the metric spaces $(C([0,1], \mathbb{R}), d_{\text{sup}})$ and $(\mathbb{R}, d)$ with $d(x, y) = |x - y|$. Is it continuous?

2.2 Topological spaces

The above Definition 2.1.5 is known as the ϵ-δ-definition of continuity. It is particularly useful not only in computations, but also since metric spaces appear very often in applications. But is this definition the most general one? Do we really *need* metric spaces to define continuity?

In the examples in Section 2, we looked at a function $f : \mathbb{R} \to \mathbb{R}$ and the two sets $S_\epsilon(a)$ and $T_\delta(a)$ given by

$$S_\epsilon(a) = \{ (x, f(x)) \in \mathbb{R}^2 \mid |f(x) - f(a)| < \epsilon \} \ ,$$
$$T_\delta(a) = \{ x \in \mathbb{R} \mid |x - a| < \delta \} \ .$$

The ϵ-δ-definition of continuity tells us that f is continuous at $a \in \mathbb{R}$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that for every $x' \in T_\delta(a)$ we have $(x', f(x')) \in S_\epsilon(a)$.

Observe that $T_\delta(a)$ is the interval $(a - \delta, a + \delta)$. Its length is 2δ. However, for the definition of continuity its length is not really important. The only thing that matters is that it stretches a little bit to the left and the right from a to detect whether f jumps in the region smaller than $a \in \mathbb{R}$ or in the one larger than a. It has to look like the blue/red interval around 6 on the x-axis in Figure 10. In other words: It only mattered that it contained an *open interval* around a.

If we want to generalise the property of being an open interval from $\mathbb{R} = \mathbb{R}^1$ to $\mathbb{R}^n$ we should look at sets, which are "open in every direction". By using these we can check for jumps in every direction. Such sets are called open balls and in fact, they can be defined for metric spaces:

Definition 2.2.1. Let (X, d) be a metric space. The ***open ball*** $B_r(x_0)$ of radius r around $x_0 \in X$ is the subset

$$B_r(x_0) = \{x \in X \mid d(x_0, x) < r\} \ .$$

A subset $U \subset X$ is called ***open with respect to the metric* d** if for every point $x \in U$ there exists $r > 0$ such that $B_r(x) \subset U$. (If the metric we are using is clear from the context, we will just say that U is ***open with respect to the metric***.)

Example 2.2.2. The open balls in $\mathbb{R}$ are precisely the open intervals (a, b) for some numbers $a, b \in \mathbb{R}$ with $a < b$.

Exercise 2.2.3. Let $U \subset \mathbb{R}$ be a subset, which is open with respect to the metric $d(x, y) = |x - y|$. Show that there are sequences of real numbers $(a_i)_{i\in\mathbb{N}}$ and $(b_i)_{i\in\mathbb{N}}$, such that $a_i < b_i$ for all $i \in \mathbb{N}$ and

$$U = \bigcup_{i\in\mathbb{N}} (a_i, b_i) \ ,$$

i.e. U is the union of countably many open intervals (a_i, b_i).

If (X, d) is a metric space, the subsets $U \subset X$ that are open with respect to the metric d have a number of interesting properties listed in the following theorem.

Theorem 2.2.4. *Let (X, d) be a metric space.*

a) The two sets X and $\emptyset$ are open with respect to the metric.

b) If $U, V \subset X$ are open with respect to the metric, then $U \cap V$ is open with respect to the metric as well.

c) Let I be a set and let $U_i \subset X$ for $i \in I$ be a family of subsets, which are open with respect to the metric. Then the union

$$U = \bigcup_{i\in I} U_i$$

is open with respect to the metric as well.

Proof. The statement in a) about X is obviously true. Since there are no points in the empty set $\emptyset$, there is nothing to check. Hence, it is open with respect to the metric.

For the proof of b) let $x \in U \cap V$ (if $U \cap V = \emptyset$ we are done). Since $x \in U$ and U is open with respect to the metric there is $r_1 > 0$ such that $B_{r_1}(x) \subset U$. Likewise, there exists $r_2 > 0$ such that $B_{r_2}(x) \subset V$. Let $r = \min\{r_1, r_2\}$. Then we have $B_r(x) \subset B_{r_i}(x)$ for $i \in \{1, 2\}$. Hence, $B_r(x) \subset U \cap V$.

For the proof of c) let $x \in U$. It is contained in at least one U_i. Since U_i is open with respect to the metric, there is $r > 0$ such that $B_r(x) \subset U_i \subset U$ and U is open with respect to the metric. □

Remark 2.2.5. It is important to note that point a) in Theorem 2.2.4 does *not* imply that arbitrary intersections of sets which are open with respect to the metric are open again. To see this, let $a, b \in \mathbb{R}$ with $a < b$ and consider the family of open intervals

$$V_n = \left(a - \tfrac{1}{n}, b + \tfrac{1}{n}\right) \ .$$

Then we have

$$\bigcap_{n \in \mathbb{N}} V_n = [a, b] \ ,$$

which is not an open interval anymore. However, Theorem 2.2.4 *does* imply that finite intersections of sets which are open with respect to the metric are open again. This can be checked using induction and it is a good exercise to do so!

Families of sets which satisfy the properties stated in Theorem 2.2.4 will become important in everything that follows. We therefore make the following definition.

Definition 2.2.6. Let X be a set and denote by $\mathcal{P}(X)$ the power set of X, i.e. the set of all subsets of X. A collection $\mathcal{T} \subset \mathcal{P}(X)$ of subsets of X is called a ***topology on*** X (or just a topology if the underlying set is clear) if it has the following properties:

a) $X \in \mathcal{T}$ and $\emptyset \in \mathcal{T}$.

b) If $U, V \in \mathcal{T}$, then $U \cap V \in \mathcal{T}$.

c) If I is a set and $U_i \in \mathcal{T}$ for each $i \in I$, then $\bigcup_{i \in I} U_i \in \mathcal{T}$.

A pair $(X, \mathcal{T})$ of a set X and a topology $\mathcal{T}$ on X is called a ***topological space***. If the topology that we mean is clear, we will often just talk about the topological space X, or even just the space X. We will call a subset $U \subset X$ with $U \in \mathcal{T}$ ***open with respect to the topology*** $\mathcal{T}$. If the topology that we mean is clear from the context, we will just say that U is an open subset of X or we will say "U is open in X".

Example 2.2.7. Let (X, d) be a metric space and let $\mathcal{T}(d)$ be the family of all subsets of X that are open with respect to the metric d. As we have seen in Theorem 2.2.4 the collection $\mathcal{T}(d)$ of subsets is a topology on X. Hence, the pair $(X, \mathcal{T}(d))$ is a topological space. We will sometimes call $\mathcal{T}(d)$ the ***metric topology***. Note that $U \subset X$ is open with respect to the metric d if and only if $U \in \mathcal{T}(d)$, i.e. U is open with respect to the metric topology.

Given a set X, we can always turn it into a topological space using one of the following two examples.

Example 2.2.8. Let X be a set and let $\mathcal{T}_{\text{ind}} = \{\emptyset, X\}$, i.e. $\mathcal{T}_{\text{ind}}$ is the family of subsets of X, which only consists of X itself and the empty set. This has the properties (a)–(c) from Def. 2.2.6 and therefore is a topology on X. It is called the ***indiscrete topology on*** X.

Exercise 2.2.9. Let X be a set and consider the topological space $(X, \mathcal{T}_{\text{ind}})$, where $\mathcal{T}_{\text{ind}}$ is the indiscrete topology from the example above. Let (Y, d_Y) be a metric space and consider it as a topological space equipped with the metric topology $\mathcal{T}(d_Y)$. Show that any continuous map $f\colon X \to Y$ has to be constant.

Example 2.2.10. Let X be a set and consider the family $\mathcal{T}_{\text{dis}} = \mathcal{P}(X)$ given by the power set of X, i.e. the family of all subsets of X. This also has the properties (a)–(c) from Def. 2.2.6 and defines a topology, which is called the ***discrete topology on*** X. By definition every subset of X is open in this topology.

To understand how open sets are related to continuity we need the definition of the preimage of a set with respect to a map. Let $f\colon X \to Y$ be an arbitrary map between sets (for example a map between metric spaces) and let $V \subset Y$ be an arbitrary subset of Y. The ***preimage*** $f^{-1}(V) \subset X$ of V under f is defined to be

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

i.e. it consists of all points in X that are mapped into V by f. Even though the notation of the preimage is similar to that of the inverse map, it is very important to note that they are completely different beasts: The inverse map only exists, if the map f is bijective, whereas the preimage always exists and yields a set instead of just a single point.

Of course, if the map f is bijective, then $f^{-1}(\{y\})$ – the preimage of the set containing just the element y – agrees with the set $\{f^{-1}(y)\}$, i.e. the set containing just the image of y under the inverse map.

Exercise 2.2.11. Let $f\colon X \to Y$ be as above and let $V_1, V_2 \subset Y$ be two subsets. Check that:

$$f^{-1}(V_1 \cup V_2) = f^{-1}(V_1) \cup f^{-1}(V_2) \quad \text{and} \quad f^{-1}(V_1 \cap V_2) = f^{-1}(V_1) \cap f^{-1}(V_2) \ . \qquad (2)$$

What is $f^{-1}(\emptyset)$? What do we get for $f^{-1}(Y)$?

We are now ready to define what we mean by a continuous map between two topological spaces.

Definition 2.2.12. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. A map $f\colon X \to Y$ is called ***continuous*** if we have

$$f^{-1}(U) \in \mathcal{T}_X \quad \text{for all } U \in \mathcal{T}_Y \ .$$

In other words: The map f is continuous if for any subset U of Y which is open with respect to the topology $\mathcal{T}_Y$ its preimage $f^{-1}(U)$ is open with respect to the topology $\mathcal{T}_X$.

Remark 2.2.13. If we allow ourselves to be a little sloppy, we could shorten the above definition to: A map $f\colon X \to Y$ is continuous if the preimages of open sets in Y are open in X.

Before we discuss why Definition 2.2.12 is sensible, let us consider the following special case: Let (X, d_X) and (Y, d_Y) be metric spaces and let $f\colon X \to Y$ be a map between them. We now have two definitions for the continuity of f:

1) Since X and Y are metric spaces, we can use the ϵ-δ-definition of continuity from Def. 2.1.5.

2) We can also consider the two topologies $\mathcal{T}(d_X)$ on X and $\mathcal{T}(d_Y)$ on Y and use Def. 2.2.12.

We would be in deep trouble if these two definitions were not consistent. Luckily, we have the following theorem:

Theorem 2.2.14. *Let (X, d_X), (Y, d_Y) be metric spaces. A map $f\colon X \to Y$ is continuous according to Def. 2.1.5, if and only if it is continuous as a map of the corresponding topological spaces $(X, \mathcal{T}(d_X))$ and $(Y, \mathcal{T}(d_Y))$, i.e. continuous according to Def. 2.2.12.*

Proof. Let us first assume that f is continuous according to Def. 2.2.12. This means that $f^{-1}(U) \subset X$ is open for every open set $U \subset Y$. Let $x \in X$ and let $\epsilon > 0$ and consider

$$V = B_\epsilon(f(x)) = \{y \in Y \mid d_Y(y, f(x)) < \epsilon\} \ .$$

The set V is an element in $\mathcal{T}(d_Y)$ and therefore an open set in Y. Thus, $f^{-1}(V)$ is open in X by hypothesis. Since we have $x \in f^{-1}(V)$, there is $\delta > 0$ such that $B_\delta(x) \subset f^{-1}(V)$ by the definition of the metric topology $\mathcal{T}(d_X)$. Hence, we have found for every $\epsilon > 0$ a $\delta > 0$ such that $x' \in B_\delta(x)$ implies $f(x') \in V$. Using the definition of $B_\delta(x)$ and V, we see that we have found for every $\epsilon > 0$ a number $\delta > 0$ such that $d_X(x', x) < \delta$ implies $d_Y(f(x'), f(x)) < \epsilon$. Therefore, f is continuous with respect to Definition 2.1.5.

To prove the other direction let us assume that f is continuous according to Definition 2.1.5. Let $U \subset Y$ be an open set. We have to show that $f^{-1}(U)$ is open in X. If $f^{-1}(U)$ is empty, it is open. Hence, we can assume $f^{-1}(U) \neq \emptyset$. Let $x \in f^{-1}(U)$ be a point in the preimage. To see that $f^{-1}(U)$ is open we have to find a radius $r > 0$, such that $B_r(x) \subset f^{-1}(U)$. Since U is open and $f(x) \in U$, there is a radius $\epsilon > 0$, such that $B_\epsilon(f(x)) \subset U$. Since f is continuous with respect to Definition 2.1.5, there exists a $\delta > 0$ corresponding to this ϵ, such that for all $x' \in B_\delta(x)$ we have that $f(x') \in B_\epsilon(f(x)) \subset U$. In particular, $B_\delta(x) \subset f^{-1}(U)$ and we can choose $r = \delta$. Since $x \in f^{-1}(U)$ was arbitrary, $f^{-1}(U)$ is a subset of X that is open with respect to $\mathcal{T}(d_X)$. □

Example 2.2.15. To illustrate how one can apply Definition 2.2.12 directly to check the continuity or non-continuity of a map, consider the topological space $(\mathbb{R}, \mathcal{T}(d))$ with the metric topology $\mathcal{T}(d)$ induced by $d(x, y) = |x - y|$. It follows from the above theorem that maps $f\colon \mathbb{R} \to \mathbb{R}$ that are continuous according to Definition 2.2.12 are continuous with respect to Definition 2.0.1. Consider the function $f\colon \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} x^2 & \text{for } x \leq 4 \ , \\ 3x - 8 & \text{for } 4 < x \leq 6 \ , \\ \frac{1}{2}x & \text{for } 6 < x \ . \end{cases} \tag{3}$$

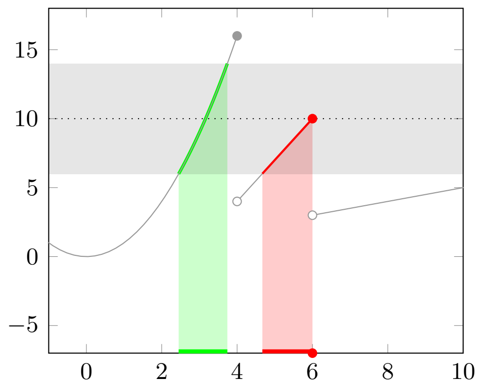


Figure 12: The graph of the non-continuous function f from (3).

It is *not* continuous. The graph of f is shown in Figure 12. According to Definition 2.2.12 there has to be at least one open set $U \subset \mathbb{R}$, such that $f^{-1}(U)$ is not open. Consider the open interval $U = (6, 14)$. The grey area in Figure 12 encloses all the points, such that $f(x) \in U$. We have

$$f^{-1}(U) = \left(\sqrt{6}, \sqrt{14}\right) \cup \left(\tfrac{14}{3}, 10\right] \ .$$

This set can also be found in Figure 12. It is the union of the red and green interval on the x-axis. In particular, $f^{-1}(U)$ is the union of two disjoint intervals, one of which is not open. It is a good exercise to check that the set $f^{-1}(U)$ is in fact also not open.

As we can see, a point where the function jumps creates a "loose end" like the red point in the graph, which will lead to a non-open preimage of an open set.

The composition of two continuous maps is again continuous as the following theorem shows.

Theorem 2.2.16. *Let $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ and $(Z, \mathcal{T}_Z)$ be three topological spaces and let $g\colon X \to Y$ and $f\colon Y \to Z$ be continuous maps. Then the composition $f \circ g\colon X \to Z$ is also continuous.*

Proof. Let $U \subset Z$ be an open subset. We have to check that $(f \circ g)^{-1}(U)$ is open in X. But we have

$$(f \circ g)^{-1}(U) = \{x \in X \mid f(g(x)) \in U\} = \left\{x \in X \mid g(x) \in f^{-1}(U)\right\} = g^{-1}(f^{-1}(U))$$

and $f^{-1}(U)$ is open in Y, because f is continuous. Hence, $g^{-1}(f^{-1}(U))$ is open in X, since g is continuous. □

2.3 Some topological vocabulary

In this section we gather some terminology concerning topological spaces. This is meant to be a vocabulary of concepts that will become important in later chapters. In all of the subsections $(X, \mathcal{T}_X)$ will always denote a topological space.

2.3.1 Closed sets and neighbourhoods

As mentioned in Definition 2.2.6 we call the elements $U \in \mathcal{T}_X$ *open* in X. There are two closely related notions, which we define now.

Definition 2.3.1. Let $(X, \mathcal{T}_X)$ be a topological space. A set $A \subset X$ is called ***closed*** if $X \setminus A \in \mathcal{T}_X$. In other words, a subset $A \subset X$ is closed if the complement $X \setminus A$ is an open set.

Remark 2.3.2. Sets are not doors! The notions *open* and *closed* are not mutually exclusive. On the contrary: Every topological space contains at least two subsets, which are closed *and* open, namely the whole space X and the empty set. We will see later, when we talk about connectedness of topological spaces, that there can be even more subsets, which are closed and open. Often there are also a lot of subsets, which are neither closed nor open. Consider for example $(\mathbb{R}, \mathcal{T}(d))$ with $d(x, y) = |x - y|$. The interval $(3, 4] \subset \mathbb{R}$ is neither closed nor open.

Exercise 2.3.3. Let $(X, \mathcal{T}_X)$ be a topological space. Show the following facts about closed sets in X using the axioms (a)–(c) from Definition 2.2.6:

a) The empty set $\emptyset$ and the whole space X are closed.

b) If $A \subset X$ and $B \subset X$ are closed, so is their union $A \cup B$.

c) Let I be a set and let $(A_i)_{i \in I}$ be a family of closed subsets of X. Then $\bigcap_{i \in I} A_i$ is closed as well.

Hint: Use De Morgan's laws.

Definition 2.3.4. Let $(X, \mathcal{T}_X)$ be a topological space and let $x \in X$ be a point. A ***neighbourhood of*** x ***in*** X is a subset $U \subset X$, such that there exists an open set $V \subset X$ with $x \in V \subset U$.

Example 2.3.5. Consider the space $\mathbb{R}^2$ as a topological space equipped with the metric topology and let $x_0 = (0, 0) \subset \mathbb{R}^2$ be the origin. The set

$$U = \left\{ x \in \mathbb{R}^2 \mid \|x\| \leq \frac{3}{2} \right\}$$

is a neighbourhood of x_0, since it contains the set $B_1(0, 0) = \{x \in \mathbb{R}^2 \mid \|x\| < 1\}$, which is open in the metric topology by definition. This situation is shown in Figure 13.

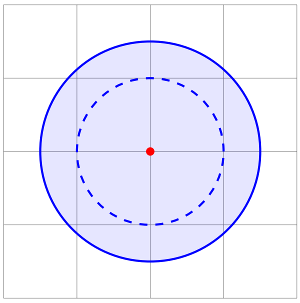


Figure 13: A disk, which is a neighbourhood of the point marked in red in $\mathbb{R}^2$.

Example 2.3.6. Consider the topological space $\mathbb{R}$ equipped with the metric topology. The interval $(-2,2) \subset \mathbb{R}$ is a neighbourhood of $\{0\} \subset \mathbb{R}$, since it is open and contains 0. However, the interval $[0,1) \subset \mathbb{R}$ is *not* a neighbourhood of $\{0\}$. If it were, there would be a set $V \subset \mathbb{R}$, which is open with respect to the metric, contained in $[0,1)$ and such that $0 \in V$. This would imply that V also contains an open interval around 0, in particular we would have $-\epsilon \in V$ for some $\epsilon > 0$, which is a contradiction.

Example 2.3.7. Let $(X, \mathcal{T}_{\text{dis}})$ be a set equipped with the *discrete topology* $\mathcal{T}_{\text{dis}}$ defined in Example 2.2.10. Let $x \in X$ be an arbitrary point in X. Note that the subset $U = \{x\} \subset X$ is open in X, since every set is open in the discrete topology. In particular, U is a neighbourhood of x in X.

This is of course different for $(\mathbb{R}^n, \mathcal{T}(d))$ with the *metric topology* associated to the metric $d(x,y) = \|x-y\|$. The set $U = \{x\}$ is not open in the topology $\mathcal{T}(d)$ (Why?) and the only open set contained in U is the empty set, which in turn does not contain x. Hence, U is not a neighbourhood of x in $(\mathbb{R}^n, \mathcal{T}(d))$.

2.3.2 Non-metrisable topological spaces

Theorem 2.2.14 only applies to maps between metric spaces. It is important to note that there are topological spaces, which can not be equipped with a metric inducing the topology. These are called ***non-metrisable topological spaces***. Hence, the notion of topological spaces is more general than that of metric spaces. We have already met one example.

Example 2.3.8. Let X be a set with at least two elements. Consider the topological space $(X, \mathcal{T}_{\text{ind}})$, where $\mathcal{T}_{\text{ind}}$ is the indiscrete topology defined in Example 2.2.8. Then we have

Lemma 2.3.9. *There is no metric $d\colon X \times X \to \mathbb{R}$, such that $\mathcal{T}(d) = \mathcal{T}_{ind}$. Therefore $(X, \mathcal{T}_{ind})$ is a non-metrisable topological space.*

Proof. To arrive at a contradiction assume that $d\colon X \times X \to \mathbb{R}$ exists, such that $\mathcal{T}(d) = \mathcal{T}_{\text{ind}}$. Let $x_0 \in X$ and consider the set $U = X \setminus \{x_0\}$. This is non-empty by assumption. Let $y \in U$. Since $y \neq x_0$, we have $d = d(x_0, y) > 0$, hence $x_0 \notin B_{d/2}(y_0)$. We have therefore shown that $U \subset X$ is open with respect to the metric d (or in other words: $U \in \mathcal{T}(d)$), but $U \notin \mathcal{T}_{\text{ind}}$. This is a contradiction. □

2.3.3 Convergence of sequences in topological spaces

In the module "Foundations of Mathematics I" you have dealt with the convergence of sequences of real (and even complex numbers): A sequence $(a_n)_{n\in\mathbb{N}}$ of real numbers converges to a limit point $a \in \mathbb{R}$ if for every $\epsilon > 0$ there is an $N \in \mathbb{N}$, such that $|a_n - a| < \epsilon$ for all $n > N$. Just like the continuity of maps, convergence is a notion, for which the actual value of the distance $|a_n - a|$ does not really matter. We just want it to become arbitrarily small. Therefore, it should come as no surprise that we can also define convergence of sequences in topological spaces.

Definition 2.3.10. Let $(X, \mathcal{T}_X)$ be a topological space and let $(a_n)_{n\in\mathbb{N}}$ with $a_n \in X$ be a sequence of points in X. We say that a_n ***converges to*** $a \in X$ if for every neighbourhood U of a in X there is an $N \in N$, such that $a_n \in U$ for all $n > N$.

Example 2.3.11. Let $(X, \mathcal{T}_{\text{dis}})$ be a set equipped with the discrete topology defined in Example 2.2.10. Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in X and suppose that it converges to a point $a \in X$. As explained in Example 2.3.7 the set $U = \{a\}$ containing just the point a is a neighbourhood of a in X. The condition for convergence now has to hold for this particular U. But this means that there is an $N \in \mathbb{N}$, such that $a_n \in \{a\}$ for all $n > N$, which is equivalent to $a_n = a$ for all $n > N$. Therefore a sequence that converges in a topological space with the discrete topology eventually becomes constant.

Exercise 2.3.12. Consider the set $\mathbb{N}_+ = \mathbb{N} \cup \{\infty\}$, i.e. the set of all natural numbers plus an additional point, which we call ∞. We define a topology $\mathcal{T}_+$ on $\mathbb{N}_+$ in the following way: All subsets of $\mathbb{N}$ are in $\mathcal{T}_+$ and all sets that contain ∞ and a complement of a finite set are in $\mathcal{T}_+$. For example, the set $\{n \in \mathbb{N} \mid n > 5\} \cup \{\infty\}$ is in $\mathcal{T}_+$, while the sets $\{2k \mid k \in \mathbb{N}\} \cup \{\infty\}$ and $\{1, 2, 3, \infty\}$ are not.

a) Show that $\mathcal{T}_+$ is in fact a topology on $\mathbb{N}_+$.

b) Show that the inclusion $\iota \colon \mathbb{N} \to \mathbb{N}_+$ is continuous.

c) Let $(Y, \mathcal{T}_Y)$ be another topological space and let $(a_n)_{n\in\mathbb{N}}$ be a sequence of points in Y. It gives rise to a continuous map $f \colon \mathbb{N} \to Y$ with $f(n) = a_n$. Show that the sequence converges if and only if there is a continuous map $F \colon \mathbb{N}_+ \to Y$, such that $F \circ \iota = f$.

The space $\mathbb{N}_+$ is also called the *one-point compactification of* $\mathbb{N}$.

2.3.4 Bases

Consider $\mathbb{R}^n$ and $\mathbb{R}^m$ as topological spaces equipped with their respective metric topologies. Let $f \colon \mathbb{R}^n \to \mathbb{R}^m$ be a map between them. According to Def. 2.2.12, f is continuous if and only if $f^{-1}(U)$ is open in $\mathbb{R}^n$ for every open set $U \subset \mathbb{R}^m$. Now observe that an arbitrary open set $U \subset \mathbb{R}^m$ can be written as the union of open balls $B_r(x)$ for suitable radii $r > 0$ and points $x \in U$ as in Def. 2.2.1. Indeed, by definition U is open if for every $x \in U$ we have a radius r_x, such that $B_{r_x}(x) \subset U$. Hence,

$$U = \bigcup_{x \in U} B_{r_x}(x) \ .$$

By (2) it therefore suffices to check that $f^{-1}(B_r(x))$ is open for any $r > 0$ and any $x \in X$ to deduce that $f^{-1}(U)$ is open for any open set $U \subset \mathbb{R}^m$. The feature of the family

$$\mathcal{B} = \{B_r(x) \subset \mathbb{R}^m \mid r > 0, x \in X\} \tag{4}$$

that we used here was that any open set $U \subset \mathbb{R}^m$ can be written as a *union* of sets in $\mathcal{B}$. This seems to make our life a lot easier. Therefore these families deserve a name:

Definition 2.3.13. Let $(X, \mathcal{T})$ be a topological space. A family $\mathcal{B} \subset \mathcal{T}$ is called a ***basis of the topology*** $\mathcal{T}$ if every open set $U \in \mathcal{T}$ is a union of elements of $\mathcal{B}$.

Example 2.3.14. As we have seen in the first paragraph, the family $\mathcal{B}$ in (4) is a basis of the metric topology $\mathcal{T}(d)$ on $\mathbb{R}^m$ with respect to the metric $d(x, y) = \|x - y\|$. In particular, we obtain that the family of open intervals

$$\mathcal{B} = \{(a, b) \subset \mathbb{R} \mid a, b \in \mathbb{R}, a < b\}$$

forms a basis for the metric topology on $\mathbb{R}$.

Example 2.3.15. The basis of a topology is not unique. If we again look at $\mathbb{R}$ equipped with the metric topology $\mathcal{T}(d)$ induced by the metric $d(x, y) = |x - y|$, we find that the family of all intervals with *rational* endpoints, i.e.

$$\mathcal{B}' = \{(a, b) \subset \mathbb{R} \mid a, b \in \mathbb{Q}, a < b\}$$

is also a basis for this topology. (Note that this is a countable set.)

Remark 2.3.16. Even though the terminology is the same, the basis of a topology has nothing to do with the basis of a vector space! The only similarity is that both describe a subset of elements generating the whole object.

If we know already that $\mathcal{B}$ is the basis of a topology $\mathcal{T}$ on a set X, then we can retrieve $\mathcal{T}$ from $\mathcal{B}$ by taking all possible unions of the elements of $\mathcal{B}$. This is equivalent to saying that $\mathcal{T}$ consists of all subsets $U \subset X$ with the property that for every $x \in U$, there is a $B \in \mathcal{B}$, such that $x \in B \subset U$. But what if we start with an arbitrary family of subsets $\mathcal{B}$ and take all possible unions of elements in $\mathcal{B}$ and call the result $\mathcal{T}_{\mathcal{B}}$? What are the conditions that ensure that $\mathcal{T}_{\mathcal{B}}$ is indeed a topology? This is what the next lemma is about.

Lemma 2.3.17. *Let X be a set and let $\mathcal{B}$ be a family of subsets of X. Let*

$$\mathcal{T}_{\mathcal{B}} = \{U \subset X \mid \textit{for all } x \in U \textit{ there exists } B \in \mathcal{B} \textit{ such that } x \in B \subset U\} \ .$$

Then $\mathcal{T}_{\mathcal{B}}$ is a topology with basis $\mathcal{B}$ if and only if $\mathcal{B}$ has the following two properties:

a) For each $x \in X$ there exists a $B \in \mathcal{B}$ with $x \in B$.

b) If $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$ then there exists a $B_3 \in \mathcal{B}$ with $x \in B_3 \subset B_1 \cap B_2$.

Proof. Let us first assume that $\mathcal{T}_\mathcal{B}$ is a topology with basis $\mathcal{B}$ and prove that a) and b) hold in this case: Since $\mathcal{T}_\mathcal{B}$ is a topology, we have $X \in \mathcal{T}_\mathcal{B}$. By the definition of $\mathcal{T}_\mathcal{B}$ this means that for all $x \in X$, there is a $B \in \mathcal{B}$ with $x \in B$, which is property a). Let $B_1, B_2 \in \mathcal{B}$. Note that B_1, B_2 are then also in $\mathcal{T}_\mathcal{B}$. Since $\mathcal{T}_\mathcal{B}$ is a topology by assumption, this implies that $B_1 \cap B_2 \in \mathcal{T}_\mathcal{B}$. But by definition of $\mathcal{T}_\mathcal{B}$ we know that there exists a $B_3 \in \mathcal{B}$ with $x \in B_3 \subset B_1 \cap B_2$, which is property b).

Let us now prove the converse statement, i.e. we assume that $\mathcal{B}$ is a family of subsets of X with properties a) and b) and want to show that $\mathcal{T}_\mathcal{B}$ is a topology. Property a) makes sure that $X \in \mathcal{T}_\mathcal{B}$. The empty set $\emptyset$ is also in $\mathcal{T}_\mathcal{B}$. To see this, note that we have to check the condition "for all $x \in \emptyset$ there exists ...", but since there is no $x \in \emptyset$, the statement is automatically true. Hence, $\mathcal{T}_\mathcal{B}$ satisfies condition a) in Def. 2.2.6.

Let $U_1, U_2 \in \mathcal{T}_\mathcal{B}$. We need to see that $U_1 \cap U_2 \in \mathcal{T}_\mathcal{B}$. If this is empty, we are done. Otherwise, let $x \in U_1 \cap U_2$. Since $U_1 \in \mathcal{T}_\mathcal{B}$, there is a $B_1 \in \mathcal{B}$, such that $x \in B_1 \subset U_1$. Likewise, there is a $B_2 \in \mathcal{B}$, such that $x \in B_2 \subset U_2$. In particular, $x \in B_1 \cap B_2$. By property b) in the statement of the lemma there is a $B_3 \in \mathcal{B}$ with $x \in B_3 \subset B_1 \cap B_2$. Altogether, we have $x \in B_3 \subset U_1 \cap U_2$, which shows $U_1 \cap U_2 \in \mathcal{T}_\mathcal{B}$. This is condition b) in Def. 2.2.6.

Let $(U_i)_{i \in I}$ be a family of sets indexed by a set I with $U_i \in \mathcal{T}_\mathcal{B}$ for all $i \in I$. Let $U = \bigcup_{i \in I} U_i$. Take a point $x \in U$. It will be in at least one set U_j for some $j \in I$. Since $U_j \in \mathcal{T}_\mathcal{B}$, we know that there exists $B \in \mathcal{B}$ with $x \in B \subset U_j \subset U$. We started with an arbitrary point $x \in U$, hence this implies that $U \in \mathcal{T}_\mathcal{B}$. We have proven that $\mathcal{T}_\mathcal{B}$ satisfies condition c) in Def. 2.2.6 and therefore defines a topology on X. □

Exercise 2.3.18. Let $X = \mathbb{R}^n$ let $\mathcal{B} = \{X\}$ be the family that contains only the element X.

a) Check that this family $\mathcal{B}$ is the basis of a topology on X.

b) We already know this topology. Which topology $\mathcal{T}_\mathcal{B}$ on $X = \mathbb{R}^n$ do you get this way?

Exercise 2.3.19. Let X be a set and let $\mathcal{B} = \{\{x\} \mid x \in X\}$ be the family of subsets of X, which contains only one-point sets. (These are sometimes also called "singletons").

a) Check that this family $\mathcal{B}$ is the basis of a topology $\mathcal{T}_\mathcal{B}$ on X.

b) We also know this topology already. Which one is it?

2.3.5 Homeomorphisms

Now that we have understood the definition of continuous maps between topological spaces, we can introduce the following equivalence relation.

Definition 2.3.20. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be two topological spaces. We say that they are ***homeomorphic*** if there are continuous maps $f\colon X \to Y$ and $g\colon Y \to X$, such that

$$f \circ g = \mathrm{id}_Y \qquad \text{and} \qquad g \circ f = \mathrm{id}_X \ . \tag{5}$$

In this case we call f and g ***homeomorphisms***. More precisely, we say that a continuous map $f\colon X \to Y$ ***is a homeomorphism*** if there is a continuous map $g\colon Y \to X$, such that (5) holds.

Theorem 2.3.21. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and let $f\colon X \to Y$ be a homeomorphism. Then $U \subset Y$ is open if and only if $f^{-1}(U) \subset X$ is open.*

Proof. Since f is continuous, we have that $f^{-1}(U)$ is open in X for any open subset $U \subset Y$. Let us prove the other direction: If $f^{-1}(U)$ is open in X, then U is open in Y. Since f is a homeomorphism, there is a continuous map $g\colon Y \to X$, such that (5) holds. Similar to the proof of Theorem 2.2.16 we have

$$U = \mathrm{id}_Y^{-1}(U) = (f \circ g)^{-1}(U) = g^{-1}(f^{-1}(U)) \ .$$

We assumed that $f^{-1}(U)$ is open and we have that g is continuous, hence $g^{-1}(f^{-1}(U))$ is open as well, but this subset agrees with U. □

Remark 2.3.22. As we can see from Theorem 2.3.21 a homeomorphism $f\colon X \to Y$ is a continuous map, that establishes a 1 : 1-correspondence between the two topologies $\mathcal{T}_Y$ and $\mathcal{T}_X$ via $U \mapsto f^{-1}(U)$. We will see an explicit construction of a homeomorphism in Example 3.1.12 below.

In Linear Algebra you look at questions of the form: How many vector spaces are there up to isomorphism? How do we distinguish the isomorphism classes? Similarly, we can ask the analogous question in Topology: How many topological spaces are there up to homeomorphism? How do we recognise spaces that are not homeomorphic?

2.3.6 A network as a topological space

Most of the examples of topological spaces we have seen so far were metric spaces. In this section we will give an example of a non-metrisable topological space, which can be used to describe a computer network. We will see that the topology indeed gives us information about the connected components and also allows us to describe loops in the network as continuous maps $\gamma\colon [0, 1] \to X$ such that $\gamma(0) = \gamma(1)$.

Consider the computer network sketched in Figure 14. It consists of four computers $c_1, \ldots, c_4$ and three of them are connected by links l_{12}, l_{13}, l_{23}. To write down a topological space that describes this setup we first have to fix an underlying set X. In our case, this set consists of all computers and all connections:

$$X = \{c_1, c_2, c_3, c_4, l_{12}, l_{23}, l_{13}\}$$

Next we have to come up with a topology on this set X, i.e. we need to give a family $\mathcal{T}_X$ of subsets $U \subset X$, which we declare to be open. In the picture the links between the

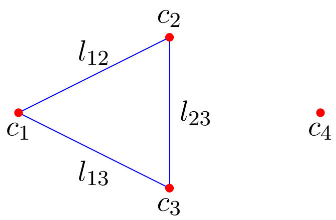


Figure 14: A computer network with four computers c_i and three links l_{jk}.

computers are drawn as straight (blue) lines. In reality they could be wireless connections or cables. We do not care about their nature. However, we want the connections to resemble open intervals. We also want an open neighbourhood of a computer c_i to consist of the computer itself and all of its outgoing links. If we write down a wish list of open sets it therefore looks like this:

$$\mathcal{B}_X = \{\{c_1, l_{12}, l_{13}\}, \{c_2, l_{12}, l_{23}\}, \{c_3, l_{13}, l_{23}\}, \{c_4\}, \{l_{12}\}, \{l_{13}\}, \{l_{23}\}\}$$

Of course this is not a topology on X. For example, the whole set X is not contained in it, neither is the empty set. However, it satisfies the properties stated in Lemma 2.3.17 that reveal it as the basis of a topology on X. Let us check this: We see that each element of X appears in at least one set in $\mathcal{B}_X$, hence it satisfies a) in Lemma 2.3.17. The intersection of two different sets in $\mathcal{B}_X$ is either empty or of the form $\{l_{jk}\}$ for suitable choices of j and k. But these sets are in $\mathcal{B}_X$. Hence, it satisfies b) in Lemma 2.3.17.

Using this lemma we obtain a topology $\mathcal{T}_X$ on X. A set U is open with respect to this topology if it can be written as a union of sets from $\mathcal{B}_X$. For example, the set

$$U = \{c_1, c_2, l_{12}, l_{13}, l_{23}, c_4\} = \{c_1, l_{12}, l_{13}\} \cup \{c_2, l_{12}, l_{23}\} \cup \{c_4\}$$

is open. The complement of this set is $\{c_3\}$, which is then closed by Definition 2.3.1. In a similar way one can show that $\{c_1\}$ and $\{c_2\}$ are also closed. The sets $\{c_1\}$, $\{c_2\}$ and $\{c_3\}$ are not open, since any open subset $U \in \mathcal{T}_X$ with $c_1 \in U$ has to contain at least one element $V \in \mathcal{B}_X$ with $c_1 \in V$. But then we must have $V = \{c_1, l_{12}, l_{13}\}$. In particular, V will also contain l_{12}, hence U will as well.

Since $\{c_1\} \subset X$ is not open, we see that $\mathcal{T}_X$ is not the discrete topology $\mathcal{T}_{\text{dis}}$ on X (remember that $\mathcal{T}_{\text{dis}}$ was the topology, where every set was open). Let us work out, which subsets of X are open and closed. The set $\{c_4\}$ is open by definition, since $\{c_4\} \in \mathcal{B}_X \subset \mathcal{T}_X$. It is also closed, since its complement

$$U_{123} = \{c_1, c_2, c_3, l_{12}, l_{13}, l_{23}\} = \{c_1, l_{12}, l_{13}\} \cup \{c_2, l_{12}, l_{23}\} \cup \{c_3, l_{13}, l_{23}\}$$

can be written as a union of sets from $\mathcal{B}_X$ and is therefore open. Thinking about this a bit further we see that the only subsets of X that are open and closed are

$$X, \emptyset, U_{123} \text{ and } \{c_4\} \ .$$

Note that the network has two connected components that consist precisely of the sets $\{c_4\}$, which is not connected to anything else, and the rest of the network, i.e. U_{123}. This is not a coincidence as we will see in Section 4.

In the beginning of this section we claimed that this was another example of a non-metrisable topological space. The reason for this is roughly because in a metric space we can always separate two points by disjoint open sets, whereas we can not do this in our space X. To make this precise consider first a metric space (Y, d). Let $y_1, y_2 \in Y$ be two points with $y_1 \neq y_2$. Then we have $d(y_1, y_2) > 0$. For $i \in \{1, 2\}$ define

$$U_i = B_{\frac{1}{4}d(y_1,y_2)}(y_i) \subset X \ .$$

Then we have $y_i \in U_i$. Let $y \in U_1 \cap U_2$. This means that $d(y, y_i) < \frac{1}{4}d(y_1, y_2)$, which implies

$$d(y_1, y_2) \leq d(y_1, y) + d(y, y_2) < \frac{1}{2}d(y_1, y_2) \ .$$

This is a contradiction since we have $d(y_1, y_2) \geq 0$. Therefore there can be no such point and $U_1 \cap U_2 = \emptyset$. To summarise: For any two distinct points $y_1, y_2 \in Y$ in a metric space (Y, d) we can always find open sets $U_1 \subset Y$ and $U_2 \subset Y$ with $y_1 \in U_1$, $y_2 \in U_2$ and $U_1 \cap U_2 = \emptyset$.

Now let us check if that can be done in the space X: Consider the two distinct points $c_1 \in X$ and $c_2 \in X$. As we have already worked out above, any open set that contains c_1 also has to contain l_{12}. But the same is true for c_2. Hence, the intersection of two open sets U_1 and U_2 with $c_1 \in U_1$ and $c_2 \in U_2$ will never be empty. This shows that there is no metric $d\colon X \times X \to \mathbb{R}$ such that the metric topology $\mathcal{T}(d)$ agrees with $\mathcal{T}_X$.

For a computer scientist it might be interesting to know if there are any loops in the network: In our example space X there are two connections from c_1 to c_3 - a direct one and another one via c_2 following first the link l_{12} and then l_{23}. In such a situation it might be important to check if data that comes via the two different paths is consistent. By following first one connection from c_1 to c_3 and then the other connection back to c_1 we obtain a loop in X. How do we describe this topologically? A loop in a topological space should be given by a path in X that starts and ends at the same point. Let us try, if we can write down a continuous map $\gamma\colon [0,1] \to X$ with $\gamma(0) = \gamma(1) = c_1$, where the topology on the unit interval $[0,1]$ should be the metric topology induced by the metric $d\colon [0,1] \times [0,1] \to \mathbb{R}$ given by $d(x,y) = |x - y|$. The best we can do in this case is a map like this:

$$\gamma\colon [0,1] \to X \quad ; \quad t \mapsto \begin{cases} c_1 & \text{if } t = 0 \text{ or } t = 1 \\ l_{12} & \text{if } 0 < t < \frac{1}{3} \\ c_2 & \text{if } t = \frac{1}{3} \\ l_{23} & \text{if } \frac{1}{3} < t < \frac{2}{3} \\ c_3 & \text{if } t = \frac{2}{3} \\ l_{13} & \text{if } \frac{2}{3} < t < 1 \end{cases}$$

A sketch of what this map does is given in Figure 15. If we let t go from 0 to 1 the point $\gamma(t)$ certainly jumps once around the loop in Figure 14. But this does not look very continuous. We will see that in fact our topology on X was chosen in such a way that this map is continuous. As we have seen above, the family $\mathcal{B}_X$ is a basis for the topology $\mathcal{T}_X$. Therefore we just need to check that $\gamma^{-1}(U)$ is open in $[0,1]$ for every $U \in \mathcal{B}_X$. For the subsets of the form $\{l_{jk}\}$ we obtain

$$\gamma^{-1}(\{l_{12}\}) = \left(0, \frac{1}{3}\right)$$
$$\gamma^{-1}(\{l_{23}\}) = \left(\frac{1}{3}, \frac{2}{3}\right)$$
$$\gamma^{-1}(\{l_{13}\}) = \left(\frac{2}{3}, 1\right)$$

All of these preimages are open intervals. The preimages of the neighbourhoods of the computers c_i look like this:

$$\gamma^{-1}(\{c_1, l_{12}, l_{13}\}) = \left[0, \frac{1}{3}\right) \cup \left(\frac{2}{3}, 1\right]$$
$$\gamma^{-1}(\{c_2, l_{12}, l_{23}\}) = \left(0, \frac{2}{3}\right)$$
$$\gamma^{-1}(\{c_3, l_{13}, l_{23}\}) = \left(\frac{1}{3}, 1\right)$$

Note that in the first case the preimage is the union of the two open balls $B_{1/3}(0)$ and $B_{1/3}(1)$ in the metric space $[0,1]$. Therefore it is open in the metric space $([0,1], d)$, even though it is not open as a subset of $\mathbb{R}$ (compare with Remark 3.1.7 in the section about subspaces). Thus we see that $\gamma^{-1}(U)$ is open for any $U \in \mathcal{B}_X$, which implies that γ is continuous.

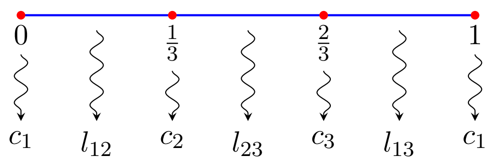


Figure 15: The map γ sends each red point to the corresponding c_i and the blue intervals inbetween to l_{jk} as shown.

3 Operations with topological spaces

Metric spaces already provide a rich source of examples for topological spaces. In this section we will learn how to create new spaces from known ones using operations like subspaces, products and quotients. We will also discuss some examples of topological spaces, like spheres S^n, disks D^n, tori $\mathbb{T}^n$ and the projective spaces $\mathbb{R}P^n$ and $\mathbb{C}P^n$, which will play a central role in everything that follows.

3.1 Subspaces of topological spaces

Let $(X, \mathcal{T}_X)$ be a topological space and let $Y \subset X$ be a subset. Let $\mathcal{T}_{Y\subset X}$ be the following family of subsets of Y:

$$\mathcal{T}_{Y\subset X} := \{U \subset Y \mid U = V \cap Y \text{ for some } V \in \mathcal{T}_X\} \ . \tag{6}$$

Lemma 3.1.1. *The family $\mathcal{T}_{Y\subset X}$ is a topology on Y.*

Proof. We have $Y = X \cap Y$ and $X \in \mathcal{T}_X$, which implies $Y \in \mathcal{T}_{Y\subset X}$. Likewise, $\emptyset = \emptyset \cap Y$ and therefore $\emptyset \in \mathcal{T}_{Y\subset X}$. Hence, $\mathcal{T}_{Y\subset X}$ satisfies condition (a) from Def. 2.2.6.

To see that it also satisfies (b), suppose that $U_1, U_2 \in \mathcal{T}_{Y\subset X}$. This means that there are sets $V_1, V_2 \in \mathcal{T}_X$, such that $U_i = V_i \cap Y$ for $i \in \{1, 2\}$. Since $\mathcal{T}_X$ is a topology on X, we have $V_1 \cap V_2 \in \mathcal{T}_X$ and $U_1 \cap U_2 = (V_1 \cap Y) \cap (V_2 \cap Y) = (V_1 \cap V_2) \cap Y$ proving (b).

To see (c) let I be a set and let $U_i \in \mathcal{T}_{Y\subset X}$ for $i \in I$. By the definition of $\mathcal{T}_{Y\subset X}$ this means that for every $i \in I$ there is $V_i \in \mathcal{T}_X$, such that $U_i = V_i \cap Y$. Since $\mathcal{T}_X$ is a topology on X, we have $\bigcup_{i\in I} V_i \in \mathcal{T}_X$. Now note that

$$\bigcup_{i\in I} U_i = \bigcup_{i\in I} (V_i \cap Y) = \left(\bigcup_{i\in I} V_i\right) \cap Y \ ,$$

which implies (c) in Def. 2.2.6. □

Definition 3.1.2. Let $(X, \mathcal{T}_X)$ be a topological space and let $Y \subset X$ be a subset. The topology $\mathcal{T}_{Y\subset X}$ defined in (6) is called the ***subspace topology on Y induced by X***. The topological space $(Y, \mathcal{T}_{Y\subset X})$ is called a ***(topological) subspace of X***. If the space X is clear from the context, we will refer to the topology $\mathcal{T}_{Y\subset X}$ just as the subspace topology.

Remark 3.1.3. Observe that the subspace topology is the topology that contains "just enough" open sets, such that the inclusion map $\iota\colon Y \to X$ is continuous. Indeed, let $V \in \mathcal{T}_X$, then we have $\iota^{-1}(V) = V \cap Y$, which is open in Y according to the definition of $\mathcal{T}_{Y\subset X}$.

Definition 3.1.4. Let $(X, \mathcal{T}_X), (Z, \mathcal{T}_Z)$ be topological spaces and let $Y \subset X$ be a subspace of X. Let $\iota_Y : Y \to X$ be the inclusion map. Let $f\colon X \to Z$ be a continuous map. We define the ***restriction of f to Y*** to be the map $f|_Y = f \circ \iota_Y : Y \to Z$. By the last remark and Thm. 2.2.16 it is continuous.

Exercise 3.1.5. Let $\mathbb{R}$ be the real line equipped with the metric $d\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ given by $d(x, y) = |x - y|$. Consider the topological space $(\mathbb{R}, \mathcal{T}(d))$ and let $\mathbb{Z} \subset \mathbb{R}$ be the subset of integers. Show that the subspace topology $\mathcal{T}_{\mathbb{Z}\subset\mathbb{R}}$ on $\mathbb{Z}$ induced by $\mathbb{R}$ agrees with the discrete topology $\mathcal{T}_{\text{dis}}$ on $\mathbb{Z}$ described in Example 2.2.10.

Exercise 3.1.6. Let (X, d_X) be a metric space and let $Y \subset X$ be a subset. Then Y is a metric space by restricting the metric, i.e. we define $d_Y = d_X|_{Y\times Y}$ and consider the metric space (Y, d_Y). Show that the induced metric topology $\mathcal{T}(d_Y)$ on Y agrees with the subspace topology $\mathcal{T}_{Y\subset X}$ on Y induced by the topological space $(X, \mathcal{T}(d_X))$.

Remark 3.1.7. Suppose that $(X, \mathcal{T}_X)$ is a topological space and $Y \subset X$ is a subspace. Let $U \subset Y$ be a set that is open in the subspace topology $\mathcal{T}_{Y \subset X}$. We can consider this set also as a subset of X, since $U \subset Y \subset X$. However, as a subset of X the set U will in general *not be open*!

To see this, let $X = \mathbb{R}$ be equipped with the metric topology induced by $d(x, y) = |x - y|$ and consider the subspace $Y = [0, 2)$. The set $U = [0, 1)$ is *open* in Y equipped with the subspace topology, since it can be written as the intersection $(-1, 1) \cap Y$ and the interval $(-1, 1) \subset X$ is open, but U is *not open* in X.

However, we have the following **important special case**: Let $Y \subset X$ be a subspace, such that Y is open in X and let $U \subset Y$ be an open subset of Y. By the definition of the subspace topology we have $U = V \cap Y$ for a subset $V \subset X$ that is open in X. But since we assumed that Y is open in X, U is then open in X as well.

Example 3.1.8. Let $\mathbb{R}^n$ be the n-dimensional euclidean space and for $v = (v_1, \dots, v_n)$ let $\|v\| = \left(\sum_{i=1}^n v_i^2\right)^{1/2}$ be the norm on $\mathbb{R}^n$. As already mentioned in Example 2.1.2 the function $d(x, y) = \|x - y\|$ defines a metric on $\mathbb{R}^n$ and we consider (as always, if not stated otherwise) $\mathbb{R}^n$ as a topological space equipped with the metric topology $\mathcal{T}(d)$. The n-dimensional sphere is defined as

$$S^n = \{v \in \mathbb{R}^{n+1} \mid \|v\| = 1\} \subset \mathbb{R}^{n+1} .$$

It is also a topological space equipped with the subspace topology induced by $(\mathbb{R}^{n+1}, \mathcal{T}(d))$.

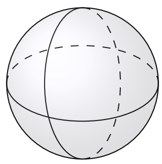


Figure 16: The 2-dimensional sphere

The next theorem tells us that we can glue together continuous maps that are given on two subspaces A and B of a topological space X as long as the two pieces A and B are either both closed or open.

Theorem 3.1.9. *Let $(X, \mathcal{T}_X), (Y, \mathcal{T}_Y)$ be topological spaces and let $A \subset X$ and $B \subset X$ be subspaces, such that $X = A \cup B$ and such that A and B are both closed or both open in X. Let $f \colon X \to Y$ be a map, such that $f|_A$ and $f|_B$ are both continuous. Then f is continuous as well.*

Proof. We will only show the case where A and B are both open. The case where the two are closed is very similar.

Let $\iota_A \colon A \to X$ be the inclusion map of A and let $\iota_B \colon B \to X$ be the corresponding map for B. Let $U \subset Y$ be open. We have to show that $f^{-1}(U)$ is open in X. Observe that

$$(f \circ \iota_A)^{-1}(U) = \{x \in A \mid f(x) \in U\} = f^{-1}(U) \cap A$$

and likewise $(f \circ \iota_B)^{-1}(U) = f^{-1}(U) \cap B$. By assumption we therefore have that the set $f^{-1}(U) \cap A$ is open in A and $f^{-1}(U) \cap B$ is open in B. But we also assumed that A and B are both open in X. Hence, by the last paragraph of Remark 3.1.7, we see that $f^{-1}(U) \cap B$ and $f^{-1}(U) \cap A$ are also open in X. Now

$$f^{-1}(U) = f^{-1}(U) \cap (A \cup B) = (f^{-1}(U) \cap A) \cup (f^{-1}(U) \cap B)$$

is the union of two open subset of X and therefore open itself. □

Example 3.1.10. Let $X = Y = \mathbb{R}$ and consider them as topological spaces equipped with the metric topology induced by $d(x, y) = |x - y|$. Consider the map

$$f : \mathbb{R} \to \mathbb{R} \quad ; \quad x \mapsto \begin{cases} x^2 & \text{for } x > 0 \\ x^3 & \text{for } x \leq 0 \end{cases}$$

The restriction $f|_{[0,\infty)}$ is given by $x \mapsto x^2$ (note that $0^2 = 0^3 = 0$) and the restriction $f|_{(-\infty,0]}$ is $x \mapsto x^3$. Both restrictions are continuous and the two sets $[0, \infty)$ and $(-\infty, 0]$ are closed in $\mathbb{R}$ (Why?). Therefore, f is continuous.

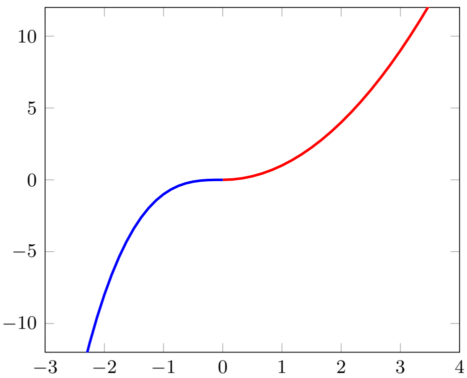


Figure 17: The graph of the continuous function f from the example.

Example 3.1.11. Let $X = Y = \mathbb{R}$ as above and consider the map

$$f : \mathbb{R} \to \mathbb{R} \quad ; \quad x \mapsto \begin{cases} x^2 + 2 & \text{for } x > 0 \\ x^3 & \text{for } x \leq 0 \end{cases}$$

The restriction $f|_{(0,\infty)}$ is given by $x \mapsto x^2 + 2$ and the restriction $f|_{(-\infty,0]}$ is still $x \mapsto x^3$ as in the last example. Both restrictions are continuous. However, $(0, \infty)$ is an open subset of $\mathbb{R}$ and $(-\infty, 0]$ is closed in $\mathbb{R}$. Therefore Theorem 3.1.9 is not applicable. In fact, f is **not** continuous at 0 as Figure 18 of the graph shows.

Example 3.1.12. Now that we have a little more vocabulary at our hands, we can give an example of a homeomorphism. Let $(\mathbb{R}^2, \mathcal{T}(d))$ be the 2-dimensional euclidean space equipped with the metric topology given by $d(x, y) = \|x - y\|$. Let

$$X = \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| + |x_2| = 1\} \subset \mathbb{R}^2$$

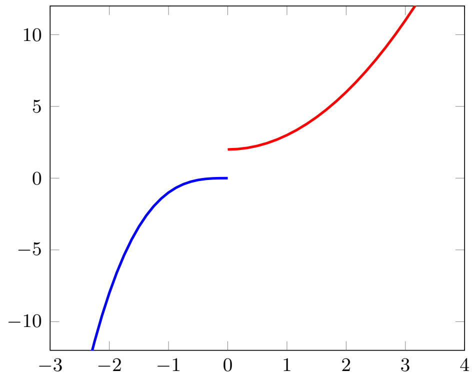


Figure 18: The graph of the function f from the example.

be equipped with the subspace topology. The space X is sketched in blue in Figure 19. Consider the following two maps

$$f\colon X \to S^1 \quad ; \quad (x_1, x_2) \mapsto \left(\frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \frac{x_2}{\sqrt{x_1^2 + x_2^2}} \right) \ ,$$
$$g\colon S^1 \to X \quad ; \quad (x_1, x_2) \mapsto \left(\frac{x_1}{|x_1|+|x_2|}, \frac{x_2}{|x_1|+|x_2|} \right) \ .$$

To see that f is continuous, we will assume that we already know that $\hat{f}\colon \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}^2$ given by $\hat{f}(x) = \frac{x}{\|x\|}$ is continuous. (In fact, this map is also differentiable and hence continuous.) Let $\iota_{S^1}\colon S^1 \to \mathbb{R}^2$ and $\iota_X\colon X \to \mathbb{R}^2 \setminus \{0\}$ be the inclusion maps. Since S^1 and X are both equipped with the respective subspace topologies as subspaces of $\mathbb{R}^2$, both maps ι_{S^1} and ι_X are continuous. We have the identity

$$\hat{f} \circ \iota_X = \iota_{S^1} \circ f$$

of maps $X \to \mathbb{R}^2$ and the left hand side is continuous, since it is the composition of two continuous maps. Hence, the right hand side is continuous as well. Let $U \subset S^1$ be open. By definition of the subspace topology, this means that there is an open subset $V \subset \mathbb{R}^2$ with $U = V \cap S^1$. Using $U = \iota_{S^1}^{-1}(V)$ we obtain that

$$f^{-1}(U) = f^{-1}(\iota_{S^1}^{-1}(V)) = (\iota_{S^1} \circ f)^{-1}(V)$$

is open in X. This proves that f is continuous. It is a good exercise to prove the continuity of g in a similar way.

For the two compositions $f \circ g$ and $g \circ f$ we obtain after a small calculation

$$(f \circ g)(x_1, x_2) = \left(\frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \frac{x_2}{\sqrt{x_1^2 + x_2^2}} \right) \ ,$$
$$(g \circ f)(x_1, x_2) = \left(\frac{x_1}{|x_1| + |x_2|}, \frac{x_2}{|x_1| + |x_2|} \right) \ ,$$

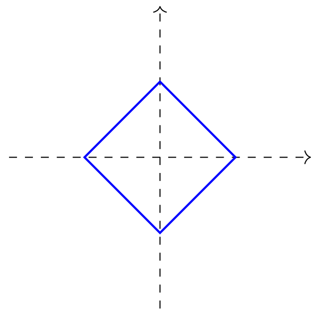


Figure 19: This subspace $X \subset \mathbb{R}^2$ is homeomorphic to S^1.

but since $\sqrt{x_1^2 + x_2^2} = \|x\| = 1$ for every point $x = (x_1, x_2) \in S^1$ and $|x_1| + |x_2| = 1$ for every $(x_1, x_2) \in X$, we have $f \circ g = \mathrm{id}_{S^1}$ and $g \circ f = \mathrm{id}_X$. Therefore X and S^1 are homeomorphic.

3.2 Products of topological spaces

Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be two topological spaces. Is there a natural way to equip the cartesian product $X \times Y$ with a topology $\mathcal{T}_{X\times Y}$? In the last section we defined a topology on a subset of a topological space in such a way that the inclusion map is continuous. For the cartesian product we also have natural maps, namely the projections $p_X : X \times Y \to X$ and $p_Y : X \times Y \to Y$. Can we find a topology on $X \times Y$ which contains just enough open sets to make these maps continuous? Let $U \subset X$ and $V \subset Y$. We define $U \times V$ to be the subset

$$U \times V = \{(x, y) \in X \times Y \mid x \in U, y \in V\} \subset X \times Y \ .$$

If we want both projections p_X and p_Y to be continuous, we need to make sure that $p_X^{-1}(U)$ is open for every open set $U \subset X$ and $p_Y^{-1}(V)$ is open for every open set $V \subset Y$. Note that $p_Y^{-1}(V) = X \times V$ and $p_X^{-1}(U) = U \times Y$ and $U \times V = p_Y^{-1}(V) \cap p_X^{-1}(U)$. Hence, our still to be defined topology $\mathcal{T}_{X\times Y}$ needs to contain the following family of subsets

$$\mathcal{B}_{X\times Y} = \{U \times V \mid U \in \mathcal{T}_X, V \in \mathcal{T}_Y\} \ . \tag{7}$$

However, this family of subsets does not form a topology.

Example 3.2.1. To see this, let us look at the following example: Let $(\mathbb{R}, \mathcal{T}(d))$ be the real line equipped with the metric topology from $d(x, y) = |x - y|$. Consider the corresponding family of subsets $\mathcal{B}_{\mathbb{R}\times\mathbb{R}}$. Let $I_1 = (0, 2)$ and $I_2 = (1, 3)$ be the open intervals from 0 to 2 and 1 to 3 respectively. The family $\mathcal{B}_{\mathbb{R}\times\mathbb{R}}$ contains the sets $I_1 \times I_1$ and $I_2 \times I_2$. If $\mathcal{B}_{\mathbb{R}\times\mathbb{R}}$ were a topology, then it would also contain their union $U = (I_1 \times I_1) \cup (I_2 \times I_2)$. The subset $U \subset \mathbb{R}^2$ is sketched in Figure 20.

Let $p_1, p_2 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be the two projections. For a subset $A \subset \mathbb{R} \times \mathbb{R}$ define $p_i(A) = \{p_i(a) \in \mathbb{R} \mid a \in A\}$ for $i \in \{1, 2\}$. Every set $A \in \mathcal{B}_{\mathbb{R}\times\mathbb{R}}$ has the property

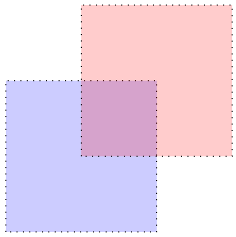


Figure 20: A union of two open sets, which is not of the form $U \times V$.

$A = p_1^{-1}(p_1(A)) \cap p_2^{-1}(p_2(A))$. Let $I_3 = (0, 3)$ be the open interval from 0 to 3. For the set $U = (I_1 \times I_1) \cup (I_2 \times I_2)$ we have

$$p_1^{-1}(p_1(U)) \cap p_2^{-1}(p_2(U)) = I_3 \times I_3 \neq U \ .$$

Hence, $U \notin \mathcal{B}_{\mathbb{R}\times\mathbb{R}}$.

As we see from the last example, $\mathcal{B}_{X\times Y}$ is not always closed under taking unions. Nevertheless, it still provides a *basis* for a topology on $X \times Y$.

Theorem 3.2.2. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and let $\mathcal{B}_{X\times Y}$ be the family of subsets defined in (7). Then it defines a basis for a topology $\mathcal{T}_{X\times Y}$ on $X \times Y$. The two projections $p_X : X \times Y \to X$ and $p_Y : X \times Y \to Y$ are continuous with respect to this topology. The topology $\mathcal{T}_{X\times Y}$ is called the* **product topology on** $X \times Y$.

Proof. To see that $\mathcal{B}_{X\times Y}$ defines the basis of a topology we need to check that it has the two properties a) and b) in Lemma 2.3.17. We have $X \in \mathcal{T}_X$ and $Y \in \mathcal{T}_Y$, therefore $X \times Y \in \mathcal{B}_{X\times Y}$, from which property a) follows easily.

Let $B_1 = U_1 \times V_1 \in \mathcal{B}_{X\times Y}$ and $B_2 = U_2 \times V_2 \in \mathcal{B}_{X\times Y}$. Suppose $B_1 \cap B_2 \neq \emptyset$ and let $x \in B_1 \cap B_2$. In particular, $x \in B_1$ and therefore $p_X(x) \in U_1$. Likewise, we get $p_X(x) \in U_2$, since $x \in B_2$. Hence, $p_X(x) \in U_1 \cap U_2$. Similarly, we obtain $p_Y(x) \in V_1 \cap V_2$. Let $B_3 = (U_1 \cap U_2) \times (V_1 \cap V_2)$ and note that $x \in B_3 \subset B_1 \cap B_2$, which is property b) in Lemma 2.3.17. Hence, $\mathcal{B}_{X\times Y}$ is the basis of a topology $\mathcal{T}_{X\times Y}$.

The continuity of $p_X : X \times Y \to X$ with respect to this topology follows if we can show that $p_X^{-1}(U) \in \mathcal{T}_{X\times Y}$ for every $U \in \mathcal{T}_X$. But $p_X^{-1}(U) = U \times Y \in \mathcal{B}_{X\times Y} \subset \mathcal{T}_{X\times Y}$. The continuity of p_Y can be seen similarly. □

Example 3.2.3. Apparently, we now have two very similar topologies on $\mathbb{R}^2$. The metric topology $\mathcal{T}(d)$ using the metric $d(x, y) = \|x - y\|$ and the product topology $\mathcal{T}_{\mathbb{R}\times\mathbb{R}}$. In fact, these two are identical, i.e. we have

$$\mathcal{T}(d) = \mathcal{T}_{\mathbb{R}\times\mathbb{R}} \ .$$

The best way to see this is sketched in Figure 21. A basis of $\mathcal{T}(d)$ is given by all open discs around every point in $\mathbb{R}^2$ of every radius $r > 0$. Each element in the basis $\mathcal{T}_{\mathbb{R}\times\mathbb{R}}$ is a union of sets of the form $I \times I$ for an open interval $I \subset \mathbb{R}$. If the interval I is bounded, the element $I \times I$ looks like a square in $\mathbb{R}^2$. If we start with a basis element in one topology and fix a point in it, then we can always find a basis element in the other topology, which lies completely inside it. This is shown in Figure 21. Therefore we can cover a basis element in one topology using just the basis elements of the other topology (note that we are allowed to take infinite unions). Therefore the two topologies generated by the two bases agree.

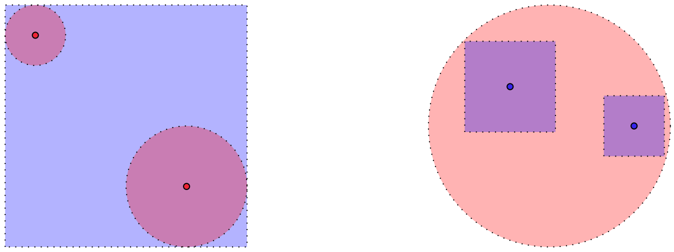


Figure 21: For each of the two topologies $\mathcal{T}(d)$ and $\mathcal{T}_{\mathbb{R}\times\mathbb{R}}$ the basis sets can be covered by basis elements of the other topology proving that the generated topologies are in fact equal.

We can also give a more formal argument based on the following observation about metric topologies.

Lemma 3.2.4. *Let X be a set and let $d_1, d_2\colon X\times X\to\mathbb{R}$ be two metrics on X, such that there exist two constants $c, c' > 0$ with*

$$c\, d_1(x,y) \le d_2(x,y) \le c'\, d_1(x,y)$$

for all $x, y \in X$. Then we have $\mathcal{T}(d_1) = \mathcal{T}(d_2)$.

Proof. We have to prove that any subset $U \subseteq X$, which is open in the topology $\mathcal{T}(d_1)$ is also open in $\mathcal{T}(d_2)$ and vice versa. Let $B_s^{d_i}(x_0)$ be the open ball centred at x_0 with radius $s > 0$ with respect to the metric d_i. Let $y \in B_s^{d_1}(x_0)$. Using our assumptions on d_1 and d_2 we obtain

$$d_2(x_0, y) \le c'\, d_1(x_0, y) < c's\ ,$$

which implies that $y \in B_{c's}^{d_2}(x_0)$. Therefore $B_s^{d_1}(x_0) \subseteq B_{c's}^{d_2}(x_0)$. Similarly, the other inequality yields $B_s^{d_2}(x_0) \subseteq B_{s/c}^{d_1}(x_0)$.

Let $U \in \mathcal{T}(d_1)$ and let $y \in U$. Since U is open with respect to the metric d_1, there is a radius r, such that $B_r^{d_1}(y) \subseteq U$. But then we have $B_{cr}^{d_2}(x_0) \subseteq B_r^{d_1}(x_0) \subset U$. Since $y \in U$ was chosen arbitrarily, this means that $U \in \mathcal{T}(d_2)$. Let $V \in \mathcal{T}(d_2)$ and $z \in V$. We have again a radius $r > 0$, such that $B_r^{d_2}(z) \subseteq V$. As in the previous case we get $B_{r/c'}^{d_1}(z) \subseteq B_r^{d_2}(z) \subset V$, which implies that V is open with respect to the metric d_1, i.e. $V \in \mathcal{T}(d_1)$. □

Let $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$ and let $d^{\max}(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$. It is easy to see that this defines a metric on $\mathbb{R}^2$. For $(z_1, z_2) \in \mathbb{R}^2$ we have

$$\max\{|z_1|, |z_2|\} \le (z_1^2 + z_2^2)^{1/2} \le \sqrt{2}\, \max\{|z_1|, |z_2|\}\ .$$

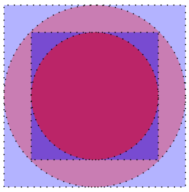


Figure 22: The two inequalities in the theorem imply that the balls with respect to the two metrics d_1 and d_2 can be stacked inside each other as shown.

which we can apply to $(z_1, z_2) = (x_1 - y_1, x_2 - y_2)$ to see that

$$d^{\max}(x, y) \leq d(x, y) \leq \sqrt{2}\, d^{\max}(x, y) \ .$$

It follows from Lemma 3.2.4 that $\mathcal{T}(d) = \mathcal{T}(d^{\max})$. However, the open balls with respect to the metric $d^{\max}$ are given by

$$B_r^{d^{\max}}(x) = I_r(x_1) \times I_r(x_2)$$

where $I_r(x)$ is the open interval $(x-r, x+r)$. In particular, we see that $B_r^{d^{\max}}(x) \in \mathcal{B}_{\mathbb{R}\times\mathbb{R}}$, i.e. the open ball is an element of the basis of the product topology. Let $U \times V \in \mathcal{B}_{\mathbb{R}\times\mathbb{R}}$ and let $y \in U \times V$. Since the open intervals form a basis for the metric topology on $\mathbb{R}$, we can find $r_1, r_2 > 0$ such that $I_{r_1}(y_1) \times I_{r_2}(y_2) \subseteq U \times V$. Let $r = \min\{r_1, r_2\}$. We have

$$B_r^{d^{\max}}(y) \subseteq I_{r_1}(y_1) \times I_{r_2}(y_2) \subseteq U \times V \ .$$

Therefore any element in $\mathcal{B}_{\mathbb{R}\times\mathbb{R}}$ can be written as a union of open balls with respect to the metric $d^{\max}$. This implies that $\mathcal{T}(d^{\max}) = \mathcal{T}_{\mathbb{R}\times\mathbb{R}}$.

Remark 3.2.5. Let $(X_i, \mathcal{T}_{X_i})$ for $i \in \{1, \ldots, n\}$ be topological spaces. Now that we know how to turn $X_1 \times X_2$ into a topological space using the topology $\mathcal{T}_{X_1\times X_2}$, we can inductively define a corresponding topology $\mathcal{T}_{X_1\times\cdots\times X_n}$ on the cartesian product

$$X_1 \times \cdots \times X_n = \prod_{i=1}^{n} X_i \ .$$

We just have to view $X_1 \times \cdots \times X_n$ as the product of the two spaces $X_1 \times \cdots \times X_{n-1}$ and X_n and construct the topology $\mathcal{T}_{X_1\times\cdots\times X_n}$ using $\mathcal{T}_{X_1\times\cdots\times X_{n-1}}$ and $\mathcal{T}_{X_n}$. Any other decomposition of $X_1 \times \cdots \times X_n$ into factors leads to the same topology on the product.

Example 3.2.6. In the introduction we asked if we can continuously deform the surface of a doughnut into a sphere. This "doughnut surface" is also called the 2-dimensional ***torus*** and – similar to the spheres – there is a torus in each dimension. We have already seen how to define the spheres S^n as a topological subspace of $\mathbb{R}^{n+1}$. The n-dimensional torus can be defined as a product of circles:

$$\mathbb{T}^n = \underbrace{S^1 \times \cdots \times S^!}_{n \text{ times}}$$

The 2-dimensional torus $\mathbb{T}^2 = S^1 \times S^1$ is sketched in Figure 23. While the circle can be parametrised by an angle $\varphi \in [0, 2\pi)$, we see from the picture that the parametrisation of the torus requires two angles. We will see in the next section how this will give us another description of the torus as a quotient space.

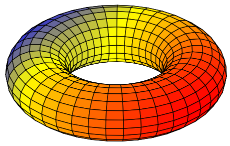


Figure 23: The 2-dimensional torus $\mathbb{T}^2$.

Theorem 3.2.7. *Let $(X, \mathcal{T}_X)$ and $(Y_i, \mathcal{T}_{Y_i})$ for $i \in \{1, \ldots, n\}$ be topological spaces and let $Y = Y_1 \times \cdots \times Y_n$ be the product space equipped with the topology $\mathcal{T}_Y = \mathcal{T}_{Y_1 \times \cdots \times Y_n}$. Let $\pi_j \colon Y \to Y_j$ for $j \in \{1, \ldots, n\}$ be the projection maps. Let $f \colon X \to Y$ be a map, such that $\pi_j \circ f \colon X \to Y_j$ is continuous for all $j \in \{1, \ldots, n\}$. Then f is continuous as well.*

Proof. Just as in (7) a basis for the product topology on Y consists of sets of the form $U_1 \times \cdots \times U_n$ for open sets $U_j \subset Y_j$, $j \in \{1, \ldots, n\}$. Therefore it suffices to check that $f^{-1}(U_1 \times \cdots \times U_n)$ is open in X. However, we have

$$U_1 \times \cdots \times U_n = \pi_1^{-1}(U_1) \cap \cdots \cap \pi_n^{-1}(U_n) \ .$$

which yields

$$\begin{aligned} f^{-1}(U_1 \times \cdots \times U_n) &= f^{-1}(\pi_1^{-1}(U_1) \cap \cdots \cap \pi_n^{-1}(U_n)) \\ &= f^{-1}(\pi_1^{-1}(U_1)) \cap \cdots \cap f^{-1}(\pi_n^{-1}(U_n)) \\ &= (\pi_1 \circ f)^{-1}(U_1) \cap \cdots \cap (\pi_n \circ f)^{-1}(U_n) \ . \end{aligned}$$

By assumption $\pi_j \circ f$ is continuous for every $j \in \{1, \ldots, n\}$. Therefore, $(\pi_j \circ f)^{-1}(U_j)$ is open in X for every open subset $U_j \subset Y_j$ and the finite intersection of those is again open. □

Remark 3.2.8. Let I be an *arbitrary* set (not necessarily finite) and let $(X_i, \mathcal{T}_{X_i})$ for $i \in I$ be topological spaces. We can also consider the product

$$X = \prod_{i \in I} X_i \ .$$

A point in the set X is a sequence $(x_i)_{i \in I}$ of points with $x_i \in X_i$. There are projection maps $\pi_j \colon X \to X_j$ for each $j \in I$ given by $\pi_j((x_i)_{i \in I}) = x_j$. In analogy to the topology on finite products we define the topology on X to be the one generated by the basis

$$\mathcal{B}_{\prod_{i \in I} X_i} = \left\{ \pi_{j_1}^{-1}(U_{j_1}) \cap \cdots \cap \pi_{j_k}^{-1}(U_{j_k}) \mid k \in \mathbb{N},\ j_1, \ldots, j_k \in I,\ U_{j_\ell} \in \mathcal{T}_{X_{j_\ell}} \text{ for all } 1 \leq \ell \leq k \right\} \ .$$

These are the *finite* intersections of preimages of open sets in the the factors X_j with respect to the projections. Note that for $I = \{1, 2\}$ we have $\pi_1^{-1}(U_1) \cap \pi_2^{-1}(U_2) = U_1 \times U_2$ and the topology generated by the basis $\mathcal{B}_{\prod_{i \in I} X_i}$ agrees with the topology $\mathcal{T}_{X_1 \times X_2}$.

Exercise 3.2.9. Rewrite the proof of Theorem 3.2.7 in such a way that it also works for products of the form discussed in Remark 3.2.8. Is it important that the basis $\mathcal{B}_{\prod_{i \in I} X_i}$ only contains finite intersections of open sets in the factors?

Exercise 3.2.10. Let $(X_i, \mathcal{T}_{X_i})$ for $i \in \{1, 2\}$ be two topological spaces and let $Y_i \subset X_i$ be a subspace equipped with the subspace topology $\mathcal{T}_{Y_i \subset X_i}$. We now have two ways to equip the product $Y_1 \times Y_2$ with a topology:

a) We could take the product topology $\mathcal{T}_{Y_1 \times Y_2}$ obtained from the two subspace topologies $\mathcal{T}_{Y_1 \subset X_1}$ and $\mathcal{T}_{Y_2 \subset X_2}$.

b) We could also consider $Y_1 \times Y_2$ as a subset of $X_1 \times X_2$ and take the subspace topology $\mathcal{T}_{Y_1 \times Y_2 \subset X_1 \times X_2}$.

Luckily, these two ways to equip $Y_1 \times Y_2$ with a topology agree. Can you prove that, i.e. can you prove that

$$\mathcal{T}_{Y_1 \times Y_2 \subset X_1 \times X_2} = \mathcal{T}_{Y_1 \times Y_2} \ ?$$

3.3 Quotients of topological spaces

One of the most interesting ways to obtain new topological spaces from known ones is by gluing them together. This is an operation that makes use of equivalence relations. The resulting topological spaces are called quotient spaces and the corresponding topology is the quotient topology. Quotients of metric spaces are not necessarily metrisable anymore. So this is an operation which makes use of the full generality of topological spaces.

3.3.1 Equivalence relations

The best way to understand equivalence relations is by looking at a sock drawer. Suppose that we want to sort our socks into different boxes according to their colour. How can we describe this mathematically? Let S be the set of socks. We introduce the following relation on the set S: We say that two socks $s_1, s_2 \in S$ are equivalent (and we write $s_1 \sim s_2$ in this case) if they have the same colour. This is an example of an equivalence relation. Observe that each element $s \in S$ now has an equivalence class $[s]$ it belongs to. This class is defined as the set of all socks $s' \in S$ that have the same colour as s, i.e.

$$[s] = \left\{ s' \in S \mid s' \sim s \right\} \ .$$

We write $S/{\sim}$ for the set of all equivalence classes. Note that each element in $S/{\sim}$ is a subset of S. In our example, each equivalence class contains all socks of a fixed colour and the number of equivalence classes is the number of boxes that we need.

The relation given by "$s_1 \sim s_2$ if s_1 and s_2 have the same colour" is a weakened version of the equality relation, since $s_1 \sim s_2$ does not mean that s_1 and s_2 are the same, but only that they agree in some aspect we have chosen. The following definition captures the properties of such relations

Definition 3.3.1. Let X be a set. A binary relation $x_1 \sim x_2$ for $x_1, x_2 \in X$ is called an ***equivalence relation*** if it has the following properties:

a) It is *reflexive*, i.e. $x \sim x$ for all $x \in X$.

b) It is *symmetric*. This means: If $x_1 \sim x_2$, then we have $x_2 \sim x_1$.

c) It is *transitive*, i.e. for $x_1, x_2, x_3 \in X$ with $x_1 \sim x_2$ and $x_2 \sim x_3$ we have that $x_1 \sim x_3$.

For each $x \in X$ we define the equivalence class of x to be the set

$$[x] = \{x' \in X \mid x' \sim x\}$$

and we denote the set of all equivalence classes by $X/\sim$.

Example 3.3.2. Our example relation has all of the properties above. A sock $s \in S$ certainly has the same colour as itself, hence $\sim$ is reflexive. If s_1 has the same colour as s_2, then s_2 has the same colour as s_1. Therefore $\sim$ is symmetric. Finally, if s_1 has the same colour as s_2, and s_2 has the same colour as s_3, then s_1 and s_3 are of the same colour, which means that $\sim$ is transitive as well.

Example 3.3.3. Let $X = \mathbb{Z}$ be the set of all integers and define $x_1 \sim x_2$ if $x_1 - x_2$ is an even number. Let us check that this is an equivalence relation: For each $x \in \mathbb{Z}$ we have that $x - x = 0$ is an even number, hence $x \sim x$ and $\sim$ is reflexive. Let $x_1, x_2 \in \mathbb{Z}$. If $x_1 - x_2$ is even, then $x_2 - x_1 = -(x_1 - x_2)$ is also even. Hence, if $x_1 \sim x_2$, then we also have $x_2 \sim x_1$, which proves that $\sim$ is symmetric. Finally, let $x_1, x_2, x_3 \in \mathbb{Z}$ with $x_1 - x_2$ even and $x_2 - x_3$ even. The sum of two even numbers is again even, therefore $x_1 - x_3 = (x_1 - x_2) + (x_2 - x_3)$ is even. This shows that $x_1 \sim x_2$ together with $x_2 \sim x_3$ implies $x_1 \sim x_3$. Hence, $\sim$ is transitive.

The set $\mathbb{Z}/\sim$ has two elements: The equivalence class of an even number consists of all even numbers and the equivalence class of an odd number consists of all odd numbers. Indeed, any two odd numbers are equivalent, since the difference of two odd numbers is even. Any two even numbers are equivalent for the same reason. But an odd number and an even number are not equivalent, since their difference is odd.

Example 3.3.4. Let $X = \mathbb{Z}$ be the set of all integers. If we define $x_1 \sim x_2$ if $x_1 - x_2$ is an odd number, then this is *not* an equivalence relation, since it is not reflexive: For $x \in \mathbb{Z}$ we have $x - x = 0$, which is not odd. Hence, $x \not\sim x$.

Example 3.3.5. Let $X = \mathbb{R}$ be the set of all real numbers. Define $x_1 \sim x_2$ if $x_1 - x_2 \in \mathbb{Z}$. This is another example of an equivalence relation and the proof is very similar to Example 3.3.3.

- reflexivity: Let $x \in \mathbb{Z}$. We have $x - x = 0 \in \mathbb{Z}$. Therefore $x \sim x$.

- symmetry: Let $x_1, x_2 \in \mathbb{Z}$. If we have $x_1 - x_2 \in \mathbb{Z}$, then we also have $x_2 - x_1 = -(x_1 - x_2) \in \mathbb{Z}$. Therefore $x_1 \sim x_2$ implies $x_2 \sim x_1$.

- transitivity: Let $x_1, x_2, x_3 \in \mathbb{Z}$. If $x_1 - x_2 \in \mathbb{Z}$ and $x_2 - x_3 \in \mathbb{Z}$, then $x_1 - x_3 = (x_1 - x_2) + (x_2 - x_3) \in \mathbb{Z}$. Hence, $x_1 \sim x_2$ and $x_2 \sim x_3$ together imply that $x_1 \sim x_3$.

Example 3.3.6. Equivalence relations can also be used to collapse subsets to a single element. Let us illustrate how this is done: Let X be a set and let $A \subset X$ be a subset. For $x_1, x_2 \in X$ we define

$$x_1 \sim x_2 \quad \text{if } (x_1 = x_2) \text{ or } (x_1 \text{ and } x_2 \text{ are both in } A)\ .$$

This is by definition reflexive. If we have $x_1, x_2 \in X$ with $x_1 \sim x_2$, then $x_1 = x_2$ or both elements are in A. Hence, $x_2 \sim x_1$ as well and $\sim$ is symmetric. Let $x_1, x_2, x_3 \in X$ with $x_1 \sim x_2$ and $x_2 \sim x_3$. If $x_1 = x_2$ or $x_2 = x_3$, then we have $x_1 \sim x_3$ as well. But the only remaining case implies that all elements are in A and therefore in particular that $x_1 \sim x_3$, which means that $\sim$ is transitive.

Let $x \in X$. If $x \notin A$, then the equivalence class $[x]$ contains just the element x. On the other hand if $x \in A$, then by definition all other elements in A are also in the equivalence class of x. Therefore we have

$$X/{\sim} = \{[x] \in X/{\sim} \ \mid x \in X \setminus A\} \cup \{[a] \in X/{\sim} \ \mid a \in A\}$$

and $\{[x] \in X/{\sim} \ \mid x \in X \setminus A\}$ is in bijection with $X \setminus A$, whereas $\{[a] \in X/{\sim} \ \mid a \in A\}$ is just a single element.

3.3.2 Quotient spaces

Let $(X, \mathcal{T}_X)$ be a topological space and let $\sim$ be an equivalence relation on X. As above we denote by $X/{\sim}$ the set of all equivalence classes of $\sim$. Is there a way to equip the set $X/{\sim}$ with a natural topology? We defined the subspace topology in such a way that the inclusion map became continuous. Likewise, the product topology was defined such that the projections are continuous. Note that the set $X/{\sim}$ also comes with a natural map. Let

$$q\colon X \to X/{\sim} \quad ; \quad x \mapsto [x]$$

be the map that sends $x \in X$ to the corresponding equivalence class $[x] \in X/{\sim}$. This map is surjective, since every equivalence class $[x]$ comes from at least one element $x \in X$.

Definition 3.3.7. Let $(X, \mathcal{T}_X)$ be a topological space and let $\sim$ be an equivalence relation on X. Let $q\colon X \to X/{\sim}$ be given by $q(x) = [x]$. The ***quotient topology*** $\mathcal{T}_{X/\sim}$ ***on*** $X/{\sim}$ is defined as follows

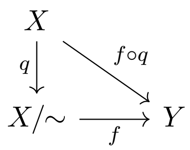
$$\mathcal{T}_{X/\sim} = \left\{ U \subset X/{\sim} \ \mid q^{-1}(U) \in \mathcal{T}_X \right\}\ .$$

Let $q\colon X \to Y$ be a surjective map between two topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$. We call q a ***quotient map*** if a $U \subset Y$ is open if and only if $q^{-1}(U) \subset X$ is open.

By definition the map $q\colon X \to X/\sim$ with $q(x) = [x]$ is a quotient map. Of course we have to check that $\mathcal{T}_{X/\sim}$ really defines a topology:

Theorem 3.3.8. *Let $(X, \mathcal{T}_X)$ be a topological space, let $\sim$ be an equivalence relation on X and let $\mathcal{T}_{X/\sim}$ be as in Definition 3.3.7. Then it defines a topology on $X/\sim$. The map q is continuous with respect to this topology.*

Moreover, if $(Y, \mathcal{T}_Y)$ is another topological space and $f\colon X/\sim \to Y$ is a map, then f is continuous if and only if $f \circ q\colon X \to Y$ is continuous.



Proof. We have to check that $X/\sim$ and $\emptyset$ are both in $\mathcal{T}_{X/\sim}$. But this is true, since $q^{-1}(\emptyset) = \emptyset$ and $\emptyset \in \mathcal{T}_X$. Likewise, $q^{-1}(X/\sim) = X$ and $X \in \mathcal{T}_X$. This proves that $\mathcal{T}_{X/\sim}$ satisfies condition a) in Def. 2.2.6. Let $U, V \in \mathcal{T}_{X/\sim}$. This means that $q^{-1}(U) \in \mathcal{T}_X$ and $q^{-1}(V) \in \mathcal{T}_X$. But then $q^{-1}(U \cap V) = q^{-1}(U) \cap q^{-1}(V) \in \mathcal{T}_X$ and hence $U \cap V \in \mathcal{T}_{X/\sim}$, which implies that $\mathcal{T}_{X/\sim}$ satisfies b) in Def. 2.2.6. Let I be a set and let $U_i \subset X/\sim$ be an open subset for each $i \in I$. We have $q^{-1}(U_i) \in \mathcal{T}_X$ for each $i \in I$. Let $U = \bigcup_{i \in I} U_i$. Using one of the properties of the preimage we get

$$q^{-1}(U) = \bigcup_{i \in I} q^{-1}(U_i) \in \mathcal{T}_X$$

But this means that $U \in \mathcal{T}_{X/\sim}$ and therefore $\mathcal{T}_{X/\sim}$ satisfies c) in Def. 2.2.6.

To see the continuity of q let $U \subset X/\sim$ be open. But then we have that $q^{-1}(U) \subset X$ is also open just by the definition of $\mathcal{T}_{X/\sim}$.

For the last statement, let $U \subset Y$ be open in Y. By definition of the quotient topology, $f^{-1}(U)$ is open in $X/\sim$ if and only if $q^{-1}(f^{-1}(U)) = (f \circ q)^{-1}(U)$ is open in X. But this means that f is continuous if and only if $f \circ q$ is continuous. $\square$

Example 3.3.9. We give an example that shows how we can apply equivalence relations to identify points in a topological space. Let $I = [0, 1] \subset \mathbb{R}$ be the closed unit interval. This is a topological subspace of $\mathbb{R}$ equipped with the metric $d(x, y) = |x - y|$. Consider the following equivalence relation on I:

$$x \sim y \qquad \Leftrightarrow \qquad (x = y) \text{ or } (x \text{ and } y \text{ are both in } \{0, 1\}) \ .$$

This is the situation in Example 3.3.6 with $X = I$ and $A = \{0, 1\}$, i.e. the equivalence relation identifies the point $0 \in I$ with $1 \in I$. As a set the equivalence classes, i.e. the points of $I/\sim$, are given by the elements of the open interval $(0, 1)$ together with the class $\{0, 1\}$ that contains the two identified points.

Let us look at the quotient topology $\mathcal{T}_{I/\sim}$ on $I/\sim$. What does an open neighbourhood $U \subset I/\sim$ of the point $x_0 = \{0, 1\} \in I/\sim$ look like? By definition the preimage of U under q has to be an open subset of I that contains the two points 0 and 1. In particular, it has to contain an open set of the form $[0, \epsilon_1) \cup (1 - \epsilon_2, 1]$ for suitable constants $\epsilon_1, \epsilon_2 > 0$. In fact, the quotient space $I/\sim$ seems to be homeomorphic to the circle S^1 as sketched in Figure 24. Let us prove that this is indeed the case!

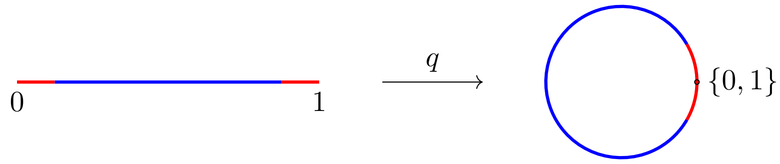


Figure 24: A neighbourhood of the point $\{0, 1\} \in I/\sim$ looks like the red part shown on the right hand side. Its preimage in I under q is shown in red on the left.

In the following we consider $S^1 \subset \mathbb{R}^2$ as a topological space equipped with the subspace topology ($\mathbb{R}^2$ carries the metric topology with $d(x, y) = \|x - y\|$). We need to define a map $f\colon I/\sim \to S^1$. Consider $\hat{f}\colon I \to S^1$ given by

$$\hat{f}\colon I \to S^1 \qquad ; \qquad x \mapsto (\cos(2\pi x), \sin(2\pi x))$$

This is a continuous map and we have $\hat{f}(0) = \hat{f}(1) = (1, 0)$. This means that there is a well-defined map

$$f\colon I/\sim \to S^1 \qquad ; \qquad [x] \mapsto (\cos(2\pi x), \sin(2\pi x))$$

where $[x] \in I/\sim$ denotes the equivalence class of x. Note that we have to pick an element $x \in [x]$ for this to make sense, but our choice does not matter as we already checked. Since we have $\hat{f} = f \circ q$ and $\hat{f}$ is continuous it follows from Theorem 3.3.8 that f is continuous. Note that the function $\arccos\colon [-1, 1] \to [0, \pi]$ is continuous and we have

$$\frac{\arccos(\cos(2\pi x))}{2\pi} = x \quad \text{for } 0 \leq x \leq \frac{1}{2}$$
$$\frac{\arccos(\cos(2\pi x))}{2\pi} = 1 - x \quad \text{for } \frac{1}{2} \leq x \leq 1$$

Therefore the map in the other direction is given by

$$g\colon S^1 \to I/\sim \qquad ; \qquad (x, y) \mapsto \begin{cases} \left[\frac{\arccos(x)}{2\pi}\right] & \text{if } y \geq 0 \\ \left[1 - \frac{\arccos(x)}{2\pi}\right] & \text{if } y < 0 \end{cases}$$

To see that this is continuous let $S^1_+ = \{(x, y) \in S^1 \mid y \geq 0\}$ and let $S^1_- = \{(x, y) \in S^1 \mid y \leq 0\}$. These are two closed subsets of S^1. By definition we have

$$g|_{S^1_+}(x, y) = \left[\frac{\arccos(x)}{2\pi}\right] \ ,$$

which we can write as $q \circ \hat{g}_+$ with $\hat{g}_+ \colon S^1_+ \to I$ given by $\hat{g}_+(x,y) = \frac{\arccos(x)}{2\pi}$. The map $\hat{g}_+$ is continuous, therefore $g|_{S^1_+}$, which is the composition with q, is continuous as well. Likewise, we have

$$g|_{S^1_-}(x,y) = \left[1 - \frac{\arccos(x)}{2\pi}\right] \ . \tag{8}$$

There is something to check here: If $y = 0$ we have $x \in \{-1, 1\}$ and $\frac{\arccos(-1)}{2\pi} = \frac{1}{2} = 1 - \frac{\arccos(-1)}{2\pi}$. Using the identification $[0] = [1]$ in the quotient space, we also have

$$\left[\frac{\arccos(1)}{2\pi}\right] = [0] = [1] = \left[1 - \frac{\arccos(1)}{2\pi}\right] \ .$$

Hence, (8) is also true for the points $(x,y) \in \{(-1,0),(1,0)\}$. Just as above, we can write this map as a composition of a continuous map with q. Therefore, $g|_{S^1_-}$ is continuous as well. Now we apply Theorem 3.1.9 to obtain that g is continous.

The only thing left to check is that $f \circ g = \mathrm{id}_{S^1}$ and $g \circ f = \mathrm{id}_{I/\sim}$. But for $(x,y) \in S^1_+$ we have

$$(f \circ g)(x,y) = (x, \sqrt{1-x^2}) = (x,y)$$

and for $(x,y) \in S^1_-$ we also obtain

$$(f \circ g)(x,y) = (x, -\sqrt{1-x^2}) = (x,y) \ .$$

That the other composition $g \circ f$ is the identity follows from our observations about the arccos-function above. Altogether we see that f provides a homeomorphism between $I/\sim$ and S^1.

Example 3.3.10. The last example is a special case of a more general observation: Consider $\mathbb{R}^n$ as a topological space equipped with the metric topology for the metric $d(x,y) = \|x - y\|$. Let

$$D^n = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\} \subset \mathbb{R}^n$$

be equipped with the subspace topology. This space is called the ***n*-dimensional disk**. Note that the $(n-1)$-dimensional sphere S^{n-1} is a subspace of D^n. The case $n = 1$ corresponds to the setting considered in Example 3.3.9. Consider the following equivalence relation on D^n:

$$x \sim y \qquad \Leftrightarrow \qquad (x = y) \ \text{ or } \ (x \text{ and } y \text{ are both in } S^{n-1}) \ .$$

As in the last example, this equivalence relation identifies the boundary sphere of the disk to a point. In analogy to the example above we expect the quotient space $D^n/\sim$ to be homeomorphic to the sphere S^n. We will prove that this is in fact the case.

Viewing the sphere S^n as a subspace of $\mathbb{R}^{n+1}$ we can consider the map

$$\hat{f} \colon D^n \to S^n \qquad ; \qquad x \mapsto \left((2\sqrt{1 - \|x\|^2})\, x,\ 2\|x\|^2 - 1\right)$$

First, we need to check that $\|\hat{f}(x)\| = 1$. This follows from the calculation

$$\begin{aligned}\|\hat{f}(x)\|^2 &= (2\sqrt{1-\|x\|^2})^2\,\|x\|^2 + (2\|x\|^2-1)^2 \\ &= 4(1-\|x\|^2)\,\|x\|^2 + (4\|x\|^4 - 4\|x\|^2 + 1) = 1\ .\end{aligned}$$

Therefore the image of $\hat{f}$ lies indeed inside the n-dimensional sphere $S^n \subset \mathbb{R}^{n+1}$. Let $\iota\colon S^n \to \mathbb{R}^{n+1}$ be the inclusion. The map $\iota \circ \hat{f}\colon D^n \to \mathbb{R}^{n+1}$ is continuous, since all the projections to the $n+1$ factors are continuous maps $D^n \to \mathbb{R}$. It follows then from the definition of the subspace topology that $\hat{f}$ is also continuous.

Observe that all points $x \in D^n$ with $\|x\| = 1$ are mapped to $(0,1)$. These are all the points that are identified by the equivalence relation. Just as in the previous example we obtain a well-defined continuous map $f\colon D^n/{\sim} \to S^n$ given by $f([x]) = \hat{f}(x)$. Again this requires a choice $x \in [x]$ and the fact that all points with $\|x\| = 1$ are mapped to the same point shows that this choice does not matter.

We will show later in the section about compactness that f is indeed a homeomorphism. For now, we will just check that it is a bijection. To see that f is injective, let $[x],[y] \in D^n/{\sim}$ with $f([x]) = f([y])$. By definition we then have $\hat{f}(x) = \hat{f}(y)$ and therefore $2\|x\|^2 - 1 = 2\|y\|^2 - 1$, which implies $\|x\| = \|y\|$. If $\|x\| \neq 1$ it follows from the equality of the first components of $\hat{f}$ that $x = y$. Since all points with $\|x\| = 1$ are identified by the equivalence relation, we therefore get $[x] = [y]$ in any case.

To see that f is surjective consider $z \in S^n \subset \mathbb{R}^{n+1}$ and write it as $z = (z_1, z_2)$ with $z_1 \in \mathbb{R}^n$ and $z_2 \in \mathbb{R}$. Any point $(z_1,z_2) \in S^n$ satisfies $\|z_1\|^2 + z_2^2 = 1$. Therefore the only point $(z_1,z_2) \in S^n$ with $z_2 = 1$ is $(0,1)$. If $e_1 \in \mathbb{R}^n$ denotes the first unit vector, then $\hat{f}(e_1) = (0,1)$. Hence, the point $(0,1)$ is in the image of f and it remains to consider the case $-1 \leq z_2 < 1$. Let

$$x = \frac{1}{\sqrt{2(1-z_2)}}\, z_1\ .$$

Then we have $\|x\|^2 = \frac{1}{2}(1+z_2)$ and $\hat{f}(x) = (z_1,z_2)$. Thus, we get $f([x]) = (z_1,z_2)$, which shows that f is surjective. Together with the above we have seen that f is a continuous bijection.

Example 3.3.11. We have already constructed the 2-torus $\mathbb{T}^2 = S^1 \times S^1$ as a product space. In this example we will see that $\mathbb{T}^2$ is homeomorphic to a quotient space of the unit square $I^2 = [0,1] \times [0,1]$. Consider the following equivalence relation on I^2:

$$\begin{aligned}(s_1,t_1) \sim (s_2,t_2) \qquad \Leftrightarrow \qquad & (s_1 = s_2 \text{ and } t_1 = t_2) \\ & \text{or } (s_1, s_2 \in \{0,1\} \text{ and } t_1 = t_2) \\ & \text{or } (s_1 = s_2 \text{ and } t_1, t_2 \in \{0,1\}) \\ & \text{or } (s_1, s_2 \in \{0,1\} \text{ and } t_1, t_2 \in \{0,1\})\end{aligned}$$

In particular, $(s,0) \sim (s,1)$ for all $s \in [0,1]$ and $(0,t) \sim (1,t)$ for all $t \in [0,1]$. In the quotient space $I^2/{\sim}$ the left edge of the square is glued to the right edge and the top edge is glued to the bottom edge in a corresponding way. The outcome of this procedure is shown in Figure 25 and we see that it is indeed homeomorphic to $\mathbb{T}^2$.

Using a similar map as in Example 3.3.9 we can also give an explicit homeomorphism: Let $g\colon [0,1] \to S^1$ be given by $g(t) = (\cos(2\pi t), \sin(2\pi t))$ and define

$$\hat{f}\colon I^2 \to S^1 \times S^1 \qquad ; \qquad (s,t) \mapsto (g(s), g(t)) \ .$$

This is continuous by Theorem 3.2.7, since the two maps $(s,t) \mapsto g(s)$ and $(s,t) \mapsto g(t)$ are continuous. Since $g(0) = g(1)$ we have $\hat{f}(s_1,t_1) = \hat{f}(s_2,t_2)$ if $(s_1,t_1) \sim (s_2,t_2)$. Just as in Example 3.3.9 we therefore have a well-defined continuous map

$$f\colon I^2/{\sim} \to S^1 \times S^1 \qquad ; \qquad [s,t] \mapsto \hat{f}(s,t) \ .$$

In the section about compactness we will learn about an elegant way to prove that f is in fact a homeomorphism. Nevertheless, in this case a direct proof is also not too difficult. We leave it as an exercise:

Exercise 3.3.12. Let $\sim_2$ be the equivalence relation on $I = [0,1]$ that identifies the two endpoints $\{0,1\}$ of the interval (see Example 3.3.9).

a) Show that $h\colon I/{\sim_2} \times I/{\sim_2} \to I^2/{\sim}$ given by $h([x],[y]) = [x,y]$ is well-defined and continuous.

b) Let $j\colon I/{\sim_2} \times I/{\sim_2} \to S^1 \times S^1$ be given by $j([x],[y]) = (g(x), g(y))$ with g as above. Show that $f \circ h = j$ and deduce that f is a homeomorphism using the result from Example 3.3.9.

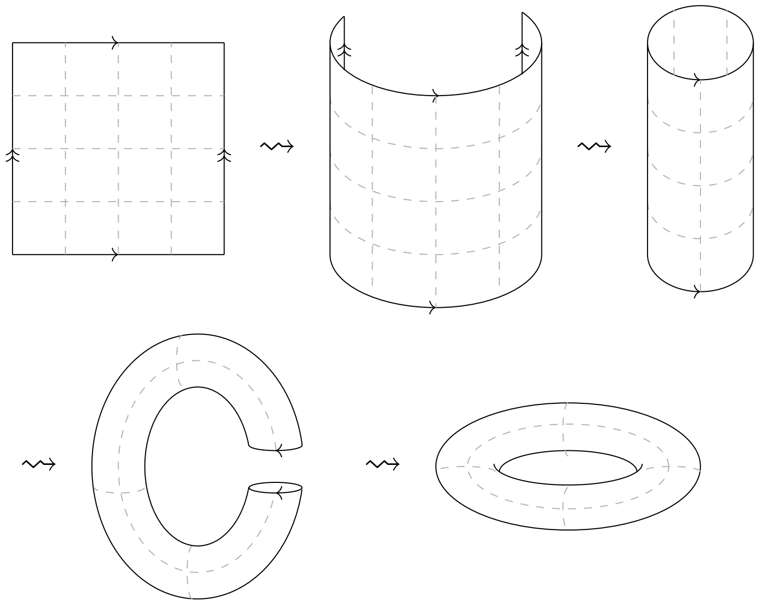


Figure 25: The torus as a quotient of the unit square.

Example 3.3.13. Instead of identifying the edges of the unit square as we did in Example 3.3.11, we can also glue edges together after twisting them first. An example of such a quotient starts again with the unit square I^2 and uses the following equivalence relation:

$$(s_1, t_1) \sim (s_2, t_2) \qquad \Leftrightarrow \qquad \begin{array}{l}(s_1 = s_2 \text{ and } t_1 = t_2) \\ \text{or } (s_1, s_2 \in \{0,1\} \text{ and } t_1 = 1 - t_2)\end{array}$$

In particular, it identifies $(0,t)$ with $(1, 1-t)$ for all $t \in [0,1]$. The quotient by this equivalence relation is sketched in Figure 26. The topological space $M = I^2/{\sim}$ obtained this way is known as the ***Möbius strip*** or the ***Möbius band***.

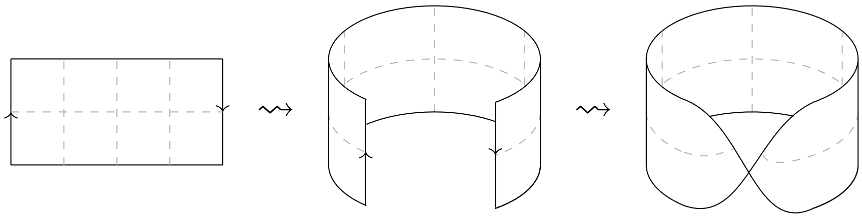


Figure 26: The Möbius strip as a quotient of the unit square.

Exercise 3.3.14. Let $M = I^2/{\sim}$ be the Möbius strip as in Exercise 3.3.13 and let

$$\partial M = \{[s,t] \in M \mid t \in \{0,1\}\} \ .$$

This is the boundary curve of M and in this exercise we will show that it is homeomorphic to a circle. In fact, we will prove that it is homeomorphic to $I/{\sim_2}$, where $\sim_2$ is the equivalence relation that identifies the two endpoints of the interval $I = [0,1]$ from Example 3.3.9.

a) Show that the map $\hat{f}\colon [0,1] \to \partial M$ given by

$$\hat{f}(x) = \begin{cases} [2x, 0] & \text{for } 0 \leq x \leq \frac{1}{2} \ , \\ [2x-1, 1] & \text{for } \frac{1}{2} < x \leq 1 \end{cases}$$

is continuous.

b) Show that $\hat{f}$ is surjective and that $\hat{f}(x) = \hat{f}(y)$ implies $x = y$ or $x, y \in \{0,1\}$.

c) Let $q\colon I \to I/{\sim_2}$ be the quotient map. Show that there is a well-defined bijective and continuous map $f\colon I/{\sim_2} \to \partial M$, such that $f \circ q = \hat{f}$.

d) Show that the map $g\colon \partial M \to S^1$ given by

$$g([s,t]) = \begin{cases} (\cos(\pi s), \sin(\pi s)) & \text{if } t = 0 \ , \\ (\cos(\pi(s+1)), \sin(\pi(s+1))) & \text{if } t = 1 \end{cases}$$

is well-defined and continuous.

e) Check that $g \circ f\colon I/{\sim_2} \to S^1$ agrees with the map $h\colon I/{\sim_2} \to S^1$ given by $h([x]) = (\cos(2\pi x), \sin(2\pi x))$. We know from Example 3.3.9 that h is a homeomorphism. Deduce that f is a homeomorphism as well.

Example 3.3.15. We can combine the equivalence relation that we used to obtain the Möbius band in Example 3.3.13 with the one that glues the two remaining edges in an orientation preserving way. This yields the equivalence relation

$$\begin{aligned}(s_1, t_1) \sim (s_2, t_2) \qquad \Leftrightarrow \qquad & (s_1 = s_2 \text{ and } t_1 = t_2) \\ & \text{or } (s_1, s_2 \in \{0, 1\} \text{ and } t_1 = t_2) \\ & \text{or } (s_1 = 1 - s_2 \text{ and } t_1, t_2 \in \{0, 1\}) \\ & \text{or } (s_1, s_2 \in \{0, 1\} \text{ and } t_1, t_2 \in \{0, 1\})\end{aligned}$$

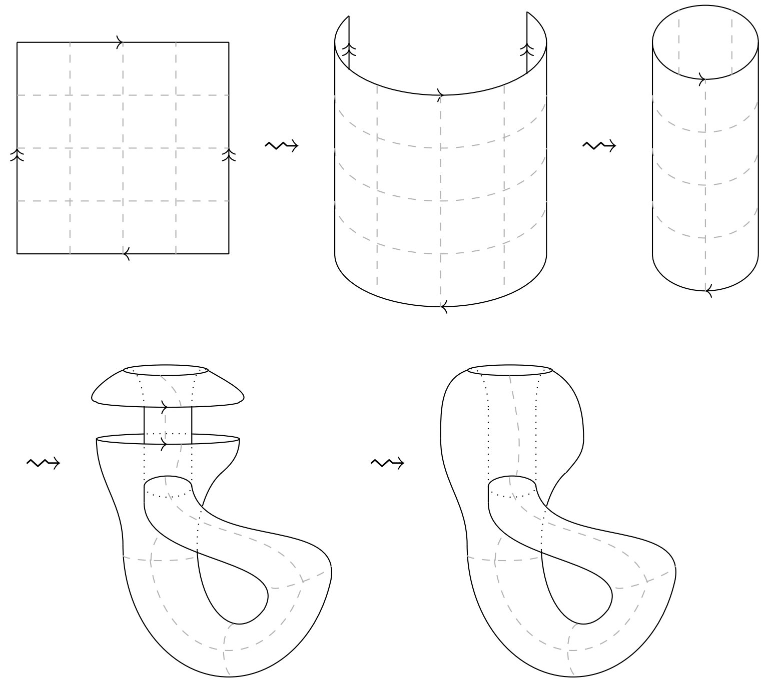


Figure 27: The Klein bottle is a quotient space of the unit square.

Its effect is shown in Figure 27. The quotient space $K = I^2/{\sim}$ we obtain in this case is called the ***Klein bottle***. It is a surface, but – as indicated by the picture – K can not be embedded into $\mathbb{R}^3$ without self-intersections. However, it can be embedded in $\mathbb{R}^4$.

There is another way to construct the Klein bottle. It can also be obtained by gluing two Möbius strips along their boundary circles. To see how that works we will go the other way: Start with K and cut it into two copies of M. This is shown in Figure 28:

Note that the red lines drawn in the second picture really describe a *circle* inside of K since the endpoints are identified. The equivalence relation after the cut shown in the third picture identifies the red parts with each other and the blue parts with each other

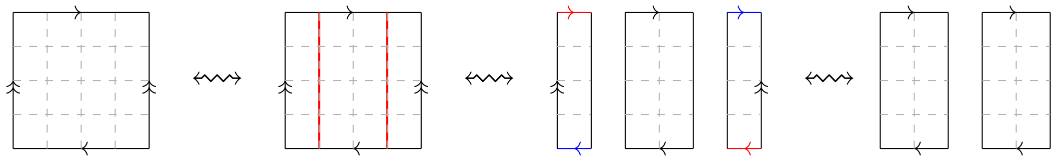


Figure 28: Cutting the Klein bottle K along the red lines yields two Möbius bands.

as shown. The fourth picture shows two Möbius strips and is obtained from the third by glueing together the two double arrowed edges in an orientation preserving way. This is compatible with the equivalence relation as it is shown in the third picture.

Example 3.3.16. Let $\mathbb{R}P^n$ be the set of all lines through the origin in $\mathbb{R}^{n+1}$. We want to use the metric topology on $\mathbb{R}^{n+1}$ to construct a topology on $\mathbb{R}P^n$. Let $\sim$ be the following equivalence relation on $\mathbb{R}^{n+1} \setminus \{0\}$:

$$v_1 \sim v_2 \qquad \Leftrightarrow \qquad \text{there is } \lambda \in \mathbb{R} \text{ such that } v_2 = \lambda\, v_1 \ .$$

Observe that we can identify the equivalence class $[v] \in (\mathbb{R}^{n+1} \setminus \{0\})/\!\sim$ of $v \in \mathbb{R}^{n+1} \setminus \{0\}$ with the line through the origin spanned by v. In particular, we obtain a bijection

$$f\colon \mathbb{R}P^n \to (\mathbb{R}^{n+1} \setminus \{0\})/\!\sim \ .$$

The right hand side is a topological space if we use the standard metric topology on $\mathbb{R}^{n+1}$, the subspace topology on $\mathbb{R}^{n+1} \setminus \{0\}$ and the corresponding quotient topology on $(\mathbb{R}^{n+1} \setminus \{0\})/\!\sim$. If we define a subset $U \subset \mathbb{R}P^n$ to be open if and only if $f(U) \subset (\mathbb{R}^{n+1} \setminus \{0\})/\!\sim$ is open we therefore obtain a natural topology $\mathcal{T}$ on $\mathbb{R}P^n$ as well. The topological space $(\mathbb{R}P^n, \mathcal{T})$ is called the n-dimensional ***real projective space***.

There also are ***complex projective spaces*** $\mathbb{C}P^n$ that are defined completely analogously: $\mathbb{C}P^n$ is the set of complex lines in $\mathbb{C}^{n+1}$ through the origin. Let $\sim$ be the equivalence relation on $\mathbb{C}^{n+1} \setminus \{0\}$ given by

$$v_1 \sim v_2 \qquad \Leftrightarrow \qquad \text{there is } \lambda \in \mathbb{C} \text{ such that } v_2 = \lambda\, v_1 \ .$$

As above we have the standard metric topology on $\mathbb{C}^{n+1}$ and the induced subspace topology on $\mathbb{C}^{n+1} \setminus \{0\}$ giving us a quotient topology on $(\mathbb{C}^{n+1} \setminus \{0\})/\!\sim$ that induces a topology on $\mathbb{C}P^n$ just as above.

Exercise 3.3.17. In this exercise we will see that we can obtain the real projective space also as the quotient space of a sphere.

a) Let $S^n \subset \mathbb{R}^{n+1}$ be the n-sphere equipped with the subspace topology. Consider the equivalence relation on S^n given by $v_1 \sim_{S^n} v_2$ if and only if there is $\lambda \in \{-1, 1\}$ such that $v_2 = \lambda\, v_1$. Show that this is an equivalence relation.

b) Let $\iota\colon S^n \to \mathbb{R}^{n+1} \setminus \{0\}$ be the inclusion map. Show that it induces a continuous bijection $h\colon S^n/\!\sim_{S^n} \to \mathbb{R}P^n$.

c) Show that for any open subset $U \subset S^n/\!\sim_{S^n}$ the set $h(U) \in \mathbb{R}P^n$ is open as well and deduce that h is a homeomorphism.

4 Connectedness

Let $\mathbb{R}$ be the real line equipped with the standard topology. Consider the topological subspace $Y = [0,1] \cup [2,3] \subset \mathbb{R}$. As pointed out in Remark 3.1.7 the subset $U = [0,1] \subset Y$ is open in the subspace topology, since $U = (-1, \frac{3}{2}) \cap Y$ and $(-1, \frac{3}{2})$ is open in $\mathbb{R}$. In a similar way we see that $V = [2,3]$ is open in Y. We have decomposed Y as the disjoint union of two open subsets. This situation was only possible, because Y consists of two disconnected components. Therefore we make the following definition:

Definition 4.0.1. Let $(X, \mathcal{T})$ be a topological space. We say that X is ***disconnected*** if there are two open subsets $U \subset X$ and $V \subset X$ such that both are non-empty, $U \cup V = X$ and $U \cap V = \emptyset$. We say that a topological space $(X, \mathcal{T})$ is ***connected*** if it is not disconnected.

Since U is the complement of V, we obtain that both sets are both open and closed in Y. In fact, we have the following alternative descriptions of connectedness:

Theorem 4.0.2. *Let $(X, \mathcal{T})$ be a topological space. The following statements are equivalent:*

a) X is connected,

b) X and $\emptyset$ are the only subsets of X, which are both open and closed,

c) X is not the union of two non-empty disjoint closed subsets.

Proof. First we show that a) implies b): Assume that X is connected. Let $U \subset X$ be a subset that is open and closed. Let $V = X \setminus U$. Since U is closed, V is an open subset of X. Moreover, we have $U \cup V = X$ and $U \cap V = \emptyset$. If U and V were both non-empty, we would have that X is disconnected contradicting our assumption. Hence, one of the two must be empty. This yields that U is either $\emptyset$ or X.

Next we prove that b) implies c): Let us assume that b) holds. We will show that if we assume that c) does not hold, this will lead to a contradiction. If c) is not true we have two non-empty closed subsets $A \subset X$ and $B \subset X$ such that $A \cup B = X$ and $A \cap B = \emptyset$. But since B is closed, $A = X \setminus B$ is open. Therefore A is a non-empty proper subset of X that is open and closed. This is a contradiction to b).

Finally, we prove that c) implies a). This is a consequence of the following fact: If U and V are non-empty open subsets of X, such that $X = U \cup V$ and $U \cap V = \emptyset$. Then $U = X \setminus V$ and $V = X \setminus U$ is also a decomposition of X into non-empty disjoint closed subsets. □

Our first example of a connected space is the closed unit interval:

Theorem 4.0.3. *Let $I = [0,1] \subset \mathbb{R}$ be equipped with the subspace topology of $(\mathbb{R}, \mathcal{T}(d))$. Then I is connected.*

Proof. Assume that $I = U \cup V$ for two non-empty open subsets $U, V \subset I$ with $U \cap V = \emptyset$. Let $x, y \in I$ with $0 < x < y < 1$ and $x \in U$, $y \in V$. Such elements exist (possibly after switching the roles of U and V), since $I = U \cup V$ and U,V are non-empty. Let

$$W = \{y' \in V \mid x < y'\} \ .$$

Note that $y \in W$. Therefore W is non-empty. Let $s = \inf(W)$ be its largest lower bound. Since $s \in I$, it has to be either in U or in V. Suppose $s \in U$. Since U is open, there is $\epsilon > 0$ such that the interval $(s - \epsilon, s + \epsilon)$ is in U. But the intersection of all such intervals with $W \subset V$ is non-empty since s is its infimum. Therefore $s \in U$ would imply that $U \cap V$ is not empty, which is a contradiction. Hence, we must have $s \in V$. But V is also open. Therefore there is $\epsilon' > 0$, such that $J = (s - \epsilon', s + \epsilon') \subset V$. In particular, none of the points in J is in U. Since $x \in U$ and $x \leq s$, we must have that x is left of J on the real line, hence $x \leq s - \epsilon'$. But then any point in $(s - \epsilon', s)$ is in W and smaller than the infimum, which is also a contradiction. Thus, I is connected. □

If $f : X \to Y$ is a homeomorphism and X is connected, then Y is connected as well. In particular, since any closed interval $[a, b] \subset \mathbb{R}$ is either homeomorphic to the unit interval or consists of a single point (in case $a = b$), all of these are connected. It turns out that any continuous image of a connected space is again connected:

Theorem 4.0.4. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and let $f : X \to Y$ be a continuous map. Suppose X is connected, then the image $f(X) \subset Y$ is also connected (in the subspace topology).*

Proof. Since $f : X \to f(X)$ is continuous, we may without loss of generality assume that $Y = f(X)$. We will show that if Y is not connected, then X is also not connected.

Suppose that Y is not connected. Then we have non-empty open subsets $U, V \subset Y$ with $U \cap V = \emptyset$ and $U \cup V = Y$. Since f is continuous, $f^{-1}(U)$ and $f^{-1}(V)$ are open with $f^{-1}(U) \cap f^{-1}(V) = \emptyset$ and $f^{-1}(U) \cup f^{-1}(V) = X$. If $f^{-1}(U) = \emptyset$, then $f^{-1}(V) = X$. This implies $V = f(X) = Y$ and therefore also $U = \emptyset$, which is a contradiction. Hence, we have that $f^{-1}(U)$ is not empty. Likewise, it follows that $f^{-1}(V)$ is non-empty. But then $f^{-1}(U)$ and $f^{-1}(V)$ are a decomposition of X into non-empty disjoint open subsets. □

Assume that we have a space X that can be written as a union of connected subspaces which all have one point in common. We would expect X to be connected again. In fact, a little more is true:

Lemma 4.0.5. *Let $(X, \mathcal{T}_X)$ be a topological space and let $C \subset X$ be a connected subspace of X. Moreover, let I be a set and let $(C_i)_{i \in I}$ be a family of connected subspaces of X, each of which intersects C. Then the subspace $Y = C \cup \bigcup_{i \in I} C_i$ is connected.*

Proof. Without loss of generality we can restrict to the subspace $Y \subset X$ and assume that $X = Y$. For the sake of contradiction suppose that $X = U \cup V$ for two non-empty open subsets $U, V \subset X$ with $U \cap V = \emptyset$ and $U \cup V = X$. For each $i \in I$ the subspace C_i must lie entirely in U or entirely in V. For if C_i meets both U and V, we would have a decomposition $C_i = (C_i \cap U) \cup (C_i \cap V)$ of C_i into two non-empty disjoint open subsets contradicting that C_i is connected. Similarly, C lies entirely in U or in V. But if C is in V, then for all $i \in I$ the subspace C_i intersects V, and hence lies entirely in V. This means that X would lie entirely in V, and U would be empty, which is a contradiction. Similarly, C can not be in U. Hence, our initial decomposition $X = U \cup V$ can not exist and X must be connected. □

Example 4.0.6. We can write the real line $\mathbb{R}$ equipped with the metric topology as a union of the intervals $I_n = [-n, n]$ for $n \in \mathbb{N}$ as follows:

$$\mathbb{R} = \bigcup_{n \in \mathbb{N}} I_n \ .$$

Now we can apply Lemma 4.0.5 with $I = \mathbb{N}$, $C_n = I_n$ for $n \in I$ and $C = I_1$. Each of the spaces I_n is homeomorphic to the unit interval and hence connected by Theorem 4.0.3. Therefore $\mathbb{R}$ is connected as well.

The open unit interval $J = (0, 1)$ can be written as an infinite union of the closed intervals $J_n = \left[\frac{1}{n+1}, 1 - \frac{1}{n+1}\right]$ as follows

$$J = \bigcup_{n \in \mathbb{N}} J_n \ .$$

For $n = 1$ we have that $J_1 = \{\frac{1}{2}\}$ is contained in all the other J_n. Moreover, each of the J_n (including J_1) is homeomorphic to the unit interval and therefore connected. Now we can apply Lemma 4.0.5 with $I = \mathbb{N}$, $C_n = J_n$ for $n \in I$ and $C = J_1$ again to see that J is connected.

Since any open interval of the form (a, b) for $a, b \in \mathbb{R}$ with $a < b$ is homeomorphic to J, we have that these are all connected.

Exercise 4.0.7. Let $a, b \in \mathbb{R}$ with $a < b$. Let $\mathbb{R}$ be equipped with the metric topology. Show that any half open interval $[a, b)$ is connected using Lemma 4.0.5.

Example 4.0.8. Let $\mathbb{R}^n$ be equipped with the metric topology $\mathcal{T}(d)$ for $d(x, y) = \|x - y\|$. For a $x \in \mathbb{R}^n \setminus \{0\}$ define $\ell_x = \{\lambda x \in \mathbb{R}^n \mid \lambda \in \mathbb{R}\}$. This is the line through the origin spanned by the vector x. In particular, ℓ_x is homeomorphic to $\mathbb{R}$ and therefore connected. Moreover, we can write $\mathbb{R}^n$ as the union of all those lines

$$\mathbb{R}^n = \bigcup_{x \in \mathbb{R}^n \setminus \{0\}} \ell_x$$

and the origin $0 \in \mathbb{R}^n$ is contained in all of them. Thus, we can apply Lemma 4.0.5 with $C = \{0\}$, $I = \mathbb{R}^n \setminus \{0\}$ and $C_x = \ell_x \subset \mathbb{R}^n$ for $x \in I$. This proves that $\mathbb{R}^n$ is connected.

Remark 4.0.9. It is also true that the product space $(X \times Y, \mathcal{T}_{X \times Y})$ of two connected spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ is connected again and although the proof is not too difficult, we will skip it here. This gives an alternative way to show that $\mathbb{R}^n$ is connected.

Example 4.0.10. By Remark 4.0.9 and Theorem 4.0.4 the quotient spaces of the unit interval we have met in the previous section are all connected, since they are continuous images of the connected space $I^2 = I \times I$ under a continuous map $q\colon I^2 \to I^2/{\sim}$. We give a small list of connected spaces:

- the spheres S^n for each $n \in \mathbb{N}$ are connected, since they are obtained as quotient spaces of the connected space I^n by collapsing the boundary to a point,
- the Klein bottle K is a quotient of I^2,
- the Möbius strip M is a quotient of I^2,
- the tori $\mathbb{T}^n = S^1 \times \cdots \times S^1$ are products of the connected space S^1 and hence are connected by Remark 4.0.9,
- for $n \geq 2$ the space $\mathbb{R}^n \setminus \{0\}$ is connected, therefore the space $\mathbb{R}P^n$ is connected as well.

Exercise 4.0.11. Let $n \geq 2$ and let $\mathbb{R}^n$ be equipped with the metric topology as above. Show that $\mathbb{R}^n \setminus \{0\}$ is connected. What goes wrong in the case $n = 1$?

4.1 Path-connected spaces

The unit interval $I = [0, 1]$ is a topological space equipped with the subspace topology of the real line $\mathbb{R}$, where $\mathbb{R}$ is equipped with its standard metric topology $\mathcal{T}(d)$ induced by the metric $d(x, y) = |x - y|$. This allows us to talk about continuous paths in a topological space X, which leads to another concept that is closely related to connectedness:

Definition 4.1.1. Let $(X, \mathcal{T}_X)$ be a topological space. A ***continuous path*** from $x_0 \in X$ to $x_1 \in X$ is a is a continuous map $\gamma\colon I \to X$, such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$.

We say that X is ***path-connected*** if for any two points $x_0, x_1 \in X$ there exists a continuous path from x_0 to x_1.

An example of a continuous path in $\mathbb{T}^2$ is shown in Figure 29. Given a path $\gamma_{01}\colon I \to X$ from x_0 to x_1 and a path $\gamma_{12}\colon I \to X$ from x_1 to x_2 we define their concatenation to be

$$(\gamma_{01} * \gamma_{12})\colon I \to X \qquad ; \qquad (\gamma_{01} * \gamma_{12})(t) = \begin{cases} \gamma_{01}(2t) & \text{for } 0 \leq t \leq \frac{1}{2} \\ \gamma_{12}(2t-1) & \text{for } \frac{1}{2} < t \leq 1 \end{cases} . \tag{9}$$

The map $\gamma_{01} * \gamma_{12}$ is continuous, which follows from Theorem 3.1.9 and runs first through the path γ_{01} and then through γ_{12}. Given a point $x \in X$ we define the constant path $c_x\colon I \to X$ to be $c_x(t) = x$. It just stays on the point x throughout the whole interval. We can also reverse a path as follows: Given $\gamma\colon I \to X$ from x_0 to x_1 the path γ^- given by $\gamma^-(t) = \gamma(1 - t)$ is a path from x_1 to x_0. It is continuous, since it can be written as the composition of γ with the continuous map $r\colon I \to I$ given by $r(t) = 1 - t$. These operations enable us to define the following equivalence relation:

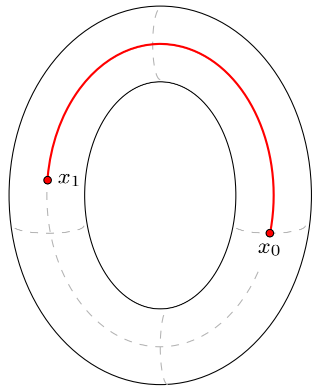


Figure 29: A continuous path from x_0 to x_1 on the torus.

Definition 4.1.2. Let $(X, \mathcal{T}_X)$ be a topological space. Let $\sim_p$ be the equivalence relation on X given by

$$x_1 \sim_p x_2 \qquad \Leftrightarrow \qquad \text{there is a continuous path from } x_1 \text{ to } x_2 \ .$$

We call the equivalence class $[x]$ of $x \in X$ the ***path-component*** of x and define $\pi_0(X) = X/{\sim_p}$ to be the ***set of path-components***.

Our considerations above show that $\sim_p$ is indeed an equivalence relation. Indeed, we have

- $\sim_p$ is reflexive, since c_x is a path from x to itself,
- $\sim_p$ is symmetric, since we can reverse the path γ from x_0 to x_1 to obtain γ^-, which is a path from x_1 to x_0,
- $\sim_p$ is transitive, since for a given path γ from x_0 to x_1 and a path γ' from x_1 to x_2 their concatenation $\gamma * \gamma'$ is a path from x_0 to x_2.

We see that X is path-connected if and only if $\pi_0(X)$ contains just one element. Note that in principle $\pi_0(X)$ could be equipped with the quotient topology and considered as a topological space. However, this is rarely done. Most of the time we will just look at the set $\pi_0(X)$.

The two notions of connectedness we discussed are closely related. As the next theorem shows, path-connected spaces are connected in the sense of Definition 4.0.1.

Theorem 4.1.3. *Let $(X, \mathcal{T}_X)$ be a path-connected topological space. Then X is also connected.*

Proof. Let $X = U \cup V$ for open subsets $U, V \subset X$ with $U \cap V = \emptyset$. For the sake of contradiction assume that U and V are both non-empty (i.e. that X is disconnected). Let $x_0 \in U$ and $x_1 \in V$. Since X is path-connected there is a path $\gamma \colon I \to X$ from x_0 to x_1. Since I is connected, the image $\gamma(I) \subset X$ is also connected by Theorem 4.0.4.

However, the two sets $U \cap \gamma(I)$ and $V \cap \gamma(I)$ provide a decomposition of $\gamma(I)$ into non-empty disjoint open subsets in the subspace topology on $\gamma(I)$. This is a contradiction. Hence, either U or V has to be empty and X is connected. □

With the last theorem in mind it may look like connectedness and path-connectedness are equivalent notions. However, as we will see in the next example, the converse of Theorem 4.1.3 is not true.

Example 4.1.4. Let $\mathbb{R}^2$ be equipped with the standard metric topology $\mathcal{T}(d)$ induced by the metric $d(x,y) = \|x-y\|$. Let $T \subset \mathbb{R}^2$ be the following subspace

$$T = \left\{(x,y) \in (0,1] \times \mathbb{R} \mid y = \sin\left(\frac{1}{x}\right)\right\} \cup \{(0,0)\} \ .$$

It consists of the graph of the function $x \mapsto \sin(\frac{1}{x})$ on $(0,1]$ and the vertical interval from -1 to 1 at $x=0$. A sketch of this topological space is shown in Figure 30. Since the oscillations for x close to 0 become so frequent that they are hard to draw, they are depicted by a blue block in Figure 30. The point $(0,0)$ is shown as a red dot. The space T is called the ***topologist's sine curve***.

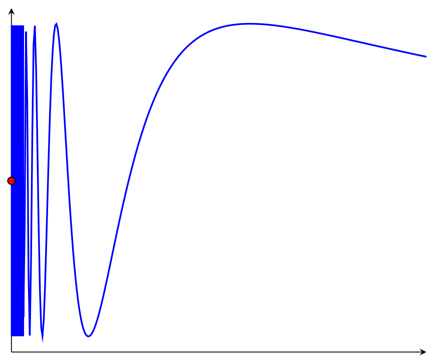


Figure 30: A sketch of the topological space T

Theorem 4.1.5. *The topological space T equipped with the subspace topology $\mathcal{T}_{T\subset\mathbb{R}^2}$ is connected, but not path-connected.*

Proof. Let $U, V \subset T$ be two open subsets with $U \cap V = \emptyset$ and $U \cup V = T$. Suppose without loss of generality that $(0,0) \in U$. Since U is open in the subspace topology and the metric topology on $\mathbb{R}^2$ agrees with $\mathcal{T}_{\mathbb{R}\times\mathbb{R}}$, the set U contains a subset of the form $U' = U'' \cap T$ for $U'' = (-\epsilon_1, \epsilon_1) \times (-\epsilon_2, \epsilon_2)$ with $\epsilon_1, \epsilon_2 > 0$. The sequence $b_n = \frac{1}{\pi n}$ converges to 0 in $\mathbb{R}$. It follows that there is $N \in \mathbb{N}$, such that $a_m = (b_m, 0) \in U''$ for $m > N$. Note that $a_m \in U'' \cap T = U' \subset U$. The subspace

$$T' = \left\{(x,y) \in (0,1] \times \mathbb{R} \mid y = \sin\left(\frac{1}{x}\right)\right\}$$

is homeomorphic to the half-open interval $(0,1]$ and in particular connected (see Exercise 4.1.6). Since the points $a_m \in T'$ lie in U for $m > N$, we obtain that $T' \subset U$. (Otherwise $U \cap T'$ and $V \cap T'$ would be a decomposition of T' into non-empty disjoint open subsets.) But since $(0,0) \in U$, we then have $T \subset U$ and $V = \emptyset$. Hence, T is connected.

To see that T is not path-connected we will show that there can be no continuous path in T from $(0,0)$ to any point $(x_0, y_0) \in T'$. Assume that $\gamma\colon I \to T$ is such a path. Since γ is continuous, we know that $\gamma^{-1}(\{(0,0)\}) \subset I$ is a closed subset of the compact set I. Hence, it is itself compact and there is a maximal value $t_0 \in I$, such that $\gamma(t_0) = (0,0)$ and $\gamma((t_0,1])$ runs through $T' = T \setminus \{(0,0)\}$. (For more details about this argument see the next section about compactness.) Since γ is continuous at t_0 there is a $\delta > 0$ with the property that $|t - t_0| < \delta$ implies $\|\gamma(t) - \gamma(t_0)\| < \frac{1}{2}$. Let $p\colon [0,1] \times \mathbb{R} \to [0,1]$ be the continuous map given by $p(x,y) = x$. In particular, $p \circ \gamma$ is continuous. Thus, $(p \circ \gamma)([t_0, t_0 + \delta))$ has to be a connected subset of $[0,1]$ that contains 0. Hence, it has contain an interval $[0, \epsilon)$ for some $\epsilon > 0$. Let

$$a_n = \frac{1}{\frac{\pi}{2} + 2\pi n}$$

Since a_n converges to 0 for $n \to \infty$ in $\mathbb{R}$, there is an $N \in \mathbb{N}$, such that $0 < a_m < \epsilon$ for all $m > N$. By the above there has to be a point $t_0 \le t_m < t_0 + \delta$ with $(p \circ \gamma)(t_m) = a_m$ for all these values of m. This implies, however, that $\gamma(t_m) = (a_m, 1)$, because $\sin(\frac{1}{a_m}) = 1$. But then

$$\|\gamma(t_m) - \gamma(t_0)\| = \sqrt{a_m^2 + 1} > \frac{1}{2} \ ,$$

which is a contradiction to the continuity of γ at t_0. Thus, such a path γ can not exist. □

Exercise 4.1.6. Let $\mathbb{R}$ be equipped with its standard topology. Let $U \subset \mathbb{R}$ be a subspace and let $h\colon U \to \mathbb{R}$ be a continuous function. Let $\Gamma(h)$ be the graph of h, which is defined as the subspace

$$\Gamma(h) = \{(x,y) \in U \times \mathbb{R} \mid y = h(x)\} \subset U \times \mathbb{R} \ .$$

a) Show that the function $f\colon \Gamma(h) \to U$ given by $f(x,y) = x$ is continuous.

b) Show that the function $g\colon U \to \Gamma(h)$ given by $g(x) = (x, h(x))$ is continuous.

c) Show that $f \circ g = \mathrm{id}_U$ and $g \circ f = \mathrm{id}_{\Gamma(h)}$.

5 Compactness

Compactness is a wonderful property that often allows us to deduce global features of a topological space from local ones. To give an example of this phenomenon let $(X, \mathcal{T}_X)$ be a topological space and let $f\colon X \to \mathbb{R}$ be a continuous function with values in $\mathbb{R}$. It is a consequence of continuity that f is locally bounded, i.e. every point $x \in X$ has a neighbourhood $U \subset X$, such that $f|_U$ is bounded. If X is compact, then f is (globally) bounded, i.e. we can find $a, b \in \mathbb{R}$ with $a < b$, such that $f(X) \subset [a, b]$. We will give a more detailed argument below.

The definition of compactness says that any cover of X by open sets has a finite subcover in the sense that only finitely many of the open sets suffice to cover all of X:

Definition 5.0.1. Let $(X, \mathcal{T}_X)$ be a topological space. A family of open subsets $(U_i)_{i\in I}$ of X is called an ***open cover of*** X if

$$\bigcup_{i\in I} U_i = X \ .$$

A topological space X is called ***compact*** if for every open cover $(U_i)_{i\in I}$ there is a finite subset $J = \{i_1, \ldots, i_n\} \subset I$, such that $\bigcup_{k=1}^n U_{i_k} = X$.

Just as connectedness, the compactness of a topological space is not only preserved by homeomorphisms, but also transfers to continuous images (compare with Theorem 4.0.4):

Theorem 5.0.2. *Let $(X, \mathcal{T}_X)$ be a compact topological space, let $(Y, \mathcal{T}_Y)$ be another topological space and let $f : X \to Y$ be a continuous map. Then $f(X)$ is compact.*

Proof. Since $f : X \to f(X)$ is continuous, we may assume without loss of generality that $Y = f(X)$. Let $(U_i)_{i\in I}$ be an open cover of Y. Since f is continuous, $f^{-1}(U_i) \subset X$ is open for all $i \in I$. Moreover, $\bigcup_{i\in I} f^{-1}(U_i) = f^{-1}\left(\bigcup_{i\in I} U_i\right) = f^{-1}(Y) = X$. Therefore, $(f^{-1}(U_i))_{i\in I}$ is an open cover of X. Since X is compact, there is $n \in \mathbb{N}$ and elements $\{i_1, \ldots, i_n\} \subset I$, such that $\bigcup_{k=1}^n f^{-1}(U_{i_k}) = X$. But then we have

$$\bigcup_{k=1}^n U_{i_k} \subseteq Y = f(X) \subseteq \bigcup_{k=1}^n U_{i_k}$$

and hence $f(X) = \bigcup_{k=1}^n U_{i_k}$. Therefore, $f(X)$ is compact. ☐

Compactness has another interesting property: It is inherited by closed subsets as the next theorem shows.

Theorem 5.0.3. *Let $(X, \mathcal{T}_X)$ be a compact topological space, let $A \subset X$ be a closed subspace (equipped with the subspace topology). Then A is compact.*

Proof. Let $(U_i)_{i\in I}$ be an open cover of A. By the definition of the subspace topology we have an open subset $V_i \subset X$ for each $i \in I$, such that $U_i = V_i \cap A$. Observe that $X \setminus A$ is open and that we have an open cover of the form:

$$X = \bigcup_{i\in I} V_i \cup (X \setminus A) \ .$$

Since X is compact, we can now find a number $n \in \mathbb{N}$ and indices $i_1, \ldots, i_n \in I$, such that $X = \bigcup_{k=1}^n V_{i_k} \cup (X \setminus A)$. This situation is shown in Figure 31. But then we must have $A = X \cap A = U_{i_1} \cup \cdots \cup U_{i_n}$, i.e. we have found a finite open subcover of $(U_i)_{i\in I}$. Hence, A is compact. ☐

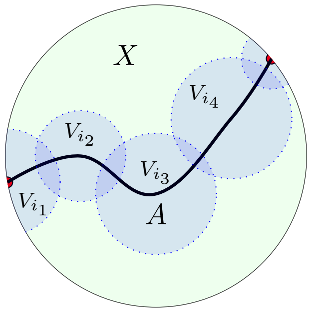


Figure 31: The disk X is compact, the line A is a closed subset. The open subsets V_{i_k} together with the complement of A form a finite open cover of X.

Conversely, it is not true that a compact subspace $A \subset X$ of a topological space X has to be closed. For example, let $(X, \mathcal{T}_X)$ be a topological space, where X is a finite set like the space from Section 2.3.6. In such a space every subset is compact. In particular, there can be compact open subsets. Nevertheless, this situation can not occur in metric spaces. The key property responsible for this has a name:

Definition 5.0.4. A topological space $(X, \mathcal{T}_X)$ is called ***Hausdorff*** (or a ***Hausdorff space***) if for any two points $x, y \in X$ with $x \neq y$ there are open subsets $U, V \subset X$ with $x \in U$, $y \in V$ and $U \cap V = \emptyset$. In other words: In a Hausdorff space any two points can be separated by disjoint open subsets.

Example 5.0.5. As already mentioned above every metric space (X, d) is Hausdorff. To see this let $x_0, x_1 \in X$ be two points with $x_0 \neq x_1$. Let $r = \frac{d(x_0,x_1)}{4} > 0$ and let $U = B_r(x_0)$ and $V = B_r(x_1)$. Suppose that $y \in B_r(x_0) \cap B_r(x_1)$. This would imply that

$$d(x_0, x_1) \leq d(x_0, y) + d(y, x_1) < 2r = \frac{d(x_0, x_1)}{2} \ ,$$

which is a contradiction since $d(x_0, x_1) > 0$. Thus, we have $B_r(x_0) \cap B_r(x_1) = \emptyset$. Since both $B_r(x_i) \subset X$ are also open, we have shown that x_0 and x_1 can be separated by disjoint open subsets. In particular, most of the spaces we have met so far, like the spheres S^n and the tori $\mathbb{T}^n$, are Hausdorff. However, the space in Section 2.3.6 is not.

As alluded to above in Hausdorff spaces compact subspaces are closed:

Theorem 5.0.6. *Let $(X, \mathcal{T}_X)$ be a Hausdorff topological space and let $A \subset X$ be a compact subspace of X. Then A is closed.*

Proof. To see that $X \setminus A$ is open we will find for each $y \in X \setminus A$ an open subset $U'_y \subset X \setminus A$ with $y \in U'_y$. Then $X \setminus A = \bigcup_{y \in X \setminus A} U'_y$ is a union of open sets and hence open itself.

Fix $y \in X \setminus A$. Since X is Hausdorff, we can find for each $x \in A$ two open subsets $U_x, V_x \subset X$ with $V_x \cap U_x = \emptyset$, $x \in V_x$ and $y \in U_x$. This construction is illustrated in Figure 32. We have an open cover of A given by $A = \bigcup_{x \in A}(V_x \cap A)$. Since A is compact, we can find an $n \in \mathbb{N}$ and a finite set of points $x_1, \ldots, x_n \in A$, such that $A \subset \bigcup_{i=1}^n V_{x_i}$.

Let $U'_y = \bigcap_{i=1}^n U_{x_i}$. This is an intersection of finitely many open sets. Therefore U'_y is open. Moreover,

$$U'_y = \bigcap_{i=1}^n U_{x_i} \subset \bigcap_{i=1}^n (X \setminus V_{x_i}) = X \setminus \left(\bigcup_{i=1}^n V_{x_i} \right) \subset X \setminus A \, . \qquad \square$$

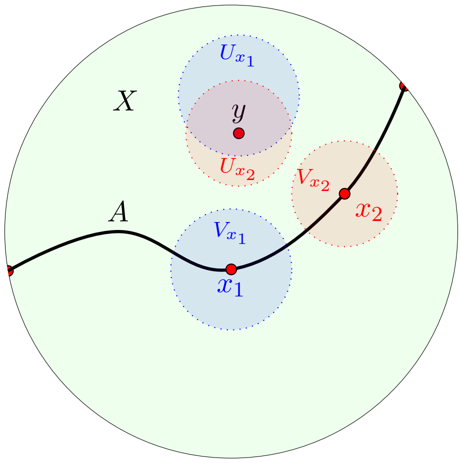


Figure 32: The disjoint subsets U_{x_i} and V_{x_i} from the proof of Theorem 5.0.6.

Example 5.0.7. Note that the half-open interval $J = [0, 1)$ is not compact. In fact, the open cover $(U_n)_{n\in\mathbb{N}}$ with

$$U_n = \left[0, 1 - \frac{1}{n}\right)$$

does not have a finite subcover. Let $\sim$ be the equivalence relation on $I = [0, 1]$ that identifies the subspace $\partial I = \{0, 1\}$ to a single point just as in Example 3.3.9, where we have shown that $I/\sim$ is homeomorphic to the circle S^1. There is a continuous bijection $f \colon J \to I/\sim$ with $f(t) = [t]$. However, the map f is *not* a homeomorphism. Suppose that it were. Then the inverse map $g = f^{-1}$ would have to be continuous. Since $g^{-1}([0, \frac{1}{2})) = f([0, \frac{1}{2}))$ is not open in $I/\sim$ (Why?), this is not the case.

As Example 5.0.7 shows there can be continuous bijections from non-compact spaces (like $[0, 1)$) to Hausdorff spaces (like $I/\sim \,\cong S^1$). It is a remarkable fact, that this can not happen, if we assume that the source space is compact, as the next theorem shows.

Theorem 5.0.8. *Let $(X, \mathcal{T}_X)$ be a compact topological space and let $(Y, \mathcal{T}_Y)$ be a Hausdorff topological space. Suppose that $f \colon X \to Y$ is a continuous bijective map between them. Then f is a homeomorphism.*

Proof. To prove the statement we need to show that the map $g = f^{-1} \colon Y \to X$ is continuous. Let $U \subset X$ be an open subset. Note that $g^{-1}(U) = f(U)$ and let $A = X \setminus U$. It suffices to check that $f(A)$ is closed, since this will imply that $f(U) = Y \setminus f(A)$ is open.

Since X is compact and A is closed, it follows from Theorem 5.0.3 that A is compact. Since f is continuous $f(A)$ is compact as well by Theorem 5.0.2. But by Theorem 5.0.6 we have that $f(A)$ is then closed. □

Another surprising property of compact spaces is that arbitrary products of compact spaces are again compact (even if the indexing set is infinite, in which case we have to use the product topology described in Remark 3.2.8). The proof of this result, the first of which is due to Tychonoff in 1930, is quite tricky and we will omit it.

Theorem 5.0.9 (Tychonoff). *Let I be a set and let $(X_i, \mathcal{T}_{X_i})$ for $i \in I$ be compact topological spaces. Then the product space*

$$X = \prod_{i \in I} X_i$$

equipped with the product topology $\mathcal{T}_{\prod_{i\in I} X_i}$ (see Remark 3.2.8) is again compact.

After we discussed all those nice consequences of compactness above, it is finally time to see some examples of compact spaces.

Example 5.0.10. Let $\mathbb{R}$ be equipped with the metric topology $\mathcal{T}(d)$ with respect to $d(x, y) = |x - y|$. Let $a, b \in \mathbb{R}$ with $a < b$. The subspace $I_{a,b} = [a, b]$ is compact. To see why this is true let $(U_i)_{i\in I}$ be an open cover of $I_{a,b}$. Let $S_x = [a, x]$ and

$$J = \{x \in [a, b] \mid S_x \text{ can be covered by a finitely many of the sets } (U_i)_{i\in I}\}$$

Note that $a \in J$. In particular, J is not empty. Let $s = \sup J$. Since $I_{a,b}$ has b as its upper bound, we have $a \leq s \leq b$. Suppose that $s < b$. Let $i_0 \in I$ be such that $s \in U_{i_0}$. Since U_{i_0} is open, we have $\epsilon > 0$ such that the open interval $(s - \epsilon, s + \epsilon)$ lies entirely in U_{i_0}. Since s is the supremum of J, we can find an $x_1 \in I_{a,b}$ with $x_1 \in (s - \epsilon, s]$ such that S_{x_1} can be covered by finitely many of the sets $(U_i)_{i\in I}$. Let $U_{i_1} \cup \cdots \cup U_{i_n} \supset S_{x_1}$ be such a finite cover. Then we have

$$\left[a, s + \frac{\epsilon}{2}\right] = S_{x_1} \cup \left(s - \epsilon, s + \frac{\epsilon}{2}\right] \subset (U_{i_1} \cup \cdots \cup U_{i_n}) \cup U_{i_0} \ ,$$

which implies that $s + \frac{\epsilon}{2} \in J$ in contradiction to the fact that $s = \sup J$. Hence, we must have $s = b$. Now pick $i_0 \in I$ with $b \in U_{i_0}$. Since U_{i_0} is open, it contains the interval $(b - \epsilon, b]$ for some $\epsilon > 0$. Since $b = \sup J$, there is an $x_1 \in (b - \epsilon, b]$, such that S_{x_1} can be covered by finitely many of the sets $(U_i)_{i\in I}$, say by $U_{i_1} \cup \cdots \cup U_{i_n}$. But then

$$[a, b] = S_{x_1} \cup (b - \epsilon, b] \subset (U_{i_1} \cup \cdots \cup U_{i_n}) \cup U_{i_0} \ ,$$

which is a finite cover. Hence, we are done.

A theorem of Heine and Borel helps us to identify all compact subspaces of $\mathbb{R}^n$ equipped with the standard topology.

Theorem 5.0.11 (Heine-Borel)**.** *Let $n \in \mathbb{N}$ and consider the topological space $(\mathbb{R}^n, \mathcal{T}(d))$ with $d(x,y) = \|x-y\|$. A subspace $A \subset \mathbb{R}^n$ is compact if and only if A is bounded and closed.*

Proof. Suppose $A \subset \mathbb{R}^n$ is bounded and closed. Since it is bounded, it is a subset of $C_k := [-k,k]^n = [-k,k] \times \cdots \times [-k,k] \subset \mathbb{R}^n$ for some $k > 0$, which is compact by Example 5.0.10 and Theorem 5.0.9. It is a closed subset of that space as well, since $A = A \cap [-k,k]^n$. By Theorem 5.0.3 it is compact.

Now suppose that $A \subset \mathbb{R}^n$ is compact. For $k > 0$ let $U_k = (-k,k)^n = (-k,k) \times \cdots \times (-k,k) \subset \mathbb{R}^n$ and note that $(U_k \cap A)_{k\in\mathbb{N}}$ provides an open cover of A, since

$$A = \bigcup_{k\in\mathbb{N}} U_k \cap A$$

and $U_k \cap A$ is open in the subspace topology of A. Using compactness and the fact that $U_k \subset U_{k+1}$ there is a $K \in \mathbb{N}$, such that $A \subset U_K \subset C_K$. This implies that A is bounded. Since $\mathbb{R}^n$ is a Hausdorff space, we obtain that A is closed by Theorem 5.0.6. □

Example 5.0.12. The above theorems give us a whole bunch of new examples of compact spaces:

- The spheres $S^n \subset \mathbb{R}^{n+1}$ are closed and bounded subsets, hence compact by Theorem 5.0.11. The same is true for the closed disks $D^n \subset \mathbb{R}^n$.
- The tori $\mathbb{T}^n = S^1 \times \cdots \times S^1$ are products of compact spaces and therefore also compact by Theorem 5.0.9.
- The Klein bottle K and the Möbius strip M are quotient spaces of the compact unit square $I^2 = I \times I$. In particular, they are compact as continuous images of a compact space by Theorem 5.0.2.
- By Exercise 3.3.17 the real projective space $\mathbb{R}P^n$ can be obtained as a quotient of the compact space S^n. By Theorem 5.0.2 it is compact.

We have constructed a continuous bijective map $f\colon D^n/\!\sim\, \to S^n$ in Example 3.3.10, where the equivalence relation identifies $S^{n-1} \subset D^n$ to a point. Since D^n is compact, the quotient space $D^n/\!\sim$ is compact as well. Moreover, S^n is a metric space, in particular it is Hausdorff. We obtain from Theorem 5.0.8 that f is in fact a homeomorphism.

Similarly, we constructed a continuous bijective map $f\colon I^2/\!\sim\, \to \mathbb{T}^2$ in Exercise 3.3.11. Theorem 5.0.8 shows again that f is a homeomorphism. Comparing the effort for this argument with the work one has to put into solving Exercise 3.3.12 should convince the reader that Theorem 5.0.8 is a powerful tool.

Let $(X, \mathcal{T}_X)$ be a compact topological space. Another consequence of compactness is that for any continuous map $f\colon X \to \mathbb{R}$ there is a point $x_0 \in X$, such that $f(x_0) = \sup\{f(x) \in \mathbb{R} \mid x \in X\}$, i.e. a point where the function achieves its maximal value. To see this note that $f(X)$ is closed and bounded by Theorem 5.0.11 and Theorem 5.0.2. Boundedness implies that $s = \sup\{f(x) \in \mathbb{R} \mid x \in X\} < \infty$. Suppose that $s \notin f(X)$.

Since $f(X)$ is closed, $\mathbb{R}\setminus f(X)$ is open and therefore there is $\epsilon > 0$, such that $(s-\epsilon, s+\epsilon) \subset \mathbb{R} \setminus f(X)$. But this would imply that $(s-\epsilon, s]$ does not contain any points from $f(X)$, which contradicts the fact that s is the supremum of $f(X)$.

Exercise 5.0.13. Let $(X, \mathcal{T}_X)$ be a topological space. Let $X_+ = X \cup \{\infty\}$, i.e. X_+ is the space X plus an additional point, which we call ∞. Define a topology $\mathcal{T}_{X_+}$ on X_+ as follows: All open subsets $U \subset X$ are also in $\mathcal{T}_{X_+}$ and if a subset $V \subset X_+$ contains ∞, then it is in $\mathcal{T}_{X_+}$ if and only if

$$V = \{\infty\} \cup X \setminus C$$

for some closed and compact subspace $C \subset X$, i.e. if it is the union of $\{\infty\}$ and the complement of a closed and compact subspace C of X. Compare this topology with the one in Exercise 2.3.12.

a) Show that $\mathcal{T}_{X_+}$ is in fact a topology on X_+.

b) Show that the topological space $(X_+, \mathcal{T}_{X_+})$ is compact with respect to this topology.

c) Show that the inclusion $X \to X_+$ is injective and continuous.

d) What happens if X was already compact from the start?

6 The fundamental group

In the first lecture we learned that the Euler characteristic of a polyhedron is a topological invariant if we are careful about what we mean by continuous deformations. In this last section we will discuss the concept of homotopy, which gives an equivalence relation on continuous maps. Two maps are homotopic if they can be continuously deformed into one another. From there we are lead to the notion of homotopy equivalence: Two spaces are homotopy equivalent if they are continuously deformable into one another. This will finally allow us to talk about topological invariants of topological spaces (and not only of polyhedra).

We will then define a group $\pi_1(X, x_0)$, which can be associated to any pair of a topological space $(X, \mathcal{T})$ and basepoint $x_0 \in X$. It is called the fundamental group of X and has the property that any continuous map $f\colon X \to Y$ induces a group homomorphism $f_*\colon \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ in such a way that homotopic maps lead to the same homomorphism. Moreover, the construction is natural in the sense that $(g \circ f)_* = g_* \circ f_*$ and $(\mathrm{id}_X)_* = \mathrm{id}_{\pi_1(X,x_0)}$. This implies that homotopy equivalent spaces have isomorphic fundamental groups. Hence, $\pi_1(X, x_0)$ is a topological invariant.

6.1 Homotopic maps

Suppose $f\colon X \to Y$ and $g\colon X \to Y$ are continuous maps between topological spaces X and Y. What would a continuous deformation of f into g look like? It should be a continuous path between the two maps f and g, i.e. a family of maps $H_t\colon X \to Y$ for each $t \in [0,1]$, such that $H_0 = f$ and $H_1 = g$ and such that H_t is continuous in t. We can think of t like the time coordinate of a movie, in which we see f for $t = 0$ (in the first frame) and g for $t = 1$ (in the last frame).

Definition 6.1.1. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces, let $I \subset \mathbb{R}$ be the unit interval $I = [0,1]$ equipped with its standard metric topology. We say that two continuous maps $f\colon X \to Y$ and $g\colon X \to Y$ are ***homotopic*** if there is a continuous map

$$H\colon X \times I \to Y \ ,$$

such that $H_0 := f$ and $H_1 = g$, where $H_t(x) := H(x,t)$. We write f is homotopic to g as: $f \sim_h g$. The map H is called a ***homotopy between*** f ***and*** g.

If the space X is the unit interval I itself, then H is a homotopy between two paths in the space Y. A sketch of such a homotopy is shown in Figure 33. This particular example of H fixes the endpoints of the paths. A more general homotopy H could also move these points around.

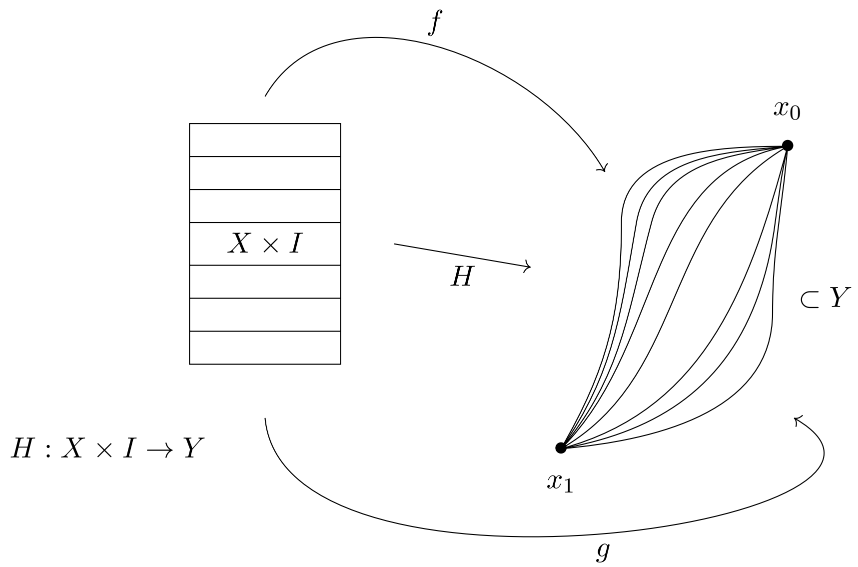


Figure 33: A homotopy between two paths f and g. In the picture we have $X = I$. The homotopy fixes x_0 and x_1.

Theorem 6.1.2. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. Being homotopic is an equivalence relation on the set of continuous maps from X to Y.*

Proof. To see that $\sim_h$ is reflexive we need to construct a homotopy $H\colon X \times I \to Y$, such that $H_0 = H_1 = f$. This holds for the continuous map $H(x,t) = f(x)$.

Suppose that we have two continuous maps f and g with $f \sim_h g$ and let $H\colon X \times I \to Y$ be a homotopy with $H_0 = f$ and $H_1 = g$. Define $H^-\colon X \times I \to Y$ by $H^-(x,t) = H(x, 1-t)$. This is continuous and satisfies $H^-_0 = H_1 = g$ and $H^-_1 = H_0 = f$. Hence, we have $g \sim_h f$. Thus, $\sim_h$ is symmetric.

Let $f, g, k\colon X \to Y$ be three continuous maps with $f \sim_h g$ and $g \sim_h k$. Let $H^{f,g}$ be a homotopy between f and g and let $H^{g,k}$ be one between g and k. Define $H^{f,k}$ by

$$H^{f,k}(x,t) = \begin{cases} H^{f,g}(x, 2t) & \text{for } 0 \leq t \leq \frac{1}{2} \;, \\ H^{g,k}(x, 2t-1) & \text{for } \frac{1}{2} < t \leq 1 \;. \end{cases}$$

Note that this is continuous when restricted to $A = X \times [0, \frac{1}{2}]$ and also over $B = X \times [\frac{1}{2}, 1]$, since $H^{f,g}_1 = g = H^{g,k}_0$. Since the two sets A and B are closed in $X \times I$, we obtain that $H^{f,k}$ is continuous by Theorem 3.1.9. By definition $H^{f,k}_0 = f$ and $H^{f,k}_1 = k$. Therefore $H^{f,k}$ is a homotopy between f and k, in particular $f \sim_h k$, and $\sim_h$ is transitive. □

Definition 6.1.3. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and let $C(X,Y)$ be the set of all continuous maps from X to Y. Define

$$[X, Y] := C(X,Y)/{\sim_h} \;.$$

The set $[X,Y]$ is called the set of homotopy classes of maps from X to Y.

Example 6.1.4. Let $\{*\}$ be the set with one element equipped with the discrete topology. This is the one-point space. Let $(X, \mathcal{T}_X)$ be another topological space. Any continuous map $f\colon \{*\} \to X$ is completely defined by $f(*) \in X$ and a homotopy $H\colon \{*\} \times I \to X$ between f and g is just a continuous path between $f(*)$ and $g(*)$. We obtain the following bijection:

$$\pi_0(X) \cong [*, X] \;,$$

where $\pi_0(X)$ is the set of all path components from Def. 4.1.2.

The circle S^1 and the cylinder $S^1 \times I$ are not homeomorphic. However, they are continuously deformable into one another if we allow to shrink the height of the cylinder down to 0. The two spaces are homotopy equivalent, which is defined as follows:

Definition 6.1.5. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be two topological spaces. We say that X ***is homotopy equivalent to*** Y (written as $X \simeq Y$) if there are two continuous maps $f\colon X \to Y$ and $g\colon Y \to X$, such that

$$f \circ g \sim_h \mathrm{id}_Y \qquad , \qquad g \circ f \sim_h \mathrm{id}_X \;.$$

If a space is homotopy equivalent to the one-point space we call it ***contractible***.

Example 6.1.6. Let $n \in \mathbb{N}$ and let $\mathbb{R}^n$ be equipped with its standard metric topology. Consider the two subspaces $\mathbb{R}^n \setminus \{0\}$ and S^{n-1}. These are homotopy equivalent. Let $f\colon S^{n-1} \to \mathbb{R}^n \setminus \{0\}$ be the inclusion and let $g\colon \mathbb{R}^n \setminus \{0\} \to S^{n-1}$ be defined by

$$g(x) = \frac{x}{\|x\|} \;.$$

The two maps are continuous and we have $g \circ f = \mathrm{id}_{S^{n-1}}$ and $(f \circ g)(x) = \frac{x}{\|x\|}$. We need to see that $f \circ g \sim_h \mathrm{id}_{\mathbb{R}^n \setminus \{0\}}$. Consider

$$H\colon (\mathbb{R}^n \setminus \{0\}) \times I \to \mathbb{R}^n \setminus \{0\} \qquad , \qquad H(x,t) = \left(\frac{1-t}{\|x\|} + t\right) x \ .$$

This is well defined, since $\left(\frac{1-t}{\|x\|} + t\right) \neq 0$ for $t \in [0,1]$. H is also continuous and we have $H_0 = f \circ g$ and $H_1 = \mathrm{id}_{\mathbb{R}^n \setminus \{0\}}$. For a sketch of what this homotopy does see Figure 34.

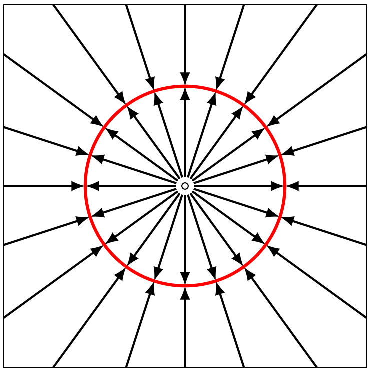


Figure 34: Sketch of the homotopy from Example 6.1.6 for $n = 2$.

Example 6.1.7. Recall the equivalence relation on the unit square I^2 described in Example 3.3.13 given by:

$$\begin{aligned}(s_1,t_1) \sim (s_2,t_2) \qquad \Leftrightarrow \qquad & (s_1 = s_2 \text{ and } t_1 = t_2) \\ & \text{or } (s_1, s_2 \in \{0,1\} \text{ and } t_1 = 1 - t_2)\end{aligned}$$

The quotient space $M = I^2/\!\sim$ is the Möbius band. We will show in this example that M is homotopy equivalent to the circle S^1. To see this we need to construct two continuous maps $f\colon M \to S^1$ and $g\colon S^1 \to M$ such that $f \circ g \sim_h \mathrm{id}_{S^1}$ and $g \circ f \sim_h \mathrm{id}_M$.

From Example 3.3.9 we already know that we can identify the circle S^1 with the quotient $I/\!\sim_{\partial I}$, where $\sim_{\partial I}$ identifies the set of endpoints $\partial I = \{0,1\}$ to a single point. We will suppress the homeomorphism $S^1 \cong I/\!\sim_{\partial I}$ in this example and just identify the two spaces. The map g is then given by

$$g\colon S^1 \to M \quad ; \quad [s] \mapsto \left[s, \frac{1}{2}\right] \ .$$

This is well-defined since

$$g([0]) = \left[0, \frac{1}{2}\right] = \left[1, \frac{1}{2}\right] = g([1]) \ .$$

Note that the map $g'\colon I \to I^2$ given by $g'(s) = (s, \frac{1}{2})$ is continuous. Let $q_M\colon I^2 \to M$ be the quotient map of M. The map $\widehat{g} = q_M \circ g'$ is the composition of two continuous maps and hence continuous. Let $q_{S^1}: I \to S^1$ be the quotient map of the circle. Note that $g \circ q_{S^1} = \widehat{g}$, so by Theorem 3.3.8 the map g is continuous.

Geometrically this map embeds S^1 onto the central circle of the Möbius strip.

The map f is defined as

$$f\colon M \to S^1 \qquad ; \qquad [s,t] \mapsto [s] \ ,$$

which is well-defined since $f([0,t]) = [0] = [1] = f([1,1-t])$. This map projects the Möbius strip down onto the circle. Similar to the argument in the last paragraph we have that $f'\colon I^2 \to I$ given by $f'(s,t) = s$ is continuous. Therefore $q_{S^1} \circ f'$ is continuous and since $f \circ q_M = q_{S^1} \circ f'$ we obtain from Theorem 3.3.8 that f is continuous as well.

We have that $f \circ g = \mathrm{id}_{S^1} \sim_h \mathrm{id}_{S^1}$. The other composition is given by

$$(g \circ f)([s,t]) = \left[s, \frac{1}{2}\right]$$

and we have to construct a homotopy $H\colon M \times I \to M$ between $g \circ f$ and id_M. Let $h\colon I \times I \to I$ be the continuous map given by

$$h(t',t) = \frac{1}{2}\left[(2t'-1)\,t + 1\right]$$

and observe that $h(1-t',t) = 1-h(t',t)$. Now we define $H\colon M \times I \to M$ by $H([s',t'],t) = [s', h(t',t)]$. This is well-defined since

$$H([0,1-t'],t) = [0,h(1-t',t)] = [0,1-h(t',t)] = [1,h(t',t)] = H([1,t'],t) \ .$$

A similar argument to the ones given above shows that H is continuous. Moreover, we have $H([s,t],0) = \left[s,\frac{1}{2}\right] = (g \circ f)([s,t])$ and $H([s,t],1) = [s,t] = \mathrm{id}_M$. Therefore, H is the homotopy we were looking for and we have $g \circ f \sim_h \mathrm{id}_M$.

Geometrically the reversed homotopy H^- shrinks down the strip onto the central circle in a way that is compatible with the equivalence relation defining M. This is sketched in Figure 35.

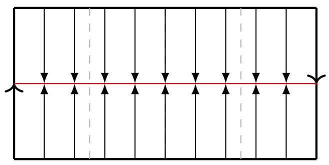


Figure 35: The homotopy between the embedding of the central circle and the identity.

Example 6.1.8. The n-dimensional disk $D^n = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\} \subset \mathbb{R}^n$ as a subspace of n-dimensional euclidean space $\mathbb{R}^n$ equipped with the standard metric topology is contractible. Let $\{*\}$ be the one-point space, let $f\colon D^n \to \{*\}$ be defined by $f(x) = *$ (We do not have any choice here!) and let $g\colon \{*\} \to D^n$ be given by $g(*) = 0$, i.e. the map that sends the point $*$ to the origin $0 \in \mathbb{R}^n$. It is easy to check that these two maps are continuous. Moreover, $f \circ g = \mathrm{id}_{\{*\}}$ and $(g \circ f)(x) = 0$. Let $H\colon D^n \times I \to D^n$ be defined by

$$H(x,t) = t\,x\ .$$

Then H is continuous and we have $H_0 = g \circ f$ and $H_1 = \mathrm{id}_{D^n}$. Therefore, $g \circ f \sim_h \mathrm{id}_{D^n}$ and $D^n \simeq \{*\}$. A sketch showing the reversed homotopy can be found in Figure 36.

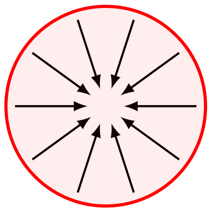


Figure 36: The contraction of the disk rescales it until it collapses into the origin.

Once of the nice properties of homotopy is that it behaves very well with respect to the composition of maps as the following theorem shows.

Theorem 6.1.9. *Let $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ and $(Z, \mathcal{T}_Z)$ be three topological spaces. Let $f_1, f_2\colon X \to Y$ be two continuous maps such that $f_1 \sim_h f_2$. Moreover, let $g_1, g_2\colon Y \to Z$ be two continuous maps such that $g_1 \sim_h g_2$. Then we also have*

$$(g_1 \circ f_1) \sim_h (g_2 \circ f_2)\ .$$

Proof. Let $H^f\colon X \times I \to Y$ be the homotopy between f_1 and f_2. Likewise, let $H^g\colon Y \times I \to Z$ be the homotopy between g_1 and g_2. Let $H^{g\circ f}\colon X \times I \to Z$ be given by

$$H^{g\circ f}(x,t) = H^g(H^f(x,t),t)\ .$$

Let $\Delta_X\colon X \times I \to X \times I \times I$ be given by $\Delta(x,t) = (x,t,t)$. This is a continuous map, since the projection to every component is continuous. Similarly, $H^f \times \mathrm{id}_I\colon X \times I \times I \to Y \times I$ given by $(H^f \times \mathrm{id}_I)(x,s,t) = (H^f(x,s),t)$ is continuous. We have that

$$H^{g\circ f} = H^g \circ (H^f \times \mathrm{id}_I) \circ \Delta_X$$

is a composition of continuous maps and therefore continuous itself. But by definition $H_0^{g\circ f} = g_1 \circ f_1$ and $H_1^{g\circ f} = g_2 \circ f_2$. Therefore $H^{g\circ f}$ is a homotopy between $g_1 \circ f_1$ and $g_2 \circ f_2$. □

6.2 The fundamental group

Let us take a closer look at the question whether the 2-dimensional sphere S^2 is homeomorphic to the 2-dimensional torus $\mathbb{T}^2$. We could approach it as follows: Fix a point $x_0 \in \mathbb{T}^2$ and let $\gamma\colon I \to \mathbb{T}^2$ be a continuous path with $\gamma(0) = \gamma(1) = x_0$, i.e. a continuous loop at x_0. If we had a homeomorphism $f\colon \mathbb{T}^2 \to S^2$, the continuous map $f \circ \gamma\colon I \to S^2$ would give a loop at $f(x_0)$ in the sphere S^2. Likewise, a continuous loop in S^2 would give one in $\mathbb{T}^2$ via the inverse of the homeomorphism inducing a bijection between the loops. Our intuition tells us that all loops in S^2 should be contractible, i.e. homotopic to the constant loop. However, there are loops in the torus, which do not look very contractible like the one running once around the hole in the middle as shown in Figure 37. Making these steps precise leads us to the definition of a powerful topological invariant called the fundamental group, which answers the above question in the negative: The space $\mathbb{T}^2$ is not even homotopy equivalent to the sphere S^2.

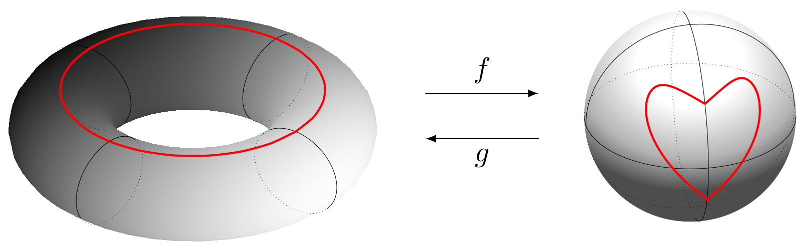


Figure 37: If there were a homeomorphism $g\colon S^2 \to \mathbb{T}^2$, it would have to map a contractible loop (right) to a non-contractible one (left).

The ideas in the first paragraph of this section lead us to studying continuous loops in topological spaces. Recall that a loop in a space X is a continuous map $\gamma\colon I \to X$, such that $\gamma(0) = \gamma(1)$. In particular, γ induces a continuous map

$$\bar{\gamma}\colon I/\!\sim\, \to X \ ,$$

where $\sim$ is the equivalence relation on I that identifies the endpoints. Since $I/\!\sim$ is homeomorphic to S^1 by Example 3.3.9, a loop can also be described by a continuous map $\bar{\gamma}\colon S^1 \to X$. For reasons that will become apparent later we restrict ourselves to loops that start and end at a chosen fixed point $x_0 \in X$. We also need to restrict the homotopies that we allow in this situation.

Definition 6.2.1. Let $(X, \mathcal{T}_X)$ be a topological space and let $x_0 \in X$ be a point in X. Then we call the pair (X, x_0) a ***pointed topological space*** and $x_0 \in X$ the ***basepoint*** of X. (Note that, even though we do not include the topology $\mathcal{T}_X$ in the notation (X, x_0), we still consider X as a topological space and not just as a set.)

Let (X, x_0), (Y, y_0) be pointed topological spaces. A continuous map $f\colon X \to Y$ is called a ***based continuous map*** (or a pointed map, or a basepoint-preserving map) if $f(x_0) = y_0$. Let $f, g\colon X \to Y$ be based continuous maps. A homotopy $H\colon X \times I \to Y$ between f and g is called a ***based homotopy*** (or a pointed homotopy or a basepoint-preserving homotopy) if $H_t(x_0) = H(x_0, t) = y_0$ for all $t \in I$. Being based homotopic is still an equivalence relation, which we denote by $f \sim_{h,+} g$.

Let $C_+((X, x_0), (Y, y_0))$ be the set of all based continuous maps between the pointed topological spaces (X, x_0) and (Y, y_0). Then we define

$$[(X, x_0), (Y, y_0)]_+ = C_+((X, x_0), (Y, y_0))/\sim_{h,+}$$

to be the based homotopy classes of based continuous maps. We will use the notation $[f]_+$ for the element in $[(X, x_0), (Y, y_0)]_+$ represented by the based continuous map $f\colon X \to Y$.

Example 6.2.2. Let $S^0 := \{-1, 1\} \subset \mathbb{R}$ be equipped with the subspace topology and choose $-1 \in S^0$ as the basepoint of S^0, i.e. $(S^0, \{-1\})$ is a pointed topological space. Let (X, x_0) be another pointed topological space. Observe that we have

$$[(S^0, \{-1\}), (X, x_0)]_+ \cong \pi_0(X) \ .$$

In fact, a based continuous map $f\colon S^0 \to X$ has to satisfy $f(-1) = x_0$ and is therefore completely determined by $f(1)$. Likewise, a based homotopy $H\colon S^0 \times I \to X$ has to satisfy $H_t(-1) = H(-1, t) = x_0$ and is therefore fixed by knowing $H_t(1) = H(1, t)$ for all $t \in I$.

Definition 6.2.3. Let $S^1 \subset \mathbb{R}^2$ be the circle equipped with the subspace topology. Let $z_0 = (1, 0) \in S^1$. We define the ***fundamental group of*** (Y, y_0) to be

$$\pi_1(Y, y_0) = [(S^1, z_0), (Y, y_0)]_+ \ .$$

Remark 6.2.4. At this point it is not clear why the set of homotopy classes of maps $\pi_1(Y, y_0)$ deserves to be called “group”. In fact, we have not yet defined the multiplication of two elements from $\pi_1(Y, y_0)$.

Remark 6.2.5. We will often identify a based continuous map $\bar{\gamma}\colon S^1 \to Y$ (representing an equivalence class $[\bar{\gamma}]_+ \in \pi_1(Y, y_0)$) with the induced continuous map $\gamma\colon I \to Y$, which satisfies $\gamma(0) = \gamma(1) = y_0$. In this picture a based homotopy between two loops γ_0 and γ_1 at y_0 is a continuous map $H\colon I \times I \to Y$ with $H_t(0) = H(0, t) = y_0$ and $H_t(1) = H(1, t) = y_0$ for all $t \in I$ and $H_0(s) = H(s, 0) = \gamma_0(s)$, $H_1(s) = H(s, 1) = \gamma_1(s)$ for all $s \in I$. Such a based homotopy is shown in Figure 38. Since these two ways of viewing loops at y_0 are in $1:1$-correspondence to each other, we will denote the equivalence class $[\bar{\gamma}]_+ \in \pi_1(Y, y_0)$ also by $[\gamma]_+$. If $g = [\gamma]_+ \in \pi_1(Y, y_0)$, then we will also say that γ represents g.

Theorem 6.2.6. *Let (X, x_0), (Y, y_0) and (Z, z_0) be pointed topological spaces. Let $f_1, f_2\colon X \to Y$ be based continuous maps such that $f_1 \sim_{h,+} f_2$. Moreover, let $g_1, g_2\colon Y \to Z$ be based continuous maps such that $g_1 \sim_{h,+} g_2$. Then we also have*

$$(g_1 \circ f_1) \sim_{h,+} (g_2 \circ f_2) \ .$$

Proof. Let $H^f\colon X \times I \to Y$ be the based homotopy between f_1 and f_2. Likewise, let $H^g\colon Y \times I \to Z$ be the based homotopy between g_1 and g_2. Let $H^{g\circ f}\colon X \times I \to Z$ be given by

$$H^{g\circ f}(x, t) = H^g(H^f(x, t), t) \ .$$

As we have seen in Theorem 6.1.9 this is continuous and satisfies $H_0^{g\circ f} = g_1 \circ f_1$, $H_1^{g\circ f} = g_2 \circ f_2$. Observe that we also have $H_t^{g\circ f}(x_0) = H^g(H^f(x_0, t), t) = H^g(y_0, t) = z_0$. Therefore, $H^{g\circ f}$ is a based homotopy between $g_1 \circ f_1$ and $g_2 \circ f_2$. □

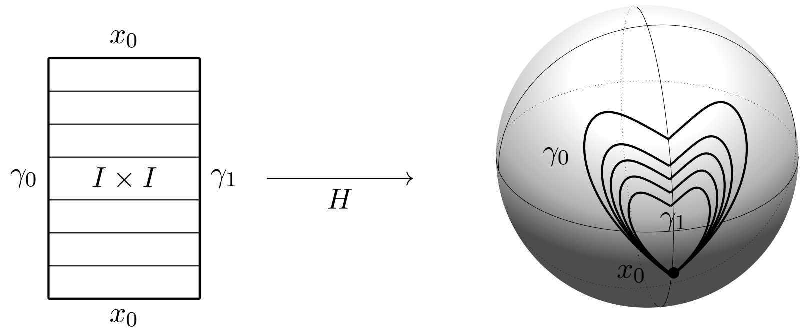


Figure 38: A based homotopy between the two loops γ_0 and γ_1 at x_0.

6.2.1 The group structure on $\pi_1(X, x_0)$

Let $(X, \mathcal{T}_X)$ be a topological space. Given a continuous path $\gamma_{01}\colon I \to X$ from $x_0 \in X$ to $x_1 \in X$ and a continuous path $\gamma_{12}\colon I \to X$ from $x_1 \in X$ to $x_2 \in X$, we defined their concatenation in equation (9) in Section 4.1 to be

$$(\gamma_{01} * \gamma_{12})\colon I \to X \qquad ; \qquad (\gamma_{01} * \gamma_{12})(s) = \begin{cases} \gamma_{01}(2s) & \text{for } 0 \le s \le \frac{1}{2} \\ \gamma_{12}(2s-1) & \text{for } \frac{1}{2} < s \le 1 \end{cases} .$$

A based loop in a pointed topological space (X, x_0) is a continuous path $\gamma\colon I \to X$ starting and ending at $x_0 \in X$ by Remark 6.2.5. In particular, the concatenation $\gamma * \gamma'$ of two such loops $\gamma, \gamma'\colon I \to X$ is always defined. We would like to use the operation $*$ to obtain the group multiplication of $\pi_1(X, x_0)$. However, the elements of $\pi_1(X, x_0)$ are given by based homotopy classes of loops. Hence, we need to check that concatenation is well-defined on those.

Theorem 6.2.7. *Let (X, x_0) be a pointed topological space. Let $g_0, g_1 \in \pi_1(X, x_0)$. Let $\gamma_0, \gamma_0', \gamma_1, \gamma_1'\colon I \to X$ be continuous loops in X at x_0 with $[\gamma_0]_+ = [\gamma_0']_+ = g_0$ and $[\gamma_1]_+ = [\gamma_1']_+ = g_1$. Then we have*

$$(\gamma_0 * \gamma_1) \sim_{h,+} (\gamma_0' * \gamma_1') .$$

In particular, $[\gamma_0 * \gamma_1]_+ = [\gamma_0' * \gamma_1']_+$.

Proof. First observe that $[\gamma_0]_+ = [\gamma_0']_+$ implies that there exists a based homotopy $H^{(0)}\colon I \times I \to X$ between γ_0 and γ_0', i.e. $H^{(0)}$ satisfies the following conditions (see Remark 6.2.5) for all $s, t \in I$:

$$H_0^{(0)}(s) = \gamma_0(s) \qquad , \qquad H_1^{(0)}(s) = \gamma_0'(s) .$$
$$H_t^{(0)}(0) = x_0 = H_t^{(0)}(1) .$$

Likewise, since $[\gamma_1]_+ = [\gamma_1']_+$, there is a based homotopy $H^{(1)}$ between γ_1 and γ_1', i.e. $H^{(1)}$ has to satisfy the conditions:

$$H_0^{(1)}(s) = \gamma_1(s) \qquad , \qquad H_1^{(1)}(s) = \gamma_1'(s) .$$
$$H_t^{(1)}(0) = x_0 = H_t^{(1)}(1) .$$

To prove the theorem we need to construct a based homotopy $H^{(01)}\colon I\times I\to X$ between $\gamma_0 * \gamma_1$ and $\gamma_0' * \gamma_1'$. Let

$$H^{(01)}(s,t) = \begin{cases} H^{(0)}(2s,t) & \text{for } 0\leq s\leq \frac{1}{2}\ , \\ H^{(1)}(2s-1,t) & \text{for } \frac{1}{2} < s \leq 1\ . \end{cases}$$

Since $H^{(0)}(1,t) = H^{(0)}_t(1) = x_0 = H^{(1)}_t(0) = H^{(1)}(0,t)$, the map $H^{(01)}$ is continuous when restricted to $A = [0,\frac{1}{2}]\times[0,1]$ and when restricted to $B = [\frac{1}{2},1]\times[0,1]$. Since A and B are closed subsets of $I\times I$ with $A\cup B = I\times I$, the map $H^{(01)}$ is continuous by Theorem 3.1.9. Moreover, by comparing the definition of $H^{(01)}$ with the one of the concatenated loops from (9) we see that

$$H^{(01)}(s,0) = H^{(01)}_0(s) = (\gamma_0 * \gamma_1)(s) \qquad , \qquad H^{(01)}(s,1) = H^{(01)}_1(s) = (\gamma_0' * \gamma_1')(s)\ .$$
$$H^{(01)}_t(0) = H^{(0)}_t(0) = x_0 = H^{(1)}_t(1) = H^{(01)}_t(1)\ .$$

But these are precisely the conditions for a based homotopy between $\gamma_0 * \gamma_1$ and $\gamma_0' * \gamma_1'$. Hence, we have $(\gamma_0 * \gamma_1) \sim_{h,+} (\gamma_0' * \gamma_1')$ and therefore also $[\gamma_0 * \gamma_1]_+ = [\gamma_0' * \gamma_1']_+$. □

By Theorem 6.2.7 we can define the multiplication of two elements $g_0 = [\gamma_0]_+ \in \pi_1(X,x_0)$ and $g_1 = [\gamma_1]_+$ to be

$$g_0\cdot g_1 := [\gamma_0 * \gamma_1]_+\ .$$

In fact, as we will see next, this multiplication map is part of a group structure on $\pi_1(X,x_0)$.

Theorem 6.2.8. *Let (X,x_0) be a pointed topological space. The set $\pi_1(X,x_0)$ is a group with respect to the following definition of multiplication, neutral element and inverse element:*

- *The multiplication of the elements $g_0 = [\gamma_0]_+$, $g_1 = [\gamma_1]_+ \in \pi_1(X,x_0)$ is given by*

$$g_0\cdot g_1 := [\gamma_0 * \gamma_1]_+\ .$$

- *The neutral element $e = [c_{x_0}]_+$ is given by the based homotopy class of the constant path $c_{x_0}\colon I\to X$ at $x_0\in X$.*

- *The inverse of the element $g = [\gamma]_+ \in \pi_1(X,x_0)$ is given by the based homotopy class of the reversed loop, i.e. $g^{-1} = [\gamma^-]_+$.*

Proof. For brevity let $G = \pi_1(X,x_0)$. We have already seen in Theorem 6.2.7 that the multiplication is well-defined. To see that the data listed above indeed defines a group we need to show that

a) for every $g_0,g_1,g_2\in G$ we have $(g_0\cdot g_1)\cdot g_2 = g_0\cdot(g_1\cdot g_2)$ (associativity),

b) for every $g\in G$ we have $e\cdot g = g\cdot e = g$ (identity element),

c) for every $g \in G$ we have $g \cdot g^{-1} = g^{-1} \cdot g = e$ (inverse element).

Before we prove that G indeed satisfies all these points, we make the following helpful observation: Suppose $\alpha \colon I \to I$ is a continuous map with $\alpha(0) = 0$ and $\alpha(1) = 1$ and let $g \in G$ be represented by $\gamma \colon I \to X$. Then we have

$$\gamma \sim_{h,+} \gamma \circ \alpha \ . \tag{10}$$

Indeed, let $H \colon I \times I \to X$ be given by $H(s,t) = \gamma(s(1-t) + t\,\alpha(s))$. This is continuous, $H_0(s) = \gamma(s)$ and $H_1(s) = \gamma(\alpha(s)) = (\gamma \circ \alpha)(s)$. Moreover, $H_t(0) = \gamma(t\alpha(0)) = \gamma(0) = x_0$, $H_t(1) = \gamma(1) = x_0$. Thus, H is a based homotopy between γ and $\gamma \circ \alpha$.

To address point a) let $g_i = [\gamma_i]_+ \in \pi_1(X, x_0)$ for $i \in \{0,1,2\}$. The element $(g_0 \cdot g_1) \cdot g_2$ is represented by the loop

$$((\gamma_0 * \gamma_1) * \gamma_2)(s) = \begin{cases} \gamma_0(4s) & \text{for } 0 \le s \le \frac{1}{4} \\ \gamma_1(4s-1) & \text{for } \frac{1}{4} < s \le \frac{1}{2} \\ \gamma_2(2s-1) & \text{for } \frac{1}{2} < s \le 1 \end{cases}$$

whereas the composition $g_0 \cdot (g_1 \cdot g_2)$ is represented by

$$(\gamma_0 * (\gamma_1 * \gamma_2))(s) = \begin{cases} \gamma_0(2s) & \text{for } 0 \le s \le \frac{1}{2} \\ \gamma_1(4s-2) & \text{for } \frac{1}{2} < s \le \frac{3}{4} \\ \gamma_2(4s-3) & \text{for } \frac{3}{4} < s \le 1 \end{cases} \ .$$

Observe that $\gamma_0 * (\gamma_1 * \gamma_2)$ is just a reparametrisation of $(\gamma_0 * \gamma_1) * \gamma_2$. In fact, let

$$\alpha \colon I \to I \qquad ; \qquad \alpha(s) = \begin{cases} 2s & \text{for } 0 \le s \le \frac{1}{4} \\ s + \frac{1}{4} & \text{for } \frac{1}{4} < s \le \frac{1}{2} \\ \frac{s}{2} + \frac{1}{2} & \text{for } \frac{1}{2} < s \le 1 \end{cases}$$

and note that α is continuous, $\alpha(0) = 0$, $\alpha(1) = 1$ and $\gamma_0 * (\gamma_1 * \gamma_2) \circ \alpha = (\gamma_0 * \gamma_1) * \gamma_2$. The map α is sketched in Figure 39. By (10) we obtain

$$[\gamma_0 * (\gamma_1 * \gamma_2)]_+ = [(\gamma_0 * \gamma_1) * \gamma_2]_+$$

and therefore also $g_0 \cdot (g_1 \cdot g_2) = (g_0 \cdot g_1) \cdot g_2$.

The identity element $e \in G$ is represented by the constant loop c_{x_0} at x_0. Let $g = [\gamma]_+ \in G$. To prove that b) holds we have to show that

$$c_{x_0} * \gamma \sim_{h,+} \gamma$$

Consider the continuous map

$$H \colon I \times I \to X \qquad ; \qquad (s,t) \mapsto (c_{x_0} * \gamma)\left(\frac{1+s}{2}\,t + (1-t)\,s\right) \ .$$

It satisfies $H_0(s) = H(s,0) = (c_{x_0} * \gamma)(s)$, $H_1(s) = (c_{x_0} * \gamma)(\frac{1+s}{2}) = \gamma(s)$, $H_t(0) = (c_{x_0} * \gamma)(\frac{t}{2}) = x_0$ and $H_t(1) = (c_{x_0} * \gamma)(1) = \gamma(1) = x_0$. Hence, H provides a based homotopy between $c_{x_0} * \gamma$ and γ. The relation $\gamma * c_{x_0} \sim_{h,+} \gamma$ is shown similarly; we leave its proof as an exercise.

Last, but not least we have to prove that c) holds: Let $g \in G$ be represented by $\gamma\colon I \to X$. Then

$$(\gamma * \gamma^-)(s) = \begin{cases} \gamma(2s) & \text{for } 0 \leq s \leq \frac{1}{2} \ , \\ \gamma(2(1-s)) & \text{for } \frac{1}{2} < s \leq 1 \ . \end{cases}$$

Define $H\colon I \times I \to X$ to be the continuous map given by

$$H(s,t) = \begin{cases} \gamma(2s(1-t)) & \text{for } 0 \leq s \leq \frac{1}{2} \ , \\ \gamma(2(1-s)(1-t)) & \text{for } \frac{1}{2} < s \leq 1 \ . \end{cases}$$

This satisfies $H_0 = \gamma * \gamma^-$ and $H_1(s) = \gamma(0) = x_0$. Moreover, $H_t(0) = \gamma(0) = x_0$ and $H_t(1) = \gamma(0) = x_0$. Altogether, H provides a based homotopy between $\gamma * \gamma^-$ and c_{x_0}. Thus, $g \cdot g^{-1} = [\gamma * \gamma^{-1}]_+ = [c_{x_0}]_+ = e$. Switching γ and γ^- we also get $g^{-1} \cdot g = e$. $\square$

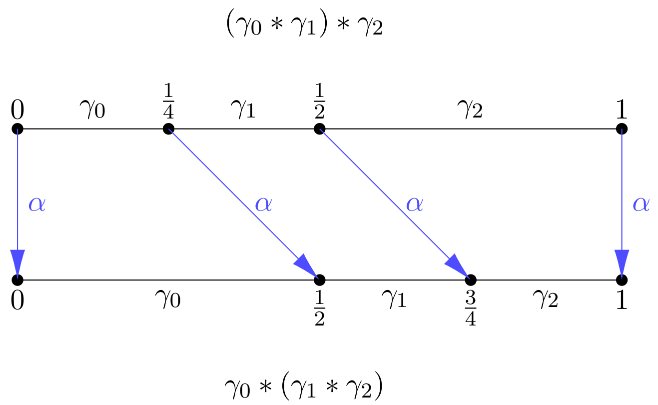


Figure 39: Sketch of the reparametrisation α of the unit interval I used to show the associativity of the multiplication in $\pi_1(X, x_0)$.

6.2.2 Functoriality and Homotopy invariance of $\pi_1(X, x_0)$

The fundamental group $\pi_1(X, x_0)$ associated to a a pointed topological space (X, x_0) has two important properties, which turn it into an invaluable object in algebraic topology: It is functorial and homotopy invariant. Functoriality refers to the following property:

Theorem 6.2.9. *Let (X, x_0), (Y, y_0) be pointed topological spaces and let $f\colon X \to Y$ be a based continuous map between them. The map f induces a group homomorphism*

$$f_*\colon \pi_1(X, x_0) \to \pi_1(Y, y_0) \ ,$$

which is defined by $f_*([\gamma]_+) = [f \circ \gamma]_+$*. If* (Z, z_0) *is another topological space and* $g\colon Y \to Z$ *a based continuous map, then we have*

$$(g \circ f)_* = g_* \circ f_* \ .$$

Moreover, $id_{X*} = id_{\pi_1(X,x_0)}$*.*

Proof. Let $\gamma, \gamma'\colon I \to X$ be continuous loops at $x_0 \in X$ with the property that $\gamma \sim_{h,+} \gamma'$. By Theorem 6.2.6 we have $f \circ \gamma \sim_{h,+} f \circ \gamma'$. Therefore f_* is well-defined.

Let $g_i = [\gamma_i]_+ \in \pi_1(X, x_0)$ for $i \in \{0, 1\}$. Then we have

$$(f \circ (\gamma_0 * \gamma_1))(s) = \begin{cases} (f \circ \gamma_0)(2s) & \text{for } 0 \leq s \leq \frac{1}{2} \ , \\ (f \circ \gamma_1)(2s-1) & \text{for } \frac{1}{2} < s \leq 1 \ . \end{cases}$$

This means $f \circ (\gamma_0 * \gamma_1) = (f \circ \gamma_0) * (f \circ \gamma_1)$. Thus,

$$\begin{aligned} f_*(g_0 \cdot g_1) &= f_*([\gamma_0 * \gamma_1]_+) = [f \circ (\gamma_0 * \gamma_1)]_+ = [(f \circ \gamma_0) * (f \circ \gamma_1)]_+ \\ &= [f \circ \gamma_0]_+ \cdot [f \circ \gamma_1]_+ = f_*(g_0) \cdot f_*(g_1) \ . \end{aligned}$$

From the definition of f_* we get $(g \circ f)_*([\gamma]_+) = [g \circ f \circ \gamma]_+ = (g_* \circ f_*)([\gamma]_+)$ and $\mathrm{id}_{X*}([\gamma]_+) = [\gamma]_+$. □

The next theorem tells us that two based continuous maps between pointed topological spaces that are based homotopic yield the same homomorphism between the fundamental groups. This fact is often referred to as the ***homotopy invariance*** of $\pi_1(X, x_0)$: The fundamental group can not distinguish between based homotopic maps.

Theorem 6.2.10. *Let* (X, x_0)*,* (Y, y_0) *be pointed topological spaces and let* $f, f'\colon X \to Y$ *be a based continuous with the property that* $f \sim_{h,+} f'$*. Then*

$$f_* = f'_* \ .$$

Proof. Let $g \in \pi_1(X, x_0)$ be represented by the loop $\gamma\colon I \to X$. By Theorem 6.2.6 we have $f \circ \gamma \sim_{h,+} f' \circ \gamma$. Therefore, $f_*(g) = f_*([\gamma]_+) = [f \circ \gamma]_+ = [f' \circ \gamma]_+ = f'_*([\gamma]_+) = f'_*(g)$. □

We can use the homotopy invariance to show that $\pi_1(X, x_0)$ is in fact a ***based homotopy invariant***. This means that two pointed topological spaces, which are based homotopy equivalent, have isomorphic fundamental groups.

Corollary 6.2.11. *Let* (X, x_0)*,* (Y, y_0) *be pointed topological spaces, let* $f\colon X \to Y$ *and* $g\colon Y \to X$ *be based continuous maps between them with the property that*

$$f \circ g \sim_{h,+} id_Y \qquad , \qquad g \circ f \sim_{h,+} id_X$$

(a **based homotopy equivalence***). Then* f *induces an isomorphism of the corresponding fundamental groups:*

$$f_*\colon \pi_1(X, x_0) \xrightarrow{\cong} \pi_1(Y, y_0) \ .$$

Proof. From Theorem 6.2.10 we obtain that $f_* \circ g_* = (f \circ g)_* = \mathrm{id}_{Y*} = \mathrm{id}_{\pi_1(Y,y_0)}$ and likewise $g_* \circ f_* = \mathrm{id}_{\pi_1(X,x_0)}$. Thus, f_* is an isomorphism with inverse g_*. □

6.3 The fundamental group of the circle

The best algebraic invariant of topological spaces is useless if there is no way to compute it. But what does "computing" mean in the case of the fundamental group? Let (X, x_0) be a fixed pointed topological space. To understand $\pi_1(X, x_0)$ we could try to prove that it is isomorphic to another group that we already know very well. For example, it could be isomorphic to the trivial group. The easiest example of a pointed space, whose fundamental group is non-trivial, is the circle. We will show that

$$\pi_1(S^1, z_0) \cong \mathbb{Z}$$

for an arbitrarily chosen basepoint $z_0 \in S^1$. Observe that rotating the circle is a homeomorphism. Any rotation $r\colon S^1 \to S^1$ induces an isomorphism $\pi_1(S^1, z_0) \cong \pi_1(S^1, r(z_0))$ by Theorem 6.2.11. Hence, we can without loss of generality assume that $z_0 = (1, 0)$. To simplify the situation further we can identify $\mathbb{R}^2$ with the complex plain $\mathbb{C}$. This maps $(1, 0) \in \mathbb{R}^2$ to $1 \in \mathbb{C}$.

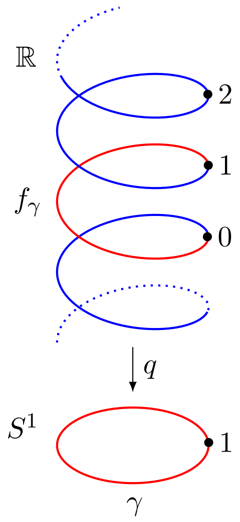


Figure 40: $\mathbb{R}$ is a covering space of S^1.

The proof rests on the following property of the continuous map $q\colon \mathbb{R} \to S^1$ given by

$$q(t) = \exp(2\pi i t)\ .$$

We will treat the following theorem as a black box and just give an indication of the proof below.

Theorem 6.3.1. *For any continuous loop $\gamma\colon I \to S^1$ with $\gamma(0) = \gamma(1) = 1$ there exists a (unique) continuous function $f_\gamma\colon I \to \mathbb{R}$ with $f_\gamma(0) = 0$ and*

$$\gamma(t) = q(f_\gamma(t)) = \exp\left(2\pi i f_\gamma(t)\right)$$

Moreover, for any based homotopy $H\colon I \times I \to S^1$ between two loops γ_0 and γ_1 with $\gamma_i(0) = \gamma_i(1) = 1$ there exists a (unique) continuous function $F_H\colon I \times I \to \mathbb{R}$ with $F_t(0) = F(0, t) = 0$ for all $t \in I$,

$$H(s, t) = q(F_H(s, t)) = \exp\left(2\pi i F_H(t)\right)\ .$$

and $F_H(s, 0) = f_{\gamma_0}(s)$, $F_H(s, 1) = f_{\gamma_1}(s)$.

To see why this theorem is true it helps to think of the real line as a helix curled up over S^1 as shown in Figure 40. The map $q\colon \mathbb{R} \to S^1$ takes a point in $\mathbb{R}$ and projects it down onto S^1. The loop γ that runs once around the circle (shown in red) lifts to the continuous path f_γ from 0 to 1 in $\mathbb{R}$. The property of q that makes this lifting work is that each point $z \in S^1$ has a neighbourhood U such that $q^{-1}(U)$ is homeomorphic to $U \times \mathbb{Z}$. Hence, if we have constructed the value of f_γ at some point $t \in I$, we can extend the construction a little further by using that each point in $\mathbb{R}$ has a neighbourhood, which is mapped homeomorphically by q onto its image in S^1. We say that $\mathbb{R}$ is a covering space of the circle.

Define $\omega_1 \colon I \to S^1$ by $\omega_1(t) = \exp(2\pi i t)$. Let $g = [\omega_1]_+ \in \pi_1(S^1, 1)$ be the element it represents in the fundamental group. Note that we have a homomorphism

$$\alpha \colon \mathbb{Z} \to \pi_1(S^1, 1) \tag{11}$$

given by $\alpha(n) = g^n$.

Theorem 6.3.2. *The homomorphism α from (11) is an isomorphism. In particular, we obtain*

$$\pi_1(S^1, 1) \cong \mathbb{Z} .$$

Proof. We will construct the inverse homomorphism of α. Let $\gamma \colon I \to S^1$ be a loop at $1 \in S^1$. By Theorem 6.3.1 there is a continuous function $f_\gamma \colon I \to \mathbb{R}$ with $f(0) = 0$ and $\gamma = q \circ f_\gamma$. Since $q(f_\gamma(1)) = \gamma(1) = 1$ and $q^{-1}(\{1\}) = \mathbb{Z} \subset \mathbb{R}$, we can associate an integer $n_\gamma \in \mathbb{Z}$ to each loop γ. This integer is called the ***degree of*** γ.

Suppose that $\gamma_0, \gamma_1 \colon I \to S^1$ are two loops based at 1 with the property that $\gamma_0 \sim_{h,+} \gamma_1$. Let $H \colon I \times I \to S^1$ be a based homotopy between them. By the second part of Theorem 6.3.1, H is of the form $H = q \circ F_H$ for some continuous function $F_H \colon I \times I \to \mathbb{R}$ with the properties listed in the Theorem. In particular,

$$q(F_H(1,t)) = H(1,t) = H_t(1) = x_0 .$$

But this means that $F_H(1,t) \in \mathbb{Z}$ for all $t \in I$. Since the interval is connected and $\mathbb{Z}$ has the discrete topology, this means that $F_H(1,t)$ is constant. This implies

$$n_{\gamma_0} = f_{\gamma_0}(1) = F_H(1,0) = F_H(1,1) = f_{\gamma_1}(1) = n_{\gamma_1} .$$

Altogether we have a well-defined map

$$\beta \colon \pi_1(S^1, 1) \to \mathbb{Z} \qquad , \qquad \beta([\gamma]_+) = n_\gamma ,$$

which is independent of the choice of $\gamma \in [\gamma]_+$. Let $\gamma_0, \gamma_1 \colon I \to S^1$ be continuous loops based at $x_0 \in X$. Let $f_{\gamma_0}, f_{\gamma_1} \colon I \to \mathbb{R}$ be the functions associated to them by Theorem 6.3.1. Define $\tilde{f}_{\gamma_0 * \gamma_1} \colon I \to \mathbb{R}$ by

$$\tilde{f}_{\gamma_0 * \gamma_1}(t) = \begin{cases} f_{\gamma_0}(t) & \text{for } 0 \leq t \leq \frac{1}{2} \\ n_{\gamma_0} + f_{\gamma_1}(t) & \text{for } \frac{1}{2} < t \leq 1 \end{cases}$$

and note that this is continuous, since $f_{\gamma_0}(1) = n_{\gamma_0} = n_{\gamma_0} + f_{\gamma_1}(0)$. Moreover, it has the same properties as $f_{\gamma_0 * \gamma_1}$ listed in Theorem 6.3.1. Since these properties uniquely characterise $f_{\gamma_0 * \gamma_1}$, we must have $\tilde{f}_{\gamma_0 * \gamma_1} = f_{\gamma_0 * \gamma_1}$. Hence, we have

$$\beta([\gamma_0]_+ \cdot [\gamma_1]_+) = \beta([\gamma_0 * \gamma_1]_+) = f_{\gamma_0 * \gamma_1}(1) = \tilde{f}_{\gamma_0 * \gamma_1}(1) = n_{\gamma_0} + n_{\gamma_1} = \beta([\gamma_0]_+) + \beta([\gamma_1]_+)$$

and β is a group homomorphism.

To see that β is injective it suffices to prove that $\ker(\beta) = \{e\}$. Suppose that $\beta([\gamma]_+) = n_\gamma = 0$. This means that the associated function f_γ satisfies $f_\gamma(0) = f_\gamma(1) = 0$, i.e. it is a loop in $\mathbb{R}$ based at 0. Let $H\colon I \times I \to \mathbb{R}$ be given by $H(s,t) = t\, f_\gamma(s)$. We have that $(q \circ H)_0 = c_{q(0)} = c_1 \in S^1$, i.e. the constant loop at $1 \in S^1$ and $(q \circ H)_1 = q \circ f_\gamma = \gamma$. Moreover, $(q \circ H)_t(0) = (q \circ H)_t(1) = q(0) = 1 \in S^1$. Therefore $q \circ H$ is a based homotopy between γ and the constant path on 1. But this implies $[\gamma]_+ = [c_1]_+ = e \in \pi_1(S^1, 1)$. Thus, β is injective.

To prove that β is surjective and that it is inverse to α it suffices to show that $\beta([\omega_1]_+) = 1 \in \mathbb{Z}$. This follows from the fact that the function $f_{\omega_1}\colon I \to \mathbb{R}$ that is associated to ω_1 by Theorem 6.3.1 is given by $f_{\omega_1}(t) = t$. Hence, $\beta([\omega_1]_+) = f_{\omega_1}(1) = 1$. $\square$