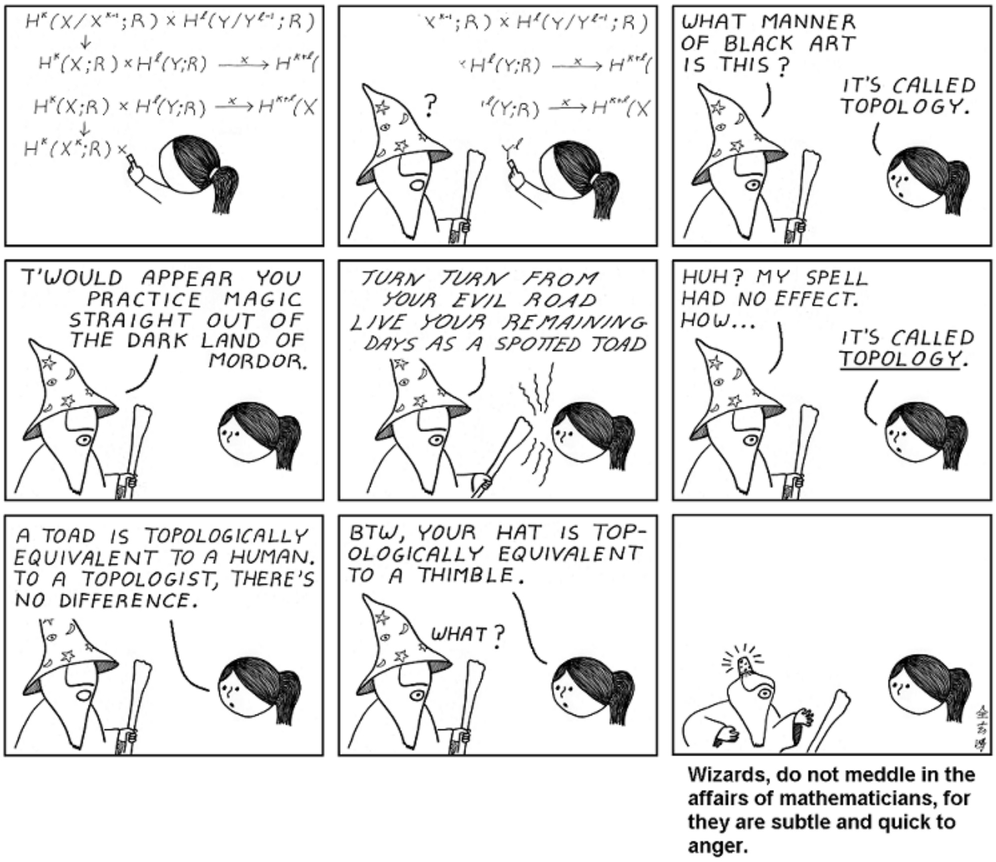
Introduction to Topology

Anthony Conway

Fall 2021



Source: https://abstrusegoose.com/253

Contents

Introduction

Informally, topology studies the properties of shapes that are preserved under continuous deformations. As it stands, this description is certainly not precise, but, as this class progresses, we will make it more formal by describing what we mean by "shape" and "continuous deformation". In a couple of weeks, we will have introduced enough terminology to understand that in fact topology is in fact the study of topological spaces up to homeomorphism.

But why study topology, and what should you expect from this class?

- Topology provides the language to formalise the notion of "closeness" when distances cannot be measured. It also formalises what we mean by "shape" and "continuous deformation". Setting up this language will be the goal of Chapter 1.
- Topology provides a more abstract take on a number of topics from real analysis. For instance, the intermediate value theorem holds much more generally than in $\mathbb{R}$ and is underpinned by the concept of connectedness. Similarly, the extreme value theorem relies on the notion of compactness.[1]
- Topology provides the vocabulary and tools to describe and compares spaces. What characteristics do the sphere S^2 and the torus T^2 share? What sets them apart? How can you prove that they are "different"? The question of distinguishing spaces will be at the heart of Chapter 2, where we will learn about the fundamental group.
- Topology is used in a variety of mathematical fields: it is not only a gateway to topics in algebraic and geometric topology, it is also used in functional analysis, algebraic geometry, logic and differential geometry, to only name a few .

On a more practical level, what is assumed throughout these notes?

- First and foremost, it is expected that the reader is comfortable with proofs.
- Next, it will be very helpful if one is familiar with notions from analysis such as continuity and convergence: since we will be learning about generalisations of these concepts, having seen them before will definitely be useful.
- Finally, a word about group theory. On the one hand, group theory is *not* a prerequisite to this class. On the other hand, a large part of this class will be devoted to the so-called fundamental group. So while there is no need to worry (we will introduce all the relevant notions at the beginning of Chapter 2), it is also indispensable to spend some time with the group theory exercises that will be sprinkled through the problem sets.

One final note before jumping into the topic: if while reading these notes, you notice mistakes, inaccuracies or typos, do not hesitate to point them out to me, by email or in person.

[1] The *intermediate value theorem* states that a continuous function $f\colon [a,b] \to \mathbb{R}$ takes on any value between $f(a)$ and $f(b)$. The *extreme value theorem* states that a continuous function $f\colon [a,b] \to \mathbb{R}$ must attain a maximum and minimum at least once.

Chapter 1

Point set topology

This first chapter decomposes into two sections. Section 1.1 introduces the essential definitions and constructions in point set topology, while Section 1.2 focuses on the more involved notions of connectedness and compactness.

1.1 Topological spaces and continuous maps

The goal of this section is to introduce the first notions of point set topology and to familiarise the reader with some typical examples and constructions. In Subsections 1.1.1 and 1.1.2, we define topological spaces, and the maps between them. While we give several abstract examples, a running theme will be to illustrate the theory with the more familiar examples from analysis and, more generally, from metric spaces; in particular this will be the focus of Subsection 1.1.3. Using the notions of bases and subbases from Subsection 1.1.4, the goal of Subsection 1.1.5 is then to build more examples of topological spaces using products and quotients. Finally, subsection 1.1.6 is concerned with the notion of convergence in arbitrary topological spaces and how it compares to the eponymous concept from analysis.

1.1.1 Topological spaces

We introduce the definition of a topological space and discuss some first examples and terminology. The main reference for this section is Section 12 of Munkres' textbook [Mun00].

Definition 1.1. A *topology* on a set X is a family $\mathcal{T}$ of subsets of X such that:

1. The empty set $\emptyset$ and X are elements of $\mathcal{T}$.
2. If $\{U_i\}_{i\in I}$ is a collection of elements of $\mathcal{T}$, then the union $\bigcup_{i\in I} U_i$ is also an element of $\mathcal{T}$.
3. If U_1 and U_2 are elements of $\mathcal{T}$, then their intersection $U_1 \cap U_2$ is also an element of $\mathcal{T}$.

A set X endowed with a topology $\mathcal{T}$ is called a *topological space*. The elements of $\mathcal{T}$ are called the *open sets of X*.

Formally, a topological space is the data of a pair $(X, \mathcal{T})$, with X a set and $\mathcal{T}$ a topology on X. In practice however, when the context is clear, we will often simply refer to X itself as a topological space or, for short, as a space.

Example 1.2. We give some examples of topological spaces.

1. For the set $X = \{a, b, c, d\}$, the family $\mathcal{T} = \{\emptyset, X, \{a\}, \{b\}, \{a, b\}\}$ is a topology.

2. For any set X, the family $\mathcal{T}_{\text{triv}} = \{\emptyset, X\}$ is a topology called the *trivial topology* on X.

3. For any set X, the family $\mathcal{T}_{\text{disc}} = \mathcal{P}(X)$ of all the subsets of X is a topology, called the *discrete topology*. With respect to this topology, all subsets of X are open. We say that X is *discrete* if it is endowed with the discrete topology.

4. Consider $X = \mathbb{R}^n$ with the usual Euclidean distance

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

where $x = (x_i)_{i=1}^n$ and $y = (y_i)_{i=1}^n$. For $x \in \mathbb{R}^n$ and $\varepsilon > 0$, use

$$B(x, \varepsilon) = \{y \in \mathbb{R}^n \mid d(x, y) < \varepsilon\}$$

to denote the ball centered at x with radius ε. The *standard topology* on $X = \mathbb{R}^n$ is

$$\mathcal{T} = \big\{U \subset \mathbb{R}^n \mid \text{ for every } x \in U \text{ there exists } \varepsilon > 0 \text{ such that } B(x, \varepsilon) \subset U\big\}.$$

We check that $\mathcal{T}$ is indeed a topology.

- The empty set belongs to $\mathcal{T}$: this is vacuously true. It is also clear that $\mathbb{R}^n \in \mathcal{T}$.
- We prove that if U_i belong to $\mathcal{T}$ for $i \in I$, then $\bigcup_{i \in I} U_i$ belongs to $\mathcal{T}$. By definition of the union, for $x \in \bigcup_{i \in I} U_i$, there is an index $i \in I$ so that $x \in U_i$. But now since $U_i \in \mathcal{T}$, the definition of $\mathcal{T}$ implies that there exists an $\varepsilon > 0$ so that $B(x, \varepsilon) \subset U_i$. As we certainly have $U_i \subset \bigcup_{i \in I} U_i$, this ball is contained in $\bigcup_{i \in I} U_i$. For every $x \in \bigcup_{i \in I} U_i$, we have therefore found an $\varepsilon > 0$ so that $B(x, \varepsilon) \subset \bigcup_{i \in I} U_i$, proving that $\bigcup_{i \in I} U_i \in \mathcal{T}$.
- We prove that if U_1, U_2 belong to $\mathcal{T}$, then $U_1 \cap U_2$ belongs to $\mathcal{T}$. Given $x \in U_1 \cap U_2$, by definition of the intersection we know that $x \in U_1$ and $x \in U_2$. By definition of $\mathcal{T}$, this means that for $i = 1, 2$, there exists an $\varepsilon_i > 0$ so that $B(x, \varepsilon_i) \subset U_i$. We take $\varepsilon < \min(\varepsilon_1, \varepsilon_2)$ and observe that $B(x, \varepsilon) \subset U_1 \cap U_2$: this follows from the inclusions $B(x, \varepsilon) \subset B(x, \varepsilon_i) \subset U_i$ for $i = 1, 2$. We conclude that $U_1 \cap U_2 \in \mathcal{T}$, as required.

Observe that by definition of $\mathcal{T}$, the balls $B(x, r)$ are open[1], and it is for this reason we often call them *open balls*. Focusing on the case $n = 1$ (i.e. $X = \mathbb{R}$), this means for instance that the interval $B(x, r) = (x - r, x + r)$ is open and more generally, so are the intervals (a, b) for any $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$; see the first problem set.

Next, we describe a slighty different, but equivalent, definition of a topology. Namely instead of requiring the intersection of two open sets to be open, one can require that finite intersections of open sets be open.

Remark 1.3. Condition 3 of Definition 1.1 is equivalent to the following condition. "If $U_1, \ldots, U_n$ are elements of $\mathcal{T}$, then their intersection, $\bigcap_{i=1}^n U_i$, is also an element of $\mathcal{T}$". This is an exercise on the first problem set.

Terminology 1.4. We introduce some additional terminology.

1. A subset $C \subset X$ is *closed in* X if $X \setminus C$ is open in X.

2. A subset $N \subset X$ is a *neighborhood of* $x \in X$ if there exists $U \in \mathcal{T}$ such that $x \in U \subset N$.

[1] Given $y \in B(x, r)$, we take $\varepsilon < r - d(x, y)$ and show that $B(y, \varepsilon) \subset B(x, r)$. For $z \in B(y, \varepsilon)$, we have

$$d(x, z) \le d(x, y) + d(y, z) \le d(x, y) + \varepsilon = r.$$

For a topological space $(X, \mathcal{T})$, the axioms of a topology imply that the empty set $\emptyset$ and X are closed, that intersections of closed subsets are closed, and that the union of finitely many closed subsets is closed; this is an exercise on the first problem set.

Example 1.5. Here are examples of the terminology introduced in Terminology 1.4:

1. The interval $[a, b]$ is closed in $\mathbb{R}$ (with the standard topology) as its complement $\mathbb{R} \setminus [a, b] = (-\infty, a) \cup (b, \infty)$ is open: it is a union of two open sets.
2. The closed interval $[x - \varepsilon, x + \varepsilon] \subset \mathbb{R}$ is a neighborhood of x as it contains the open set $B(x, \varepsilon) = (x - \varepsilon, x + \varepsilon)$ which itself contains x.

The concept of a neighborhood will be discussed further during the first active learning session. We nevertheless record a fact that we will use frequently (a proof will be given in the Active learning session 1.7).

Remark 1.6. A subset $U \subset X$ of a topological space is open if and only if it is a neighborhood of each of its points.

Active learning 1.7. In the first active learning session, we will learn some further terminology and illustrate it with examples. Here is a quick summary of what will be covered, without details.

- Let Y be a set. For two topologies $\mathcal{T}_1, \mathcal{T}_2$ on Y, with $\mathcal{T}_1 \subset \mathcal{T}_2$, we say that $\mathcal{T}_1$ is *coarser* than that $\mathcal{T}_2$ and that $\mathcal{T}_2$ is *finer* than $\mathcal{T}_1$. Compare the four topologies on $Y = \{1, 2\}$.
- Let $(X, \mathcal{T})$ be a topological space and let $A \subset X$ be a subset.
  - The *interior* of A in X is the subset:
  $$\mathring{A} = \{x \in X \mid A \text{ is a neighborhood of } x \in X\}.$$
  We will prove that $\mathring{A} = \bigcup_{U \text{ open}, A,\ U \subset A} U$ and deduce that $\mathring{A}$ is the largest open set included in A. In particular $\mathring{A}$ is open and is contained in A. Additionally, one sees that A is open if and only if it is a neighborhood of each of its points.
  - The *closure* of A in X is the subset $\overline{A} \subset X$ defined by
  $$\overline{A} = \{x \in X \mid X \setminus A \text{ is not a neighbhorhood of } x \in X\}.$$
  We will prove that $\overline{A} = \bigcap_{C \text{ closed},\ C \supset A} C$ and deduce that $\overline{A}$ is the smallest closed set in X containing A. In particular, $\overline{A}$ is closed and contains A.
  - The *boundary* of $A \subset X$ is defined as $\partial A = \overline{A} \setminus \mathring{A}$.

  We will discuss these concepts in the case where $X = \mathbb{R}$ and $A = [0, 1)$.

1.1.2 Continuous maps

We introduce continuous maps between topological spaces, as well as homeomorphisms and the subspace topology. The main reference for this section is [Mun00, Section 18].

Definition 1.8. Let X, Y be two topological spaces. A map $f \colon X \to Y$ is *continuous* if for every open set $U \subset Y$, the subset $f^{-1}(U) \subset X$ is open.

We start with some first remarks about continuous maps.

Remark 1.9. Let X, Y, Z be topological spaces.[2]

[2]This remark can be summarised by saying that topological spaces as objects and continuous maps as morphisms form a *category*.

1. The identity map $\mathrm{id}_X\colon (X,\mathcal{T}) \to (X,\mathcal{T})$ is continuous: if $U \subset X$ is open, then $\mathrm{id}_X^{-1}(U) = U$ is also open in X. Here, it matters that the domain and target are endowed with the same topology: in general $\mathrm{id}_X\colon (X,\mathcal{T}_1) \to (X,\mathcal{T}_2)$ is continuous if and only if $\mathcal{T}_1$ is finer than $\mathcal{T}_2$.

2. If $f\colon X \to Y$ and $g\colon Y \to Z$ are continuous, then $g \circ f\colon X \to Z$ is also continuous: if $U \subset Z$ is open, then $(g \circ f)^{-1}(U) = g^{-1}(f^{-1}(U))$ is also open in X.

Before giving examples of continuous functions, we describe some equivalent characterisations of continuity.

Proposition 1.10. *For a map $f\colon X \to Y$ between topological spaces, the following are equivalent:*

1. *the map f is continuous;*
2. *for every closed subset $C \subset Y$, the set $f^{-1}(C)$ is closed in X;*
3. *for every $x \in X$, the following property holds: for every neighborhood V of $f(x) \in Y$, the set $f^{-1}(V) \subset X$ is a neighborhood of $x \in X$.*

Proof. The equivalence between the first two statements is an exercise on the first problem set. It remains to prove the equivalence between the first and the third items.

We prove that (1) $\Rightarrow$ (3). Assume that f is continuous and fix a neighborhood V of $f(x) \in Y$. By definition of a neighborhood, there is an open set $U \subset Y$ with $f(x) \in U \subset V$. It follows that $f^{-1}(U) \subset X$ is an open set with $x \in f^{-1}(U) \subset f^{-1}(V)$, and so $f^{-1}(V)$ is a neighborhood of $x \in X$.

Finally, we prove that (3) $\Rightarrow$ (1). Let $U \subset Y$ be an open subset. As U is open, it is a neighborhood of each of its points; recall Remark 1.6. In particular, U is a neighborhood of $f(x) \in U$ for every $x \in f^{-1}(U)$. By condition (3), this implies that $f^{-1}(U)$ is a neighborhood of every $x \in f^{-1}(U)$. This means that $f^{-1}(U)$ is open and therefore f is continuous, as required. □

If $f\colon X \to Y$ satisfies the third property of Proposition 1.10 for $x \in X$, we say that f *is continuous at x*. The equivalence between the first and third item of Proposition 1.10 therefore says that f is continuous if and only if f is continuous at every $x \in X$.

Next, we move on to some first examples of continuous maps.

Example 1.11. Let X and Y be topological spaces.

1. Constant maps are continuous: if $f\colon X \to Y$ is the constant map with value $y \in Y$, then for an open set $U \subset Y$, the inverse image $f^{-1}(U)$ equals X if U contains y and is empty otherwise. In either case, $f^{-1}(U)$ is open in Y.

2. If a space X is discrete, then every map $f\colon X \to Y$ is continuous: if $U \subset Y$ is open, then so is $f^{-1}(U) \subset X$ because all subsets of a discrete space are open.

3. If a set Y is endowed with the trivial topology, then every map $f\colon X \to Y$ is continuous: the topology on Y is $\{\emptyset, Y\}$ and both $f^{-1}(Y) = X$ and $f^{-1}(\emptyset) = \emptyset$ are open in X.

4. If one endows $\mathbb{R}^n$ and $\mathbb{R}^m$ with the standard topology, then we will see in Subsection 1.1.3 that a map $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is continuous at $x \in \mathbb{R}^n$ if and only if for every $x \in \mathbb{R}^n$ and every $\varepsilon > 0$, there exists a $\delta > 0$ such that if $d(x,y) < \delta$, then $d(f(x), f(y)) < \varepsilon$. In other words, in Euclidean space, the notion of continuity from Definition 1.8 matches the one seen during a first analysis course. Thus, for instance, the function $f\colon \mathbb{R} \to \mathbb{R}, x \mapsto ax + b$ is continuous for any $a, b \in \mathbb{R}$.

As in many fields of mathematics, we introduce some terminology to specify when two objects (in this case topological spaces) are thought of as being "the same".

Definition 1.12. A map $f\colon X \to Y$ between topological spaces is a *homeomorphism* if f is continous, bijective and $f^{-1}\colon Y \to X$ is continuous. If such a map exists, then the spaces X and Y are called *homeomorphic*, and this is denoted $X \cong Y$.

It might seem redundant to ask for the inverse of a continuous bijective map to also be continuous: one could hope that this condition is automatic. As the next example shows however, this is not the case.

Remark 1.13. A bijective continuous map need not be a homeomorphism. For example, for any set X, the map $f = \mathrm{id}_X\colon (X, \mathcal{T}_{\text{disc}}) \to (X, \mathcal{T}_{\text{triv}})$ is continuous and bijective, however its inverse $f^{-1} = \mathrm{id}_X\colon (X, \mathcal{T}_{\text{triv}}) \to (X, \mathcal{T}_{\text{disc}})$ is not continuous if X contains more than one element. Indeed, as X has more than one element, one can pick a subset $U \subset X$ with $U \neq \emptyset, X$ (an element of the discrete topology); it follows that $U = \mathrm{id}_X(U) \notin \{\emptyset, X\} = \mathcal{T}_{\text{triv}}$.

Instead of immediately giving examples of homeomorphisms, we first widen our range of topological spaces and of continuous maps.

Definition 1.14. Let $(X, \mathcal{T}_X)$ be a topological space. The *induced topology* on a subset $A \subset X$ is

$$\mathcal{T}_A = \{U \subset A \mid \text{ there exists an open set } V \subset X \text{ with } U = A \cap V\}.$$

It can be checked that $\mathcal{T}_A$ is a topology on A (this is an exercise on the first problem set), and we say that A is a *subspace* of X.

Example 1.15. Consider the set $X = \mathbb{R}$ with the standard topology, as well as the subspaces $A = [0, \infty)$, and $U = [0, 1)$, so that $U \subset A \subset X$.

1. $U = [0, 1)$ is open in $A = [0, \infty)$, as $[0, 1) = (-1, 1) \cap A$ with $(-1, 1)$ open in $\mathbb{R}$.
2. $U = [0, 1)$ is not open in $X = \mathbb{R}$, since U is not a neighborhood of $0 \in U$ (recall from Remark 1.6 that U is open if and only if it is a neighborhood of all of its points).

Example 1.15 illustrates the following point. Declaring a set to be open only makes sense if the ambient topological space is clear: one should say "A is open in X", instead of "A is open".

Finally, we list some additional properties of continuous maps.

Proposition 1.16. *Let X, Y be topological spaces, and let $f\colon X \to Y$ be a continuous map.*

1. *If $A \subset X$ is a subspace, then the inclusion $j\colon A \to X$ is continuous.*
2. *The restriction $f|_A\colon A \to Y$ is continuous.*
3. *If $B \subset Y$ is a subspace with $\mathrm{im}(f) \subset B$, then the map $g\colon X \to B$ obtained from f by restricting the target is continuous.*
4. *If $Y \subset Z$ is a subspace, then the function $h\colon X \to Z$ obtained from f by extending the target is continuous.*

Proof. This is an exercise on the first problem set. □

Example 1.17. Here are more examples of (non-)homeomorphisms:

1. The function $f\colon \mathbb{R} \to \mathbb{R}, x \mapsto ax + b$ is a homeomorphism: it is continuous and bijective with continuous inverse $g\colon \mathbb{R} \to \mathbb{R}, y \mapsto (y - b)/a$. Here, to see that f and g are continuous, recall from Example 1.11 that the "general" definition of continuity from Definition 1.8 coincides with the ε-δ definition seen during an analysis course.
2. Consider $(-1, 1)$ with the topology induced by $\mathbb{R}$ and $f\colon \mathbb{R} \to (-1, 1), x \mapsto \frac{x}{1+|x|}$. We argue that f is a homeomorphism: the map f is bijective, with inverse $g\colon (-1, 1) \to \mathbb{R}$ given by $g(y) = \frac{y}{1-|y|}$, and using the knowledge from a first analysis course, these two maps are known to be continuous.

3. We conclude with a more satisfactory example of a continuous bijective map that is not a homeomorphism (a first example appeared in Example 1.13). Endow $\mathbb{R}$ and $\mathbb{R}^2$ with the standard topology, and consider the subspaces $[0,1) \subset \mathbb{R}$ and

$$S^1 = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} \subset \mathbb{R}^2,$$

as well as the map $f\colon [0,1) \to S^1, t \mapsto (\cos(2\pi t), \sin(2\pi t))$. We argue that f is continuous and bijective, but not a homeomorphism. First, f is continuous because it is obtained by restricting the domain and target of the continuous map $\mathbb{R} \to \mathbb{R}^2, t \mapsto (\cos(2\pi t), \sin(2\pi t))$; here we used the second and third points of Proposition 1.16. Verifying that f is bijective is not overly difficult and so, instead, we argue that its inverse $g := f^{-1}\colon S^1 \to [0,1)$ is not continuous. Indeed, $U = [0,1/4)$ is a neighborhood of $g(1,0) = f^{-1}(1,0) = 0$, but $g^{-1}(U) = f(U)$ is not a neighborhood of $(1,0) \in S^1$.[3] Thus by Proposition 1.10, g is not continuous (at $(1,0) \in S^1$).

1.1.3 Metric spaces

We introduce metrics and describe how they induce topologies. We focus in particular on the various metrics on $\mathbb{R}^n$ and the topologies they induce. The main reference for this section is [Mun00, Sections 20 and 21].

Definition 1.18. Let X be a set. A map $d\colon X \times X \to [0,\infty)$ is called a *distance* or a *metric* if it satisfies the three following properties for all $x,y,z \in X$:

1. non-degeneracy: $d(x,y) = 0$ if and only if $x = y$;
2. symmetry: $d(x,y) = d(y,x)$;
3. triangle inequality: $d(x,z) \leq d(x,y) + d(y,z)$.

A set X endowed with a metric is called a *metric space.*

Formally, a metric space is the data of a pair (X,d), but when the distance d is clear from the context, we often refer to X as a metric space.

Terminology 1.19. Let (X,d) be a metric space. Given $x \in X$ and $\varepsilon > 0$, the *(open) ball of radius ε centered at x* is the set

$$B_d(x,\varepsilon) := \{y \in X \mid d(x,y) < \varepsilon\}.$$

Again, when the distance is clear, we drop it from the notation, and simply write $B(x,\varepsilon)$.

Example 1.20. Here are some examples of metric spaces.

1. The set $X = \mathbb{R}$ can be endowed with the metric $d(x,y) = |x-y|$ in which case, the balls are the intervals $B(x,\varepsilon) = (x-\varepsilon, x+\varepsilon)$.
2. More generally, given $p \in \mathbb{R}_{\geq 1}$, the set $X = \mathbb{R}^n$ can be equipped with the metric d_p, defined for $x = (x_i)_{i=1}^n$ and $y = (y_i)_{i=1}^n$ by the formula

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

[3] Assume for a contradiction that $f(U)$ is a neighborhood of $(1,0) \in S^1$. Thus, there exists an open set $V \subset S^1$ with $(1,0) \in V \subset f(U)$. By definition of the induced topology, we know that $V = W \cap S^1$ with $W \subset \mathbb{R}^2$ open. By definition of the standard topology on $\mathbb{R}^2$, there is an $\varepsilon > 0$ so that $(1,0) \in B((1,0),\varepsilon) \subset W$ and so

$$(1,0) \in B((1,0),\varepsilon) \cap S^1 \subset V \subset f(U) = f([0,1/4)).$$

On the other hand, we can find $y \in B((1,0),\varepsilon) \cap S^1$ so that $f^{-1}(y) \in [1/4,1)$, contradicting the inclusion $B((1,0),\varepsilon) \cap S^1 \subset f([0,1/4))$.

The triangle inequality is proved using Minkowski's inequality, from analysis. For $p = 2$, one obtains the usual Euclidean metric. For $p = \infty$, one gets a metric by setting

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

3. Any non-empty set X can be endowed with the *discrete metric*

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

As we alluded to, a metric gives rise to a topology.

Definition 1.21. Given a metric space (X, d), the *topology induced by the metric* d is

$$\mathcal{T}_d = \{U \subset X \mid \text{ for every } x \in U, \text{ there exists } \varepsilon > 0 \text{ such that } B_d(x, \varepsilon) \subset U\}.$$

The fact that $\mathcal{T}_d$ is a topology is proved in the exact same way as in Example 1.2.

Remark 1.22. Here are some remarks about the topology induced by a metric.

1. We check that open balls are in fact open, i.e. that they belong to $\mathcal{T}_d$. Given $y \in B(x, r)$, we set $\varepsilon = r - d(x, y)$ and show that $B(y, \varepsilon) \subset B(x, r)$. For $z \in B(y, \varepsilon)$, we have

$$d(x, z) \le d(x, y) + d(y, z) \le d(x, y) + \varepsilon = r.$$

2. We check that a subset $V \subset X$ is a neighborhood of $x \in X$ if and only if there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subset V$. If $V \subset X$ is a neighborhood of $x \in X$, there exists an open set $U \subset X$ with $x \in U \subset V$ and so, by definition of $\mathcal{T}_d$, there exists $\varepsilon > 0$ so that $B(x, \varepsilon) \subset U \subset V$. Conversely, if there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subset V$, then we have $x \in B(x, \varepsilon) \subset V$, with $B(x, \varepsilon) \subset X$ open (by the first item), proving that V is a neighborhood of x.

Example 1.23. Here are some examples of topologies induced by a metric:

1. By definition, the Euclidean metric d_2 induces the standard topology on $X = \mathbb{R}^n$.
2. For an arbitrary set X, the discrete metric d induces the discrete topology. To see this, we must show that $\mathcal{T}_d = \mathcal{P}(X)$. Since $\mathcal{T}_d \subset \mathcal{P}(X)$, we need only show $\mathcal{T}_d \supset \mathcal{P}(X)$. The empty set $U = \emptyset \in \mathcal{T}_d$ belongs to $\mathcal{P}(X)$. If $U \subset X$ is a non-empty open set, then for any $x \in U$, and any $0 < \varepsilon < 1$, we have $B_d(x, \varepsilon) = \{x\} \subset U$, proving that $U \in \mathcal{T}_d$.

Given the first item of Example 1.23, it is interesting to wonder about the topologies induced by the d_p for $p \ne 2$. Do we get the same result for all p or do the d_p induce different topologies? To answer this question, we introduce some terminology.

Definition 1.24. Two metrics d and d' on a set X are *equivalent* if there exist positive constants c, c' such that for all $x, y \in X$

$$cd(x, y) \le d'(x, y) \le c'd(x, y).$$

A motivation for introducing this notion lies in the following proposition.

Proposition 1.25. *Equivalent metrics induce the same topology.*

Proof. Assume that d and d' are equivalent metrics on a set X. We show that $\mathcal{T}_d = \mathcal{T}_{d'}$. We first prove the inclusion $\mathcal{T}_d \subset \mathcal{T}_{d'}$. Let $U \in \mathcal{T}_d$ be an open set. For every $x \in U$, there exist an $r > 0$ such that $B(x, r) \subset U$. Since d and d' are equivalent, there exist positive constants c, c' such that $cd(x, y) \le d'(x, y) \le c'd(x, y)$ for every $y \in X$. This implies that

$$B_{d'}(x, cr) \subset B_d(x, r) \subset B_{d'}(x, c'r).$$

This first inclusion implies that $B_{d'}(x, cr) \subset B_d(x, r) \subset U$ which implies that $U \in \mathcal{T}_{d'}$. The proof of the inclusion $\mathcal{T}_d \supset \mathcal{T}_{d'}$ proceeds similarly. ☐

Next, we note that the converse to Proposition 1.25 is false.

Remark 1.26. We present pairs of metrics that induce the same topology and yet are not equivalent. Given a metric space (X, d), consider $\overline{d}(x, y) = \frac{d(x,y)}{1+d(x,y)}$. It is an exercise on the second problem set to show that $\overline{d}$ is indeed a metric that induces the same topology as d. It is then a further exercise to prove that these metrics are in general not equivalent.

Example 1.27. Here are some examples of equivalent metrics.

1. On $X = \mathbb{R}^n$, the metrics d_p are seen to all be equivalent by noting that for all $p \in [1, \infty)$, and for all $x, y \in \mathbb{R}^n$, the following inequalities hold (we omit the details):
$$d_\infty(x, y) \leq d_p(x, y) \leq n^{1/p} d_\infty(x, y).$$
Thanks to Proposition 1.25, we deduce that the metrics d_p therefore all induce the same topology as d_2, i.e. they all induce the standard topology.

2. On $X = \mathbb{R}$, the standard metric and the discrete metric are not equivalent. This can be shown directly, but one can also use the contrapositive of Proposition 1.25: as the induced topologies are distinct, the metrics cannot be equivalent.

We conclude this subsection on metric spaces by proving a result that we already mentioned in Example 1.11: in a metric space, the definition of continuity is equivalent to the more familiar "ε, δ"-definition from a first course in analysis.

Proposition 1.28. *Let (X, d_X) and (Y, d_Y) be two metric spaces, let $f\colon X \to Y$ be a map, and let $x \in X$. The following assertions are equivalent:*

1. *The map f is continuous at x.*
2. *For every $\varepsilon > 0$, there exists $\delta > 0$ such that for every $x' \in X$ with $d_X(x, x') < \delta$, one has $d_Y(f(x), f(x')) < \varepsilon$.*

Proof. We check that (1) $\Rightarrow$ (2). We saw in Proposition 1.10 that f is continuous at x if and only if for every neighborhood $V \subset Y$ of $f(x)$, the set $f^{-1}(V) \subset X$ is a neighborhood of x. As we saw in Remark 1.22, this is in turn equivalent to asking that for every $V \subset Y$ with $B_{d_Y}(f(x), \varepsilon) \subset V$ for some $\varepsilon > 0$, there exists a $\delta > 0$ such that $B_{d_X}(x, \delta) \subset f^{-1}(V)$. *In particular,* taking $V = B_{d_Y}(f(x), \varepsilon)$, we get that for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $B_{d_X}(x, \delta) \subset f^{-1}(B_{d_Y}(f(x), \varepsilon))$. This statement is equivalent to (2): for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every $x' \in X$ with $d_X(x, x') < \delta$, one has $d_Y(f(x), f(x')) < \varepsilon$."

We now check that (2) $\Rightarrow$ (1). In the proof of (1) $\Rightarrow$ (2), all statements were equivalences except for the penultimate sentence. We show that this implication is also an equivalence. Thus we suppose that for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $B_{d_X}(x, \varepsilon) \subset f^{-1}(B_{d_Y}(f(x), \varepsilon))$ and we fix a $V \subset Y$ such that $B_{d_Y}(f(x), \varepsilon) \subset V$ for some $\varepsilon > 0$. By hypothesis, there exists $\delta > 0$ such that $B_{d_X}(x, \delta) \subset f^{-1}(B_{d_Y}(f(x), \varepsilon))$ which is itself in $f^{-1}(V)$. We have therefore shown that (1) $\Leftrightarrow$ (2) and this concludes the proof of the proposition. □

1.1.4 Bases and subbases

At this point, we have introduced the mathematical objects (topological spaces) we wish to study and the maps between them (continuous maps) and, additionally, we have seen how these definitions relate to notions studied in analysis when working in $\mathbb{R}^n$. On the other hand, we lack constructions of further examples of topological spaces. To remedy this, we need to a way to specify a topological space without having to enumerate all the open sets in its topology. This is the point of bases and subbases, and the main reference for this section is [Mun00, Section 13].

Definition 1.29. Let $(X, \mathcal{T})$ be a topological space.

1. A subset $\mathcal{B} \subset \mathcal{T}$ is a *basis* for the topology $\mathcal{T}$ if every element of $\mathcal{T}$ can be written as a union of elements of $\mathcal{B}$.

2. A subset $S \subset \mathcal{T}$ is a subbasis of $\mathcal{T}$ if the set of finite intersections of elements of S forms a basis for $\mathcal{T}$.

If $\mathcal{B}$ is a basis for a topology, then this topology is entirely determined by $\mathcal{B}$: elements of $\mathcal{T}$ are given by unions of its elements (and similarly for subbases). For this reason, we often say that $\mathcal{T}$ *is generated* by $\mathcal{B}$ or S and we write $\mathcal{T}_{\mathcal{B}}$ and $\mathcal{T}_S$.

Remark 1.30. Here are some remarks concerning bases and subbases:

1. Every basis is a subbasis.

2. A subset $\mathcal{B} \subset \mathcal{T}$ is a basis for a topology $\mathcal{T}$ if and only if for every $U \in \mathcal{T}$ and every $x \in U$, there exists $B \in \mathcal{B}$ such that $x \in B \subset U$.

3. Bases and subbases provide a simpler way to check that a map is continuous. A map $f\colon X \to Y$ is continuous for a topology $\mathcal{T}_S$ on Y if and only if $f^{-1}(U) \subset X$ is open for every $U \in S$[4]; this is a exercise on the second problem set.

Example 1.31. Here are some examples of bases and subbases.

1. Set $X = \{0, 1, 2\}$. A basis for the topology $\mathcal{T} = \{\emptyset, X, \{0\}, \{0, 1\}, \{0, 2\}\}$ on X is given by $\mathcal{B} = \{\{0\}, \{0, 1\}, \{0, 2\}\}$ while a subbasis is given by $S = \{\{0, 1\}, \{0, 2\}\}$.

2. For any set X, the set $\mathcal{B} = \{\{x\} \mid x \in X\}$ of all singletons in X is a basis for the discrete topology on X.

3. For any set X, the set $\mathcal{B} = \{X\}$ is a basis for the trivial topology on X.

4. For $X = \mathbb{R}^n$ the set of open balls is a basis for the standard topology. More generally, on a metric space (X, d), the set of all open balls

$$\mathcal{B} = \{B(x, r) \mid x \in X, r > 0\}$$

is a basis for the topology $\mathcal{T}_d$ on X. Indeed, if $U \in \mathcal{T}_d$, then for every $x \in U$, there exists an $r(x) > 0$ with $B(x, r(x)) \subset U$. This implies the equality

$$U = \bigcup_{x \in U} B(x, r(x))$$

which shows that $\mathcal{B}$ is indeed a basis for $\mathcal{T}_d$.

Next, we provide an equivalent definition of a basis for a topology, which is useful in practice to specify a topology on a set X without having to describe all its open sets.

Proposition 1.32. *For a set X, a subset $\mathcal{B} \subset \mathcal{P}(X)$ is a basis for a topology on X if and only if the two following conditions hold:*

1. *the set X is covered by the elements of $\mathcal{B}$, i.e. $X = \bigcup_{B \in \mathcal{B}} B$;*

2. *for every $B_1, B_2 \in \mathcal{B}$ and every $x \in B_1 \cap B_2$, there exists $B \in \mathcal{B}$ with $x \in B \subset B_1 \cap B_2$.*

Proof. We prove the forward direction: we assume that $\mathcal{B}$ is a basis for a topology $\mathcal{T}$ and verify that it satisfies conditions (1) and (2). Since $\mathcal{B}$ is a basis for $\mathcal{T}$, every element of $\mathcal{T}$ is covered by elements of $\mathcal{B}$. In particular $X \in \mathcal{T}$, is a union of elements of $\mathcal{B}$. For the second condition,

[4] The same is true for bases.

if $B_1, B_2 \in \mathcal{B} \subset \mathcal{T}$, then we know that $B_1 \cap B_2 \in \mathcal{T}$ and thus $B_1 \cap B_2$ is covered by elements of $\mathcal{B}$ which gives the second point.

We now prove the backward direction. Assuming that $\mathcal{B}$ satisfies (1) and (2), we show that the family $\mathcal{T}$ of unions of elements of $\mathcal{B}$ is a topology (by definition $\mathcal{B}$ will then be a basis for $\mathcal{T}$). We verify that $\mathcal{T}$ satisfies the three axioms of a topology. For the first axiom, by convention, the empty set is an empty union and so belongs to $\mathcal{T}$, while $X \in \mathcal{T}$ by (1). The second axiom is immediate: a union of elements of $\mathcal{B}$ is still a union of elements of $\mathcal{B}$. To prove the third axiom, assume that $B, B' \in \mathcal{T}$ so that $B = \bigcup_i B_i$ and $B' = \bigcup_i B'_i$ with $B_i, B'_i \in \mathcal{B}$; our aim is to show that $B \cap B' = \cup_{i,j} B_i \cap B'_j \in \mathcal{T}$. This reduces to showing that the intersection of two elements $B_1, B_2 \in \mathcal{B}$ still belongs to $\mathcal{T}$. This now follows from the second point: since for every $x \in B_1 \cap B_2$, there exists a $B(x) \in \mathcal{B}$ with $x \in B(x) \subset \mathcal{B}_1 \cap \mathcal{B}_2$, we can write $B_1 \cap B_2 = \bigcup_{x \in B_1 \cap B_2} B(x) \in \mathcal{T}$, and we are done. □

Using Proposition 1.32, we deduce a similar characterisation of subbases.

Corollary 1.33. *A subset $S \subset \mathcal{P}(X)$ is a subbasis of a topology on X if and only if $X = \bigcup_{U \in S} U$.*

Proof. We prove the forward direction. If S is a subbasis for a topology $\mathcal{T}$, then every element of $\mathcal{T}$ is a union of finite intersections of elements of S. Applying this to $X \in \mathcal{T}$, we can write $X = \bigcup_{i \in I} \bigcap_{i_j=1}^{n_i} V_{i_j}$ with $V_{i_j} \in S$. This latter set is contained in $\bigcup_{U \in S} U$ and so we get $X = \bigcup_{U \in S} U$, as required.

We prove the converse: we assume that $X = \bigcup_{U \in S} U$ and prove that S is a subbasis for a topology. Using $\mathcal{B}$ to denote the family of all finite intersections of elements of S, it is equivalent to show that $\mathcal{B}$ is a basis for a topology $\mathcal{T}$ (it will then follow that S is a subbasis for $\mathcal{T}$). We check that $\mathcal{B}$ satisfies the two conditions of Proposition 1.32. The first follows from the sequence of inclusions

$$X = \bigcup_{U \in S} U \subset \bigcup_{B \in \mathcal{B}} B \subset X.$$

To see the second, note that by definition of $\mathcal{B}$, the intersection of two elements $B_1, B_2 \in \mathcal{B}$ is again an element of $\mathcal{B}$. Thus for $x \in B_1 \cap B_2$, we can pick $B = B_1 \cap B_2 \in \mathcal{B}$ and have $x \in B \subset B_1 \cap B_2$. This concludes the proof. □

1.1.5 The product and quotient topology

Armed with the notions of bases and subbases, we can now describe how to endow products and quotients with a topology. This will prove useful to exhibit familiar shapes such as the torus and the sphere as topological spaces. In both cases, we will begin by recalling some notions from set theory. The main reference for this section is [Mun00, Sections 15 and 22].

The product topology

There are two natural topologies that one can put on products of spaces: the first is called the *box topology*, while the second is called the *product topology*. Despite the fact that both topologies coincide on finite products, the definition of the box topology is easier to remember. As we shall see however, for arbitrary products, the product topology is preferrable: the box topology lacks several desirable properties.

Recall that the *product* of the sets $X_1, X_2, \ldots, X_n$ is the set

$$\prod_{i=1}^{n} X_i = X_1 \times X_2 \times \ldots \times X_n = \{(x_1, x_2, \ldots, x_n) \mid x_i \in X_i \text{ for } i = 1, \ldots, n\}. \tag{1.1}$$

More generally, the *product* of a collection $\{X_i\}_{i \in I}$ indexed by a set I is the set

$$\prod_{i \in I} X_i = \left\{ x \colon I \to \bigcup_{i \in I} X_i \;\middle|\; x(i) \in X_i \text{ for all } i \in I \right\}. \tag{1.2}$$

Observe that if $I = \{1, \dots, n\}$, then (1.2) recovers the definition of the product given in (1.1). Additionally, elements of a product as in (1.2) are denoted $x = (x_i)_{i \in I}$ and there are projections

$$\pi_j \colon \prod_{i \in I} X_i \to X_j$$
$$x \mapsto x(j) = x_j.$$

Furthermore, if $X_i = X$ for every $i \in I$, then we often write the product as X^I (indeed, X^I coincides with the set $\mathrm{Map}(I, X)$ of maps from I to X).

We now define what will be a basis for a topology on this product.

Construction 1.34. Consider the subset $\mathcal{B}' \subset \mathcal{P}\left(\prod_{i \in I} X_i\right)$ defined by

$$\mathcal{B}' := \left\{ \prod_{i \in I} U_i \subset \prod_{i \in I} X_i \;\middle|\; U_i \subset X_i \text{ is open for all } i \in I \right\}.$$

We verify that this is a basis by using the criterion from Proposition 1.32. The first condition holds because $\prod_{i \in I} X_i$ belongs to $\mathcal{B}'$. The second condition follows promptly from the set theoretical equality $(\prod_{i \in I} U_i) \cap (\prod_{i \in I} V_i) = \prod_{i \in I} (U_i \cap V_i)$.

Definition 1.35. Given a family of topological spaces $\{X_i\}_{i \in I}$, the *box topology* on the product $\prod_{i \in I} X_i$ is the topology generated by the basis $\mathcal{B}'$.

Example 1.36. Here are some examples of the box topology:

1. If $\{X_i\}_{i \in I}$ is a family of discrete spaces, then the box topology on $\prod_{i \in I} X_i$ is the discrete topology. Indeed, to see that $\mathcal{P}\left(\prod_{i \in I} X_i\right) \subset \mathcal{T}_{\mathcal{B}'}$, note that any set $U \subset \prod_{i \in I} X_i$ can be written as
$$U = \bigcup_{x \in U} \{x\} = \bigcup_{x \in U} \prod_{i \in I} \{x_i\}.$$
2. If $\mathbb{R}$ is endowed with the standard topology, then the box topology on $\mathbb{R}^n$ coincides with the standard topology; this is an exercise on the second problem set.

The next example shows the limitations of the box topology on infinite products.

Example 1.37. Consider a countable family of copies of $\mathbb{R}$ with the standard topology: $X_i = \mathbb{R}$ for $i \in \mathbb{N}$. We show that using the box topology, the map $f \colon \mathbb{R} \to \mathbb{R}^{\mathbb{N}}, t \mapsto (t, t, \dots)$ is not continuous. Indeed, for $U_i = (-\frac{1}{i}, \frac{1}{i}) \subset X_i$, the set $U := \prod_{i \in \mathbb{N}} U_i \subset \mathbb{R}^{\mathbb{N}}$ is open (since each $U_i \subset X_i$ is open), but $f^{-1}(U) \subset \mathbb{R}$ is not open since

$$f^{-1}(U) = \{t \in \mathbb{R} \mid t \in U_i \text{ for every } i \in \mathbb{N}\} = \bigcap_{i \in \mathbb{N}} U_i = \{0\}.$$

Each of the components of the map f from Example 1.37 is continuous and we therefore expect f itself to be continuous under a reasonable topology on $\mathbb{R}^{\mathbb{N}}$. The conclusion is that the box topology is not adequate for arbitrary products as it may contain too many open sets. For instance, in Example 1.37, we want a topology on $\mathbb{R}^{\mathbb{N}}$ for which $\prod_{i \in \mathbb{N}} (-\frac{1}{i}, \frac{1}{i})$ is not open.

Construction 1.38. Let $\{X_i\}_{i \in I}$ be a family of spaces and recall that $\pi_j \colon \prod_{i \in I} X_i \to X_j$ denotes the j-th projection. Consider the subset $S \subset \mathcal{P}(\prod_{i \in I} X_i)$ defined by

$$S = \left\{ \pi_j^{-1}(U_j) \subset \prod_{i \in I} X_i \;\middle|\; j \in I, U_j \subset X_j \text{ open} \right\}.$$

Note that by definition, elements of S are of the form

$$\pi_j^{-1}(U_j) = U_j \times \prod_{i \in I \setminus \{j\}} X_i \tag{1.3}$$

with $j \in I$ and $U_j \subset X_j$ an open set. The family S of subsets satisfies the subbasis criterion of Corollary 1.33 (by (1.3), we have $\prod_{i\in I} X_i = \pi_j^{-1}(X_j) \in S$ for any j). The basis corresponding to S is given by finite intersections of elements of S:

$$\mathcal{B} = \left\{ \prod_{j\in J} U_j \times \prod_{i\in I\setminus J} X_i \;\middle|\; J \subset I \text{ finite, } U_j \subset X_j \text{ open for all } j \in J \right\}.$$

Definition 1.39. Given a family of topological spaces $\{X_i\}_{i\in I}$, the *product topology* on $\prod_{i\in I} X_i$ is the topology generated by $\mathcal{B}$.

Remark 1.40. Let $\{X_i\}_{i\in I}$ be a collection of sets.

1. If the index set I is finite, then the basis $\mathcal{B}$ for the box topology and the basis $\mathcal{B}'$ for the product topology are equal; thus the box topology and the product topology agree for finite products. In general however, the product topology is coarser that the box topology.

2. The subbasis S consists of the subsets that must be open for the projections π_j to be continuous. In other words, the product topology is the coarsest topology for which all projections are continuous.

3. We check that for every set X and every family of maps $\{f_i \colon X \to X_i\}_{i\in I}$ there exists a unique map $f \colon X \to \prod_{i\in I} X_i$ so that $\pi_i \circ f = f_i$ for all $i \in I$. For existence: for $x \in X$, define $f(x)_i := f_i(x)$; uniqueness is not hard to verify.

   In fact, if X is a space, then the map f is continuous for the product topology if and only if all the f_i are continuous. Indeed, if f is continuous, then $f_i = \pi_i \circ f$ is also continuous (because π_i is continuous for each $i \in I$); conversely if each f_i is continuous, then for each $U := \pi_i^{-1}(U_i) \in S$ with $U_i \subset X_i$ open,

   $$f^{-1}(U) = f^{-1}\pi_i^{-1}(U_i) = (\pi_i \circ f)^{-1}(U) = f_i^{-1}(U_i) \subset X$$

   is open in X because $U_i \subset X_i$ is open and f_i is continuous (here recall that by Remark 1.30, this suffices to show that f is continuous).

Example 1.41. We give some examples involving the product topology.

1. We argue that the map $f \colon \mathbb{R} \to \mathbb{R}^{\mathbb{N}}, t \mapsto (t, t, \ldots)$ from Example 1.37 is continuous when $\mathbb{R}^{\mathbb{N}}$ is endowed with the product topology. This is a consequence of Remark 1.40 above: the fact that each $f_i \colon \mathbb{R} \to \mathbb{R}$ is given by the continuous map $f_i(t) = t$ implies that f itself is continuous.

2. If $\{X_i\}_{i\in I}$ is a family of discrete spaces, then the product topology on $\prod_{i\in I} X_i$ is in general not the discrete topology. For instance, it is an exercise on the third problem set to show that $X = \prod_{i\in I}\{0, 1\}$ is discrete if and only if I is finite.

The quotient topology

Before defining the quotient topology, we first recall the notion of the quotient of a set by an equivalence relation.

Remark 1.42. Let X be a set.

- A *(binary) relation* on X is a subset $\mathcal{R} \subset X \times X$. Instead of writing $(x, y) \in \mathcal{R}$, we write $x\mathcal{R}y$. Examples to have in mind are $x\mathcal{R}_= y$ if and only if $x = y$ or $x\mathcal{R}_\leq y$ if and only if $x \leq y$, or if $X = \mathbb{R}$, and $n \in \mathbb{Z}$ is fixed, another example is $x\mathcal{R}_n y$ if and only if $x - y$ is divisible by n.

- A relation on a set X is an *equivalence relation* if

1. $\mathcal{R}$ is reflexive: $x\mathcal{R}x$ for all $x \in X$;
2. $\mathcal{R}$ is symmetric: if $x\mathcal{R}y$ then $y\mathcal{R}x$ for all $x, y \in X$;
3. $\mathcal{R}$ is transitive: if $x\mathcal{R}y$ and $y\mathcal{R}z$, then $x\mathcal{R}z$ for all $x, y, z \in X$.

For equivalence relations, one usually writes $\sim$ instead of $\mathcal{R}$. The reader can check that $\mathcal{R}_=$ and $\mathcal{R}_n$ are examples of equivalence relations.

- The *equivalence class* of $x \in X$ is defined as
$$[x] := \{x' \in X \mid x \sim x'\}.$$
Note that by reflexivity, x belongs to its equivalence class. Furthermore, by symmetry and transitivity, any two equivalence classes are either equal or disjoint.

- The *quotient set* of X by an equivalence relation $\sim$ is defined as
$$X/\!\sim\, = \{[x] \mid x \in X\}.$$
There is also a projection that maps each element of X to its equivalence class:
$$\begin{aligned} \pi\colon X &\to X/\sim \\ x &\mapsto [x]. \end{aligned}$$

Next, we move on to the main definition of this subsection.

Definition 1.43. For a topological space $(X, \mathcal{T})$ and an equivalence relation $\sim$ on X, *the quotient topology* on $X/\!\sim$ is
$$\{U \subset X/\!\sim \,\mid \pi^{-1}(U) \in \mathcal{T}\}.$$
The quotient topology is indeed a topology, the finest such that the projection $\pi\colon X \to X/\!\sim$ is continuous; this is an exercise on the third problem set.

To understand a quotient space $X/\!\sim$, one often describes a more familiar space Y homeomorphic to it. Therefore, before giving examples of quotient spaces, we provide conditions for a map $X \to Y$ to produce a homeomorphism $X/\!\sim\, \to Y$.

Remark 1.44. In what follows, $\sim$ is an equivalence relation on a set X.

- If a map $f\colon X \to Y$ satisfies $f(x) = f(x')$ for all $x \sim x' \in X$, then there exists a unique map $g\colon X/\sim\, \to Y$ such that $f = g \circ \pi$ (indeed, define $g([x]) := f(x)$ for $[x] \in X/\!\sim$). In this setting, we say that f *descends to the quotient* or that g is *induced by* f. Note that
  - the induced map g has the same image as f;
  - the induced map g is injective if and only if $f(x) = f(x') \Rightarrow x \sim x'$ for every $x, x' \in X$

  Summarising, a map $f\colon X \to Y$ induces a bijection $g\colon X/\sim\, \to Y$ if f is surjective and if, for every $x, x' \in X$, we have " $x \sim x'$ if and only if $f(x) = f(x')$".

- Assume that X, Y are spaces and that $f\colon X \to Y$ is a map that descends to $g\colon X/\!\sim\, \to Y$. We verify that f is continuous if and only if g is continuous. This follows from the following observation: given an open set $U \subset Y$, the set $g^{-1}(U) \subset X/\!\sim$ is open if and only if $\pi^{-1}(g^{-1}(U)) \subset X$ is open if and only if $f^{-1}(U) \subset X$ is open.

- Assume that X, Y are spaces and that $f\colon X \to Y$ is a map that descends to $g\colon X/\!\sim\, \to Y$. If f is open[5], then g is open: for $U \subset X/\!\sim$ open the set
$$g(U) = \{g([x]) \mid [x] \in U\} = \{f(x) \mid \pi(x) \in U\} = f(\pi^{-1}(U))$$
is also open because π is continuous and f is open.

[5]Recall from the first problem set that a map $\varphi\colon X \to Y$ is *open* if the image of every open set is open and that, for φ bijective, φ^{-1} is continuous if and only if φ is open.

Example 1.45. Here are some examples of quotient spaces.

1. Consider $X = \mathbb{R}^2$ equipped with the standard topology and the equivalence relation $(x_1, x_2) \sim (y_1, y_2)$ if and only if $x_1^2 + x_2^2 = y_1^2 + y_2^2$. The equivalence classes are the circles centered at the origin, as well as the origin. We argue that $X/\sim$ is homeomorphic to $[0, \infty)$ with the standard topology. A quick verification using Remark 1.44 shows that the continuous map $f \colon \mathbb{R}^2 \to \mathbb{R}$ given by $f(x) := ||x|| = \sqrt{x_1^2 + x_2^2}$ descends to a continuous bijection $g \colon \mathbb{R}^2/\sim \to [0, \infty)$. It remains to show that g^{-1} is continuous, i.e. that g is open. Using Remark 1.44, it suffices to show that f is open. As we saw on the second problem set, this reduces to showing that the image of any open ball $B(x, r)$ is open in $[0, \infty)$. This can now be verified explicitly as

$$f(B(x,r)) = \begin{cases} (||x|| - r, ||x|| + r) & \text{if } ||x|| \geq r \\ [0, ||x|| + r) & \text{if } ||x|| < r \end{cases}$$

is open in $[0, \infty)$. This concludes the proof of the fact that $X/\sim$ is homeomorphic to $[0, \infty)$.

2. Every subset $A \subset X$ of a set X induces an equivalence relation on X via $x \sim_A y \Leftrightarrow x = y$ or $x, y \in A$. The quotient set is denoted X/A and the projection $\pi \colon X \to X/A$ identifies all the elements of A to a single point. Informally, "X/A crushes $A \subset X$ to a single point".

   Let us for instance consider the case $A = \{0, 1\} \subset [0, 1] = X$. We argue that X/A is homeomorphic to $S^1 \subset \mathbb{R}^2 \cong \mathbb{C}$ with the standard topology. A quick verification using Remark 1.44 shows that the continuous map $f \colon [0, 1] \to S^1, t \mapsto e^{2\pi i t}$ descends to a continuous bijection $g \colon [0, 1]/\{0, 1\} \to S^1$. It remains to show that g^{-1} is continuous, i.e. that g is open.[6] We therefore have to show that if $U \subset [0, 1]/\{0, 1\}$ is such that $\pi^{-1}(U) \subset [0, 1]$ is open, then $g(U) = f(\pi^{-1}(U)) \subset S^1$ is open, i.e. is a neighborhood of each of its points. We therefore fix $z \in f(\pi^{-1}(U)) \subset S^1$, so that $z = f(t) = e^{2\pi i t}$ for some $t \in \pi^{-1}(U) \subset [0, 1]$ and wish to show that $f(\pi^{-1}(U))$ is a neighborhood of z.

   - Assume that $t \neq 0, 1$. As $\pi^{-1}(U)$ and $(0, 1)$ are open in $[0, 1]$, so is $\pi^{-1}(U) \cap (0, 1)$. Therefore, there exists $\varepsilon > 0$ such that $(t - \varepsilon, t + \varepsilon) \subset \pi^{-1}(U) \cap (0, 1)$. Since $z = f(t) \in f((t - \varepsilon, t + \varepsilon)) \subset f(\pi^{-1}(U))$, it suffices to show that $f((t - \varepsilon, t + \varepsilon))$ is open in S^1, as this would show that $f(\pi^{-1}(U))$ is indeed a neighborhood of $z = f(t)$. This follows because we can write $f((t - \varepsilon, t + \varepsilon)) = S^1 \cap \ell^{-1}((t - \varepsilon, t + \varepsilon))$ where $\ell = \frac{1}{2\pi} \log \colon \mathbb{C} \setminus [0, \infty) \to (0, 1]$ is a continuous branch of the complex logarithm.
   - Assume that $t = 0$ or $t = 1$. As $\pi^{-1}(U)$ is open in $[0, 1]$ and necessarily contains $\{0, 1\}$, it must contain $[0, \varepsilon) \cup (1 - \varepsilon, 1]$ for some $\varepsilon > 0$. Adapting the argument above, one can verify that $f([0, \varepsilon) \cup (1 - \varepsilon, 1]) \subset S^1$ is open, which together with the inclusion $z = f(t) \in f([0, \varepsilon) \cup (1 - \varepsilon, 1]) \subset f(\pi^{-1}(U))$ implies that $f(\pi^{-1}(U))$ is indeed a neighborhood of $z = 1 \in S^1$.

3. Generalising the previous example, if $X = D^n$ is the unit disc and $A = \partial D^n$ is its boundary, then X/A is homeomorphic to the sphere S^n. We proved this for $n = 1$ above and we will check the general statement in Subsection 1.2.2 below.

4. Consider $X = [0, 1] \times [0, 1]$ equipped with the standard topology and the equivalence relation: $(x, y) \sim (x', y')$ if "$x = x'$ and $\{y, y'\} = \{0, 1\}$" or if "$\{x, x'\} = \{0, 1\}$ and $y = y'$, see Figure 1.1. We argue that $X/\sim$ is homeomorphic to the torus $S^1 \times S^1$. Use $\pi \colon [0, 1] \to [0, 1]/\{0, 1\}$ to denote the canonical projection. A verification using Remark 1.44 shows that the continuous map

$$\begin{aligned} f \colon [0, 1] \times [0, 1] &\to [0, 1]/\{0, 1\} \times [0, 1]/\{0, 1\} \\ (x, y) &\mapsto (\pi(x), \pi(y)) \end{aligned}$$

[6] This time, one cannot reduce this to proving that f is open because it is not: $f([0, \varepsilon)) \subset S^1$ is not open.

descends to a continuous bijection $g\colon X/{\sim} \to [0,1]/\{0,1\}\times[0,1]/\{0,1\}$. It turns out that g^{-1} is continuous, although this time we will not give the details (this will however follow from results in Section 1.2.2). Since we proved above that $[0,1]/\{0,1\} = S^1$, we therefore deduce that $X/{\sim} \cong [0,1]/\{0,1\} \times [0,1]/\{0,1\} \cong S^1 \times S^1$, as claimed.

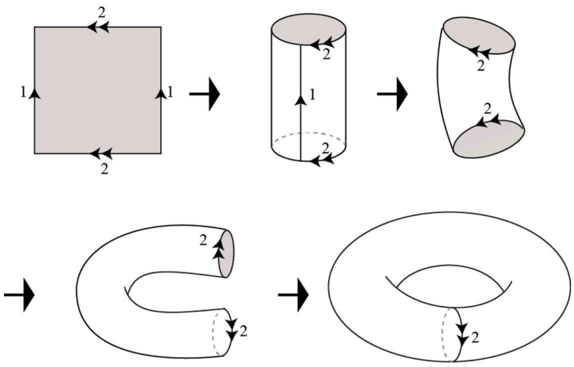


Figure 1.1: The upper left of this figure shows the equivalence relation on $X = [0,1] \times [0,1]$ described in the fourth item of Example 1.45. Ranging from the upper left to the lower right, one sees why the quotient space is a torus. This figure comes from "Survey of Graph Embeddings into Compact Surfaces" by Sophia N. Potoczak, https://digitalcommons.library.umaine.edu/etd/2155/

Active learning 1.46. In the second active learning session, we will discuss the following notions:

1. the Möbius band;
2. the Klein bottle;
3. the real projective plane $\mathbb{R}P^2$;
4. the (compact orientable) surface of genus g;
5. the definition of a (topological) manifold: a *a topological n-manifold* is a topological space M that satisfies the following three conditions:
   (a) M is *locally homeomorphic to* $\mathbb{R}^n$: for every $x \in M$ there is an open set U containing x and a homeomorphism $\psi\colon U \to \mathbb{R}^n$.
   (b) M is *Hausdorff*: for every $x \neq y \in X$, there exists disjoint open sets $U, V \subset X$ with $x \in U$ and $y \in V$.; we repeated this definition a few days later in Definition 1.52.
   (c) M is *second countable*: the topology on M has a countable basis.

   A 2-manifold is called a *surface*. We explained why the Klein bottle, the sphere, and the surface of genus g are all examples of surfaces. We also mentioned the notion of a manifold with boundary, an example of which is the Möbius band. Additionally, we saw that the *connected sum* $M_1 \# M_2$ of manifolds M_1 and M_2 is defined as

$$(M_1 \setminus D_1^n) \sqcup (M_2 \setminus D_2^n)/\sim,$$

   where the equivalence relation identifies x with $\varphi(x)$ for every $x \in S^{n-1}$, for some (fixed) homeomorphism $\varphi\colon S^{n-1} \to S^{n-1}$.[7] For example, we have $\Sigma_1 \# \Sigma_1 \cong \Sigma_2$ (and more generally $\Sigma_g \# \Sigma_h \cong \Sigma_{g+h}$).

[7]The fact that this definition does not depend on the choice of φ is a very hard theorem, far beyond the scope of this class. The interested reader can nonetheless look at [FNOP19, Theorem 5.11] to get a sense of the proof which relies on the so-called "Annulus Theorem".

1.1.6 Sequences and convergence

We conclude this first section by revisiting another concept from analysis: convergence. The main reference for this section is [Mun00, Sections 17 and 21].

Before proceeding however, we introduce a notion that could have been discussed in Subsection 1.1.4 when we were studying bases.

Definition 1.47. Let X be a space and let $x \in X$. A family $\mathcal{B}_x$ of neighborhoods of x is a *neighbhourhood basis* of x if for every neighborhood V of x, there exists $B \in \mathcal{B}_x$ with $B \subset V$.

Example 1.48. Here are examples of neighbhourhood bases:

1. If X is discrete and $x \in X$, then $\mathcal{B}_x = \{\{x\}\}$ is a neighborhood basis of $x \in X$.

2. If X is equipped with the trivial topology and $x \in X$, then $\mathcal{B}_x = \{X\}$ is a neighborhood basis of $x \in X$.

3. If (X, d) is a metric space and $x \in X$, then by definition of the topology $\mathcal{T}_d$, the set $\mathcal{B}_x = \{B_d(x, \varepsilon) \mid \varepsilon > 0\}$ of open balls centered at $x \in X$ is a neighborhood basis of $x \in X$.

A *sequence* in a set X is a map $\mathbb{N} \to X, n \mapsto x_n$. Sequences are often denoted $(x_n)_{n \in \mathbb{N}}$ or simply (x_n).

Definition 1.49. Let (x_n) be a sequence in a topological space X. An element $x \in X$ is called *a limit of* (x_n) *in* X if for every open set $U \subset X$ containing x, there exists a positive integer $N > 0$ such that $x_n \in U$ whenever $n \geq N$. In this case, we say that (x_n) *converges towards* x in X and write $x_n \to x$.

Remark 1.50. Before giving some examples, we note that that the use of neighborhood bases simplifies several verifications. In what follows X is a topological space.

1. Suppose $\mathcal{B}_x$ is a neighborhood basis of $x \in X$. We argue that a sequence (x_n) converges to $x \in X$ if and only if for every $B \in \mathcal{B}_x$, there exists an $N > 0$ such that $n \geq N$ implies $x_n \in B$.

   To prove the forward direction, we suppose (x_n) converges to $x \in X$. As $B \in \mathcal{B}_x$ is a neighborhood of x, there is an open set $U \subset X$ with $x \in U \subset B$. By definition of convergence, there exists $N > 0$ so that $n \geq N$ implies $x_n \in U \subset B$, which proves the forward implication. To prove the converse, start with an open set $U \subset X$ containing x. As U is open, it is a neighborhood of $x \in U$ and so there exists $B \in \mathcal{B}_x$ with $x \in B \subset U$. By assumption, there exists $N > 0$ such that $n \geq N$ implies $x_n \in B \subset U$, proving the other implication.

2. Assume that S is a subbasis for the topology on X. It can be proved that $x_n \to x \in X$ if and only if for every $U \in S$ containing x, there exists an $N > 0$ such that $n \geq N$ implies $x_n \in U$. This is an exercise on the fourth problem set.

Example 1.51. We study the notion of convergence in various topological spaces.

1. We show that if (X, d) is a metric space, then for $(X, \mathcal{T}_d)$, we recover the usual notion of convergence from analysis. In Example 1.48, we saw that the set $\mathcal{B}_x = \{B_d(x, \varepsilon) \mid \varepsilon > 0\}$ of open balls centered at x is a neighborhood basis of $x \in X$. By the first item of Remark 1.50, a sequence (x_n) converges to $x \in X$ if and only if for every $\varepsilon > 0$ there exists an $N > 0$ so that $n \geq N$ implies $x_n \in B_d(x, \varepsilon)$; i.e

$$\forall \varepsilon > 0, \ \exists N > 0 \text{ such that } n \geq N \Rightarrow d(x, x_n) < \varepsilon.$$

2. We argue that in a discrete space X, the only convergent sequences are the ones that become constant starting from a certain index; such sequences are called *stationary*. To see this, assume that a sequence (x_n) converges to $x \in X$, and apply the definition of convergence with the open set $U = \{x\}$: there exists an $N > 0$ such that $n \geq N$ implies $x_n \in U$, i.e. $x_n = x$.

3. Endow a set X with the cocountable topology[8]

$$\mathcal{T}_c = \{U \subset X \mid X \setminus U \text{ is countable or } U = \emptyset\}.$$

We argue that, as for the discrete topology, the convergent sequences are the stationary ones. To see this, assume that (x_n) converges to x and consider $U := (X \setminus (x_n)_{n\in\mathbb{N}}) \cup \{x\}$. By definition of $\mathcal{T}_c$, note that U is an open set that contains x. Thus, as x_n converges to x, there exists an $N > 0$ such that $n \geq N$ implies $x_n \in U$, i.e. $x_n = x$ for $n \geq N$.

4. If X is endowed with the trivial topology, then all sequences converge to all points in X: one can only apply the definition of convergence with $U = X$.

Despite being pathological, this last example reveals that sequences may converge towards several distinct limits. We now study a class of spaces in which a limit, if it exists, is unique.

Definition 1.52. A space X is *Hausdorff* if for every $x \neq y \in X$, there exists disjoint open sets $U, V \subset X$ with $x \in U$ and $y \in V$.

Note that being Hausdorff is a topological property: for homeomorphic spaces $X \cong Y$, one sees that X is Hausdorff if and only if Y is Hausdorff. The next result shows that in such a space, limits are unique.

Proposition 1.53. *Let (x_n) be a sequence in a Hausdorff space X. If x_n converges to $x \in X$ and (x_n) converges to $y \in X$, then $x = y$.*

Proof. We assume that (x_n) converges to x and show that if $y \neq x$, then (x_n) does not converge to y. As $x \neq y$ and X is Hausdorff, there are disjoint open sets $U \subset X$ and $V \subset X$ with $x \in U$ and $y \in V$. As x_n converges to $x \in X$ and U is an open set containing x, there exists an $N > 0$ such that $n \geq N$ implies $x_n \in U$. As U and V are disjoint, this means that for every $n \geq N$, we have $x_n \notin V$. As $V \subset X$ is an open set containing y, the sequence (x_n) therefore does not converge to $y \in X$. □

Example 1.54. Here are some examples of Hausdorff spaces.

1. Every discrete space is Hausdorff: take $U = \{x\}$ and $V = \{y\}$ in Definition 1.52.

2. If a set with at least two elements is endowed with the trivial topology, then it is not Hausdorff.

3. If (X, d) is a metric space, then $(X, \mathcal{T}_d)$ is Hausdorff. Indeed, if $x \neq y \in X$ are distinct, then $r := \frac{d(x,y)}{2}$ is positive, we set $U := B(x, r)$ and $V := B(y, r)$ and use the triangle inequality to verify that these two open sets are disjoint (if $z \in U$, then $d(x, y) \leq d(x, z) + d(z, y)$ gives $r < d(x, y) - d(x, z) \leq d(z, y)$ and therefore $z \notin V$).

Remark 1.55. We record how being Hausdorff interacts with various operations on spaces.

1. If X is Hausdorff and $Y \subset X$ is a subspace, then Y is Hausdorff. Indeed, if $x \neq y \in Y$, then since X is Hausdorff, there are disjoint open sets $U, V \subset X$ respectively containing x, y and the conclusion follows by noting that $x \in U \cap Y$ and $y \in V \cap Y$, where $U \cap Y$ and $V \cap Y$ are disjoint open sets in Y.

[8] Recall that a set X is *countable* if there is an injective map $X \hookrightarrow \mathbb{N}$.

2. If $\{X_i\}_{i\in I}$ is a family of Hausdorff spaces, then $\prod_{i\in I} X_i$ is Hausdorff, both in the box topology and product topology. On the other hand, if X is Hausdorff and $\sim$ is an equivalence relation on X, then $X/\sim$ is in general not Hausdorff. This is an exercise on the fourth problem set.

We saw in Proposition 1.53 that in a Hausdorff space, limits are unique. It is natural to ask if the converse is true, i.e. if being Hausdorff is the optimal condition. As we will see in the next Active learning session, the answer to this question is negative in general.[9] On the other, the converse is true provided more is assumed of X.

Definition 1.56. A space X is *first countable* if every $x \in X$ admits a countable basis $\mathcal{B}_x$ of neighborhoods.

We will learn in Active learning 1.57 that metric spaces are first countable but we now also present (without giving all the details) an amusing example of a first countable space that is not Hausdorff, namely the *line with two origins*. This space, which we denote X, is obtained as the quotient the subspace $\mathbb{R} \times \{\pm 1\}$ of $(\mathbb{R}^2, \mathcal{T}_{\text{std}})$ by the equivalence relation $\left(\begin{smallmatrix} x \\ -1 \end{smallmatrix}\right) \sim \left(\begin{smallmatrix} x \\ 1 \end{smallmatrix}\right)$ for $x \neq 0$. The "two origins" are the classes $[\left(\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right)]$ and $[\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)]$. We outline why X is first countable but not Hausdorff without delving into the full details. Firstly, using $\pi \colon \mathbb{R} \times \{\pm 1\} \to X$ to denote the projection, we see that a countable neighbhoorhood basis of $x \in X$ is obtained by considering $\{\pi((x - \frac{1}{n}, x + \frac{1}{n}) \times \{1\})\}_{n \in \mathbb{N}_{>0}}$ for $x \neq \pi(\left(\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right))$ and $\pi((-\frac{1}{n}, \frac{1}{n}) \setminus \{0\} \times \{-1\} \cup \{\left(\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right)\})_{n \in \mathbb{N}_{>0}}$ for $x = \pi(\left(\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right))$. Secondly, one notes that in X, the sequence $\left(\pi\left(\begin{smallmatrix} 1/k \\ 1 \end{smallmatrix}\right)\right)_k$ converges to both $\pi\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ and $\pi\left(\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right)$ because $\pi\left(\begin{smallmatrix} 1/k \\ 1 \end{smallmatrix}\right) = \pi\begin{pmatrix} 1/k \\ -1 \end{pmatrix}$. In particular Proposition 1.53 implies that X is not Hausdorff.

Here is a summary of what was discussed in the third active learning session.

Active learning 1.57. As mentioned above, it turns out that if X is first countable, then X being Hausdorff is equivalent to limits being unique provided they exist [Mun00, Sections 21 and 30]. In fact, this condition is also useful to study the following concepts:

- A subspace $A \subset X$ is *sequentially closed* if for any a sequence (a_n) in A with $a_n \to x$ for some $x \in X$, one has $x \in A$.
- A map $f \colon X \to Y$ between topological spaces is *sequentially continuous at* $x \in X$ if for every sequence (x_n) in X with $x_n \to x$, one has $f(x_n) \to f(x) \in Y$. The map f is *sequentially continuous* if it is sequentially continuous at every $x \in X$.

These concepts are related to the familiar notions of closedness and continuity as follows.

1. Closed subsets are sequentially closed.
2. Continuous functions are sequentially continuous.
3. Metric spaces are first countable.
4. For first countable spaces, the converses of the first and second items hold.

   Theorem 1.58. *Let X be a first countable topological space.*

   (a) X is Hausdorff if and only if convergent sequences admit a unique limit;
   (b) $A \subset X$ is closed if and only if A is sequentially closed;
   (c) $f \colon X \to Y$ is continuous at $x \in X$ if and only if f is sequentially continuous at x.

   The proof of the first item of this theorem makes use of the following lemma.

   Lemma 1.59. *If X is first countable, then every $x \in X$ admits a neighborhood basis $\mathcal{B}_x = (B_n)_n$ where each B_n is open, the B_n are decreasing (i.e. $B_n \supset B_{n+1}$ for every n) and for every sequence (x_n) with $x_n \in B_n$ for every n, the sequence x_n converges to x.*

[9] For example $\mathbb{R}$ with the cocountable topology is not Hausdorff (this holds more generally for uncountably infinite sets with the cocountable topology) but has the unique limit property because convergent sequences in $(\mathbb{R}, \mathcal{T}_c)$ are stationary; recall Example 1.51.

1.2 Connectedness and compactness

We describe two important properties that a topological space can possess: connectedness and compactness. Informally, a space is connected if it is made of only one piece, while a space is compact if it is "not too big". These notions will allow us to tell certain spaces apart and also lead to generalisations of some well known results from analysis such as the intermediate value theorem.

1.2.1 Connectedness

After discussing the definition of connectedness, we study how it behaves under various operations on spaces and list several examples of connected spaces. Applications include the intermediate value theorem as well as proving that $\mathbb{R}$ and $\mathbb{R}^n$ are not homeomorphic for $n > 1$. The main reference is [Mun00, Section 23-25].

Connected spaces

We define the notion of connectedness and study its first properties.

Definition 1.60. A topological space X is *connected* if for every disjoint open sets $U, V \subset X$ with $X = U \cup V$, one has $U = \emptyset$ or $V = \emptyset$.

Intuitively, a space is connected if it is only made of one piece.

Remark 1.61. Here are some remarks about connectedness:

1. For homeomorphic spaces X and Y, note that X is connected if and only if Y is connected: this follows because the definition of connectedness only involves open sets.

2. Here is an equivalent characterisation of connectedness: a space X is connected if and only if the only subsets of X that are both open and closed are $\emptyset$ and X (such subsets are called *clopen*). To see this, note that if X is not connected, then $X = U \cup V$, with $U, V \subset X$ non-empty disjoint open sets, which are therefore both clopen. Conversely, if $U \subset X$ is clopen (with $U \neq \emptyset, X$), then writing $X = U \sqcup (X \setminus U)$ shows that X is not connected.

3. When we say that a subspace $Y \subset X$ is connected, it is always with respect to the induced topology. One can check that $Y \subset X$ is connected if and only if $Y \cap U \cap V = \emptyset$ and $Y \subset U \cup V$ implies $Y \subset U$ or $Y \subset V$ for any open sets $U, V \subset X$.

Example 1.62. Here are some examples of connected and disconnected spaces:

1. Every set endowed with the trivial topology is connected.

2. A set X endowed with the discrete topology is not connected provided $\#X > 1$. Indeed, if $\#X > 1$, then the open sets $U = \{x\}$ and $V = X \setminus \{x\}$ form a non-trivial decomposition of X.

3. The subspace $Y = [-1, 0) \cup (0, 1]$ of $X = \mathbb{R}$ is not connected. The open sets $U = [-1, 0)$ and $V = (0, 1]$ of Y form a non-trivial decomposition of Y.

Proving that a space is not connected involves decomposing it non-trivially into a union of two open sets; this is often not overly difficult. On the other hand, it is frequently harder to prove that a space is connected. For this reason, we start with some general results about connectedness, before focusing on the case $\mathbb{R}$.

First, we give an (intuitive) criterion for a union of connected subspaces to be connected.

Proposition 1.63. *If $\{A_i\}_{i \in I}$ is a family of connected subspaces of a space X with $\cap_{i \in I} A_i$ non-empty, then the subspace $\cup_{i \in I} A_i$ is connected.*

Proof. To show that $Y = \bigcup_{i \in I} A_i$ is connected, consider two open sets $U, V \subset X$ with $Y \cap U \cap V = \emptyset$ and $Y \subset U \cup V$; the goal is to show that $Y \subset U$ or $Y \subset V$; recall the third item of Remark 1.61.

Since the intersection is non-empty, we can fix $x \in X$ such that $x \in A_i$ for all $i \in I$. As $x \in \bigcap_{i \in I} A_i \subset A_i \subset Y \subset U \cup V$, we have $x \in U$ or $x \in V$. Assume without loss of generality that $x \in U$ for every $i \in I$. Next, for every $i \in I$, we have $A_i \cap U \cap V \subset Y \cap U \cap V = \emptyset$ and $A_i \subset Y \subset U \cup V$, the connectedness of A_i implies that $A_i \subset U$ or $A_i \subset V$. But since $x \in A_i \cap U$ and $A_i \cap U \cap V = \emptyset$, we must have the inclusion $A_i \subset U$ for every $i \in I$ and therefore the inclusion $Y = \bigcup_{i \in I} A_i \subset U$. □

The next result implies (in particular) that the closure of a connected subspace is connected.

Proposition 1.64. *If $A \subset X$ is a connected subspace of a space X, then every subspace $B \subset X$ with $A \subset B \subset \overline{A}$ is also connected.*

Proof. Let $U, V \subset X$ be two open subsets with $B \cap U \cap V = \emptyset$ and $B \subset U \cup V$; we must show that B is included in one of these open sets. The inclusion $A \subset B$ implies that $A \cap U \cap V \subset B \cap U \cap V = \emptyset$ and $A \subset B \subset U \cup V$. Since A is connected, we have $A \subset U$ or $A \subset V$. Without loss of generality, let us suppose that $A \subset U$ and therefore $A \cap V = A \cap U \cap V = \emptyset$. We now check the inclusion $B \subset U$. Given $x \in B \subset U \cup V$, it suffices to show that x does not belong to V: this would imply that $x \in U$ because $B \subset U \cup V$. Assume for a contradiction that $x \in V$, so that x does not belong to the closed set $X \setminus V$ which contains A (because A and V are disjoint). We deduce that

$$x \notin \bigcap_{C \text{ closed }, C \supset A} C = \overline{A}.$$

This contradicts our assumption that $B \subset \overline{A}$. We therefore have shown that $x \in U$ which implies that $B \subset U$ and concludes the proof of the proposition. □

The next result shows that connectedness is preserved under continuous maps.

Proposition 1.65. *If $f : X \to Y$ is continuous map with X connected, then $f(X)$ is connected.*

Proof. We prove the contrapositive. Assume that $f(X)$ is not connected, so that there are two disjoint open non-empty subsets $U, V \subset f(X)$ with $U \cup V = f(X)$. We will prove that X is not connected. By Proposition 1.16, the map $g : X \to f(X) =: Z$ obtained from f by restricting its target is also continuous. To prove that X is not connected, we will show that $X = U' \cup V'$, where $U' = g^{-1}(U)$ and $V' := g^{-1}(V)$, and check that U', V' are non-empty disjoint open subsets of X.

First, as g is surjective and U, V are non-empty, so are U', V'. Next, since g is continous and U, V are open so are U' and V'. Using the properties of inverse images, we then check that

$$U' \cap V' = g^{-1}(U) \cap g^{-1}(V) = g^{-1}(U \cap V) = g^{-1}(\emptyset) = \emptyset,$$
$$U' \cup V' = g^{-1}(U) \cup g^{-1}(V) = g^{-1}(U \cup V) = g^{-1}(Z) = X.$$

Therefore U', V' are disjoint open sets whose union is X, contradicting the fact that X is connected. □

Next, we study how connectedness behaves under products.

Proposition 1.66. *Let $X_1, \ldots, X_n$ be topological spaces. The product $X_1 \times \ldots \times X_n$ is connected if and only if each X_i is connected.*

Proof. We first suppose that the product is connected. Since, each $\pi_i : X_1 \times \ldots \times X_n \to X_i$ is continuous, Proposition 1.65 implies that its image, namely X_i, is connected.

To prove the converse, we suppose that each X_i is connected and wish to show that the product is also connected. We proceed by induction on $n \geq 1$. Using the homeomorphism

$$X_1 \times \ldots \times X_{n-1} \times X_n \cong (X_1 \times \ldots \times X_{n-1}) \times X_n$$

from the third problem set, it suffices to prove the following assertion: if X and Y are connected, then so is $X \times Y$. So assume that X, Y are connected, fix $b \in Y$ and consider the decomposition

$$X \times Y = \bigcup_{x \in X} (X \times \{b\}) \cup (\{x\} \times Y) =: \bigcup_{x \in X} T_x.$$

By Proposition 1.63, if we are able to show that each T_x is connected and that $\bigcap_{x\in X} T_x \neq \emptyset$, then it will follow that $X \times Y$ is connected. The intersection $\bigcap_{x\in X} T_x$ contains $X \times \{b\}$ and is therefore non-empty. To see that T_x is connected, we apply Proposition 1.63 by noting that the subspace $X \times \{b\} \subset X \times Y$ is connected (since it is homeomorphic to the connected space X[10]), that $\{x\} \times Y$ is connected (for the analogous reason) and that $X \times \{b\} \cap \{x\} \times Y = \{(x, b)\}$ is non-empty. We have therefore shown that $X \times Y = \bigcup_{x\in X} T_x$ is connected. □

Remark 1.67. The same result holds for arbitrary products provided one uses the product topology. On the other hand, it turns out that a product of connected spaces need not be connected for the box topology.

The fact that the quotient space of a connected set is connected is an exercise on the fifth problem set, but we record the result here for later reference.

Proposition 1.68. *If X is a connected topological space and $\sim$ is an equivalence relation on X, then the quotient space $X/\sim$ is connected.*

Connectedness in $\mathbb{R}$

The promised generalisation of the intermediate value theorem could already be proved at this point. However, in order to understand why it actually generalises the classical result in $\mathbb{R}$, we must prove the intuitive result that (a, b) is connected.

Theorem 1.69. *Every open interval $(a, b) \subset \mathbb{R}$ is connected.*

Proof. Assume for a contradiction that (a, b) is not connected. By definition, this means that there are non-empty, disjoint open set $U, V \subset (a, b)$ with $(a, b) = U \cup V$. As U, V are non-empty, we can chose $u \in U$ and $v \in V$. As U and V are disjoint, u and v are distinct and without loss of generality, we assume $u < v$. Now consider the subset $S \subset (a, b)$ given by

$$S = \{s \in (a, b) \mid [u, s] \subset U\}.$$

Note that S is bounded (as it is included in (a, b)) and non-empty (as it contains u). As a consequence, S admits a supremum $s_0 = \sup S$. [11] The result will now follow from the two following steps:

1. prove that $s_0 \in (a, b)$;
2. prove that $s_0 \notin U$ and $s_0 \notin V$.

Using these two facts, we get $s_0 \notin U \cup V = (a, b) \ni s_0$, leading to the required contradiction. Before proving each of these steps, we record a claim for later use.

Claim. *Every $x \in V$ with $u < x$ is an upper bound for S.*

Proof. Assume for a contradiction that this is not the case: there exists $s \in S$ with $x < s$. We then have $x \in (u, s) \subset [u, s] \subset U$, and so $x \in U \cap V$, which is impossible as $U \cap V = \emptyset$. □

[10] This verification is left to the reader

[11] Given a set $A \subset \mathbb{R}$, we say that $x \in \mathbb{R}$ is an *upper bound* for A if $x \geq a$ for every $a \in A$. The *supremum* of A (if it exists) is its smallest upper bound.

We now carry out the first step of our strategy: we show that $s_0 \in (a, b)$. We have the inequalities $a < u$ (as $u \in (a, b)$), $u \leq s_0$ (because s_0 is an upper bound of $S \ni u$) and $s_0 \leq v$ (because s_0 is the smallest upper bound of S and v is such an upper bound, by the Claim) and $v < b$ (because $v \in (a, b)$). Summarising, as required, we have proved

$$a < u \leq s_0 \leq v < b.$$

Next, we carry out the second step in our strategy: we prove that $s_0 \notin U$ and $s_0 \notin V$.

Assume for a contradiction that $s_0 \in U$. As U is open, there exists $\varepsilon > 0$ such that $(s_0 - \varepsilon, s_0 + \varepsilon) \subset U$. Note that $[u, s_0) \subset U$ (otherwise, we would have an $x \in (u, s_0) \cap V$ which, by the Claim, would be an upper bound of S that is smaller than s_0). Thus we have the inclusion $[u, s_0 + \varepsilon/2] \subset U$ which implies that $s_0 + \varepsilon/2$ belongs to S. This is impossible because s_0 is an upper bound of S. This proves that $s_0 \notin U$.

Assume for a contradiction that $s_0 \in V$. As V is open, there exists $\varepsilon > 0$ such that $(s_0 - \varepsilon, s_0 + \varepsilon) \subset V$ and $u < s_0 - \varepsilon$. By the Claim, $s_0 - \varepsilon/2$ is an upper bound of S, which is impoosible because s_0 is the smallest such upper bound. This proves that $s_0 \notin V$.

Our two steps are proved and, as explained above, we now have $s_0 \in (a, b) = U \cup V$, but $s_0 \notin U$ and $s_0 \notin V$. This is a contradiction and we conclude that (a, b) is connected, as required. □

Corollary 1.70. *The Euclidean line* $\mathbb{R}$ *is connected, as are the intervals* $(a, b), [a, b), (a, b], [a, b], (-\infty, a), (-\infty, a](a, \infty)$ *and* $[a, \infty)$.

Proof. As we saw in an exercise on the first problem set, the open intervals of $\mathbb{R}$ are homeomorphic. As (a, b) is connected, by Theorem 1.69, the same goes for the other intervals such as $(-\infty, a), (a, \infty)$ and $\mathbb{R}$. For the remaining intervals, we use Proposition 1.64: for instance $B = [a, b)$ is connected because $A \subset B \subset \overline{A}$ with $A = (a, b)$ connected, and similarly for the other cases. □

Finally, we prove the generalised intermediate value theorem.

Theorem 1.71. *Let* $f \colon X \to \mathbb{R}$ *be a continuous map with* X *connected, and let* $a, b \in X$. *For every* $r \in \mathbb{R}$ *with* $f(a) < r < f(b)$, *there exists* $c \in X$ *with* $f(c) = r$.

Proof. Assume for a contradiction that there is no $c \in X$ with $f(c) = r$. In other words, we suppose that $r \notin f(X)$. This implies that

$$f(X) = (f(X) \cap (-\infty, r)) \cup (f(X) \cap (r, \infty)) =: U \cup V,$$

where U and V are disjoint (because $r \notin f(X)$), non-empty (since $f(a) \in U$ and $f(b) \in V$) and open in $f(X)$ (by definition of the subspace topology). It follows that $f(X)$ is not connected, and by Proposition 1.65, this implies that X is not connected. This is the required contradiction. □

Path-connectedness

We define another notion of connectedness, known as path-connectedness. Checking for path-connectedness is easier than checking for connectedness and since path-connectedness implies connectedness, we will be able to deduce that several familiar spaces are indeed connected.

Definition 1.72. Let X be a space and let $x, y \in X$. A *path* in X from x to y is a continuous map $\gamma \colon [0, 1] \to X$ such that $\gamma(0) = x$ and $\gamma(1) = y$. A space X is *path-connected* if for every $x, y \in X$, there exists a path in X from x to y.[12]

Remark 1.73. For homeomorphic spaces X and Y, the reader can check that X is path-connected if and only if Y is path-connected.

[12] Paths could be just as well defined as continuous maps $[a, b] \to X$: two points are joined by a path $[0, 1] \to X$ if and only if they are joined by a path $[a, b] \to X$ for any $a < b$.

The next proposition relates path-connectedness to connectedness.

Proposition 1.74. *If a space X is path-connected, then it is connected.*

Proof. Assume for a contradiction that X is not connected: there are two non-empty disjoint open sets $U, V \subset X$ with $X = U \cup V$. Since U and V are non empty, we can choose $x \in U$ and $y \in V$ and since X is path-connected, there exists a path $\gamma\colon [0,1] \to X$ from x to y. Note that

$$[0,1] = \gamma^{-1}(X) = \gamma^{-1}(U \cup V) = \gamma^{-1}(U) \cup \gamma^{-1}(V) =: U' \cup V',$$

where it can be checked that $U', V' \subset [0,1]$ are two disjoint non-empty open subsets of $[0,1]$. This contradicts Corollary 1.70, according to which $[0,1]$ is connected. □

Active learning 1.75. The topics that will be covered during the fourth active learning session include:

- If a subspace $X \subset \mathbb{R}^n$ is convex[13], then it is path-connected, and thus connected.
- For every $n > 1$ and every $x \in \mathbb{R}^n$, the space $\mathbb{R}^n \setminus \{x\}$ is path-connected, and thus connected.
- For every $n > 0$, the sphere $S^n = \{x \in \mathbb{R}^{n+1} | ||x|| = 1\}$ is path-connected, and thus connected.
- For $n > 1$, the spaces $\mathbb{R}$ and $\mathbb{R}^n$ are not homeomorphic.
- The converse of Proposition 1.74 is false, but there is a partial converse. Namely, we learnt that a space X is *locally path-connected* if for every $x \in X$ and every neighborhood $U \ni x$, there exists a path-connected neighorhood V with $x \in V \subset U$, and we then proved the tollowing result:

  Proposition 1.76. *If X is connected and locally path-connected, then it is path-connected.*

  We also mentioned that the *topologist's sine curve* $X = \{(0,0)\} \cup \{(t, \sin(\frac{1}{t}) \mid t \in (0,1]\} \subset \mathbb{R}^2$ is an example of a connected space that is not path-connected.

1.2.2 Compactness

After discussing the definition of compactness, we study how it behaves under various operations on spaces. We then focus on the case of intervals in $\mathbb{R}$ before moving to $\mathbb{R}^n$ where we recover the familiar characterisation of compact subspaces as those that are closed and bounded. The main reference is [Mun00, Section 26-28].

1.2.2.1 Compact spaces

To define compactness, some terminology is needed. A family $\mathfrak{U} = \{U_i\}_{i \in I}$ of open sets of a space X is called an *open cover* if $X = \bigcup_{i \in I} U_i$. A *finite open subcover* of $\mathfrak{U}$ is a finite subset of $\mathfrak{U}$ that forms an open cover of X.

Definition 1.77. A space X is *compact* if every open cover of X admits a finite open subcover.

Remark 1.78. Here are some remarks about compactness.

1. Compactness is a topological notion: for homeomorphic spaces X and Y, we have that X is compact if and only if Y is compact.

[13] A subset $X \subset \mathbb{R}^n$ is *convex* if for every $x, y \in X$, the segment $[x, y] = \{tx + (1-t)y \mid t \in [0,1]\}$ is included in X.

2. For future proofs, it is helpful to spell out the definition: a space X is compact if for every family $\mathcal{U} = \{U_i\}_{i\in I}$ with $U_i \subset X$ open and $X = \bigcup_{i\in I} U_i$, there are indices $i_1, \ldots, i_m \in I$ such that $X = U_{i_1} \cup \ldots \cup U_{i_m}$.

3. A subspace $Y \subset X$ is compact if and only if for every family $\mathcal{U} = \{U_i\}_{i\in I}$ with $U_i \subset X$ open for every $i \in I$, there are indices $i_1, \ldots, i_m \in I$ such that $Y \subset U_{i_1} \cup \ldots \cup U_{i_m}$. This can be checked using the definition of compactness and the definition of the subspace topology.

Example 1.79. Here are some examples of compact spaces:

1. If a space has only finitely many elements in its topology, then it is compact. In particular, any set endowed with the trivial topology is compact.

2. A set X endowed with the discrete topology is compact if and only if it is finite. Indeed, if X is finite, then $(X, \mathcal{T}_{\text{disc}})$ is compact by the first example, while for the converse, if X is not finite, then $\mathcal{U} = \{\{x\}\}_{x\in X}$ does not admit a finite open subcover and so X is not compact.

3. The Euclidean line $\mathbb{R}$ with the standard topology is not compact: indeed the open cover $\{(n-1, n+1)\}_{n\in\mathbb{Z}}$ does not admit a finite open subcover. In particular, using the first item of Remark 1.78, we deduce that $(a, b), (a, \infty)$ and $(-\infty, a)$ are also not compact.

4. The subspace $(0, 1] \subset \mathbb{R}$ is not compact: the open cover $\{(\frac{1}{n}, 1]\}_{n\in\mathbb{N}}$ does not admit a finite subcover. In particular, using the first item of Remark 1.78, we deduce that $(a, b], [a, b), (-\infty, a]$ and $[a, \infty)$ are also not compact.

5. The closed and bounded interval $[a, b] \subset \mathbb{R}$ is compact as we shall see in Theorem 1.88.

6. It then follows from the sixth problem set that $S^n, \mathbb{R}P^2$ as well as the torus $S^1 \times S^1$ and the Möbius band $\mathcal{M}$ are all compact.

7. The subspace $X = \{0\} \cup \{1/n \mid n \in \mathbb{N}\}$ of $\mathbb{R}$ is compact. Let $\{U_i\}_{i\in I}$ be an open cover of $X \subset \mathbb{R}$: each $U_i \subset \mathbb{R}$ is open and $X \subset \bigcup_{i\in I} U_i$. Since $0 \in X$, there exists an index $i_1 \in I$ such that $0 \in U_{i_1}$. But now, by definition of X, the set $X \setminus (X \cap U_{i_1})$ is finite. Therefore, there are indices $i_2, \ldots, i_m \in I$ such that $X \setminus (X \cap U_{i_1}) \subset U_{i_2} \cup \ldots \cup U_{i_m}$ i.e. $X \subset U_{i_1} \cup U_{i_2} \cup \ldots \cup U_{i_m}$. We have therefore found the required finite open subcover.

The next two result give a criterion for a subspace of a compact space to be compact.

Proposition 1.80. *If X is compact and $Y \subset X$ is closed, then Y is compact.*

Proof. Let $\mathcal{U} = \{U_i\}_{i\in I}$ be an open cover of $Y \subset X$, i.e. a family of open sets $U_i \subset X$ with $Y \subset \bigcup_{i\in I} U_i$ (recall Remark 1.78). As $Y \subset X$ is closed, $X \setminus Y$ is open and so $(X \setminus Y) \cup \{U_i\}_{i\in I}$ is an open cover of X. As X is compact, there are indices $i_1, \ldots, i_m \in I$ such that $X = X \setminus Y \cup (U_{i_1} \cup \ldots \cup U_{i_m})$. This means that Y is contained in $U_{i_1} \cup \ldots \cup U_{i_m}$. We have therefore found our finite open subcover. □

Proposition 1.81. *If X is Hausdorff and $Y \subset X$ is a compact subspace, then Y is closed in X.*

Proof. We will show that $X \setminus Y$ is open, i.e. that $X \setminus Y$ is a neighborhood of each of its points. We therefore fix $x \in X \setminus Y$ and aim to find an open set $V \subset X$ with $x \in V \subset X \setminus Y$.

For every $y \in Y$, we have $y \neq x$ and therefore, since X is Hausdorff, there are disjoint open sets $V_y \ni x$ and $U_y \ni y$. Since the $\{U_y\}_{y\in Y}$ form an open cover of the compact set Y, there are elements $y_1, \ldots, y_m \in Y$ such that $Y \subset U_{y_1} \cup \ldots \cup U_{y_m}$. We check that $V := V_{y_1} \cap \ldots \cap V_{y_m}$ satisfies the required properties.

Firstly, V is open in X since it is an intersection of finitely many open sets of X. Next, V contains x, since each of the V_{y_i} contains x. Finally, we have

$$V \cap Y = \bigcap_{i=1}^{m} V_{y_i} \cap Y \subset \bigcap_{i=1}^{m} V_{y_i} \cap \bigcup_{j=1}^{m} U_{y_j} \subset \bigcup_{j=1}^{m} (V_{y_j} \cap U_{y_j}) = \emptyset,$$

whence the inclusion $V \subset X \setminus Y$. Thus $X \setminus Y$ is a neighorhood of every $x \in X \setminus Y$, which proves that $X \setminus Y$ is open i.e. that Y is closed, as required. $\square$

The next result shows that compactness is preserved under continuous maps.

Theorem 1.82. *If $f\colon X \to Y$ is continuous with X compact, then $f(X)$ is compact.*

Proof. Let $\mathcal{U} = \{U_i\}_{i\in I}$ be an open cover of $f(X)$, i.e. a familly of open sets $U_i \subset f(X)$ such that $f(X) \subset \bigcup_{i\in I} U_i$. As f is continuous, $f^{-1}(U_i) \subset X$ is open for every $i \in I$ and we have

$$X = f^{-1}(f(X)) = f^{-1}\left(\bigcup_{i\in I} U_i\right) = \bigcup_{i\in I} f^{-1}(U_i).$$

Thus $\{f^{-1}(U_i)\}_{i\in I}$ is an open cover of the compact space X and so there are indices $i_1, \ldots, i_m \in I$ so that $X = f^{-1}(U_{i_1}) \cup \ldots \cup f^{-1}(U_{i_m}) = f^{-1}(U_{i_1} \cup \ldots \cup U_{i_m})$. Applying f to both sides, we conclude that $f(X) \subset U_{i_1} \cup \ldots \cup U_{i_m}$ which gives the required finite subcover. $\square$

We deduce a useful criterion for a continuous bijective map to be a homeomorphism.

Corollary 1.83. *A continuous bijective map $f\colon X \to Y$ between a compact space X and a Hausdorff space Y is a homeomorphism.*

Proof. We must show that $f^{-1}\colon Y \to X$ is continuous. By Proposition 1.10, it suffices to prove that $f(C) \subset Y$ is closed for every closed subset $C \subset X$. As C is closed in a compact space X, Proposition 1.80 implies that C is compact. As f is continuous, $f(C)$ is compact by Proposition 1.82. As $f(C)$ is compact in the Hausdorff space Y, it is closed by Proposition 1.81. $\square$

Remark 1.84. We make some remarks about Corollary 1.83.

1. In Subsection 1.1.2, we saw examples of continuous bijective functions that are not homeomorphisms; let us revisit these examples in light of Corollary 1.83.
    - For $\mathrm{id}_X\colon (X, \mathcal{T}_{\text{disc}}) \to (X, \mathcal{T}_{\text{triv}})$, while $(X, \mathcal{T}_{\text{disc}})$ is compact for finite X, the target $(X, \mathcal{T}_{\text{triv}})$ is not Hausdorff when X has more than one element.
    - For $\exp\colon [0, 1) \to S^1$, while S^1 is Hausdorff, $[0, 1)$ is not compact.
2. In Subsection 1.1.5, we had to do some work to prove that the exponential map induces a homeomorphism $\exp\colon [0, 1]/\{0, 1\} \to S^1$. Thanks to Corollary 1.83, this is now automatic since S^1 is Hausdorff and $[0, 1]/\{0, 1\}$ is compact (because $[0, 1]$ is compact and the quotient of a compact space remains compact, see the fifth problem set).

Next, we study how compactness behaves under products.

Theorem 1.85. *A product $X_1 \times \cdots \times X_n$ is compact if and only if each the X_i is compact.*

Proof. First, we assume that the product is compact. Since each projection $\pi_i\colon X_1 \times \ldots \times X_n \to X_i$ is continuous, Theorem 1.82 ensures that $X_i = \mathrm{im}(\pi_i)$ is compact.

To prove the converse, we assume that each X_i is compact and show that the product is compact. As in the proof of Proposition 1.66, an induction argument on $n \geq 1$ reduces the problem to proving the following assertion: if X and Y are compact, then so is $X \times Y$. To prove this, we need an intermediary lemma.

Lemma 1.86 (The Tube Lemma)**.** *Let X be a space and let Y be a compact space. If $N \subset X \times Y$ is an open set containing $\{x_0\} \times Y$ for some $x_0 \in X$, then there exists an open set $U \subset X$ containing x_0 with $U \times Y \subset N$.*

Proof. As $N \subset X \times Y$ is open and contains $\{x_0\} \times Y$, by definition of the box topology,[14] there are open sets $U_y \subset X$ and $V_y \subset Y$ with $(x_0, y) \in U_y \times V_y \subset N$. The family $\{V_y\}_{y \in Y}$ forms an open cover of the compact space Y. Extracting a finite subcover, we deduce that there are $y_1, \ldots, y_m \in Y$ with $Y = \bigcup_{i=1}^m V_{y_i}$. We set $U := U_{y_1} \cap \ldots \cap U_{y_m}$ and verify that it satisfies the required properties.

First, U is open in X since it is the intersection of a finite number of open sets in X. Next, U contains x_0 because each U_y contains x_0. Finally, we have

$$U \times Y = \bigcap_{i=1}^{m} U_{y_i} \times \bigcup_{j=1}^{m} V_{y_j} \subset \bigcup_{j=1}^{m} (U_{y_j} \times V_{y_j}) \subset N.$$

We have therefore obtained the required U and this concludes the proof of the Tube Lemma. □

We now return to the proof of Theorem 1.85: we have compact spaces X, Y and wish to prove that $X \times Y$ is compact. Let $\mathcal{W} = \{W_i\}_{i \in I}$ be an open cover of $X \times Y$. For any $x \in X$, the family $\mathcal{W}$ is also an open cover of the subspace $\{x\} \times Y \subset X \times Y$. Since $\{x\} \times Y$ is homeomorphic to the compact space Y, it is itself compact. Thus there are indices $i_1(x), \ldots, i_m(x)$ such that $\{x\} \times Y \subset W_{i_1(x)} \times \ldots \times W_{i_m(x)} =: N(x)$. Since Y is compact, we apply the Tube Lemma to the open set $N(x) \subset X \times Y$ to obtain an open set $U(x) \subset X$ containing x with $U(x) \times Y \subset N(x)$. In particular, the family $\{U(x)\}_{x \in X}$ forms an open cover of the compact space X and so there are elements $x_1, \ldots, x_n \in X$ such that $X = U(x_1) \cup \ldots \cup U(x_n)$. This way, we obtain

$$X \times Y = \bigcup_{j=1}^{n} U(x_j) \times Y \subset \bigcup_{j=1}^{n} N(x_j) \subset \bigcup_{j=1}^{n} W_{i_1(x_j)} \cup \ldots \cup W_{i_m(x_j)}.$$

We have found a finite open subcover of $\mathcal{W}$, which concludes the proof that $X \times Y$ is compact. □

Remark 1.87. Theorem 1.85 generalises to arbitrary products provided one uses the product topology (the result is incorrect with the box topology, take $X_i = \{0, 1\}$ with the discrete topology for each i). The proof that if $\prod_{i \in I} X_i$ is compact, then each X_i is compact works as in Theorem 1.85. The proof of the converse is harder and is known as *Tychonoff's Theorem*, the interested reader can consult [Mun00, Chapter 5] for details.

1.2.2.2 Compact subspaces of $\mathbb{R}$

As when we were studying connectedness, we have now proved several properties of compact spaces without having many concrete examples at our disposal. The next result corrects this state of affairs and will lead to a complete characterisation of the compact subsets of $\mathbb{R}^n$.

Theorem 1.88. *For every $a < b \in \mathbb{R}$, the interval $[a, b]$ is compact.*

Proof. Let $\mathcal{U} = \{U_i\}_{i \in I}$ be an open cover of $[a, b]$, i.e. a family of open sets $U_i \subset \mathbb{R}$ with $[a, b] \subset \bigcup_{i \in I} U_i$. The goal is to show that $[a, b]$ admits a finite open subcover of $\mathcal{U}$. To do that, we will consider the following subset of $(a, b]$:

$$C = \{x \in (a, b] \mid [a, x] \text{ admits a finite open subcover of } \mathcal{U}\}.$$

We describe the plan of the proof and then give the details:

1. We show that C is non-empty and bounded which implies it admits a supremum $c := \sup C$.
2. We show that $c \in C$.
3. We argue that $c = b$.

[14] Recall that for finite products, the box topology coincides with the product topology

Once we will have proved these three steps, the conclusion follows readily: since $b = c \in C$, by definition of C, we deduce that $[a, b]$ admits a finite open subcover of $\mathcal{U}$, which implies it is compact. We now carry out each of the three aforementioned steps.

Firstly, we prove that C is non-empty and bounded, so that $c := \sup C$ exists. The fact that C is bounded is clear: it is included in $(a, b]$. We argue that C is non-empty. Since $a \in [a, b] \subset \bigcup_{i \in I} U_i$, there exists an index $j \in I$ such that $a \in U_j$. As U_j is open, there exists $\varepsilon > 0$ such that $(a - \varepsilon, a + \varepsilon) \subset U_j$. In particular, we have $[a, a + \varepsilon/2] \subset U_j$. Thus, $a + \varepsilon/2$ is an element of C, which is therefore non-empty.

Secondly, we show that $c \in C$. We have the inequalities $a < c$ (because c is an upper bound for C and a is a strict lower bound for C) and $c \leq b$ (because b is an upper bound for C and c is the smallest such upper bound). As a consequence, we have $c \in (a, b]$ and it remains to show that $[a, c]$ admits a finite open subcover of $\mathcal{U}$. As $c \in (a, b] \subset \bigcup_{i \in I} U_i$, there exists an index $k \in I$ such that $c \in U_k$. Since $U_k \subset \mathbb{R}$ is open, and $a < c$, there exists $\varepsilon > 0$ such that $a > c - \varepsilon$ and $(c - \varepsilon, c + \varepsilon) \subset U_k$. Next, note that there exists $d \in C$ with $d \in (c - \varepsilon, c]$: otherwise, if $d < c - \varepsilon$ for every $d \in C$, then $c - \varepsilon$ would be an upper bound of C, smaller than c. Since $d \in C$, by definition of C, the interval $[a, d]$ admits a finite subcover of $\mathcal{U}$: there are indices $i_1, \ldots, i_m \in I$ so that $[a, d] \subset U_{i_1} \cup \ldots \cup U_{i_m}$. Since $d \in (c - \varepsilon, c]$, this implies that $[a, c] = [a, d] \cup (c - \varepsilon, c] \subset U_{i_1} \cup \ldots \cup U_{i_m} \cup U_k$, so $[a, c]$ admits admits a finite open subcover of $\mathcal{U}$, which concludes the proof that $c \in C$.

Thirdly, we show that $c = b$. We already know that $c \leq b$, so assume for a contradiction that $c < b$. As above, there exists an index $\ell \in I$ and $\varepsilon > 0$ so that $[c, c + \varepsilon/2] \subset U_\ell$ and $c + \varepsilon < b$. We then have the inclusion $[a, c + \varepsilon/2] = [a, c] \cup [c, c + \varepsilon/2] \subset U_{i_1} \cup \ldots \cup U_{i_m} \cup U_\ell$. Since $c + \varepsilon < b$, we also have $c + \varepsilon/2 \in (a, b]$ from which we deduce that $c + \varepsilon/2 \in C$. This is impossible, because c is an upper bound for C.

We have therefore proved the three steps mentioned above and, as we already discussed, this implies that $[a, b]$ is compact. ☐

Next, we state the promised generalisation of the extreme value theorem.

Theorem 1.89. *If $f\colon X \to \mathbb{R}$ is a continuous map with X a compact space, then there exists $x_{min}, x_{max} \in X$, such that $f(x_{min}) \leq f(x) \leq f(x_{max})$ for every $x \in X$.*

Proof. Consider $Y := f(X)$ as a subspace of $\mathbb{R}$. As X is compact and f is continuous, Y is compact by Theorem 1.82.

Claim. *There exists $M \in Y$ such that $y \leq M$ for every $y \in Y$.*

Proof. Assume for a contradiction that for every $M \in Y$ there is an $y \in Y$ such that $M < y$. This implies that the family $\{(-\infty, y)\}_{y \in Y}$ is an open cover of the compact space Y. From this, we deduce that there are $y_1, \ldots, y_m \in Y$ such that $Y \subset \bigcup_{i=1}^m (-\infty, y_i)$. On the other hand, the element $y_0 := \max\{y_1, \ldots, y_m\}$ of Y does not belong to $\bigcup_{i=1}^m (-\infty, y_i) = (-\infty, y_0)$, which contradicts $Y \subset \bigcup_{i=1}^m (-\infty, y_i)$. ☐

Returning to the proof of Theorem 1.89, there exists an $M \in Y = f(X)$ such that $y \leq M$ for every $y \in Y$. Thus, there is an $x_{\max} \in X$ such that $M = f(x_{\max})$ with the property that $f(x) \leq f(x_{\max})$ for every $x \in X$. The existence of $x_{\min}$ can be proved analogously. ☐

Recall that a subset $X \subset \mathbb{R}^n$ is *bounded* (for the standard Euclidean metric d_2) if there exists a $R > 0$ such that $||x|| < R$ for every $x \in X$; here we write $||x|| = d_2(x, 0) = \sqrt{\sum_{i=1}^n x_i^2}$ for the norm induced by d_2. In other words, $X \subset \mathbb{R}^n$ is *bounded* if $X \subset B(0, R)$ for some $R > 0$.

Theorem 1.90. *A subspace of $\mathbb{R}^n$ is compact if and only if it is closed and bounded for the Euclidean metric.*

Proof. We first suppose that $A \subset \mathbb{R}^n$ is compact and prove it is closed and bounded. Since $\mathbb{R}^n$ is Hausdorff, Proposition 1.81 implies that A is closed in $\mathbb{R}^n$. As A is compact and $A \subset \mathbb{R}^n = \bigcup_{r>0} B(0,r)$, there exists $r_1, \ldots, r_m > 0$ such that $A \subset B(0,r_1) \cup \ldots \cup B(0,r_m)$. We deduce that $A \subset B(0,R)$ with $R = \max\{r_1, \ldots, r_m\}$ so that A is indeed bounded.

We now prove the converse: we assume that $A \subset \mathbb{R}^n$ is closed and bounded for the Euclidean metric and prove that A is compact. By assumption, there exists $R > 0$ so that $A \subset B(0,R) \subset [-R,R]^n$. By Theorem 1.88, $[-R,R]$ is compact, so $[-R,R]^n$ is compact by Theorem 1.85. Thus A is closed in the compact space $[-R,R]^n$ and is therefore compact by Proposition 1.80. □

Remark 1.91. Here are some remarks concerning Theorem 1.90.

1. One of the two implications from Theorem 1.90 holds much more generally: if (X,d) is a metric space and $A \subset X$ is compact, then A is closed in $(X, \mathcal{T}_d)$ and is bounded for the distance d; the proof is identical as the one from Theorem 1.90.

2. In an arbitrary metric space, the converse is false in general: take $(X, \mathbb{R}^n)$ and $\overline{d_2} = d_2/(1+d_2)$. This metric is bounded and induces the standard topology on $\mathbb{R}^n$ (we mentioned this in Remark 1.26). Thus, with respect to this metric, $A = \mathbb{R}^n$ is bounded; it is also closed in $\mathbb{R}^n$, but $\mathbb{R}^n$ is not compact.

Finally, we briefly describe how compactness can be characterised using sequences. Recall that a *subsequence* of a sequence $(x_n)_{n\in\mathbb{N}}$ is a sequence $(y_i)_{i\in\mathbb{N}}$ with $y_i = x_{n_i}$, where $n_1 < n_2 < \ldots$ is an increasing sequence of positive integers.

Definition 1.92. A space X is *sequentially compact* if every sequence in X admits a convergent subsequence.

The notion of sequential compactness is perhaps familiar from analysis: the Bolzano-Weierstrass theorem says that every bounded sequence in $\mathbb{R}^n$ has a convergent subsequence. In particular, this implies that $[a,b] \subset \mathbb{R}$ is sequentially compact. On the other hand, $\mathbb{R}$ is not sequentially compact as the sequence $x_n = n$ does not admit any convergent subsequence. For these two examples, we see that the same assertions hold for compactness in place of sequential compactness. It is therefore natural to ask about the relationship between these two notions. The answer again involves the notion of first countability that we saw in Definition 1.56.

Theorem 1.93. *Let X be a topological space.*

1. *If X is first countable and compact, then it is sequentially compact.*

2. *If X is a sequentially compact metric space, then it is compact.*

In particular, for metric spaces, sequential compactness is equivalent to compactness.

In the active learning session, we will study the proof of the first item of Theorem 1.93.

Active learning 1.94. The topics that will be covered during the fifth active learning session include:

1. Sequential compactness for spaces endowed with the trivial and discrete topology.

2. We proved the first item of Theorem 1.93: if X is a compact and first countable space, then it is sequentially compact. Given a topological space X, we learnt that $y \in X$ is an *accumulation point* for a sequence $(x_n) \subset X$ if every open set $U \ni x$ contains infinitely many points of the sequence (x_n). We proved that if X is compact, then every sequence admits an accumulation point; we then used Lemma 1.59 to prove that if X is first countable and if (x_n) admits y as an accumulation point, then (x_n) converges to y.

Chapter 2

Algebraic topology

We now come to the second part of this course whose main topic is the fundamental group of a space. Roughly speaking, to every topological space X, we will associate a group $\pi_1(X)$, called the *fundamental group of* X, that consists of all simple closed curves on X "up to deformation". As was mentioned in the introduction, the fundamental group will allow us to tell more spaces apart. In particular, we will be able to show that the sphere is not homeomorphic to the torus, a fact that we were unable to prove using only compactness and connectedness.

Before we get started, two disclaimers are in order. First, we now assume familiarity with group theory. While group theory is not a prerequisite for this class, we have been building up some knowledge in each of the problem sets; some of the material from these exercises is collected in Subsection 2.1.1 below. The second disclaimer concerns continuous maps and homeomorphisms. While in Chapter 1, we strove to write down most homeomorphisms explicitly, this will not be the case any more. For example, from now on statements of the form "a square is homeomorphic to a circle" or "a sphere with a little bump is homeomorphic to a sphere" will be taken for granted.

2.1 The fundamental group

The aim of this section is to define the fundamental group $\pi_1(X)$ of a space X. Section 2.1.2 gives the definition of $\pi_1(X)$ while Subsections 2.1.3 and 2.1.4 are respectively concerned with the fundamental group of spheres and homotopy equivalences. First however we collect some group theoretic notions from the problem sets.

2.1.1 A bit of group theory

The fundamental group of a space is, unsurprisingly, a group and, as a consequence, we need some familiarity with elements of group theory. Since group theory is not a prerequisite of this course, several basic basic notions were given as exercises. We now recall some of these concepts.

Definition 2.1.

- In problem set 0, it was mentioned that a *group* $(G, \cdot)$ is the data of a set G together with a *group law* $\cdot : G \times G \to G$ that satisfies the following three axioms:
  1. the group law is associative: $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$ for every $g_1, g_2, g_2 \in G$;
  2. the group law admits an element e_G, known as the *identity element* such that $g \cdot e_G = g$ and $e_G \cdot g = g$ for every $g \in G$;[1]

[1] The identity element is unique: if e_G, e'_G are identity elements, then $e_G = e_G \cdot e'_G = e'_G$.

3. for every $g \in G$, there is an element $h \in G$, known as *the inverse of* g, such that $g \cdot h = e_G$ and $h \cdot g = e_G$.[2]

In practice, one often writes e instead of e_G as well as xy instead of $x \cdot y$. Additionally, the inverse of g is denoted g^{-1}. Also, because of the associativity axiom, one can write xyz without any parentheses.

- In problem set 4, it was mentioned that a group G is *abelian* if $gh = hg$ for every $g, h \in G$.
- In problem set 2, it was mentioned that given two groups G, H, a map $f\colon G \to H$ is called a *group homomorphism* if $f(gg') = f(g)f(g')$ for every $g, g' \in G$. A bijective homomorphism is called an *isomorphism.*
- In problem set 3, it was mentioned that a *subgroup* $H \leq G$ is a subset $H \subset G$ such that $h_1 h_2^{-1} \in H$ for every $h_1, h_2 \in H$. Note that if H is a subgroup of G, then $e_G \in H$ and if $h \in H$, then $h^{-1} \in H$. Examples of subgroups include the kernel $\ker(f)$ and image $\operatorname{im}(f)$ of a group homomorphism $f\colon G \to J$:

$$\begin{aligned} \ker(f) &= \{g \in G \mid f(g) = e_J\}, \\ \operatorname{im}(f) &= \{h \in J \mid h = f(g) \text{ for some } g \in G\}. \end{aligned}$$

- In problem set 4, it was mentioned that a subgroup N is *normal* if $gng^{-1} \in N$ for every $g \in G$ and every $n \in N$. This is often denoted as $gNg^{-1} \subset N$, where

$$gNg^{-1} = \{gng^{-1} \mid n \in N\}.$$

Observe that if N is normal, then $gNg^{-1} = N$ for every $g \in G$: the inclusion $gNg^{-1} \subset N$ holds by definition, while for the reverse inclusion, as N is normal, we have $g^{-1}ng \in N$ for every $g \in G$ and $n \in N$ and thus $n \in gNg^{-1}$, as required.

Note also that every subgroup of an abelian group is normal, and the kernel of a homomorphism is a normal subgroup.

- In problem set 4, it was also mentioned that if $N \leq G$ is a subgroup, then "$g \sim h$ if and only if $gh^{-1} \in N$" defines an equivalence relation on G. If $N \trianglelefteq G$ is normal, then the *quotient group* of G by N, denoted G/N, consists of the resulting equivalence classes and we saw in problem set 5 that setting $[g] * [h] := [gh]$ turns G/N into a group. If N is not normal, the set G/N is nevertheless defined and the *index* of N in G is $[G : N] = |G/N|$.
- In problem set 6, the "first isomorphism theorem" was mentioned: for any group homomorphism $f\colon G \to H$, the following map is an isomorphism:

$$\begin{aligned} G/\ker(f) &\to \operatorname{im}(f) \\ [g] &\mapsto f(g). \end{aligned}$$

- An exercise on problem set 7 shows that given a subset $R \subset G$, there exists a subgroup $\langle R \rangle$ known as the *smallest subgroup containing* R (it is the intersection of all subgroups containing R). Similarly on problem set 8, we show that there exists a normal subgroup $\langle\langle R \rangle\rangle$ known as the *smallest normal subgroup* containing R (it is the intersection of all normal subgroups containing R).
- In problem set 7, it was also mentioned that if G_1 and G_2 are groups, then their *product* is the group $G_1 \times G_2$ with the group law $(g_1, g_2) \cdot (h_1, h_2) = (g_1 h_1, g_2 h_2)$.

[2] The inverse of g is also unique: if h_1, h_2 are inverses of g, then

$$h_1 = h_1 \cdot e_G = h_1 \cdot (g \cdot h_2) = (h_1 \cdot g) \cdot h_2 = e_G \cdot h_2 = h_2.$$

Since subgroups will play an important role in the study of covering spaces, we work through a few examples.

Example 2.2. We list the subgroups of $\mathbb{Z}$ and S_3.

1. We prove that all the subgroups of $\mathbb{Z}$ are of the form $n\mathbb{Z}$ for some $n \in \mathbb{Z}_{\geq 0}$; they are necessarily all normal since $\mathbb{Z}$ is abelian. A direct verification of the axioms shows that $n\mathbb{Z}$ is a subgroup of $\mathbb{Z}$, so we assume that $H \leq G$ is a subgroup and prove that it equals $n\mathbb{Z}$ for some $n \in \mathbb{Z}_{\geq 0}$. If H is trivial, there is nothing to prove as $H = \{0\} = 0\mathbb{Z}$. If H is non-trivial, we let $n \in H \cap \mathbb{Z}_{>0}$ be the smallest positive integer of H. We prove that $n\mathbb{Z} = H$. For the inclusion $n\mathbb{Z} \subset H$, note that $nm = \operatorname{sgn}(m)(n + \ldots + n)$ and use that H is closed under addition and taking inverses. For the inclusion $n\mathbb{Z} \supset H$, use long division to write $k \in H$ as $k = nq + r$ with $0 \leq r < n$; since $r = k - nq \in H$, by minimality of n, we have $r = 0$ and therefore $k = nq \in n\mathbb{Z}$.

2. Consider the symmetric group S_3 that consists of all bijections of $\{1, 2, 3\}$. Its elements are $\{\mathrm{id}, (12), (13), (23), (123), (132)\}$, where (12) denotes the function f such that $f(1) = 2, f(2) = 1$ and $f(3) = 3$ (and similarly for (13) and (23)) and (123) denotes the function g such that $g(1) = 2, g(2) = 3$ and $g(3) = 1$ (and similarly for (132)). The subgroups of S_3 are

$$\{\mathrm{id}, \{\mathrm{id}, (12)\}, \{\mathrm{id}, (13)\}, \{\mathrm{id}, (23)\}, \{\mathrm{id}, (123), (132)\}, S_3\}.$$

   In this list the subgroups with two elements are not normal, while all the others are.

2.1.2 The definition of the fundamental group

The aim of this section is to define the fundamental group of a space X. Informally, this group consists of all loops in X "up to deformation". To make this precise, we need to define the notion of a deformation, a concept more formally known as homotopy. Note that throughout this section, we will use I as a shorthand for $[0, 1]$. The main references for this section are [Hat02, Chapter 1.1] and [Mun00, Sections 51 and 52].

Recall from Definition 1.72 that a *path* in a space X is a continuous map $\gamma\colon [0, 1] \to X$. The *composition* of two paths γ_0 and γ_1 with $\gamma_0(1) = \gamma_1(0)$ is the path $\gamma_0 \cdot \gamma_1$ defined by

$$\gamma_0 \cdot \gamma_1(t) = \begin{cases} \gamma_0(2t) & \text{if } t \in [0, \frac{1}{2}], \\ \gamma_1(2t-1) & \text{if } t \in [\frac{1}{2}, 1]. \end{cases}$$

The fact that $\gamma_0 \cdot \gamma_1$ is continuous follows from the following lemma that was already proved in problem set 5, when we studied path-connected commponents of a space X.

Lemma 2.3. *Let X be a topological space and assume that $X = \bigcup_{i \in I} X_i$ where each X_i is closed and I is finite. If $f\colon X \to Y$ is a map such that each restriction $f|_{X_i}\colon X_i \to Y$ is continuous, then f is continuous.*

Our main interest will lie in *loops* which are paths γ satisfying $\gamma(0) = \gamma(1)$. Given two loops γ_0 and γ_1 in X that are *based* at the same point $x_0 \in X$ (i.e. such that $\gamma_0(0) = \gamma_1(0) = x_0$), the composition $\gamma_0 \cdot \gamma_1$ is once again a loop based at x_0. Our aim is now to turn the set of all loops in X based at x_0 into a group. To make this possible we formalise the notion of "loops up to deformation".

Definition 2.4. Let X be topological space and let $x_0, x_1 \in X$. Two paths $\gamma_0, \gamma_1\colon I \to X$ from x_0 to x_1 are (endpoint preserving) *homotopic* if there is a continuous map $F\colon I \times I \to X$, called a *homotopy*, such that

1. $F(x, 0) = \gamma_0(x)$ for all $x \in I$,

2. $F(x, 1) = \gamma_1(x)$ for all $x \in I$,

3. $F(0, t) = x_0$ and $F(1, t) = x_1$ for every $t \in I$.

If two based paths γ_0 and γ_1 are homotopic, then we write $\gamma_0 \simeq \gamma_1$.

Remark 2.5. Here is some notation and remarks about homotopies.

1. We often write $f_t(x)$ instead of $F(x, t)$. This way, it becomes clear that each $f_t : I \to X$ is itself a path from x_0 to x_1, and that f_0 coincides with γ_0, while f_1 coincides with γ_1. A schematic illustrated the concept of a homotopy can be seen in Figure 2.1.

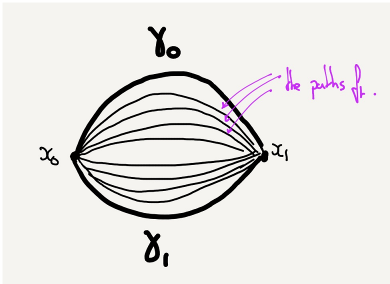


Figure 2.1: A schematic of a homotopy $f_t(x) = F(x, t)$ between paths γ_0 and γ_1 from x_0 to x_1.

2. The composition of homotopic paths gives homotopic paths, i.e. if $f_0 \simeq f_1$ and $g_0 \simeq g_1$ are homotopic paths (with $f_0(1) = g_1(0)$ and $f_1(1) = g_1(0)$), then $f_0 \cdot g_0$ is homotopic to $f_1 \cdot g_1$. Indeed if $f_0 \simeq f_1$ via f_t and $g_0 \simeq g_1$ via g_t, then the required homotopy is obtained by taking the composition of each f_t with each g_t:

$$H(x, t) := h_t(x) = \begin{cases} f_t(2x) & \text{if } x \in [0, \frac{1}{2}], \\ g_t(2x - 1) & \text{if } x \in [\frac{1}{2}, 1]. \end{cases}$$

The fact that H is continuous follows from Lemma 2.3 because H is defined on the union of two closed sets ($I \times I = [0, \frac{1}{2}] \times I \cup [\frac{1}{2}, 1] \times I$) and its restriction to each closed set is continuous: the restriction of H to $[0, \frac{1}{2}] \times I$ is $(x, t) \mapsto (2x, t) \mapsto f_t(2x)$, while its restriction to $[\frac{1}{2}, 1] \times I$ is $(x, t) \mapsto (2x - 1, t) \mapsto g_t(2x - 1)$, both of which are continuous, being the composition of continuous maps).

Example 2.6. Any two paths $g, h : I \to \mathbb{R}^n$ are homotopic: indeed, the required homotopy is $F(x, t) = f_t(x) = (1 - t)g(x) + th(x)$ for $x, t \in I$. Since g and h are both continuous, one can check that F is continuous and one verifies that $f_0 = g$, $f_1 = h$ as well as $f_t(0) = x_0$ and $f_t(1) = x_1$.

In order to define the fundamental group, recall from Remark 1.42 that an *equivalence relation* $\sim$ on a set X is required to satisfy, for all $x, y, z \in X$, the conditions $x \sim x$ (reflexivity), if $x \sim y$, then $y \sim x$ (symmetry) and if $x \sim y$ and $y \sim z$, then $x \sim z$ (transitivity).

Proposition 2.7. *Let X be a topological space and let $x_0, x_1 \in X$. The relation "γ_0 is homotopic to γ_1" is an equivalence relation on the set of paths from x_0 to x_1.*

Proof. To prove reflexivity, we must show that a path f is homotopic to itself: take the constant homotopy $f_t(x) = f(x)$ for every $t \in I$. To prove symmetry, we note that if $f_0 \simeq f_1$ via f_t, then $f_1 \simeq f_0$ via f_{1-t}. To prove transitivity, we assume that $f_0 \simeq f_1$ via f_t, write $g_0 := f_1$ and assume that $g_0 \simeq g_1$ via g_t; the fact that f_0 is homotopic to g_1 follows by consider the following homotopy:

$$H(x, t) = h_t(x) = \begin{cases} f_{2t}(x) & \text{if } t \in [0, \frac{1}{2}], \\ g_{2t-1}(x) & \text{if } t \in [\frac{1}{2}, 1]. \end{cases}$$

The map H is continuous thanks to Lemma 2.3: it is defined on the union of two closed sets ($I \times I = I \times [0, \frac{1}{2}] \cup I \times [\frac{1}{2}, 1]$) and its restriction to each of these closed sets is continuous. We have therefore proved reflexivity, symmetry and transivity and this concludes the proof of the proposition. □

Let X be a space and let $x_0 \in X$. Since Proposition 2.7 shows that being homotopic defines an equivalence relation on the set of loops in X based at x_0, we can consider the set $\pi_1(X, x_0)$ of homotopy classes of loops based at x_0. The composition of two paths based at x_0 is again a path based at x_0 and for $[f], [g] \in \pi_1(X, x_0)$, we define $[f] \cdot [g] := [f \cdot g]$; this is well defined by the second item of Remark 2.5. When the context is clear, we omit the group law from the notation: for instance, we will often write $[f][g]$ instead of $[f] \cdot [g]$.

The next result proves that $(\pi_1(X, x_0), \cdot)$ is a group.

Theorem 2.8. *Let X be a topological space and let $x_0 \in X$. The set $\pi_1(X, x_0)$ of all homotopy classes of loops in X based at x_0 is a group under the composition of paths.*

Proof. We prove that the identity element is the constant path $c_{x_0} : I \to X$ at x_0, i.e. $c_{x_0}(x) = x_0$ for every $x \in I$. We must show that $[c_{x_0}][f] = [f] = [f][c_{x_0}]$ for every $[f] \in \pi_1(X, x_0)$, i.e. that $c_{x_0} \cdot f \simeq f \simeq f \cdot c_{x_0}$ for every loop $f : I \to X$ based at x_0. In fact we can prove a more general statement: if $f : I \to X$ is a path from x_0 to x_1, then $c_{x_0} \cdot f \simeq f$ and $f \cdot c_{x_1} \simeq f$, where c_{x_i} denotes the constant path at x_i. To prove that $c_{x_0} \cdot f \simeq f$, first observe that

$$c_{x_0} \cdot f(x) = \begin{cases} x_0 & \text{if } x \in [0, \frac{1}{2}] \\ f(2x - 1) & \text{if } x \in [\frac{1}{2}, 1] \end{cases}$$

can be rewritten as $f \circ \varphi$, where $\varphi : I \to I$ is defined as $\varphi(x) = 0$ for $x \in [0, \frac{1}{2}]$ and $\varphi(x) = 2x - 1$ for $x \in [\frac{1}{2}, 1]$. Note that $\varphi \simeq \mathrm{id}_I$ via the homotopy $\varphi_t(x) = tx + (1-t)\varphi$ and it follows that $c_{x_0} \cdot f = f \circ \varphi \simeq f \circ \mathrm{id}_I = f$ via the homotopy $f \circ \varphi_t$. The fact that $f \cdot c_{x_1} \simeq f$ is proved analogously by considering $\psi : I \to I$ defined by $\psi(x) = 2x$ for $x \in [0, \frac{1}{2}]$ and $\psi(x) = 1$ for $x \in [\frac{1}{2}, 1]$.

We now prove associativity. If f, g, h are loops in X based at x_0, then we must prove that $f \cdot (g \cdot h) \simeq (f \cdot g) \cdot h$. Again, we prove this more generally, when f, g, h are paths from x_0 to x_1. Applying the definition of path composition, we obtain

$$((f \cdot g) \cdot h)(x) = \begin{cases} f(4x) & \text{if } x \in [0, \frac{1}{4}], \\ g(4x - 1) & \text{if } x \in [\frac{1}{4}, \frac{1}{2}], \\ h(4x - 2) & \text{if } x \in [\frac{1}{2}, 1], \end{cases} \quad \text{and} \quad (f \cdot (g \cdot h))(x) = \begin{cases} f(2x) & \text{if } x \in [0, \frac{1}{2}], \\ g(4x - 2) & \text{if } x \in [\frac{1}{2}, \frac{3}{4}], \\ h(4x - 3) & \text{if } x \in [\frac{3}{4}, 1]. \end{cases}$$

Next, notice that $f \cdot (g \cdot h) = ((f \cdot g) \cdot h) \circ \varphi$ where $\varphi : I \to I$ is defined by $\varphi(x) = \frac{x}{2}$ for $x \in [0, \frac{1}{2}]$, $\varphi(x) = x - \frac{1}{4}$ for $x \in [\frac{1}{2}, \frac{3}{4}]$ and $\varphi(x) = 2x - 1$ for $x \in [\frac{3}{4}, 1]$. The required homotopy is now given by $((f \cdot g) \cdot h) \circ \varphi_t$, where $\varphi_t(x) = tx + (1 - t)\varphi$.

Finally, given a loop $f : I \to X$ based at x_0, we prove that the inverse of $[f]$ is $[\overline{f}]$, where $\overline{f}$ is the loop at x_0 defined by $\overline{f}(x) = f(1 - x)$. Once again, we prove a more general statement: if $f : I \to X$ is a path from x_0 to x_1, then $f \cdot \overline{f} \simeq c_{x_0}$ and $\overline{f} \cdot f \simeq c_{x_1}$. The homotopy $f \cdot \overline{f} \simeq c_{x_0}$ is given by

$$h_t(x) = \begin{cases} f(2xt) & \text{if } x \in [0, \frac{1}{2}], \\ \overline{f}(1 - 2t + 2xt) & \text{if } x \in [\frac{1}{2}, 1]. \end{cases}$$

Switching the roles of f and $\overline{f}$ gives the homotopy required to establish that $\overline{f} \cdot f \simeq c_{x_1}$. This concludes the proof of the theorem. □

Definition 2.9. Let X be a space and let $x_0 \in X$ be a basepoint. The *fundamental group* of X, denoted $\pi_1(X, x_0)$, is the set of all homotopy classes of loops in X based at x_0, with composition of paths as the group law.

Remark 2.10. Here are some remarks about the fundamental group, and more precisely, about its dependence on the basepoint.

1. If X is a space and $x_0 \in X$ is a basepoint, then $\pi_1(X, x_0) = \pi_1(C(x_0), x_0)$, where $C(x_0)$ is the path component of X containing x_0; this is an exercise on the seventh problem set. In particular if $x_0, x_1 \in X$ do not lie in the same path component of X, then there is no relationship between $\pi_1(X, x_0)$ and $\pi_1(X, x_1)$.

2. If x_0, x_1 belong to the same path component of X, then there is a path h from x_0 to x_1 in X and the assignment $f \mapsto h \cdot (f \cdot \overline{h})$ defines a map from the set of loops based at x_1 to the set of loops based at x_0. If $f \simeq g$, then $h \cdot f \cdot \overline{h} \simeq h \cdot g \cdot \overline{h}$ and it is an exercise on the seventh problem set to show that

$$\begin{aligned} \beta_h \colon \pi_1(X, x_1) &\to \pi_1(X, x_0) \\ [f] &\mapsto [h \cdot f \cdot \overline{h}]. \end{aligned}$$

   is a group isomorphism and that $\beta_h([f])$ only depends on the homotopy class of h.

3. The previous point shows that if X is path-connected, then up to isomorphism, the fundamental group does not depend on the choice of a basepoint. For this reason, if X is path-connected, then we sometimes write $\pi_1(X)$ instead of $\pi_1(X, x_0)$.

Definition 2.11. A space X is *simply-connected* if it is path-connected and $\pi_1(X) = 1$.

2.1.3 The fundamental groups of spheres

Our aim is to compute the fundamental group of some familiar spaces, such as Euclidean space, spheres and tori. To achieve this in a timely maner, we delay the technical part of the proof that $\pi_1(S^1) \cong \mathbb{Z}$ to Subsection 2.3.2. We continue to use I as a shorthand for $[0, 1]$. The main references for this section are [Hat02, Chapter 1.1] and [Mun00, Sections 54 and 59].

We start with the fundamental group of Euclidean space.

Example 2.12. For every $n \geq 1$, the group $\pi_1(\mathbb{R}^n)$ is trivial: any two loops f_0, f_1 are homotopic via $f_t(x) = (1-t)f_0(x) + tf_1(x)$, as we saw in Example 2.6. More generally, $\pi_1(A)$ is trivial for any convex subset $A \subset \mathbb{R}^n$; the proof is identical.

Next, we record a characterisation of simple connectedness for later use.

Proposition 2.13. *A space X is simply-connected if and only if there is a unique homotopy class of paths connecting any two points in X.*

Proof. This is an exercise on the seventh problem set. □

We now describe the fundamental group of each sphere S^n for $n \geq 2$.

Proposition 2.14. *For $n \geq 2$, the fundamental group of the n-sphere S^n is trivial : $\pi_1(S^n) = 1$.*

Proof. Fix a basepoint $x_0 \in S^n$. We must show that any loop $f \colon I \to S^n$ based at x_0 is homotopic to the constant loop at x_0, i.e. that f is *nullhomotopic*. We first prove that this is the case when f is not surjective. If f is not surjective, then there is an $x \in S^n$ such that $x \notin \operatorname{im}(f)$. It follows that f factors as $f \colon I \to S^n \setminus \{x\} \to S^n$. Using the stereographic projection (mentioned in the seventh problem set), we know that $S^n \setminus \{x\} \cong \mathbb{R}^n$. Since $\mathbb{R}^n$ is simply-connected (this is Example 2.12) so is $S^n \setminus \{x\} \cong \mathbb{R}^n$: indeed using Proposition 2.13, one verifies that if a space is homeomorphic to a simply-connected space then it is itself simply-connected. Thus, since f factors as $f \colon I \to S^n \setminus \{x\} \to S^n$ with $S^n \setminus \{x\}$ simply-connected, f is indeed nullhomotopic, thus proving the statement when f is not surjective.

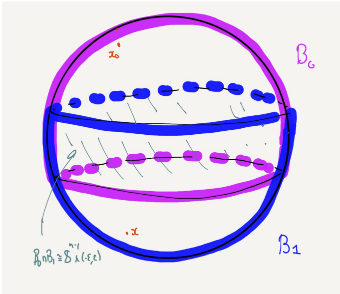


Figure 2.2: Decomposing the sphere S^n as a union of two open balls B_0 and B_1.

We now prove that every $f\colon I \to S^n$ is homotopic to a non surjective map, as this will conclude the proof of the proposition. Write S^n as a union of two slightly enlarged open hemispheres B_0, B_1 so that $x_0 \in B_0$ and $B_0 \cap B_1 \cong S^{n-1} \times (-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$, as in Figure 2.2. Additionally, fix an arbitrary point $x \in B_1$ that is not equal to x_0. We want to homotope f so that it misses x.

Claim. *There is a subdivision $0 = s_0 < s_1 < \ldots < s_m < s_{m+1} = 1$ of $[0,1]$ so that $f(s_i) \in B_0 \cap B_1$ for each i and $f([s_{i-1}, s_i])$ belongs to either B_0 or B_1 for $i = 1, \ldots, m+1$.*

Proof. Since f is continuous and both B_0 and B_1 are open, for every $s \in I$, there is an open set V_s so that $f(V_s)$ is either in B_0 or in B_1.[3] Making V_s smaller if necessary, we can assume that $V_s = (a_s, b_s)$ is an open interval (except at the at the endpoints). These open intervals cover the compact set I and we therefore obtain a finite open subcover $[a_0, b_0), \ldots, (a_i, b_i), \ldots, (a_m, b_m]$ with $a_0 = 0$ and $b_m = 1$. We can assume without loss of generality that these intervals are not subsets of one another (if $(a_r, b_s) \subset (a_s, b_s)$, removing (a_r, b_s) from the open cover still results in an open cover) and that consecutive intervals (a_u, b_u) and (a_{u+1}, b_{u+1}) are mapped to distinct B_i (otherwise replace these intervals by the new interval (a_w, b_w) with $a_w := a_u$ and $b_w = b_{u+1}$). We then obtain the required subdivision by taking s_i to lie in the interval (a_i, b_{i-1}) for $i = 1, \ldots, m$ and setting $s_0 := 0, s_{m+1} := 1$, as illustrated in Figure 2.3. □

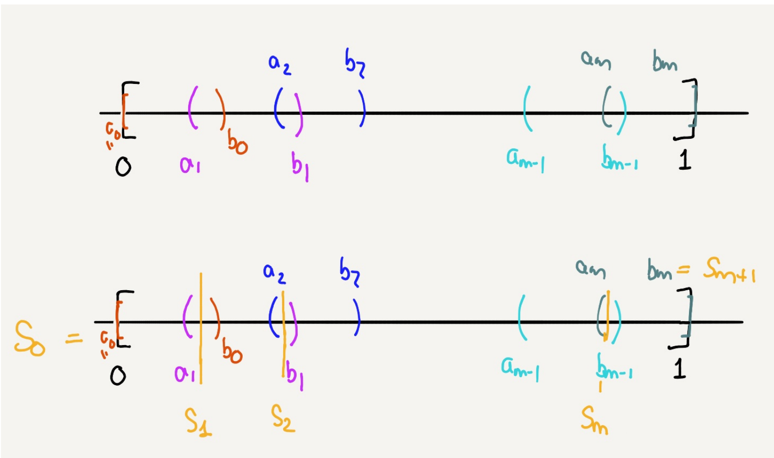


Figure 2.3: The partition of $[0,1]$ is obtained from the open cover given by $\{(a_s, b_s)\}_{s=0}^{m}$.

We now define $f_i := f|_{[s_{i-1}, s_i]}\colon [s_{i-1}, s_i] \to S^n$ so that $f = f_1 \cdots f_{m+1}$. Now consider each j such that $\operatorname{im}(f_j) \subset B_1$. Since $f(s_j)$ belongs to $B_0 \cap B_1 \cong S^{n-1} \times (-\varepsilon, \varepsilon)$ and since this space is path-connected for $n \geq 2$, we can find a path $g_j\colon I \to B_0 \cap B_1$ from $f(s_{j-1})$ to $f(s_j)$. Since B_1

[3] Indeed, every $s \in I$ is either contained in the open set $f^{-1}(B_0)$ or in the open set $f^{-1}(B_1)$.

is simply-connected,[4] we can homotope the path f_j to the path g_j (thanks to Proposition 2.13); this process is illustrated in Figure 2.4. As a consequence, we obtain a map $g\colon I \to S^n$ that is homotopic to f and such that $x \notin \operatorname{im}(g)$, as required.

We have therefore proved that f is homotopic to a non-surjective map g, which is itself null-homotopic as we saw during the first part of the proof. □

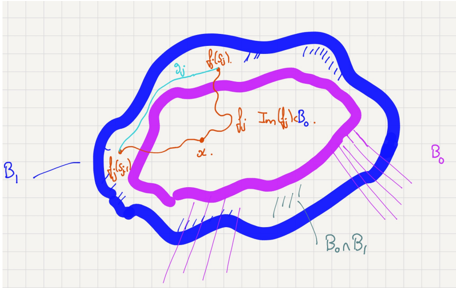


Figure 2.4: A schematic illustrating how we deform the loop f to ensure that it misses x.

The case of the circle S^1 is different and more difficult than S^n for $n \geq 2$.

Theorem 2.15. *For $n \in \mathbb{Z}$, consider the loop $\omega_n\colon I \to S^1$ based at $x_0 := (1,0)$ defined by $\omega_n(x) = e^{2\pi i n x}$. The map $\Phi\colon \mathbb{Z} \to \pi_1(S^1, x_0), n \mapsto \omega_n$ is a group isomorphism.*

Proof. First, we give a more convenient (but equivalent) definition of Φ. Consider the continuous map $p\colon \mathbb{R} \to S^1, x \mapsto e^{2\pi i x}$ and observe that $\widetilde{\omega}_n\colon I \to \mathbb{R}, x \mapsto nx$ satisfies $p \circ \widetilde{\omega}_n = \omega_n$, i.e. $\widetilde{\omega}_n$ is a *lift* of ω_n from S^1 to $\mathbb{R}$. Since any path $\widetilde{f}\colon I \to \mathbb{R}$ from 0 to n is homotopic to $\widetilde{\omega}_n$ (recall Example 2.12), we have $\omega_n = p \circ \widetilde{\omega}_n \simeq p \circ \widetilde{f}$ and, we deduce that Φ can be defined as $\Phi(n) := [p \circ \widetilde{f}]$ for any path $\widetilde{f}\colon I \to \mathbb{R}$ from 0 to n.

Next, we prove that Φ is a group homomorphism. Write $\tau_m\colon \mathbb{R} \to \mathbb{R}$ for the translation given by $\tau_m(x) = x + m$. Since $\widetilde{\omega}_m \cdot (\tau_m \circ \widetilde{\omega}_n)$ is a path in $\mathbb{R}$ from 0 to $m + n$, using our new definition of Φ, and the definition of p, we obtain the following sequence of equalities in $\pi_1(S^1)$:

$$\Phi(m+n) = [p \circ (\widetilde{\omega}_m \cdot (\tau_m \circ \widetilde{\omega}_n))] = [(p \circ \widetilde{\omega}_m) \cdot (p \circ \tau_m \circ \widetilde{\omega}_n)] = [\omega_n \cdot \omega_n] = [\omega_n][\omega_n] = \Phi(m)\Phi(n).$$

We have therefore shown that Φ is a homomorphism and it remains to prove that Φ is a bijection. This will rely on two facts whose proof will be given in Section 2.3.

1. For every path $f\colon I \to S^1$ with $f(0) = x_0 \in S^1$ and every $\widetilde{x}_0 \in p^{-1}(x_0)$, there is a unique lift $\widetilde{f}\colon I \to \mathbb{R}$ of f with $\widetilde{f}(0) = \widetilde{x}_0$

2. For every homotopy $f_t\colon I \to S^1$ of paths with $f_t(0) = x_0$ and every $\widetilde{x}_0 \in p^{-1}(x_0)$, there exists a unique lift $\widetilde{f}_t\colon I \to \mathbb{R}$ of f_t with $\widetilde{f}_t(0) = \widetilde{x}_0$ for every $t \in I$.

We now prove that Φ is surjective. Assume that $[f] \in \pi_1(S^1, x_0)$ is represented by a loop $f\colon I \to S^1$ based at $x_0 \in \mathbb{R}^2$. Set $\widetilde{x}_0 := 0 \in \mathbb{R}$ so that, using the first fact, f lifts to a path $\widetilde{f}\colon I \to \mathbb{R}$ with $\widetilde{f}(0) = \widetilde{x}_0$. Furthermore, since $(p \circ \widetilde{f})(1) = f(1) = x_0$, we deduce that $\widetilde{f}(1) \in p^{-1}(\{x_0\}) = \mathbb{Z} \subset \mathbb{R}$ and therefore $\widetilde{f}(1) = n$ for some $n \in \mathbb{Z}$. Therefore, since $\widetilde{f}$ is a path from $\widetilde{x}_0 = 0$ to n, using the new definition of the Φ, we deduce that $\Phi(n) = [p \circ \widetilde{f}] = [f]$ and therefore Φ is surjective.

Finally, we prove that Φ is injective. We suppose that $m, n \in \mathbb{Z}$ satisfy $\Phi(m) = \Phi(n)$ and prove that $m = n$. Since $\Phi(m) = \Phi(n)$, we have $[\omega_m] = [\omega_n]$, i.e. $\omega_m \simeq \omega_n$ via a homotopy $f_t\colon I \to S^1$

[4] B_1 is simply-connected because it is homeomorphic to a convex space in $\mathbb{R}^n$ and those are simply-connected (by Remark 2.12); here we are again using Proposition 2.13 to deduce that if a space is homeomorphic to a simply-connected space, then it is itself simply-connected.

with $f_0 = \omega_m$ and $f_1 = \omega_n$. By the second fact, we can lift this homotopy to a homotopy $\widetilde{f}_t \colon I \to \mathbb{R}$ with $\widetilde{f}_t(0) = \widetilde{x}_0$ for every t. Since $\widetilde{f}_0$ (resp. $\widetilde{f}_1$) is a lift of ω_m (resp. ω_n) starting at $\widetilde{x}_0 = 0$, the uniqueness of the first fact implies that $\widetilde{f}_0 = \widetilde{\omega}_m$ and $\widetilde{f}_1 = \widetilde{\omega}_n$. Thus, we deduce that $\widetilde{f}_t$ is a homotopy between $\widetilde{\omega}_m$ and $\widetilde{\omega}_n$. Since homotopies preserve endpoints of paths, injectivity follows:

$$m = \widetilde{\omega}_m(1) = \widetilde{f}_0(1) = \widetilde{f}_1(1) = \widetilde{\omega}_n(1) = n.$$

This concludes the proof, apart from the two facts that we will prove in Section 2.3. □

Example 2.16. In the seventh problem set, it was proved that for spaces X and Y, and basepoints $x_0 \in X$ and $y_0 \in Y$, we have $\pi_1(X \times Y, (x_0, y_0)) \cong \pi_1(X, x_0) \times \pi_1(Y, y_0)$. Combining this fact with Theorem 2.15, we deduce that the fundamental group of the torus is $\pi_1(S^1 \times S^1) = \mathbb{Z}^2$. As another example, this time also using Proposition 2.14, we also obtain that $\pi_1(S^1 \times S^2) = \mathbb{Z}$.

2.1.4 Invariance under homotopy equivalences

Our goal is to prove that if two spaces are homeomorphic, then their fundamental groups are isomorphic. In fact, we will prove something stronger: the fundamental group is invariant under homotopy equivalence, a notion that will be introduced shortly. Using these results, we are able to calculate the fundamental group of some additional spaces, such as the Möbius band. The main references for this section are [Hat02, Chapter 1.1] and [Mun00, Section 58].

Both of these aforementioned invariance results follow from a more general construction that associates to a continuous map $f \colon X \to Y$ a homomorphism $f_* \colon \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$.

Construction 2.17. Any continuous map $f \colon X \to Y$ induces a map from the set loops in X based at $x_0 \in X$ to the set of loops in Y based at $f(x_0) \in Y$: to a loop $\gamma \colon I \to X$ based at x_0, associate the loop $f_*(\gamma) := f \circ \gamma \colon I \to Y$ based at $f(x_0)$. Note that if γ_0 is homotopic to γ_1 via a homotopy γ_t, then $f \circ \gamma_0$ is homotopic to $f \circ \gamma_1$ via the homotopy $f \circ \gamma_t$. As a consequence, $f \colon X \to Y$ induces a map

$$\begin{aligned} f_* \colon \pi_1(X, x_0) &\to \pi_1(Y, f(x_0)) \\ [\gamma] &\mapsto [f \circ \gamma]. \end{aligned}$$

Keeping the same notation as in Construction 2.17, the next proposition shows that f_* is in fact a homomorphism, which we therefore call the *homomorphism induced by* f.

Proposition 2.18. *If $f \colon X \to Y$ is a continuous map and $x_0 \in X$ is a basepoint, then the induced map $f_* \colon \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ is a group homomorphism. Additionally, we have $(id_X)_* = id_{\pi_1(X)}$ and, if $g \colon Y \to Z$ is a second continuous map, then $(g \circ f)_* = g_* \circ f_*$.*

Proof. We verify that f_* is a homomorphism by using the definition of the induced map as well as the definition of the group law in $\pi_1(X, x_0)$ and $\pi_1(Y, f(x_0))$:

$$f_*([\gamma_0][\gamma_1]) = f_*([\gamma_0 \cdot \gamma_1]) = [f \circ (\gamma_0 \cdot \gamma_1)] = [(f \circ \gamma_0) \cdot (f \circ \gamma_1)] = [f \circ \gamma_0][f \circ \gamma_1] = f_*([\gamma_0]) f_*([\gamma_1]).$$

The additional verifications concerning induced homomorphisms are left to the reader. □

The invariance of the fundamental group by homeomorphisms now follows promptly.

Proposition 2.19. *If $f \colon X \to Y$ is a homeomorphism and $x_0 \in X$ is a basepoint, then the induced map $f_* \colon \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ is an isomorphism.*

Proof. We already know from Proposition 2.18 that f_* is a homomorphism, so it suffices to prove that f_* is bijective. Consider the inverse $f^{-1} \colon Y \to X$ of f so that $f \circ f^{-1} = \mathrm{id}_Y$ and $f^{-1} \circ f = \mathrm{id}_X$. We now deduce from Proposition 2.18 that $(f^{-1})_* \circ f_* = (f^{-1} \circ f)_* = (\mathrm{id}_X)_* = \mathrm{id}_{\pi_1(X)}$. Similarly, one obtains $f_* \circ (f^{-1})_* = \mathrm{id}_{\pi_1(Y)}$ so that f_* is an isomorphism with inverse $(f^{-1})_*$. □

As we saw in Chapter 1, compactness and connectedness can be used to distinguish spaces. For instance, during Active learning session 1.75, we were able to show that $\mathbb{R}$ is not homeomorphic to $\mathbb{R}^n$ for $n \geq 2$. However, we were unable to prove that $\mathbb{R}^3$ is not homeomorphic to $\mathbb{R}^2$ and that the sphere S^2 is not homeomorphic to the torus T^2. The use of the fundamental group allows for quick proofs of these facts.

Example 2.20. Here are some applications of Proposition 2.19.

1. The sphere S^2 is not homeomorphic to the torus T^2 because $\pi_1(S^2) = 1$, while $\pi_1(T^2) = \mathbb{Z}^2$.
2. For any $x \in \mathbb{R}^n$, we have

$$\pi_1(\mathbb{R}^n \setminus \{x\}) = \begin{cases} \mathbb{Z} & \text{if } n = 2, \\ 1 & \text{if } n \neq 2. \end{cases}$$

   To see this, recall from Active learning 1.75 that $\mathbb{R}^n \setminus \{x\}$ is homeomorphic to $S^{n-1} \times \mathbb{R}_{>0}$ and then use Proposition 2.14, Theorem 2.15 and Proposition 2.19.

3. We argue that $\mathbb{R}^2$ is not homeomorphic to $\mathbb{R}^n$ for $n \neq 2$. Assume for a contradiction that a homeomorphism $\mathbb{R}^2 \to \mathbb{R}^n$ exists. For any $x_0 \in \mathbb{R}^2$, this homeomorphism restricts to a homeomorphism $f \colon \mathbb{R}^2 \setminus \{x_0\} \to \mathbb{R}^n \setminus \{f(x_0)\}$. By Proposition 2.19, f induces an isomorphism $\pi_1(\mathbb{R}^2 \setminus \{x_0\}) \cong \pi_1(\mathbb{R}^n \setminus \{x_0\})$. Since $n \neq 2$, the previous example then implies that $\mathbb{Z}$ is the trivial group, a contradiction.

So far we have not described the fundamental group of spaces such as the Möbius band. To remedy this, we establish a property of the fundamental group that is important in its own right: its invariance under homotopy equivalences. To define this later notion, we start by generalising the concept of a homotopy of paths from Section 2.1.2.

Definition 2.21. Let X and Y be topological spaces.

1. Two continuous maps $f, g \colon X \to Y$ are *homotopic*, denoted $f \simeq g$, if there exists a continuous map $F \colon X \times [0, 1] \to Y$ such that $F(x, 0) = f(x)$ and $F(x, 1) = g(x)$ for all $x \in X$.
2. A continuous map $f \colon X \to Y$ is a *homotopy equivalence* if there is a continuous map $g \colon Y \to X$ such that $g \circ f \simeq \mathrm{id}_X$ and $f \circ g \simeq \mathrm{id}_Y$. In this case, X and Y are said to be *homotopy equivalent*.

We frequently write $f_t(x)$ instead $F(x, t)$, just as we did for homotopies of paths.

Remark 2.22. Here are some remarks concerning homotopies.

1. The definition of a homotopy from Definition 2.21 (nearly) generalises Definition 2.4 concerning path homotopies: we recover the notion of a path homotopy by taking $X = [0, 1]$ and additionally requiring that all of the f_t fix the endpoints of the paths.
2. A homeomorphism $f \colon X \to Y$ is a homotopy equivalence: the inverse f^{-1} of f satisfies the stronger condition that $f \circ f^{-1} = \mathrm{id}_Y$ and $f^{-1} \circ f = \mathrm{id}_X$.

Example 2.23. Here are some examples of homotopy equivalent spaces:

1. For any point $x_0 \in \mathbb{R}^n$, the identity map $\mathrm{id}_{\mathbb{R}^n}$ is homotopic to c_{x_0}, the constant map at x_0 (i.e. $c_{x_0}(x) := x_0$ for every $x \in \mathbb{R}^n$): the homotopy is given by $f_t(x) = (1 - t)x + tx_0$. The same reasoning applies to any convex subset $X \subset \mathbb{R}^n$.
2. The space $\mathbb{R}^n$ is homotopy equivalent to a point (such a space is called *contractible*) because $f \colon \mathbb{R}^n \to \{x_0\}, x \mapsto x_0$ is a homotopy equivalence for any $x_0 \in \mathbb{R}^n$. To see this, consider the inclusion map $g \colon \{x_0\} \to \mathbb{R}^n$ and note that $f \circ g = \mathrm{id}_{\{x_0\}}$ while $g \circ f = c_{x_0} \simeq \mathrm{id}_{\mathbb{R}^n}$ thanks to the first point of this example. For instance, an interval is homotopy equivalent to a point as illustrated on the left hand side of Figure 2.5.

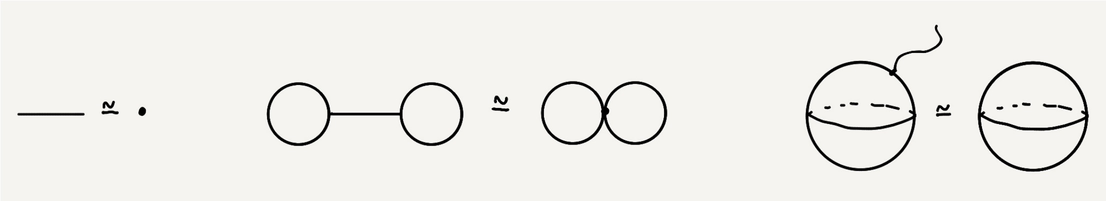


Figure 2.5: This figure illustrates some homotopy equivalences; more on this in Active learning session 2.26 and in later sections.

3. The cylinder and the Möbius band are homotopy equivalent to a circle; this is an exercise on the eight problem set.

4. For a tree [5] T that is a subgraph of a connected graph G, the projection $G \to G/T$ is a homotopy equivalence. This explains the central figure of Figure 2.5 as well as the fact that the letter "A" is homotopy equivalent both to a triangle and to the letter "O".

5. The reader can verify that the continuous map $f\colon D^2 \setminus \{0\} \to \partial D^2 = S^1, re^{\pi i\theta} \mapsto e^{\pi i \theta}$ is a homotopy equivalence: the inclusion $g\colon \partial D^2 \subset D^2$ satisfies $f \circ g = \mathrm{id}_{\partial D^2}$ and $g \circ f \simeq \mathrm{id}_{D^2 \setminus \{0\}}$.

6. The once punctured torus is homotopy equivalent to wedge of two circles $S^1 \vee S^1$, i.e. the quotient of the disjoint union of two copies of S^1 where we identify one point in each circle. More generally, the once punctured surface of genus g is homotopy equivalent to a wedge of $2g$ circles; see Figure 2.6 for an illustration when $g = 2$.

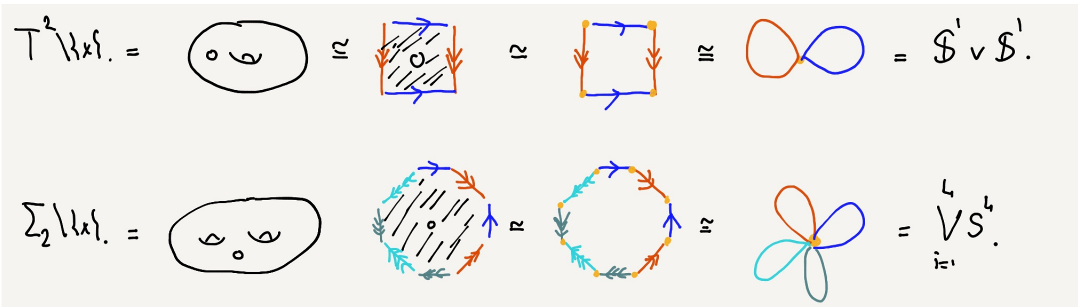


Figure 2.6: This figure illustrates the fact that $T^2 \setminus \{x\} \simeq S^1 \vee S^1$ and $\Sigma_2 \setminus \{x\} \simeq \bigvee_{i=1}^4 S^1$. More generally, one has $\Sigma_g \setminus \{x\} \simeq \bigvee_{i=1}^{2g} S^1$.

Proposition 2.19 established the homeomorphism invariance of the fundamental group. The next result is a generalisation of this fact: the fundamental group is in fact invariant under homotopy equivalence.

Proposition 2.24. *If $\varphi\colon X \to Y$ is a homotopy equivalence and $x_0 \in X$ is a basepoint, then the induced homomorphism $\varphi_*\colon \pi_1(X, x_0) \to \pi_1(X, \varphi(x_0))$ is an isomorphism.*

Proof. We must show that φ_* is bijective. We start off with a claim.

Claim. *If $\{\varphi_t\colon X \to Y\}_{t\in[0,1]}$ is a homotopy and $h\colon I \to Y$ is defined by $h(t) = \varphi_t(x_0)$, then we have the equality $(\varphi_0)_* = \beta_h \circ (\varphi_1)_*$, where $\beta_h\colon \pi_1(Y, \varphi_1(x_0)) \to \pi_1(Y, \varphi_0(x_0))$ is the isomorphism defined by $\beta_h([\gamma]) := [h \cdot \gamma \cdot \overline{h}]$.*

[5] A tree is a graph in which any two vertices are connected by a unique edge. Equivalently, a tree is a contractible graph.

Proof. Given $t \in I$, consider the homotopy $h_t \colon I \to Y, x \mapsto h(tx)$ between the constant path $h_0 = c_{\varphi_0(x_0)}$ and $h_1 = h$. For a loop $\gamma \colon I \to X$ based at x_0, the definition of the induced map leads to the required equality in $\pi_1(Y, \varphi_0(x_0))$

$$(\varphi_0)_*([\gamma]) = [\varphi_0 \circ \gamma] = [h \cdot (\varphi_1 \circ \gamma) \cdot \overline{h}] = (\beta_h \circ (\varphi_1)_*)([\gamma]),$$

where the second equality holds because $h_t \cdot (\varphi_t \circ \gamma) \cdot \overline{h}_t \colon I \to Y$ is a homotopy of loops based at $\varphi_0(x_0)$ between $h_0 \cdot (\varphi_0 \circ \gamma) \cdot \overline{h}_0 = \varphi_0 \circ \gamma$ and $h_1 \cdot (\varphi_1 \circ \gamma) \cdot \overline{h}_1 = h \cdot (\varphi_1 \circ \gamma) \cdot \overline{h}$. □

As $\varphi \colon X \to Y$ is a homotopy equivalence, there is a continuous map $\psi \colon Y \to X$ with $\varphi \circ \psi \simeq \mathrm{id}_Y$ and $\psi \circ \varphi \simeq \mathrm{id}_X$. By Proposition 2.18, the composition $\varphi \circ \psi \circ \varphi$ now induces a homomorphism

$$\varphi_* \circ \psi_* \circ \varphi_* \colon \pi_1(X, x_0) \xrightarrow{\varphi_*} \pi_1(Y, \varphi(x_0)) \xrightarrow{\psi_*} \pi_1(X, \psi(\varphi(x_0))) \xrightarrow{\varphi_*} \pi_1(Y, \varphi\psi\varphi(x_0)).$$

Applying the claim to the homotopy $\psi \circ \varphi \simeq \mathrm{id}_X$ implies that $\psi_* \circ \varphi_* = (\psi \circ \varphi)_* = \beta_h \circ (\mathrm{id}_X)_* = \beta_h$ is an isomorphism and therefore $\varphi_* \colon \pi_1(X, x_0) \to \pi_1(Y, \varphi(x_0))$ is injective. Applying the claim to the homotopy $\varphi \circ \psi \simeq \mathrm{id}_Y$ implies that $\varphi_* \circ \psi_*$ is an isomorphism and so ψ_* is injective. Since ψ_* is injective and $\psi_* \circ \varphi_*$ is an isomorphism, we conclude that $\varphi_* \colon \pi_1(X, x_0) \to \pi_1(Y, \varphi(x_0))$ is surjective. Since we already proved it is injective, we conclude it is an isomorphism, as required. □

Example 2.25. Proposition 2.24 allows us to calculate the fundamental group of some additional spaces as well as to prove that some spaces are not homotopy equivalent.

1. The fundamental group of the Möbius band and that of the cylinder are isomorphic to $\mathbb{Z}$: by Example 2.23 both spaces are homotopy equivalent to a circle and the claim then follows from Proposition 2.24. For the cylinder $S^1 \times [0, 1]$, one can also note that

$$\pi_1(S^1 \times [0, 1]) \cong \pi_1(S^1) \times \pi_1([0, 1]) \cong \pi_1(S^1) \cong \mathbb{Z}.$$

2. The fundamental group a punctured genus g surface is isomorphic to $\pi_1(\vee_{i=1}^{2g} S^1)$. The description of this group is the topic of the next section.

Active learning 2.26. The topics that will be covered during the sixth active learning session include:

1. The definition of retractions and deformation retractions. A *retraction* of a space X onto a subspace $A \subset X$ is a continuous map $r \colon X \to X$ such that $r(X) = A$ and $r|_A = \mathrm{id}_A$. A *deformation retraction* of a space X onto a subspace $A \subset X$ is a homotopy $r_t \colon X \to X$ such that $r_0 = \mathrm{id}_X$, $r_t|_A = \mathrm{id}_A$ for every $t \in [0, 1]$ and $r_1(X) = A$; we say that X *deformation retracts* on A.

2. The effect of retractions and deformation retractions on the fundamental group: if X retracts to a subspace A (and if $x_0 \in A \subset X$ is a basepoint), then the inclusion $\iota \colon A \to X$ induces an injective homomorphism $\iota_* \colon \pi_1(A, x_0) \to \pi_1(X, x_0)$; if X deformation retracts to A, then ι_* is an isomorphism.

3. The closed unit disc D^2 and the Möbius band do not retract onto their boundary.

4. Brouwer's fixed point theorem for continuous maps $D^2 \to D^2$: every continuous map $f \colon D^2 \to D^2$ admits a fixed point, i.e. a point $x \in D^2$ such that $f(x) = x$.

5. We also learnt about graphs and homotopy equivalences.

   (a) A *graph* (in topology) is a space X obtained from a discrete set of points X^0 by attaching to it a collection of closed intervals $\{I_\alpha\}_{\alpha \in A}$. In other words, X is a quotient space

$$X = \left(X^0 \sqcup \bigsqcup_{\alpha \in A} I_\alpha \right) / \sim,$$

where for each $\alpha \in A$, each end point of I_α is declared to be equivalent to an element of X^0. The image of X^0 in X is called the *vertex set* of X and the image of the $\{I_\alpha\}$ in X is called the *edge set* of X.

(b) A *subgraph* $A \subset X$ of a graph X is a graph whose vertex and edge sets are subsets of the vertex and edge sets of X; we endow A with the subspace topology.

(c) If $A \subset X$ is a contractible subgraph, then the projection $X \to X/A$ is a homotopy equivalence; see [Hat02, page 11 and Proposition 0.17] as well as Subsection 2.2.3. Note that this criterion explains the central homotopy equivalence in Figure 2.5. Combining this result with the notions from the next section, we will be abe to calculate the fundamental group of any graph.

2.2 Van Kampen's theorem

Despite knowing how to compute the fundamental group of some familiar spaces, we are still missing several, such as the genus g closed surface (for $g \geq 2$) and the wedges of spheres. Tackling these examples requires van Kampen's theorem whose statement demands a bit more group theory.

2.2.1 Free groups, free products and group presentations

In this group theoretic subsection, we define free groups, free products and group presentations. The main references for this section are [Hat02, Chapter 1.2] and [Mun00, Sections 68 and 69].

Definition 2.27. Given a set X, we write $X^{-1} = \{x^{-1} \mid x \in X\}$ for the set of formal inverses of the elements of X and $X^{\pm 1} = X \cup X^{-1}$.

1. A *word* in X is a sequence of elements $x_1, \ldots, x_n \in X^{\pm 1}$ that we write as $w = x_1 \cdots x_n$.[6] The set of words in X is denoted by $W(X)$.

2. A word $x_1 \cdots x_n$ is *reduced* if $x_i \neq x_{i+1}^{-1}$ for $i = 1, \ldots, n-1$; by convention the empty word (the word with no symbols) is reduced.

3. An *elementary reduction* consists of deleting a subword of the form xx^{-1}.

4. A *reduction* of a word w is a reduced word w' obtained from w by a sequence of elementary reductions.

Example 2.28. Let $X = \{a, b\}$ so that $X^{-1} = \{a^{-1}, b^{-1}\}$ and $X^{\pm 1} = \{a, b, a^{-1}, b^{-1}\}$. The word $bb^{-1}abaa^{-1}$ is not reduced, but after performing two elementary reductions, we obtain the word ab as a reduction.

Proposition 2.29. *Every word w in X admits a unique reduction that we denote by $\overline{w}$.*

Proof. We claim that if w_1, w_2 are words obtained by an elementary reduction on the word w, then there exists a word w_0 that either satisfies $w_0 = w_1 = w_2$ or is an elementary reduction of both w_1 and w_2. There are two cases to consider. If w_1 and w_2 are obtained by disjoint reductions, i.e. if $w = u_1 y_1 y_1^{-1} u_2 y_2 y_2^{-1} u_3, w_1 = u_1 u_2 y_2 y_2^{-1} u_3$ and $w_2 = u_1 y_1 y_1^{-1} u_2 u_3$ (here $u_1, u_2, u_3 \in W(X)$ and $y_1, y_2 \in X^{\pm 1}$), then we can take $w_0 = u_1 u_2 u_3$; if w_1 and w_2 are obtained by overlapping reductions, i.e. $w = u_1 y y^{-1} y u_2, w_1 = u_1 y u_2 = w_2$, then we can take $w_0 = w_1 = w_2$. This concludes the proof of the claim.

We now prove the proposition by induction on the length $|w|$ of the word w. If $|w| = 0$, then w is the empty word and there is nothing to prove. For the induction step, pick two reductions $w \to w_1' \to \ldots \to w_m'$ and $w \to w_1'' \to \ldots \to w_n''$, where each arrow denotes an elementary reduction. By the claim, we know that there is a word w_0 that is a either reduction of both w_1' and w_1'' or satisfies

[6] Sometimes we write words as $w = y_1^{\varepsilon_1} \cdots y_n^{\varepsilon_n}$ where $y_i \in X$ and $\varepsilon_i \in \{-1, 1\}$.

$w_1' = w_0 = w_1''$. Next, we pick a reduction $w_0 \to \ldots \to w_k$ of w_0 so that w_m' and w_k are both reductions of w_1', while w_n'' and w_k are both reductions of w_1''. Since the words w_1' and w_1'' are both shorter than w, the induction hypothesis implies that $w_m' = w_k$ and $w_n'' = w_k$, and so $w_m' = w_n''$, as required. □

Use $F(X)$ to denote the set of all reduced words in X and define the *concatenation* of two reduced words $u, v \in F(X)$ as $u \cdot v = \overline{uv}$.

Theorem 2.30. *The set $F(X)$ of all reduced words in X is a group under concatenation.*

Proof. Associativity follows from Proposition 2.29 : $\overline{\overline{uv}w} = \overline{u\overline{vw}}$ because both words are reductions of the word uvw. The empty word is the neutral element and one verifies that the inverse of a reduced word $w = x_1 \cdots x_n$ is the reduced word $w^{-1} = x_n^{-1} \cdots x_1^{-1}$. □

Definition 2.31. The *free group* $F(X)$ on a set X is defined as the group of reduced words in X, where the group law is given by concatenation.

Remark 2.32. Here are some remarks concerning the free group on a set X.

1. Since each $x \in X$ can be viewed as a reduced word, one gets an injection $X \overset{\iota}{\hookrightarrow} F(X)$. Additionally, note that by definition, X generates $F(X)$, i.e. $F(X) = \langle X \rangle$.

2. If $f \colon X \to G$ is a map to a group G, then there exists a unique homomorphism $\widetilde{f} \colon F(X) \to G$ such that $\widetilde{f} \circ \iota = f$: define $\widetilde{f}$ as $\widetilde{f}(y_1^{\varepsilon_1} \cdots y_n^{\varepsilon_n}) := f(y_1)^{\varepsilon_1} \cdots f(y_n)^{\varepsilon_n}$. The take away is that to specify a homomorphism $F(X) \to G$, it suffices to define a map $X \to G$. Here, it is convenient to depict the equality $\widetilde{f} \circ \iota = f$ using a *commutative diagram*:

$$\begin{array}{ccc} X & \xrightarrow{f} & G. \\ {\scriptstyle\iota}\downarrow & \nearrow_{\exists!\widetilde{f}} & \\ F(X) & & \end{array}$$

3. For every map $\varphi \colon X \to Y$, there is a unique group homomorphism $\varphi_* \colon F(X) \to F(Y)$ such that $\varphi_* \circ \iota_X = \iota_Y \circ \varphi$:

$$\begin{array}{ccc} X & \xrightarrow{\varphi} & Y \\ {\scriptstyle\iota_X}\downarrow & & \downarrow{\scriptstyle\iota_Y} \\ F(X) & \xrightarrow{\exists!\varphi_*} & F(Y). \end{array}$$

To see this, apply the second point to the map $\iota_Y \circ \varphi \colon X \to F(Y)$.

4. If $\varphi \colon X \to Y$ is a bijection, then the induced map $\varphi_* \colon F(X) \to F(Y)$ is an isomorphism with inverse $(\varphi^{-1})_*$. The equality $(\varphi^{-1})_* \circ \varphi_* = \mathrm{id}_{F(X)}$ follows from the uniqueness statement in the second point and the fact that $(\varphi^{-1})_* \circ \varphi_*$ and id_X satisfy $((\varphi^{-1})_* \circ \varphi_*) \circ \iota_X = \iota_X$ and $\mathrm{id}_{F(X)} \circ \iota_X = \iota_X$. The proof that $\varphi_* \circ (\varphi^{-1})_* = \mathrm{id}_{F(Y)}$ is identical.

Example 2.33. Here are some examples of free groups.

1. The free group on $X = \{a\}$ is $F(X) = \{a^n \mid n \in \mathbb{Z}\}$ and this is isomorphic to $\mathbb{Z}$ via the isomorphism $F(X) \to \mathbb{Z}, a^n \to n$.

2. The free group on $X = \{a, b\}$ consists of all reduced words in a, b, a^{-1}, b^{-1}; we denote it by F_2 and refer to it as *the* free group on "the" set of 2 elements. This terminology does not lead to any confusion because, as we noted in Remark 2.32, a bijection $X \to Y$ induces a group isomorphism $F(X) \cong F(Y)$.

3. More generally, we write F_n for the free group on any set with n elements.

Next we describe free products of groups. Given a family $\{G_\alpha\}_{\alpha\in A}$ of groups, a *word in the* $\{G_\alpha\}_{\alpha\in A}$ is a word of the form $g_1g_2\cdots g_n$, where each g_i is an element of some G_α. An *elementary reduction* of such a word consists either of removing an instance of an identity element (in any of the G_α) or of replacing a subword of the form gh, with $g,h\in G_\alpha$, by its product $g{\cdot}h\in G_\alpha$. A word is *reduced* if no such elementary reduction is possible; equivalently $g_1g_2\cdots g_n$ is reduced if $g_i\in G_{\alpha_i}\setminus\{e_{G_{\alpha_i}}\}$ for each $i=1,\dots,n$ and $\alpha_i\neq\alpha_{i+1}$ for each $i=1,\dots,n-1$. Once again, one can prove that every word in the $\{G_\alpha\}_{\alpha\in A}$ admits a unique reduction; the proof is omited. Concatenation followed by reduction is again seen to endow the set of all such reduced words with the structure of a group.

Definition 2.34. The *free product* of a family $\{G_\alpha\}_{\alpha\in A}$ of groups is the group

$$\mathop{\ast}_{\alpha\in A} G_\alpha=\big\{g_1g_2\cdots g_m\mid m\geq 0,\ g_i\in G_{\alpha_i}\setminus\{e_{G_{\alpha_i}}\},\ \alpha_i\neq\alpha_{i+1}\big\}$$

where composition is obtained by concatenation followed by reduction.

Remark 2.35. Here are some remarks concerning free products.

1. For any $\alpha\in A$, every $g\in G_\alpha$ determines an element in $\ast_{\alpha\in A}G_\alpha$, and so there are injective group homomorphisms $\iota_\alpha\colon G_\alpha\to\ast_{\alpha\in A}G_\alpha$ for each $\alpha\in A$.

2. For every group G and every family $\{f_\alpha\colon G_\alpha\to G\}_{\alpha\in A}$ of group homomorphisms, there exists a unique group homomorphism $f\colon \ast_{\alpha\in A}G_\alpha\to G$ such that $f\circ\iota_\alpha=f_\alpha$. If $g_1g_2\cdots g_m\in\ast_{\alpha\in A}G_\alpha$ with $g_i\in G_{\alpha_i}\setminus\{e_{G_{\alpha_i}}\}$ for $i=1,\dots,m$, f is defined as

$$f(g_1g_2\cdots g_m):=f_{\alpha_1}(g_1)\cdots f_{\alpha_m}(g_m).$$

3. It follows from the definition that $G\ast\{e\}\cong G$, where $\{e\}$ denotes the trivial group.

4. An exercise on the ninth problem set shows that if X is a set and if $G_\alpha=\mathbb{Z}$ for every $\alpha\in X$, then $\ast_{\alpha\in X}G_\alpha\cong F(X)$. In particular, $\mathbb{Z}\ast\mathbb{Z}=F_2$ and $\mathbb{Z}\ast F_n\cong F_{n+1}$.

Next, we explain how a group can be described by specifying a set of generators together with the relations between these generators. Let X be a set, and let $R\subset F(X)$ be a set of reduced words in X. Recall from the eigth problem set that $\langle\langle R\rangle\rangle$ denotes the smallest normal subgroup of $F(X)$ containing R and consider the following quotient group:

$$\langle X|R\rangle:=F(X)/\langle\langle R\rangle\rangle.$$

The elements of X are called the *generators* of $\langle X|R\rangle$ and the elements of R are called *relators*. For $r\in R$, we often refer to the equation $r=e_{\langle X|R\rangle}$ as a *relation*.

Definition 2.36. Let X be a set and let $R\subset F(X)$ be a subset. We say that $\langle X|R\rangle$ is a *presentation* of a group G if there is an isomorphism $\langle X|R\rangle\cong G$.

Informally, elements of $\langle X|R\rangle$ are words in X modulo the relations imposed by R.

Remark 2.37. Here are some remarks on group presentations.

1. While every group G admits a presentation (take $X=G$ and $R=\ker(\widetilde{\mathrm{id}}\colon F(G)\to G)$; the first isomorphism theorem shows that $\widetilde{\mathrm{id}}$ induces an isomorphism $F(G)/R\cong G$), presentations of G are highly non unique: for instance if $\langle X|R\rangle$ presents G, then so does $\langle X,x\mid R,x\rangle$.

2. If G is presented by $\langle X|R\rangle$, then to specify a homomorphism $G\to H$, it suffices to define a map $\varphi\colon X\to H$ such that $\widetilde{\varphi}(R)=\{e_H\}$, where $\widetilde{\varphi}\colon F(X)\to H$ denotes the unique extension of φ to $F(X)$; recall Remark 2.32. Indeed requiring that $\widetilde{\varphi}(R)=\{e_H\}$ ensures that $\widetilde{\varphi}$ vanishes on $\langle\langle R\rangle\rangle$ and therefore descends to the quotient $G=F(X)/\langle\langle R\rangle\rangle$.

Example 2.38. Here are presentations of some familar groups:

1. The free group F_n is presented by $\langle x_1, \dots, x_n \mid \emptyset\rangle$ (and in particular $\mathbb{Z} = F_1 = \langle x|\emptyset\rangle$): indeed, by definition, $\langle x_1, \dots, x_n \mid \emptyset\rangle = F(\{x_1, \dots, x_n\})/\langle\emptyset\rangle \cong F_n$.

2. For $d > 1$, the group $\mathbb{Z}_d$ is presented by $\langle x \mid x^d\rangle$. To see this, set $X = \{x\}$ so that, by Remark 2.37, the map $X \to \mathbb{Z}_d, x \mapsto 1 \bmod d$ induces a homomorphism $\langle x \mid x^d\rangle \to \mathbb{Z}_d$ which the reader can verify is an isomorphism. A more concrete way to see this that $\langle x \mid x^d\rangle = \mathbb{Z}_d$ is to observe that $\langle x \mid x^d\rangle$ has d elements, namely $1, x, \dots, x^{d-1}$ and to observe that the map $x^i \mapsto i \bmod d$ is an isomorphism.

3. The group $\mathbb{Z}^n$ is presented by $\langle x_1, \dots, x_n \mid [x_i, x_j] \text{ for all } i < j\rangle$, where we write $[x_i, x_j] := x_i x_j x_i^{-1} x_j^{-1}$. To see this, note that any element of $\langle x_1, \dots, x_n \mid [x_i, x_j] \text{ for all } i < j\rangle$ can be written as $x_1^{m_1} \cdots x_n^{m_n}$ (the relations imply that the x_i all commute with each other) and $x_1^{m_1} \cdots x_n^{m_n} \mapsto (m_1, \dots, m_n)$ is the required isomorphism. Another proof involves using Remark 2.32 to extend the map $\{x_1, \dots, x_n\} \to \mathbb{Z}^n, x_i \mapsto e_i$ [7] to a homomorphism $\varphi\colon F_n \to \mathbb{Z}^n$, to observe that $\ker(\varphi) = \langle\langle [x_i, x_j] \text{ for all } i < j\rangle\rangle$ and to conclude using the first isomorphism theorem.

4. On the ninth problem set, it will be proved that if $\{G_\alpha\}_{\alpha\in A}$ is a family of groups where each G_α has a presentation of the form $G_\alpha = \langle X_\alpha \mid R_\alpha\rangle$ and if $X_\alpha \cap X_\beta = \emptyset$, then the free product $\ast_{\alpha\in A} G_\alpha$ is presented by $\langle \bigsqcup_{\alpha\in A} X_\alpha | \bigsqcup_{\alpha\in A} R_\alpha\rangle$. So, for example, combining this fact with the first example, $\mathbb{Z}_2 \ast \mathbb{Z}_3$ is presented by $\langle a, b \mid a^2, b^3\rangle$.

5. On the ninth problem set, it will be checked that S_3 is presented by $\langle x, y \mid x^2, y^2, (xy)^3\rangle$.

6. The *abelianisation* of a group G is the group $G^{\mathrm{ab}} = G/[G, G]$, where $[G, G] = \langle\{[g, h] \mid g, h\}\rangle$ is normal and is called the *commutator subgroup*; recall that $[g, h] := ghg^{-1}h^{-1}$. One verifies that G^{ab} is abelian and that if $G \cong H$, then $G^{\mathrm{ab}} \cong H^{\mathrm{ab}}$. On the eleventh problem set, it will be shown that if $G = \langle X|R\rangle$, then $G^{\mathrm{ab}} = \langle X|R, \{[x, y]\}_{x,y\in X}\rangle$. One deduces for instance that $F_n^{\mathrm{ab}} = \mathbb{Z}^n$ and therefore that $F_n \cong F_m$ if and only if $n = m$.

2.2.2 Van Kampen's theorem

The aim of this section is to describe $\pi_1(X, x_0)$, when $X = \bigcup_{\alpha\in A} A_\alpha$ is a union of path-connected open subsets $A_\alpha \subset X$ each of which contains the basepoint $x_0 \in X$. This result is known either as van Kampen's theorem or as the Seifert-van Kampen theorem. The main references for this section are [Hat02, Chapter 1.2] and [Mun00, Section 70].

Continuing with the notation from above, for each $\alpha \in A$, the inclusion $A_\alpha \hookrightarrow X$ induces a group homomorphism $j_\alpha\colon \pi_1(A_\alpha, x_0) \to \pi_1(X, x_0)$; this follows from Proposition 2.18. By Remark 2.37, the collection of these maps extends to a homomorphism

$$\begin{aligned}\Phi\colon \mathop{\ast}_{\alpha\in A} \pi_1(A_\alpha, x_0) &\to \pi_1(X, x_0)\\ [\gamma_1]\cdots[\gamma_m] &\mapsto [j_{\alpha_1}(\gamma_1)]\cdots[j_{\alpha_m}(\gamma_m)],\end{aligned}$$

where γ_i is a loop in some A_{α_i} based at x_0 for $i = 1, \dots, m$. The less difficult part of van Kampen's theorem states that Φ is surjective if the intersections $A_\alpha \cap A_\beta$ are path-connected. The first isomorphism theorem implies that Φ induces an isomorphism $\overline{\Phi}\colon \ast_{\alpha\in A} \pi_1(A_\alpha)/\ker(\Phi) \xrightarrow{\cong} \pi_1(X)$. The more difficult part of the theorem describes $\ker(\Phi)$ explicitly. To make sense of this kernel, we write

$$i_{\alpha\beta}\colon \pi_1(A_\alpha \cap A_\beta, x_0) \to \pi_1(A_\alpha, x_0)$$

for the homomorphism induced by the inclusion $A_\alpha \cap A_\beta \subset A_\alpha$.

[7] Here $e_i \in \mathbb{Z}^n$ denotes the vector with 1 in the i-th position and 0 elsewhere.

Theorem 2.39 (van Kampen's Theorem)**.** *Assume that a space X is the union of path-connected open sets A_α each containing the basepoint $x_0 \in X$. If each intersection $A_\alpha \cap A_\beta$ is path-connected, then the homomorphism*

$$\Phi\colon \mathop{\ast}_{\alpha\in A} \pi_1(A_\alpha, x_0) \to \pi_1(X, x_0)$$

is surjective. If in addition, each intersection $A_\alpha \cap A_\beta \cap A_\gamma$ is path-connected, then $\ker(\Phi)$ *equals the normal subgroup N generated by all elements of the form $i_{\alpha\beta}(x) i_{\beta\alpha}(x)^{-1}$ for $x \in \pi_1(A_\alpha \cap A_\beta, x_0)$ and $\alpha, \beta \in A$. In particular Φ induces an isomorphism*

$$\overline{\Phi}\colon \mathop{\ast}_{\alpha\in A} \pi_1(A_\alpha, x_0)/N \xrightarrow{\cong} \pi_1(X, x_0).$$

Informally, the quotienting out by N means that in $\pi_1(X, x_0)$, we have identified $i_{\alpha\beta}(\gamma)$ with $i_{\beta\alpha}(\gamma)$ for every loop $\gamma \subset A_\alpha \cap A_\beta$. This is very intuitive: when we glue spaces together some loops become identified.

Remark 2.40. Here are a couple of remarks concerning van Kampen's theorem.

1. We often apply van Kampen when X is a union of only two spaces, i.e. $X = A_1 \cup A_2$ (with $A_1, A_2 \subset X$ as in Theorem 2.39) in which case the condition involving triple intersections is vacuously satisfied. In this case, Theorem 2.39 states that if $A_1 \cap A_2$ is path-connected, then $\pi_1(X) = (\pi_1(A_1) * \pi_1(A_2))/N$ where N is the smallest normal subgroup containing the $i_{12}(x) i_{21}(x)^{-1}$ with $x \in \pi_1(A_1 \cap A_2)$.

2. We describe the statement of van Kampen's theorem in terms of group presentations. Namely, we assume that $X = A_1 \cup A_2$ as in Theorem 2.39 with $A_1 \cap A_2$ path-connected. Suppose that we have group presentations

$$\begin{aligned} \pi_1(A_1) &= \langle u_1, \dots, u_k \mid \alpha_1, \dots, \alpha_l \rangle, \\ \pi_1(A_2) &= \langle v_1, \dots, v_m \mid \beta_1, \dots, \beta_n \rangle, \\ \pi_1(A_1 \cap A_2) &= \langle w_1, \dots, w_p \mid \gamma_1, \dots, \gamma_q \rangle. \end{aligned}$$

   Van Kampen's theorem then implies that $\pi_1(X)$ is presented by

$$\pi_1(X, x) = \langle u_1, \dots, u_k, v_1, \dots, v_m \mid \alpha_1, \dots, \alpha_l, \beta_1, \dots, \beta_n, i_{12}(w_1) = i_{21}(w_1), \dots, i_{12}(w_p) = i_{21}(w_p) \rangle.$$

   To see this, recall how to obtain the presentation of a free product from Example 2.33 and, additionally, observe that a presentation of G/N is obtained by adding the relators of N to the presentation of G.

3. If $X = A_1 \cup A_2$ with $A_1, A_2 \subset X$ as in Theorem 2.39 and we additionally assume that $A_1 \cap A_2$ is simply-connected, then $\pi_1(X) = \pi_1(A_1) * \pi_1(A_2)$.

The take away from Remark 2.40 is the following recipe for calculating (a presentation of) $\pi_1(X)$ with $X = A_1 \cup A_2$ as in van Kampen's theorem: calculate $\pi_1(A_1), \pi_1(A_2), \pi_1(A_1 \cap A_2)$ and then study the maps i_{12} and i_{12}. We illustrate this with some examples.

Example 2.41. Here are some applications of van Kampen's theorem:

1. We prove that $\pi_1(S^n) = 1$ for $n \geq 2$, recovering the result of Proposition 2.14. Write S^n as the union of two slightly enlarged open hemispheres A_1 and A_2 as in Figure 2.2 so that $A_1 \cap A_2 \cong S^{n-1} \times (-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$. Since A_1, A_2 are open and each of A_1, A_2 and $A_1 \cap A_2$ are path-connected (for the latter we used $n \geq 2$), van Kampen's theorem implies that $\pi_1(S^n) = (\pi_1(A_1) * \pi_1(A_2))/N$. Since A_1 and A_2 are simply connected, this implies that $\pi_1(S^n)$ is trivial, as claimed.

2. We prove that $\pi_1(S^1 \vee S^1) \cong F_2$. Consider the decomposition of $S^1 \vee S^1$ into two circles with small appendages as in Figure 2.7. We denote these spaces by A_1 and A_2, note that they are path-connected and open in $S^1 \vee S^1$ and are homotopy equivalent to circles. Since the intersection $A_1 \cap A_2$ is simply connected (in particular path-connected), van Kampen's theorem applies and we deduce that $\pi_1(S^1 \vee S^1) = \pi_1(S^1) * \pi_1(S^1)$. Using the last item of Remark 2.32, we deduce that $\pi_1(S^1 \vee S^1) = \mathbb{Z} * \mathbb{Z} = F_2$, as claimed.

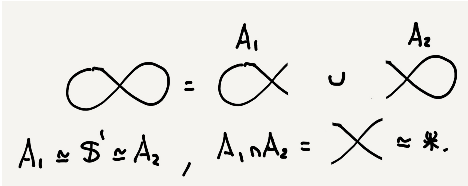


Figure 2.7: Applying van Kampen to $S^1 \vee S^1$.

3. A more general statement holds. Let $\{X_\alpha\}_{\alpha \in A}$ be a family of path-connected spaces and let $x_\alpha \in X_\alpha$ for each $\alpha \in A$. Assume that each x_α is a deformation retract of an open neighborhood U_α in X_α. Then $\pi_1(\bigvee_{\alpha \in A} X_\alpha) = *_{\alpha \in A}\, \pi_1(X_\alpha)$. This will be proved during Active learning session 2.42.

4. We prove that $\pi_1(\Sigma_g) \cong \langle a_1, b_1 \ldots, a_g, b_g \mid [a_1, b_1] \cdots [a_g, b_g] \rangle$. We show the result for $g = 2$; the general statement follows by induction. Write Σ_2 as the union of two punctured tori (call them A_1 and A_2) so that the intersection is $S^1 \times (-\varepsilon, \varepsilon)$ for some small $\varepsilon > 0$; see Figure 2.8. Each of the A_i deformation retracts to $S^1 \vee S^1$ and therefore $\pi_1(A_i) = F_2$. Since $A_1 \cap A_2$ deformation retracts to a circle, we have $\pi_1(A_1 \cap A_2) = \mathbb{Z}$. We now pick generators for these groups as in Figure 2.8, so that $\pi_1(A_1) = \langle [a], [b] \rangle, \pi_1(A_2) = \langle [c], [d] \rangle$ and $\pi_1(A_1 \cap A_2) = \langle \gamma \rangle$. This is a *choice* and a different choice will lead to a different presentation for $\pi_1(\Sigma_2)$. From now on, we write a, b, c, d instead of $[a], [b], [c], [d]$ to avoid overloading the notation and to avoid confusion with commutators. [8]

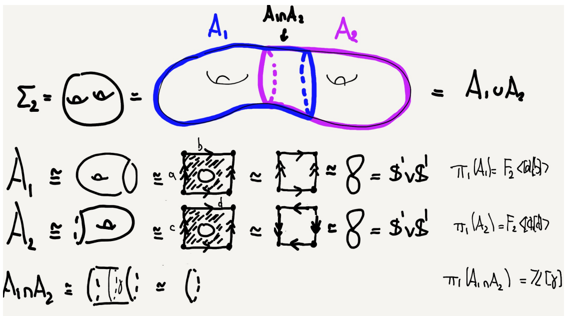


Figure 2.8: Applying van Kampen to the surface Σ_2: the decomposition

To describe N, we study i_{12} and i_{21}. The domain of both of these maps is $\pi_1(A_1 \cap A_2) = \mathbb{Z}$, so it suffices to understand the image of the generator, that we called γ. In A_i, the image of γ is the boundary of the punctured solid tori and we see see that $i_{12}(\gamma) = [a, b]$; similarly $i_{21}(\gamma) = [c, d]^{-1}$; see Figure 2.9.

Van Kampen's theorem now implies that $\pi_1(\Sigma_2) = \langle a, b, c, d \mid [a, b] = [c, d]^{-1} \rangle$, i.e. as required $\pi_1(\Sigma_2) = \langle a, b, c, d \mid [a, b][c, d] \rangle$

[8] A little care is needed with basepoints in this example, see e.g. the proof of Proposition 2.45.

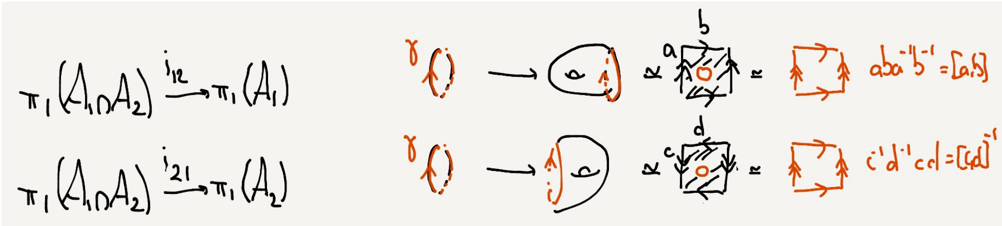


Figure 2.9: Applying van Kampen to the surface Σ_2: understanding the identifications

5. We prove that if $n \geq 3$ and if M_1 and M_2 are two path-connected n-manifolds, then $\pi_1(M_1 \# M_2) = \pi_1(M_1) * \pi_1(M_2)$.[9] Here, recall from Active learning session 1.46 that $M_1 \# M_2$ denotes the connected sum and is defined as

$$(M_1 \setminus D_1^n) \sqcup (M_2 \setminus D_2^n)/\sim,$$

where the equivalence relation identifies x with $\varphi(x)$ for every $x \in S^{n-1}$, for some (fixed) homeomorphism $\varphi\colon S^{n-1} \to S^{n-1}$. View M as $A_1 \cup A_2$ where A_i deformation retracts to $M_i \setminus D_i^n$ and $A_1 \cap A_2 \cong S^{n-1} \times (-\varepsilon, \varepsilon)$. Since $n \geq 3$, this latter space is simply-connected and van Kampen's theorem implies that $\pi_1(M_1 \# M_2) = \pi_1(M_1 \setminus D_1^n) * \pi_1(M_2 \setminus D_2^n)$. Now a second application of van Kampen shows that $\pi_1(M_i \setminus D_i^n) = \pi_1(M_i)$ for $n \geq 3$ and so the result follows.

6. In the tenth and eleventh problem sets, we will apply van Kampen's theorem to calculate the fundamental groups of the Klein bottle and of real projective space $\mathbb{R}P^2$; see also Subsection 2.2.3.

Active learning 2.42. The topics that will be covered during the seventh active learning session include:

1. We used the example of a circle to see why it is necessary that the A_α be path connected and open.

2. A discussion of the necessity of the triple intersections being path-connected can be found in [Hat02, p.144].

3. Let $\{X_\alpha\}_{\alpha \in A}$ be a family of path-connected spaces and let $x_\alpha \in X_\alpha$ for each $\alpha \in A$. We saw that if x_α is a deformation retract of an open neighborhood U_α in X_α, then van Kampen's theorem implies that

$$\pi_1\left(\bigvee_{\alpha \in A} X_\alpha\right) = \mathop{\ast}_{\alpha \in A} \pi_1(X_\alpha).$$

The details can be found in [Hat02, Example 1.21].

4. We used van Kampen to calculate the fundamental group of a graph using maximal trees; a second proof of the result can be found on the eleventh problem set. In more detail, we recalled that a *tree* is a contractible graph, that a tree in a graph X is *maximal* if it contains all the vertices of X and that if X is a connected graph with maximal tree T, then $\pi_1(X)$ is free on the number of edges of $X \setminus T$.

[9]We did not have time to get to this example in class, but I am leaving it in the notes. Note also that the result if false for $n = 2$ by the previous point: $\pi_1(\Sigma_2) = \pi_1(\Sigma_1 \# \Sigma_1) \neq \pi_1(\Sigma_1) * \pi_1(\Sigma_1)$. Time permitting we will instead calculate the fundamental group of the space obtained from a torus $T^2 \cong S^1 \times S^1$ by gluing a punctured torus to one of the S^1 factors.

2.2.3 The fundamental group and cell attachments

Several familiar spaces can be obtained by gluing a single n-disc to a simpler space. Understanding the effect of such a "cell attachment" on the fundamental group allows for quick descriptions of the fundamental group of surfaces. The main references for this section are [Hat02, Chapter 1.2] and [Mun00, Sections 72 and 73].

We start by making the notion of attaching a disc to a space more precise.

Definition 2.43. Given a space X and a continuous map $\varphi\colon \partial D^n \to X$, the *space obtained by attaching an n-cell to X along φ* is the quotient space

$$X \cup_\varphi D^n := (X \sqcup D^n)/\sim$$

under the equivalence relation $x \sim \varphi(x)$ for every $x \in \partial D^n$. In this setting, we often call φ the *attaching map*. The image of the open ball $\mathrm{Int}(D^n)$ inside $X \cup_\varphi D^n$ is denoted by e^n and is called an *n-cell*.[10]

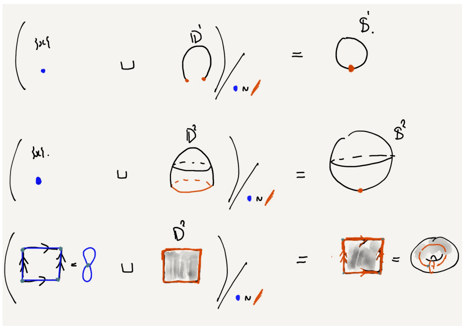


Figure 2.10: Examples of cell attachments.

Example 2.44. Here are some examples of spaces obtained by attaching an n-cell.

1. The circle S^1 can be obtained by attaching a 1-cell to a point $\{x\}$ along the constant map $\partial D^1 \to \{x\}$. This is described in the first picture of Figure 2.10.
2. More generally, the n-sphere can be obtained by attaching a single n-cell to a point $\{x\}$ along the constant map $\partial D^n \to \{x\}$. This is described in the second picture of Figure 2.10.
3. The torus can be obtained from the wedge $S^1 \vee S^1$ of two circles by attaching a 2-cell as described in the third picture of Figure 2.10.

We now want to calculate $\pi_1(X \cup_\varphi D^n)$ for X a path-connected space. The case $n = 2$ differs from the case $n \geq 3$: given a basepoint $s_0 \in \partial D^2 = S^1$, the continuous map $\varphi\colon \partial D^2 \to X$ defines a loop in X based at $x_0 := \varphi(s_0)$ and therefore determines an element in $\pi_1(X, x_0)$ that we denote by $[\varphi]$; the situation is illustrated in Figure 2.11.

[10] The terminology is slightly abusive as one is attaching the closed ball D^n and not the open ball e^n, but this usage is nonetheless common.

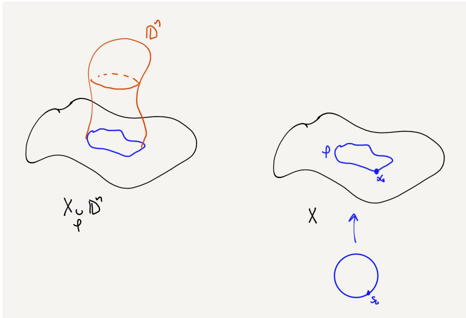


Figure 2.11: Attaching a 2-cell along a loop representing an element of $\pi_1(X)$.

Proposition 2.45. *Let X be a path-connected space, and let $\varphi\colon \partial D^n \to X$ be a continuous map.*

1. *if $n \geq 3$, then the inclusion induces an isomorphism $\pi_1(X) \cong \pi_1(X \cup_\varphi D^n)$.*
2. *if $n = 2$, then the inclusion induces an isomorphism $\pi_1(X, x_0)/\langle\langle[\varphi]\rangle\rangle \cong \pi_1(X \cup_\varphi D^n, x_0)$.*

Proof. Pick a point $p \in e^n \subset Y \setminus X$ and decompose $Y := X \cup_\varphi D^n$ as $Y = A \cup B$ where $A := Y \setminus \{p\}$ and $B = e^n$. Note that $A, B \subset Y$ are open and path-connected. Additionally, observe that A deformation retracts to X, while B is contractible and $A \cap B$ deformation retracts to S^{n-1}; all of this is illustrated in Figure 2.12.

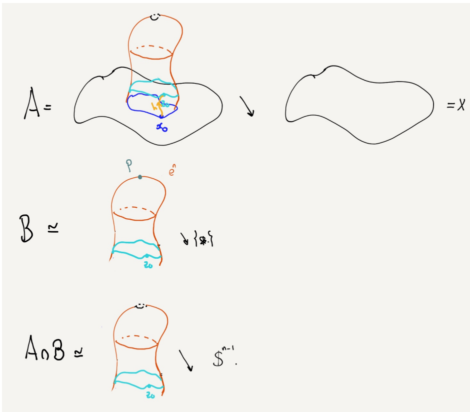


Figure 2.12: The decomposition of $Y := X \cup_\varphi D^n$ as a union of A and B.

In order to apply van Kampen's theorem, we need a basepoint that lies both in A and B, so we pick $z_0 \in A \cap B$ slightly above x_0. For $n \geq 3$, since $A \cap B \simeq S^{n-1}$ is simply-connected, van Kampen's theorem immediately yields the required outcome:

$$\pi_1(Y, x_0) \cong \pi_1(Y, z_0) \cong \pi_1(A, z_0) \cong \pi_1(A, x_0) \cong \pi_1(X, x_0).$$

We now focus on the case $n = 2$. This time, we additionally join x_0 and z_0 by a path $h \subset A$ and consider a loop $\varphi' \subset A \cap B$ based at z_0 as in Figure 2.12, so that φ' is homotopic in A to the loop $\beta_h(\varphi) = h \cdot \varphi \cdot \overline{h}$. Thus, under the change of basepoint isomorphism β_h from Remark 2.10, we have

$$\frac{\pi_1(A, x_0)}{\langle\langle[\varphi]\rangle\rangle} \cong \frac{\pi_1(A, z_0)}{\langle\langle[\varphi']\rangle\rangle}.$$

So far, if we use $j\colon \pi_1(A \cap B, z_0) \to \pi_1(A, z_0)$ to denote the inclusion map, then by van Kampen's theorem we have the isomorphism

$$\pi_1(Y, x_0) \cong \pi_1(Y, z_0) = \frac{\pi_1(A, z_0)}{\langle\langle \mathrm{im}(j)\rangle\rangle}.$$

It remains to analyze $\mathrm{im}(j)$. We see that j maps the generator of $\pi_1(A \cap B, z_0) = \pi_1(S^1) = \mathbb{Z}$ to $[\varphi'] \in \pi_1(A, z_0)$. It follows that $\mathrm{im}(j) = \langle\langle[\varphi']\rangle\rangle$, and

$$\pi_1(Y, x_0) \cong \pi_1(Y, z_0) \cong \frac{\pi_1(A, z_0)}{\langle\langle[\varphi']\rangle\rangle} \cong \frac{\pi_1(A, x_0)}{\langle\langle[\varphi]\rangle\rangle} \cong \frac{\pi_1(X, x_0)}{\langle\langle[\varphi]\rangle\rangle}.$$

This concludes the proof when $n = 2$ and therefore the proof of the proposition. □

Example 2.46. Using Proposition 2.45, we give a second calculation of the fundamental group of several familiar surfaces.

1. For the genus g surface Σ_g, we have $\pi_1(\Sigma_g) = \langle a_1, b_1 \ldots, a_g, b_g \mid [a_1, b_1] \cdots [a_g, b_g]\rangle$ because Σ_g is obtained from $\vee_{i=1}^{2g} S^1$ by attaching a single 2-cell as described in Figure 2.13.
2. For the real projective plane $\mathbb{R}P^2$, we have $\pi_1(\mathbb{R}P^2) = \langle a \mid a^2\rangle = \mathbb{Z}_2$ because $\mathbb{R}P^2$ is obtained from S^1 by attaching a single 2-cell as described in Figure 2.13.
3. The case of the Klein bottle is an exercise on problem set 9.

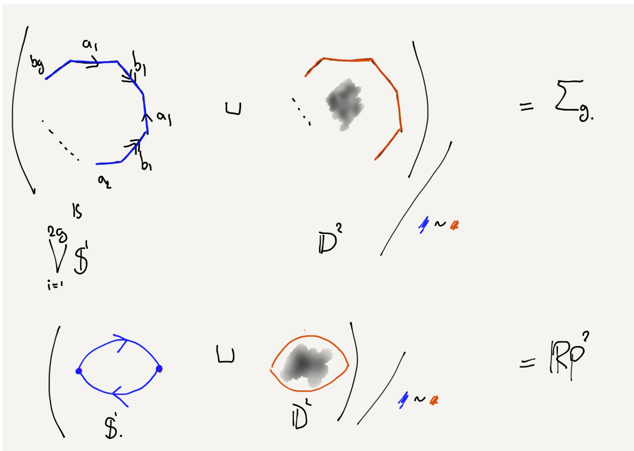


Figure 2.13: The genus g surface Σ_g and $\mathbb{R}P^2$ are obtained by attaching a 2-cell.

Remark 2.47. Here are some additional remarks on cell attachments:

1. Naturally, we can attach more than one n-cell at a time: given a space X, we pick a collection $\{\varphi_\alpha : \partial D^n_\alpha \to X\}_{\alpha \in A}$ of continuous maps and consider $\left(X \sqcup \bigsqcup_{\alpha \in A} D^n_\alpha\right) / \sim$ where the equivalence relation identifies x with $\varphi_\alpha(x)$ for every $x \in D^n_\alpha$. For example, the torus T^2 is obtained from $X^0 = \{x\}$ by attaching two 1-cells via the constant map, yielding $X^1 = S^1 \vee S^1$, and then attaching a 2-cell to X^1 as described in Example 2.44.
2. A *CW complex of dimension n* (sometimes also called an *n-complex* or a *cell complex of dimension n*) is a space X obtained inductively as follows: X^0 is a set of points with the discrete topology, X^k is obtained from X^{k-1} by attaching k-cells for $1 \leq k \leq n$ and $X = X^n$. The space X^k is called the *k-skeleton* of X. So, for instance, a graph is a 1-complex (recall Active learning Session 2.26) and, in the previous example, we endowed the torus with the structure of a CW complex of dimension 2.

By paying attention to the basepoints, Proposition 2.45 can be generalised to the case where multiple n-cells are attached. Before describing the result, we briefly introduce some notation.

Construction 2.48. Let X be a path-connected space and let $\{\varphi_\alpha : \partial D^n_\alpha \to X\}_{\alpha \in A}$ be a family of continuous maps. If $s_0 \in S^1$ is a basepoint, then φ_α determines a loop based at $\varphi_\alpha(s_0)$ that we also denote by φ_α. Choose a basepoint $x_0 \in X$ and a path γ_α in X from x_0 to $\varphi_\alpha(s_0)$ for each α. Then $\gamma_\alpha \varphi_\alpha \overline{\gamma}_\alpha$ is a loop in X based at x_0.

Proposition 2.49. *Assume that $Y = \left(X \sqcup \bigsqcup_{\alpha \in A} D^n_\alpha\right) / \sim$ is obtained from a path-connected space X by attaching n-cells along a family $\{\varphi_\alpha : \partial D^n_\alpha \to X\}_{\alpha \in A}$ of continuous maps.*[11]

1. *If $n \geq 3$, then the inclusion $X \hookrightarrow Y$ induces an isomorphism $\pi_1(X) \xrightarrow{\cong} \pi_1(Y)$.*
2. *If $n = 2$, then the inclusion $X \hookrightarrow Y$ induces an isomorphism $\pi_1(X)/N \cong \pi_1(Y)$, where N is normally generated by the loops $\gamma_\alpha \varphi_\alpha \overline{\gamma_\alpha}$ described in Construction 2.48.*

Proof. In order to apply van Kampen's theorem, we work with a space Z that deformation retracts onto Y. Namely, we consider the space $Z = (Y \sqcup \bigsqcup_\alpha R_\alpha)/ \sim$ obtained from Y by attaching ribbons $(I \times I)_{\alpha \in A}$ to Y along the paths $(\gamma_\alpha)_{\alpha \in A}$, as depicted in Figure 2.14. [12] Additionally, we let $\{p_\alpha\}_{\alpha \in A}$ be a collection of points with $p_\alpha \in e^n_\alpha \subset Z \setminus X$.

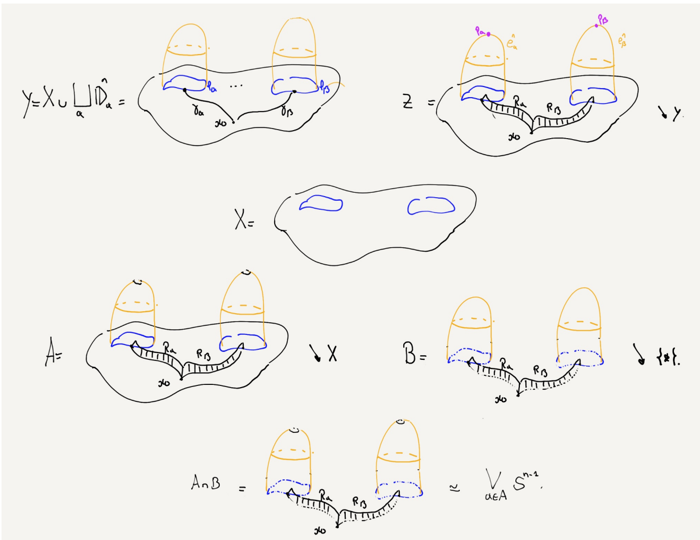


Figure 2.14: The decomposition of $Y := X \cup_{\varphi_\alpha} \bigcup_{\alpha \in A} D^n$ as a union of A and B.

We now apply van Kampen's theorem to $Z = A \cup B$, where $A = Z \setminus \bigcup_{\alpha \in A} \{p_\alpha\}$ and $B = Z \setminus X$. Essentially, as the space A deformation retracts to X, the space B is contractible, and $A \cap B$ is homotopic equivalent to $\bigvee_{\alpha \in A} S^{n-1}$, the result now follows from van Kampen's theorem. We now give a more details,but omit the detailed discussion of basepoints which can be carried out as in the proof of Proposition 2.45. For $n \geq 3$, since $A \cap B$ is simply-connected, this is immediate:

$$\pi_1(Y) = \pi_1(Z) = \pi_1(A) = \pi_1(X).$$

For $n = 2$, note that on fundamental groups, the inclusion $j : A \cap B \hookrightarrow A \simeq X$ takes the generator of each free factor of $\pi_1(A \cap B) = \ast_{\alpha \in A} \mathbb{Z}$ to $[\gamma_\alpha \varphi_\alpha \overline{\gamma_\alpha}] \in \pi_1(A, x_0)$, by definition of the cell attachment. Van Kampen's theorem now implies that

$$\pi_1(Y) = \pi_1(Z) = \frac{\pi_1(A)}{\langle\langle \mathrm{im}(j) \rangle\rangle} = \frac{\pi_1(A, x_0)}{[\{\gamma_\alpha \varphi_\alpha \overline{\gamma_\alpha}]\}_{\alpha \in A}} = \frac{\pi_1(X, x_0)}{[\{\gamma_\alpha \varphi_\alpha \overline{\gamma_\alpha}]\}_{\alpha \in A}}.$$

This concludes the proof for $n = 2$ and therefore the proof of the proposition. □

[11] Depending on the year, we might not discuss Proposition 2.49 during class.

[12] A more convincing picture can also be found in [Hat02, page 50].

Active learning 2.50. The topics that will be covered during the eigth active learning session include:

1. We saw that $\mathbb{R}P^n$ admits a cell structure with exactly one k-cell for $k = 0, \ldots, n$ and that $\pi_1(\mathbb{R}P^n) = \mathbb{Z}_2$ for $n \geq 2$. The key fact was that $\mathbb{R}P^n \cong \mathbb{R}P^{n-1} \cup_\varphi D^n$, where $\varphi\colon \partial D^n = S^{n-1} \to \mathbb{R}P^{n-1}$ is the projection map (recall that $\mathbb{R}P^{n-1} = S^{n-1}/x \sim -x$). By induction, this fact implies both the claim about the cell structure of $\mathbb{R}P^n$ and that $\pi_1(\mathbb{R}P^n) = \mathbb{Z}_2$ for $n > 1$ (use $\pi_1(\mathbb{R}P^2) = \mathbb{Z}_2$ and Proposition 2.45). To prove the fact recall from the third p-set that $\mathbb{R}P^n$ is D^n with antipodal points of ∂D^n identified. Since ∂D^n with antipodal points identified is $\mathbb{R}P^{n-1}$, the fact follows.

2. For every group G, there exists a 2-complex X_G with $\pi_1(X_G) = G$. Pick a presentation $\langle \{g_\alpha\}_{\alpha \in A} | \{r_\beta\}_{\beta \in B} \rangle$ for the group G. Take X^0 to be a single point, attach a 1-cell to X^0 for each generator, resulting in $X^1 = \vee_{\alpha \in A} S^1$ and then obtain $X_G := X^2$ by adding one 2-cell along a loop representing each relator. The fact that $\pi_1(X_G) = G$ now follows from Proposition 2.49.

3. We learnt briefly about knots $K \subset S^3$ and the group $\pi_1(S^3 \setminus K)$: a knot K is a simple closed (smooth) curve $K \subset S^3$; two knots are *isotopic* if there is a homeomorphism $H\colon S^3 \to S^3$ (that is orientation-preserving) such that $H(K) = K$.[13] The *knot group* of a knot K is $G(K) = \pi_1(S^3 \setminus K)$. If K and K' are isotopic, then $G(K) \cong G(K')$: a homeomorphism $H\colon S^3 \to S^3$ with $H(K) = K'$ restricts to a homeomorphism $S^3 \setminus K \cong \S^3 \setminus K'$ and the result now follow from the homeomorphism invariance of the fundamental group (recall Proposition 2.19).

4. We presented (without proof) Wirtinger's algorithm to obtain a presentation of the knot group $G(K) = \pi_1(S^3 \setminus K)$ from a knot diagram D of K. This presentation has one generator for each strand of the diagram D and one relation for each crossing of D (the relations depends on the sign of the crossing).

2.2.4 The proof of van Kampen's theorem

This section is devoted to the proof of van Kampen's theorem and follows [Hat02, pages 44-46]. It was emphasised during class that this proof is not considered crucial for the class: the proof is somewhat uninspiring and it is much more important to understand how to apply van Kampen's theorem than to understand the intricacies of the proof.

Proof of Theorem 2.39. We prove the first item, namely that if $A_\alpha \cap A_\beta$ is path-connected for every $\alpha, \beta \in A$, then the homomorphism $\Phi\colon \ast_{\alpha \in A} \pi_1(A_\alpha, x_0) \to \pi_1(X, x_0)$ is surjective. Given a loop $f\colon I \to X = \bigcup_{\alpha \in A} A_\alpha$ based at x_0 we must define loops $h_1, \ldots, h_m$ with h_i a loop in A_{α_i} based at x_0 so that $f \simeq h_1 \cdots h_m$.

Claim. *There is a subdivision $0 = s_0 < s_1 < \ldots < s_m < s_{m+1} = 1$ of $[0, 1]$ so that $f([s_{i-1}, s_i]) \subset A_\alpha = A_i$ for $i = 1, \ldots, m + 1$.*

Proof. The proof is identical to that in the claim in of proof of Proposition 2.14 but we recall the argument. Since f is continuous and the A_α are open, for every $s \in I$, there is an open set V_s so that $f(V_s)$ is in some A_α. Making V_s smaller if necessary, we can assume that $V_s = (a_s, b_s)$ is an open interval. These open intervals cover the compact set I and we therefore obtain a finite open subcover $(a_0, b_0), \ldots, (a_m, b_m)$ with $a_0 = 0$ and $b_m = 1$. We can assume without loss of generality that these intervals are not subsets of one another (if $(a_r, b_s) \subset (a_s, b_s)$, removing (a_r, b_s) from the open cover still results in an open cover) and that consecutive intervals (a_u, b_u) and $(a_{u+1} b_{u+1})$

[13] A more frequent definition of K and K' being (ambient) isotopic is that there is a family $F_t\colon S^3 \to S^3$ of homeomorphisms with $F_0 = \mathrm{id}$ and $F_1(K) = K'$ and such that $(x, t) \mapsto F_t(x)$ is continuous; the fact that both definitions are equivalent is not obvious.

are mapped to distinct A_α (otherwise replace these intervals by the new interval (a_w, b_w) with $a_w := a_u$ and $b_w = b_{u+1}$). We then obtain the required subdivision by taking s_i to lie in the interval (a_i, b_{i-1}) for $i = 1, \ldots, m$ and setting $s_0 := 0, s_{m+1} := 1$, as illustrated in Figure 2.3. □

Now define $f_i := f|_{[s_{i-1},s_i]} \colon [s_{i-1}, s_i] \to A_i$ so that by construction $f = f_1 \cdot f_2 \cdots f_m$. By hypothesis $A_i \cap A_{i+1}$ is path-connected for every i and so there is a path $g_i \colon I \to A_i \cap A_{i+1}$ from x_0 to $f(s_i)$. Here we used that $x_0 \in A_\alpha$ for every $\alpha \in A$. This way, we have

$$f = f_1 \cdot f_2 \cdots f_m \simeq (f_1 \cdot \overline{g}_1)(g_1 f_2 \overline{g}_2) \cdots (g_{m-1} f_{m-1}) =: h_1 \cdot h_2 \cdots h_m$$

with h_i a loop in A_i. This concludes the proof of the surjectivity of Φ.

We now prove the second assertion.[14] Namely, we prove that if $A_\alpha \cap A_\beta \cap A_\gamma$ is path connected for every α, β, γ, then $\ker(\Phi)$ equals the normal subgroup N generated by all elements of the form $i_{\alpha\beta}(x) i_{\beta\alpha}(x)^{-1}$ for $x \in \pi_1(A_\alpha \cap A_\beta, x_0)$, where $\alpha, \beta \in A$. In fact, the inclusion $N \subset \ker(\Phi)$ is not overly challenging as, for $w \in \pi_1(A_\alpha \cap A_\beta, x_0)$, we have

$$(j_\alpha \circ i_{\alpha\beta})(w))(j_\beta \circ i_{\beta\alpha})(w))^{-1} = j_\beta \circ i_{\beta\alpha}(w)(j_\beta \circ i_{\beta\alpha}(w))^{-1} = 1.$$

The real challenge is therefore to prove the inclusion $N \supset \ker(\Phi)$. Equivalently, one must prove that the induced homomorphism $\overline{\Phi} \colon *_{\alpha \in A} \pi_1(A_\alpha, x_0)/N \twoheadrightarrow \pi_1(X, x_0)$ is injective. To achieve this, we introduce some terminology.

- A *factorisation* of $[f] \in \pi_1(X, x_0)$ is a formal product of the form $[f_1] \cdots [f_m]$ such that
  1. f_i is a loop in some A_α based at x_0 and $[f_i] \in \pi_1(A_\alpha, x_0)$ for $i = 1, \ldots, m$.
  2. f is homotopic to $f_1, \ldots, f_m$ in X.

  Put differently, a factorisation of $[f]$ is a possibly non-reduced word in $*_{\alpha \in A} \pi_1(A_\alpha, x_0)$ whose image by Φ is $[f]$. The surjectivity of Φ implies that every $[f] \in \pi_1(X, x_0)$ admits a factorisation.

- Two factorisations are *equivalent* if they are related by a finite number of the following moves and their inverses:
  1. if $[f_i]$ and $[f_{i+1}]$ are in the same $\pi_1(A_\alpha, x_0)$, replace $[f_i][f_{i+1}]$ by $[f_i \cdot f_{i+1}]$;
  2. if f_i is a loop in $A_\alpha \cap A_\beta$, and one is considering $[f_i]$ as an element of $\pi_1(A_\alpha, x_0)$, then consider it as an element of $\pi_1(A_\beta, x_0)$ instead.

  Observe that the homomorphism $\overline{\Phi} \colon *_{\alpha \in A} \pi_1(A_\alpha, x_0)/N \twoheadrightarrow \pi_1(X, x_0)$ is injective if and only if two factorisations of any $[f] \in \pi_1(X, x_0)$ are equivalent.

Given an arbitrary $[f] \in \pi_1(X, x_0)$, we prove that any two factorisations of $[f]$ are equivalent, as this will therefore conclude the proof of the theorem. Let $[f_1] \cdots [f_k]$ and $[f_1'] \cdots [f_l']$ be factorisations of $[f]$. In particular, we can pick a homotopy $F \colon I \times I \to X$ between $f_1 \cdots f_k$ and $f_1' \cdots f_l'$. Using a compactness argument, one can find subdivisions $0 = s_0 < s_1 < \ldots < s_m = 1$ and $0 = t_0 < t_1 < \ldots < t_n = 1$ so that the rectangle $[s_{i-1}, s_i] \times [t_{j-1}, s_j]$ satisfies $F([s_{i-1}, s_i] \times [t_{j-1}, s_j]) \subset A_\alpha$ for some α. One can furthermore assume that

- the s subdvision of $[0, 1]$ is a subdivision of the partitions defined by the products $f_1 \cdots f_k$ and $f_1' \cdots f_l'$; see Figure 2.15.
- one can assume that $n \geq 3$ is odd.
- after perturbing the vertical sides of some of the rectangles $[s_{i-1}, s_i] \times [t_{j-1}, s_j]$, we can arrange that each point of $I \times I$ belongs to at most three rectangles. We call these new rectangles $[s_{i-1}, s_i] \times [t_{j-1}, s_j]$ and still arrange that $F(R_{ij}) \subset A_\alpha$ for some α; see Figure 2.16

[14]Depending on the year, we might not prove the second assertion during class.

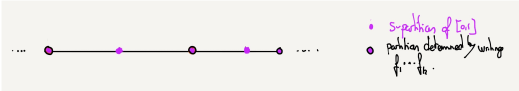


Figure 2.15: One can assume that the s-partition of $[0, 1]$ is a subpartition of the partition of $[0, 1]$ obtained by writing $f_1 \cdots f_k$.

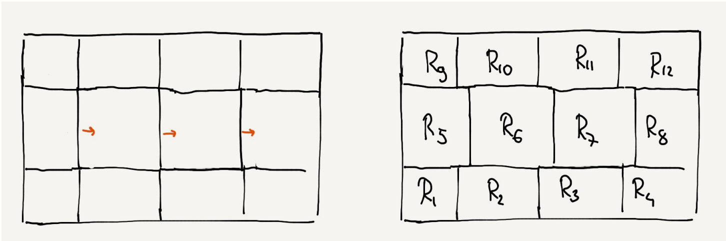


Figure 2.16: Each point of $I \times I$ belongs to at most three rectangles

For $r = 0, \ldots, mn$, we now let γ_r be the path in $I \times I$ that separates the rectangles $R_1, \ldots, R_r$ from the rectangles $R_{r+1}, \ldots, R_m$, as illustrated in Figure 2.17. This way, the composition $F \circ \gamma_r$ is a loop in X based at x_0. Here is now the plan:

1. We associate a factorisation to each $[F \circ \gamma_r]$; this factorisation depends on certain choices but the dependence goes away once we consider the factorisation up to equivalence.

2. We prove that independently of the aforementioned choices, the factorisation for $[F \circ \gamma_r]$ is equivalent to the factorisation for $[F \circ \gamma_{r+1}]$ for each r.

3. We show that the factorisation $[f_1] \cdots [f_k]$ is equivalent to the factorisation for $[F \circ \gamma_0]$, again independently of the choices that lead to the factorisation $[F \circ \gamma_0]$. The same argument (which we omit) proves the factorisation $[f'_1] \cdots [f'_r]$ is equivalent to the factorisation for $[F \circ \gamma_{mn}]$.

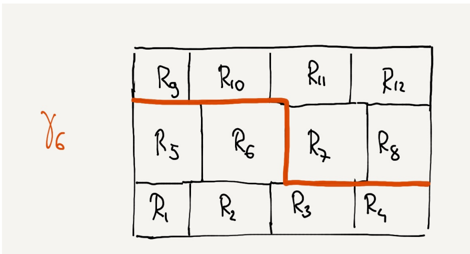


Figure 2.17: The path γ_6 that separates the rectangles $R_1, \ldots, R_6$ from the rectangles $R_7, \ldots, R_{12}$.

We carry out the first step. We show how to associate a factorisation to $[F \circ \gamma_r]$ for every $r = 0, \ldots, mn$. Consider the vertices v that the path γ_r encounters. By construction, we know that $F(v) \in A_\alpha \cap A_\beta \cap A_\gamma$ for some α, β, γ. Since this triple intersection is path-connected, there is a path g_v from $F(v)$ to x_0. We now break up γ into paths from one vertex to the next, i.e. $\gamma_r = \delta_1 \cdots \delta_z$, as illustrated in Figure 2.18. Using the definition of equivalence, we now obtain

$$\begin{aligned}[F \circ \gamma_r] &= [((F \circ \delta_1) \cdot \overline{g}_{v_1}) \cdot (g_{v_1} \cdot (F \circ \delta_2) \cdot \overline{g}_{v_2}) \cdots (g_{v_z} \cdot (F \circ \delta_z))] \\ &= [(F \circ \delta_1) \cdot \overline{g}_{v_1}] \cdot [g_{v_1} \cdot (F \circ \delta_2) \cdot \overline{g}_{v_2})] \cdots [g_{v_z} \cdot (F \circ \delta_z))].\end{aligned}$$

Here observe that each defines a loop in $A_\alpha \cap A_\beta$ and it is the choice of thinking of the class of this loop as either lying in $\pi_1(A_\alpha)$ or in $\pi_1(A_\beta)$ that defines a factorisation $[F \circ \gamma_r]$. By definition however, different choices lead to equivalent factorisations; recall the second move. The process is illustrated in Figure 2.18.

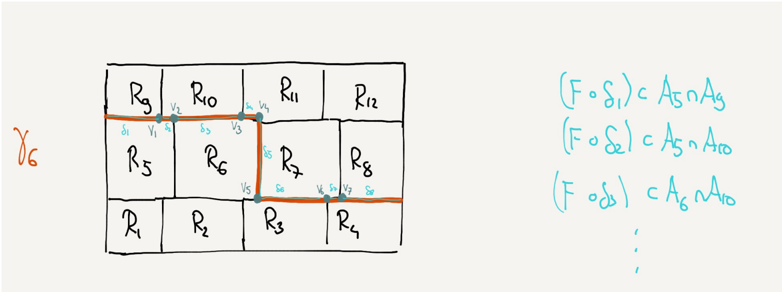


Figure 2.18: We break up the path γ_6 as $\gamma_6 = \delta_1 \cdots \delta_8$.

We now prove that regardless of the choices made above, the factorisation of $[F \circ \gamma_r]$ is equivalent to the factorisation of $[F \circ \gamma_{r+1}]$ for $r = 0, \ldots, mn-1$. For each r, the paths γ_r and γ_{r+1} differ only on the edges of the rectangle R_{r+1}: indeed γ_{r+1} takes the right then down route, while γ_r takes the down then left route, as illustrated in Figure 2.19. Using move 2 repeatedly does the trick: when the paths agree, we can just use move 2 to assume that the loops belong to the same A_α, while when the paths disagree with use move 2 together with a homotopy over the rectangle R_{r+1}.

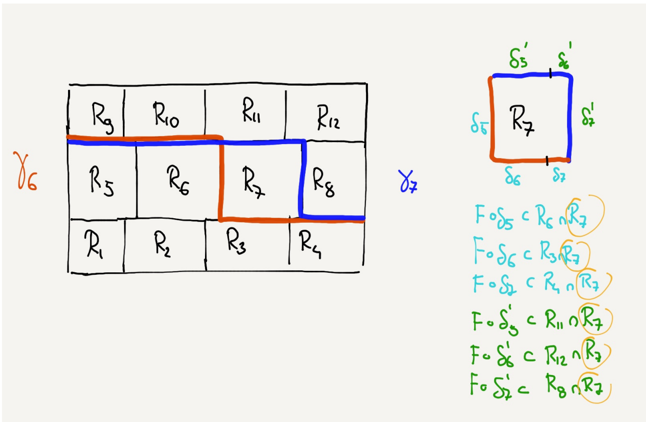


Figure 2.19: The factorisation of $[F \circ \gamma_r]$ is equivalent to the factorisation of $[F \circ \gamma_{r+1}]$

We carry out the third step[15] we prove that regardless of the choices mentioned above, the factorisation of $[F \circ \gamma_0]$ is equivalent to the factorisation $[f] = [f_1] \cdots [f_k]$; the proof that the factorisation of $[F \circ \gamma_{mn+1}]$ is equivalent to the factorisation $[f] = [f_1'] \cdots [f_l']$ is similar as so we omit it. The idea is to go from the factorisation $[f_1] \cdots [f_k]$ to a factorisation of $[F \circ \gamma_0]$ using our two moves. We chose the s-subdivision of $I \times I$ so that it is a subdivision of the partition arising from the composition $[f_1] \cdots [f_k]$. A vertex v that does not belong to this partition separates R_j from R_{j+1} for some j, say with $F(R_j) \subset A_j$ and $F(R_{j+1}) \subset A_{j+1}$; see Figure. With this notation, we have $F(v) \in A_j \cap A_{j+1} \cap A_v$. Since $A_j \cap A_{j+1} \cap A_v$ is path-connected, we again pick a path g_v from $F(v)$ to x_0. Starting from $[f_1] \cdots [f_k]$, we decompose each f_i as a product of paths δ_i between the vertices v of the s-subdivision, insert the g_v-paths when needed leading to a factorisation of $[F \circ \gamma_0]$ (we used the first of the two moves); see Figure. This factorisation might not agree with with our previously chosen factorisation of $[F \circ \gamma_0]$, but it will be equivalent to it, as we already explained.

[15]In class, we omitted the proof of the third step, but it is included here in case some readers are curious.

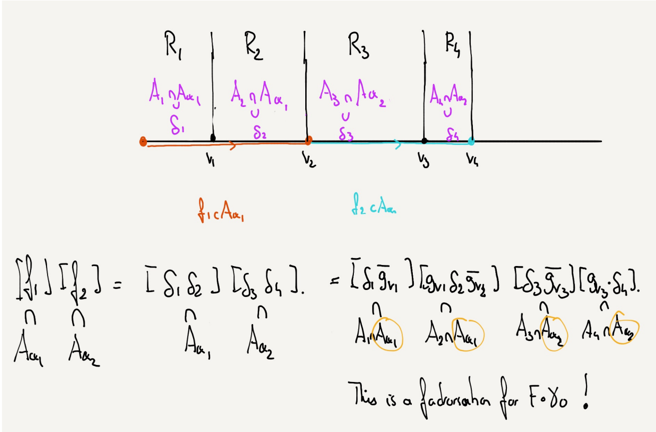


Figure 2.20: The factorisation of $[F \circ \gamma_0]$ is equivalent to the factorisation $[f] = [f_1] \cdots [f_k]$.

This concludes the proof of the fact that $[f_1] \cdots [f_k]$ and $[f_1'] \cdots [f_l']$ are equivalent factorisations of $[f]$ and therefore proves that $\overline{\Phi}$ is injective which concludes the proof of the theorem. □

2.3 Covering spaces

The definition of a covering space is easier to motivate after the fact: so many examples can be given that it makes the underlying notion worthwhile to study. From a more pragmatic viewpoint, we implicitly used (without proof) an idea from the theory of covering spaces when we proved that $\pi_1(S^1) = \mathbb{Z}$ and so it is worth completing that proof and investigating its generalisations.

Additionally, the theory of covering spaces itself will have a very nice geometric application: given a (nice enough) space X it will allow us to associate a unique (up to an appropriate notion of homeomorphism) topological space X_H to each subgroup H of $\pi_1(X)$. In other words, understanding the covering spaces of X will help us understand the subgroups of $\pi_1(X)$. [16]

If time permits, we will also see a group theoretic application of this topological theory: every subgroup of a free group is free.

2.3.1 Definition and examples

In this subsection, we define the notion of a covering space and give multiple examples. References include [Mun00, Section 53] and [Hat02, Chapter 1.3].

Definition 2.51. A *cover* of a space X is space $\widetilde{X}$ together with a continuous map $p\colon \widetilde{X} \to X$ satisfying the following property: for every $x \in X$, there is an open set $U \subset X$ containing x such that $p^{-1}(U) = \bigsqcup_\alpha U_\alpha$, where $U_\alpha \subset \widetilde{X}$ is an open set such that $p|_{U_\alpha} : U_\alpha \to U$ is a homeomorphism.

The map p is often called the *covering map*, X is called the *base space*, $\widetilde{X}$ is called the *total space*, and the U_α are called the *sheets* of $\widetilde{X}$ over U; we also say that U is *evenly covered*.

Informally, the total space $\widetilde{X}$ lies "above" the base space X and locally, $\widetilde{X}$ looks like a disjoint union of copies of X, though globally that may not be true. The way we intuitively think of cover spaces is illustrated in Figure 2.21.

[16] Here is a an additional motivation from knot theory. Suppose we want to distinguish two knots $K, K' \subset S^3$, i.e. prove they are not isotopic. In Active learning 2.50, we learnt that this can be done by proving that the knot groups $\pi_1(S^3 \setminus K)$ and $\pi_1(S^3 \setminus K')$ are not isomorphic. Unfortunately, in practice this is difficult and so instead, one studies the subgroups of the knot group, i.e. the covering spaces of $S^3 \setminus K$. Summarising, covering spaces can be used to tell knots apart...but this is outside the scope of this class.

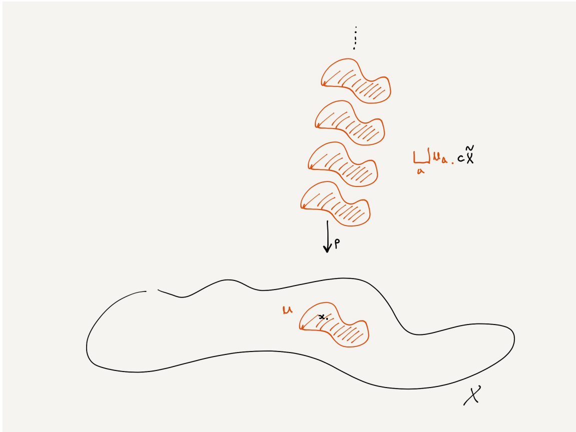


Figure 2.21: This is an informal picture illustrating how we think of covering spaces: locally $\widetilde{X}$ looks like a disjoint union of copies of X. The slogan is that a covering space locally looks like a stack of pancakes.

Example 2.52. Here are some examples of covering spaces.

1. The *trivial cover* of a space X is $\mathrm{id}_X \colon X \to X$. Disjoint unions of copies of X also form a cover of X: map each disjoint copy of X to X using the identity.
2. The map $\mathbb{C} \to \mathbb{C}, z \mapsto z^n$ induces a map $p_n \colon S^1 \to S^1$ that is a covering map. For ease of notation, we write $\mathbb{C} \cong \mathbb{R}^2$ and think of p_n as the map $p_n(\cos(2\pi\theta), \sin(2\pi\theta)) = (\cos(2\pi n\theta), \sin(2\pi n\theta))$. Assume that $(x, y) \in S^1$ with $y > 0$ (the proof is similar for (x, y) belonging to other subsets of S^1). Observe that $U := S^1 \cap \mathbb{R}^2_{y>0} \ni (x, y)$ is an evenly covered open subset of S^1:
$$\begin{aligned} p_n^{-1}(U) &= \{(\cos(2\pi\theta), \sin(2\pi\theta)) \mid \sin(2\pi n\theta) > 0 \ \theta \in [0, 1)\} \\ &= \bigsqcup_{k=0}^{n-1} \left\{ (\cos(2\pi\theta), \sin(2\pi\theta)) \mid \theta \in \left(\frac{2k}{2n}, \frac{2k+1}{2n} \right) \right\} \\ &=: \bigsqcup_{k=0}^{n} U_k. \end{aligned}$$
A direct verification shows that p_n restricts to a homeomorphism $U_k \xrightarrow{\cong} U$. An illustration of the situation for $n = 2$ can be found in Figure 2.22.
3. The same proof shows that the exponential map $\exp \colon \mathbb{R} \to S^1, \theta \mapsto e^{2\pi i\theta}$ is a covering map; again an illustration can be found in Figure 2.22.
4. On the eleventh problem set, we will show that if $p_X \colon \widetilde{X} \to X$ and $p_Y \colon \widetilde{Y} \to Y$ are covering spaces, then so is $p_X \times p_Y \colon \colon \widetilde{X} \times \widetilde{Y} \to X \times Y$. In particular, note that $\exp \times \exp \colon \mathbb{R}^2 \to S^1 \times S^1$ is a covering space of the torus and so is the infinite cylinder $\mathbb{R} \times S^1 \to S^1 \times S^1$.

Our aim is now to understand these examples from a different perspective and to obtain even more examples.

Definition 2.53.

1. A group G *acts* on a set Y if there is a map $\cdot \colon G \times Y \to Y$ such that $e_G \cdot x = x$ for every $x \in Y$ and $g \cdot (h \cdot x) = (gh) \cdot x$ for every $x \in Y$ and every $g, h \in G$.

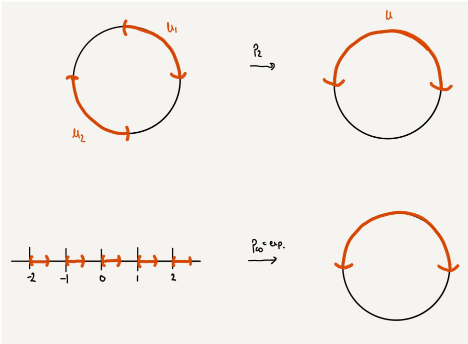


Figure 2.22: Illustration of the cover $p_2 \colon S^1 \to S^1$ and $\exp \colon \mathbb{Z} \to S^1$.

2. A group G *acts by homeomorphisms* on a topological space Y if G acts on Y as a set and the map $g\cdot \colon Y \to Y, x \mapsto g \cdot x$ is a homeomorphism for every $g \in G$.

3. If G acts on Y by homeomorphisms, then the *orbit space* of this action is the quotient space $Y/G = Y/\sim$ where $x \sim y$ if and only if $y = gx$ for some $g \in G$.

Example 2.54. We describe examples of groups acting by homeomorphisms.

1. The finite cyclic group $\mathbb{Z}_n$ acts by homeomorphisms on S^1 by $k \cdot z = e^{2\pi i k/n} z$; i.e. by appropriate rotations. A quick verification shows that $z \mapsto z^n$ descends to a homeomorphism $S^1/\mathbb{Z}_n \xrightarrow{\cong} S^1$ and that the composition $S^1 \xrightarrow{\pi} S^1/\mathbb{Z}_n \cong S^1$ coincides with the map p_n described in Example 2.52; in particular the projection $\pi \colon S^1 \to S^1/\mathbb{Z}_n \cong S^1$ is a covering map.

2. The infinite cyclic group $\mathbb{Z}$ acts by homeomorphisms on $\mathbb{R}$ by $k \cdot x = x + k$, i.e. by translations. A quick verification shows that $t \mapsto e^{2i\pi t}$ descends to a homeomorphism $\mathbb{R}/\mathbb{Z} \xrightarrow{\cong} S^1$ and that the composition $\mathbb{R} \xrightarrow{\pi} \mathbb{R}/\mathbb{Z} \cong S^1$ coincides with the map exp described in Example 2.52; in particular the projection $\pi \colon \mathbb{R} \to \mathbb{R}/\mathbb{Z} \cong S^1$ is a covering map.

3. The group $\mathbb{Z}^2$ acts by homeomorphisms on $\mathbb{R}^2$ by $(k, \ell) \cdot (x, y) = (x + k, y + \ell)$, i.e. by translations; see Figure 2.23 A quick verification shows that $(t, s) \mapsto (e^{2i\pi t}, e^{2i\pi s})$ descends to a homeomorphism $\mathbb{R}^2/\mathbb{Z}^2 \xrightarrow{\cong} S^1 \times S^1$ and that the composition $\mathbb{R}^2 \xrightarrow{\pi} \mathbb{R}^2/\mathbb{Z}^2 \cong S^1 \times S^1$ coincides with the map $\exp \times \exp$ described in Example 2.52; in particular the projection $\pi \colon \mathbb{R}^2 \to \mathbb{R}^2/\mathbb{Z}^2 \cong S^1 \times S^1$ is a covering map.

4. The group $\mathbb{Z}_2 = \{\pm 1\}$ acts by homeomorphisms on the n-sphere $S^n \subset \mathbb{R}^{n+1}$ by the antipodal action $-1 \cdot x = -x$. By definition, real projective space $\mathbb{R}P^n = S^n/\mathbb{Z}_2$ is the orbit space of this action. Proposition 2.56 shows that the projection $\pi \colon S^n \to S^n/\mathbb{Z}_2 \cong \mathbb{R}P^n$ is a covering map.

5. More examples with fewer formulas are illustrated in Figure 2.23.

We now explain how to verify that the projection $Y \to Y/G$ is a covering map, under a mild condition on the action.

Definition 2.55. An action of a group G on a space Y is *properly discontinuous* if for every $y \in Y$ there exists an open set $U \subset Y$ containing y such that $g(U) \cap U = \emptyset$ for every $g \in G \setminus \{e_G\}$.

It can be verified that each of the actions by homeomorphisms described in Example 2.54 is properly discontinuous. Note that if the action is by homeomorphisms, then the condition in Definition 2.55 can equivalently be written as $g(U) \cap h(U) = \emptyset$ for every distinct $g, h \in G$. The next proposition shows that many covering spaces arise by using orbit spaces of group actions.

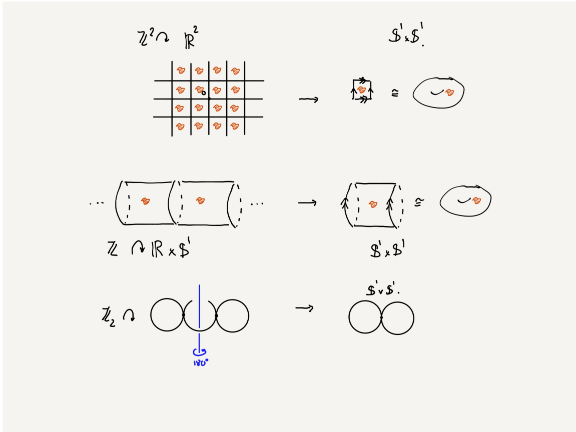


Figure 2.23: $\mathbb{Z}^2$ acts on $\mathbb{R}^2$ with orbit space $S^1 \times S^1$, $\mathbb{Z}$ acts on the infinite cylinder $\mathbb{R} \times S^1$ with orbit space $S^1 \times S^1$ and the last picture depicts the action of $\mathbb{Z}_2$ on a graph with orbit space $S^1 \vee S^1$.

Proposition 2.56. *Assume that a group G acts on a space Y properly discontinuously by homeomorphisms. Then the projection map $\pi\colon Y \to Y/G$ is a covering map.*

Proof. We claim that π is open. Given an open subset $U \subset Y$, we must show that $\pi(U) \subset Y/G$ is open or equivalently show that $\pi^{-1}(\pi(U)) \subset Y$ is open. This follows from the fact that

$$\pi^{-1}(\pi(U)) = \{y \in Y \mid \pi(y) \in \pi(U)\} = \bigcup_{g \in G} g(U) \tag{2.1}$$

is open because U is open and G acts by homeomorphisms.

Next, for every $\pi(y) \in Y/G$, we must find an open set $V \subset Y/G$ containing $\pi(y)$ such that V is evenly covered. Since $y \in Y$ and the action is properly discontinuous, there is an open set $U \subset Y$ containing y such that $g_0(U) \cap g_1(U) = \emptyset$ for every distinct $g_0, g_1 \in G$. We now set $V := \pi(U)$ and check it is evenly covered. Using (2.1) and the property of U, we see that $\pi^{-1}(V) = \bigsqcup_{g \in G} g(U)$. Next a quick verification shows that $\pi|_{g(U)}\colon g(U) \to \pi(U)$ is a homeomorphism for every $g \in G$. This concludes the proof that π is a covering map. □

As we mentioned, above, each of the actions in Example 2.54, the actions is properly discontinuous and so Proposition 2.56 implies that the corresponding quotient maps $Y \to Y/G$ are coverings.

Remark 2.57. Here are some remarks concerning covering spaces:[17]

1. Covering maps are local homeomorphisms; here recall that a map $f\colon Y \to Z$ is a *local homeomorphism* if for every $y \in Y$, there exists an open set $V \subset Y$ containing y such that $f(V) \subset Z$ is open an $f|_V\colon V \to f(V)$ is a homeomorphism. We now prove that a covering map $p\colon \widetilde{X} \to X$ is a local homeomorphim. If $\widetilde{x} \in \widetilde{X}$, then $x := p(\widetilde{x})$ is contained in an open set U that is evenly covered, $p^{-1}(U) = \bigsqcup_\alpha \widetilde{U}_\alpha$. We can now take V to be the open set $\widetilde{U}_\alpha$ that contains $\widetilde{x}$: by definition $p|_{\widetilde{U}_\alpha}\colon \widetilde{U}_\alpha \to U$ is a homeomorphism.

2. The cover $p\colon \widetilde{M} \to M$ of an topological n-manifold M is a topological n-manifold; in particular covering spaces of surfaces are surfaces. This follows from the following facts:

   (a) If M is locally homeomorphic to $\mathbb{R}^n$, then $\widetilde{M}$ is locally homeomorphic to $\mathbb{R}^n$. This essentially follows because covering maps are local homeomorphisms and being locally homeomorphic to $\mathbb{R}^n$ is a local property.

[17] This remark was not covered during class, but I'll leave it here for the curious reader.

(b) If M is Hausdorff, then so is $\widetilde{M}$. This is an exercise on the eleventh problem set.

(c) If M admits a countable basis, then so does $\widetilde{M}$; the proof is omitted.

3. The cover of an n-dimensional CW complex is an n-dimensional CW-complex; in particular any covering space of a graph is a graph; the proof is omitted.

2.3.2 Lifting properties

Informally, an idea of covering space theory is that the total space often has an easier fundamental group than the base space. Thus, very informally the idea is that questions concerning the topology of X "lift" it to easier questions in $\widetilde{X}$. To make this more precise, we recall the definition of a lift that was briefly mentioned during the proof of Theorem 2.15.

Definition 2.58. A *lift* of a continunous map $f\colon Y \to X$ is a continuous map $\widetilde{f}\colon Y \to \widetilde{X}$ satisfying $p \circ \widetilde{f} = f$.

The goal of this subsection is to show that paths in X can always be lifted to paths in $\widetilde{X}$, and similarly for homotopies. In particular, we will prove the two unproved facts that we used in Theorem 2.15 when we showed that $\pi_1(S^1) = \mathbb{Z}$. References include [Mun00, Section 53] and [Hat02, Chapter 1.3].

We start off with an example to illustrate the concept of a lift.

Example 2.59. Consider the infinite cylinder $\widetilde{X} = \mathbb{R} \times S^1$ which covers the torus $X = S^1 \times S^1$ and the map $f\colon I \to X$ with image the loop $\mu \subset T$ at x_0 illustrated in Figure 2.24. This figure shows an infinite number of lifts of f, but there is only one lift that start at the point $\widetilde{x}_0$.

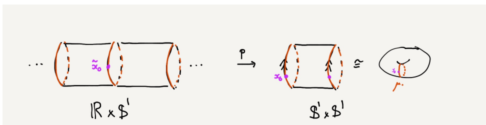


Figure 2.24: The left hand side of the figure shows lifts of the path $f\colon I \to X = S^1 \times S^1$ (with image μ) starting at x_0 to the cover $\widetilde{X} = \mathbb{R} \times S^1$. There are infinitely many such lifts, but only one that starts at $\widetilde{x}_0$.

The following *path lifting property* generalizes Example 2.59 by showing that any path in a space X can be lifted to a cover $\widetilde{X}$. As one also guess from this example, there is a uniqueness if the path is required to start at a fixed point in the cover. In fact, a similar result also holds for homotopies.

Proposition 2.60. *Let $p\colon \widetilde{X} \to X$ be a cover, let $x_0 \in X$ and fix a point $\widetilde{x}_0 \in \widetilde{X}$ with $p(\widetilde{x}_0) = x_0$.*

1. *Every path $f\colon [0,1] \to X$ starting at x_0 lifts to a unique path $\widetilde{f}\colon [0,1] \to \widetilde{X}$ starting at $\widetilde{x}_0$.*
2. *For every path homotopy $f_t\colon [0,1] \to X$, with the path f_t starting at x_0, there exists a unique path homotopy $\widetilde{f}_t\colon [0,1] \to \widetilde{X}$ lifting f_t with $f_t(0) = \widetilde{x}_0$ for every $t \in [0,1]$.*

For $X = S^1$ and $\widetilde{X} = \mathbb{R}$, Proposition 2.60 contains the two unproved facts that we used in Theorem 2.15 to show that $\pi_1(S^1) = \mathbb{Z}$. In fact, Proposition 2.60 will follow from the following more general homotopy lifting property.

Proposition 2.61. *Let $p\colon \widetilde{X} \to X$ be a cover. Given a continuous map $F\colon Y \times I \to X$ and a lift $\widetilde{F}_0\colon Y \times \{0\} \to \widetilde{X}$ of f_0, there exists a unique map $\widetilde{F}\colon Y \times I \to \widetilde{X}$ such that $\widetilde{F}(-,t)$ lifts $F(-,t)$ for every $t \in [0,1]$ and $\widetilde{F}(-,0) = \widetilde{F}_0$.*[18]

[18] The notation in the statement is convenient for the proof. However, it is more helpful to think of the proposition as saying that given a homotopy $f_t\colon Y \to X$ and a lift $g\colon Y \to \widetilde{X}$ of f_0, there exists a unique homotopy $\widetilde{f}_t$ lifting f_t such that $\widetilde{f}_0 = g$.

Before proving Proposition 2.61, we first describe how it recovers Proposition 2.60.

Proof of Proposition 2.60 assuming Proposition 2.61. The first item follows by taking $Y = \{*\}$ in Proposition 2.61; we therefore focus on the second item. Let $f_t\colon I \to X$ be a homotopy with $f_t(0) = x_0$ and $\widetilde{x}_0 \in p^{-1}(x_0)$. This defines a map $F\colon I \times I \to X$ by $F(x,t) = f_t(x)$. Using the first item, we have a unique lift $\widetilde{F}_0\colon I \times \{0\} \to \widetilde{X}$ with $f_0 = \widetilde{F}|_{I\times\{0\}}$ and $\widetilde{F}_0(0) = \widetilde{x}_0$. Applying Proposition 2.61 with $Y = I$, we have a unique $\widetilde{F}\colon I \times I \to \mathbb{R}$ such that $p \circ \widetilde{F} = F$ and $\widetilde{F}|_{I\times\{0\}} = \widetilde{F}_0$. Next note that $\widetilde{F}|_{\{0\}\times I}$ (resp. $\widetilde{F}|_{\{1\}\times I}$) are two paths that lift the constant path c_{x_0} so by the uniqueness part of the first statement, they are constant paths $\widetilde{F}(0,-) \equiv c_{\widetilde{x}_0}$ and $\widetilde{F}(1,-) \equiv c_{\widetilde{x}_0'}$. By setting $\widetilde{f}_t(x) = \widetilde{F}(x,t)$, we have found the homotopy $\widetilde{f}_t$ required by the second item. □

We now prove the more general result.

Proof of Proposition 2.61. Here is the plan of the proof.

- Step 1: We prove that for every $y_0 \in Y$, there is an open set $N \subset Y$ containing y_0 and a lift $\widetilde{F}\colon N \times I \to \widetilde{X}$ of $F\colon N \times I \to X$.
- Step 2: We prove the uniqueness part of the proposition when $Y = \{*\}$ is a point.

We now explain how these two steps allow us to conclude the proof of the proposition. By the first step, for every $y_0 \in Y$, there exists an open set $N \subset Y$ containing y_0 and a continuous map $\widetilde{F}\colon N \times I \to \widetilde{X}$ such that $p \circ \widetilde{F} = F|_{N\times I}$. We define $\widetilde{F}\colon Y \times I \to \widetilde{X}$ using each of these $\widetilde{F}$. By uniqueness of $\widetilde{F}|_{\{y\}\times I}$, they agree on all intersections and so $\widetilde{F}$ is well defined. The map is continuous because it is defined by continuous maps on an open cover of $Y \times I$. Finally, $\widetilde{F}$ is unique because it is unique on each of the $\{y\} \times I$.

We now carry out our plan.

- Step 1: We must prove that for every $y_0 \in Y$, there is a neighborhood $N \subset Y$ of y_0 and a lift $\widetilde{F}\colon N \times I \to \widetilde{X}$ of $F\colon N \times I \to X$.
  - Step 1.1: Given $y_0 \in Y$, we first construct the neighborhood $N \subset Y$ of y_0.
    For every $t \in [0,1]$, since we have $F(y_0,t) \in X = \bigcup_\alpha U_\alpha$ there exists an α with $F(y_0,t) \in U_\alpha$. Since $F^{-1}(U_\alpha) \subset Y \times I$ is an open set containing (y_0,t), by definition of the product topology, there exists an open set $y_0 \in N_t \subset Y$ and an open interval $t \in (a_t,b_t) \subset I$ with $F(N_t \times (a_t,b_t)) \subset U_\alpha$ and $(y_0,t) \in N_t \times (a_t,b_t)$; at the endpoints we take half open intervals instead. It follows that the family $(N_0 \times [0,b_0)) \cup \{N_t \times (a_t,b_t)\}_{t\in(0,1)} \cup (N_1 \times (a_1,1])$ forms an open cover of the compact space $\{y_0\} \times I$ and so it admits a finite open subcover $(N_0 \times [0,b_0)) \cup \{N_i \times (a_i,b_i)\}_{i=1}^{m-1} \cup (N_1 \times (a_m,1])$. Picking $t_i \in [a_i, b_{i-1}]$, we get a partition $0 = t_0 < t_1 < \ldots < t_m = 1$ of $[0,1]$ and we set
    $$N := \bigcap_{i=1}^{m} N_i.$$
    By construction, $N \subset Y$ is an open set containing $y_0 \in Y$ with $F(N \times [t_i,t_{i+1}]) \subset U_\alpha =: U_i$ for every $i = 0,1,\ldots,m-1$.
  - Step 1.2: We build $\widetilde{F}\colon N \times I \to \widetilde{X}$ with $p \circ \widetilde{F} = F|_{N\times I}$ and $\widetilde{F}_0 = \widetilde{F}|_{N\times\{0\}} = F|_{N\times[0,t_0]}$.
    By induction, we construct $\widetilde{F}$ on $N \times [0,t_i]$ for $i = 0,\ldots,m$: once we get to $i = m$, we will have achieved the goal of the first step. The basis of the induction is clear: by assumption, we already have a $\widetilde{F}\colon Y \times \{0\} \to \widetilde{X}$ lifting F and we can restrict it to $N \times \{0\}$. For the induction step, we have $\widetilde{F}\colon N \times [0,t_i] \to \widetilde{X}$ and we want to extend it to $N \times [0,t_{i+1}]$; so it suffices to define $\widetilde{F}$ on $N \times [t_i,t_{i+1}]$. By induction, $\widetilde{F}$ is already

defined on (y_0, t_i) and by construction of N, we have $p \circ \widetilde{F}(y_0, t_i) = F(y_0, t_i) \in U_i$. By definition of a covering space, there exists an open set $\widetilde{U}_i \subset \widetilde{X}$ such that $\widetilde{F}(y_0, t_i) \in \widetilde{U}_i$ and $p| \colon \widetilde{U}_i \to U_i$ is a homeomorphism. We can assume that $\widetilde{F}(N \times \{t_i\}) \subset \widetilde{U}_i$ (if necessary, make N smaller by replacing it by $N \times \{t_i\}$ with $(N \times \{t_i\}) \cap (\widetilde{F}|_{N \times \{t_i\}})^{-1}(\widetilde{U}_i))$. We then conclude the induction and therefore the first step by setting

$$\widetilde{F}|_{N \times [t_i, t_{i+1}]} := p^{-1}|_{U_i} \circ F|_{N \times [t_i, t_{i+1}]}.$$

- Step 2: We prove the uniqueness statement in the case where $Y = \{*\}$ is a point.

  In other words, given continuous maps $\widetilde{F}, \widetilde{F}' \colon I \to \widetilde{X}$ that lift $F \colon I \to X$ and with $\widetilde{F}(0) = \widetilde{F}'(0)$, we must show that $\widetilde{F} \equiv \widetilde{F}'$. As in the first step, we can find a partition $0 = t_0 < t_1 < \ldots < t_m = 1$ of $[0,1]$ so that $F([t_i, t_{i+1}]) \subset U_i$. We now prove the statement by induction: we show that $\widetilde{F}|_{[0,t_i]} = \widetilde{F}'|_{[0,t_i]}$ for $i = 0, \ldots, m$ (the case $i = m$ concludes the second step). For $i = 0$ this is our assumption that $\widetilde{F}(0) = \widetilde{F}'(0)$. We carry out the induction step by showing that $\widetilde{F}|_{[t_i,t_{i+1}]} \equiv \widetilde{F}'_{[t_i,t_{i+1}]}$. The inclusion $F([t_i, t_{i+1}]) \subset U_i$ implies that $\widetilde{F}([t_i, t_{i+1}]) \subset p^{-1}(U_i) = \bigsqcup_\beta \widetilde{U}_{i_\beta}$, and similarly for $\widetilde{F}'$. As $[t_i, t_{i+1}]$ is connected, we have $\widetilde{F}([t_i, t_{i+1}]) \subset \widetilde{U}_{i_\beta}$ and $\widetilde{F}'([t_i, t_{i+1}]) \subset \widetilde{U}_{i_{\beta'}}$. Since, by induction $\widetilde{F}(t_i) = \widetilde{F}'(t_i)$, we deduce that $\beta = \beta'$. We therefore have $p|_{\widetilde{U}_{i_\beta}} \circ \widetilde{F} = p|_{\widetilde{U}_{i_\beta}} \circ \widetilde{F}' \colon [t_i, t_{i+1}] \to \widetilde{U}_{i_\beta} \cong U_i$ and since $p|_{\widetilde{U}_{i_\beta}} \colon \widetilde{U}_{i_\beta} \to U_i$ is a homeomorphism (in particular it is injective), we deduce that $\widetilde{F}_{[t_i,t_{i+1}]} \equiv \widetilde{F}'_{[t_i,t_{i+1}]}$. This concludes the induction and therefore the proof the second step.

We have therefore proved the two steps which we saw were enough to conclude the proof of the proposition. □

2.3.3 The subgroup associated to a covering space

Let X be a space and let $x_0 \in X$. In this subsection, we associate a subgroup of $\pi_1(X, x_0)$ to each covering space. of X. As we shall see in Proposition 2.63, this subgroup encodes the loops in the base space that lift to loops the total space. References include [Mun00, Section 53] and [Hat02, Chapter 1.3].

Given a space X and $x_0 \in X$, from now on, we will write "let $p \colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ be a covering space" instead of "let $p \colon \widetilde{X} \to X$ be a covering space, and let $\widetilde{x}_0 \in p^{-1}(x_0)$."

Definition 2.62. Given a space X and $x_0 \in X$, the *subgroup of $\pi_1(X, x_0)$ corresponding to the covering space $p \colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$* is the subgroup $p_*(\pi_1(\widetilde{X}, \widetilde{x}_0)) \leq \pi_1(X, x_0)$.

The next proposition describes the key properties of this subgroup.

Proposition 2.63. *Let $p \colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ be a cover.*

1. *The induced map $p_* \colon \pi_1(\widetilde{X}, \widetilde{x}_0) \to \pi_1(X, x_0)$ is injective and, in particular the $\pi_1(\widetilde{X}, \widetilde{x}_0)$ is isomorphic to its image, the subgroup $p_*(\pi_1(\widetilde{X}, \widetilde{x}_0)) \leq \pi_1(X, x_0)$.*
2. *The subgroup of $\pi_1(X, x_0)$ associated to the cover can be described as*

$$\{[\gamma] \in \pi_1(X, x_0) \mid \gamma \textit{ lifts to a loop } \widetilde{\gamma} \textit{ in } \widetilde{X} \textit{ based at } \widetilde{x}_0\}.$$

Proof. To prove the first assertion, given a loop $\widetilde{\gamma} \colon I \to \widetilde{X}$ based at $\widetilde{x}_0$, we assume that $p_*([\widetilde{\gamma}]) = 1 \in \pi_1(X, x_0)$ and prove that $[\widetilde{\gamma}] = 1 \in \pi_1(\widetilde{X}, \widetilde{x}_0)$. By assumption, we know that $p \circ \widetilde{\gamma}$ is homotopic to the constant path c_{x_0} in X.[19] Apply Proposition 2.60 to lift the homotopy, resulting in a homotopy from $\widetilde{\gamma}$ to the constant path $c_{\widetilde{x}_0}$, This shows that $[\widetilde{\gamma}]=1$, as required.

We prove the second assertion. By definition, $p_*(\pi_1(\widetilde{X},\widetilde{x}_0))$ consists of the $[\gamma]\in\pi_1(X,x_0)$ that can be written as $[p\circ\widetilde{\gamma}]$ for some loop $\widetilde{\gamma}\subset X$ based at $\widetilde{x}_0$. This means that γ is homotopic to $p\circ\widetilde{\gamma}$ which is a loop that lifts to a loop at $\widetilde{x}_0$ (namely $\widetilde{\gamma}$). Lifting this homotopy, we deduce that γ itself lifts to a loop at $\widetilde{x}_0$.[20]

□

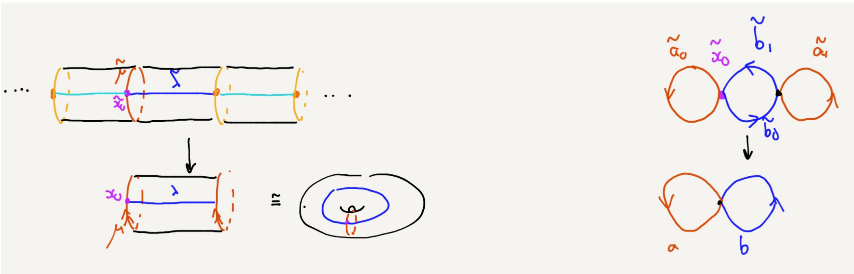


Figure 2.25: The covers mentioned in Example 2.64.

Example 2.64. We now describe the subgroup associated to various covering spaces.

1. For $p_n\colon S^1\to S^1, z\mapsto z^n$, the associated subgroup of $\pi_1(S^1)=\mathbb{Z}$ is $n\mathbb{Z}\leq\mathbb{Z}$, and for $\exp\colon\mathbb{R}\to S^1, t\mapsto e^{2\pi it}$, the associated subgroup of $\pi_1(S^1)=\mathbb{Z}$ is the trivial group $1\leq\mathbb{Z}$. Therefore we see that every subgroup of $\pi_1(S^1)=\mathbb{Z}$ arises as a covering space of S^1.

2. Consider the cover $\mathrm{id}\times\exp\colon S^1\times\mathbb{R}\to S^1\times S^1$ of the torus. We additionally consider the loops $\mu,\lambda\subset S^1\times S^1$ and the basepoint $x_0\in S^1\times S^1$ illustrated on the left hand side of Figure 2.25. With this notation, we have $\pi_1(S^1\times S^1,x_0)=\mathbb{Z}\mu\oplus\mathbb{Z}\lambda$ and the subgroup associated to the cover $\exp\times\mathrm{id}\colon\mathbb{R}\times S^1\to S^1\times S^1$ (also illustrated in Figure 2.25) is $H:=p_*(\mathbb{Z}[\widetilde{\mu}])=\mathbb{Z}[\mu]\oplus 0\leq\mathbb{Z}\mu\oplus\mathbb{Z}\lambda$.[21] As an illustration of Proposition 2.63, we see that λ does not lift to a loop and indeed $\lambda\notin H$.

3. Consider the cover $\widetilde{X}$ of $X=S^1\vee S^1$ illustrated on the right hand side of Figure 2.25. Using Active learning 2.42 or problem set 11, we see that $\pi_1(\widetilde{X},\widetilde{x}_0)=\langle\widetilde{a}_0,\widetilde{b}_0\widetilde{b}_1,\widetilde{b}_0\widetilde{a}_1\overline{\widetilde{b}_0}\rangle$. We deduce that the associated subgroup of $\pi_1(S^1\vee S^1,x_0)=F_2$ is $H:=\langle a,b^2,bab^{-1}\rangle$. As an illustration of Proposition 2.63, we see that b does not lift to a loop and indeed $b\notin H$.

Proposition 2.63 also leads to prove a fact that seemed intuitive in each of the examples we encountered so far: if $p\colon\widetilde{X}\to X$ is a covering space, then the preimage $p^{-1}(x)$ of every point $x\in X$ has the same cardinality. For later use, note that we call $p^{-1}(x)$ the *fiber* above x.

Proposition 2.65. *Every fiber of a covering space $p\colon(\widetilde{X},\widetilde{x}_0)\to(X,x_0)$, with $\widetilde{X}$ and X path-connected, has the same cardinality, namely the index of the subgroup $p_*(\pi_1(\widetilde{X},\widetilde{x}_0))$ in $\pi_1(X,x_0)$: for every x, we have*

$$|p^{-1}(x)|=[\pi_1(X,x_0)\colon p_*(\pi_1(\widetilde{X},\widetilde{x}_0))].$$

We call this quantity the degree *of the cover.*

[19] Here are some more details. Lifting the homotopy $p\circ\widetilde{\gamma}\simeq c_{x_0}$ gives a homotopy $\widetilde{\gamma}_t$ from $\widetilde{\gamma}$ to some $\widetilde{\gamma}_1$. Since $\widetilde{\gamma}$ is a loop at $\widetilde{x}_0$ and since homotopies are endpoint preserving $\widetilde{\gamma}_1$ is a loop at $\widetilde{x}_0$ lifting c_{x_0}. Since $\widetilde{\gamma}_1$ and $c_{\widetilde{x}_0}$ are two such lifts, they must be equal.

[20] Here are some more details: lifting the homotopy $\gamma\simeq p\circ\widetilde{\gamma}$ gives a homotopy $\widetilde{\gamma}_t$ from a lift $\widetilde{\gamma}_0$ of γ to some $\widetilde{\gamma}_1$. It follows that $\widetilde{\gamma}_1$ and $\widetilde{\gamma}$ are lifts of $p\circ\widetilde{\gamma}$ that start at $\widetilde{x}_0$ and therefore must be equal. Since $\widetilde{\gamma}$ is a loop at $\widetilde{x}_0$, it follows that $\widetilde{\gamma}_1$ and therefore $\widetilde{\gamma}_0$ are also loops at $\widetilde{x}_0$.

[21] When G and H are abelian, the custom is to write $G\oplus H$ instead of $G\times H$.

Proof. We first prove that $|p^{-1}(x)| = |p^{-1}(a)|$ for every $x, a \in X$ provided X is connected. To show this, we consider the set $\mathcal{O}_x = \{y \in X \mid |p^{-1}(y)| = |p^{-1}(x)|\}$ and aim to prove that $\mathcal{O}_x = X$ Since $\mathcal{O}_x$ is non-empty and X is connected, it suffices to show that $\mathcal{O}_x$ is both open and closed. To prove that $\mathcal{O}_x$ is open, we argue that it is a neighborhood of each of its points. If $y \in \mathcal{O}_x$, then there exists an evenly covered open set $U \subset X$ containing y. We argue that $U \subset \mathcal{O}_x$ (this will imply that $\mathcal{O}_x$ is a neighborhood of y). We must show that for every $z \in U$, we have $|p^{-1}(z)| = |p^{-1}(y)|$. This is because $p^{-1}(U) = \bigsqcup_{\alpha \in A} U_\alpha$ and $p|_{U_\alpha} : U_\alpha \to U$ is a homeomorphism: for each $\alpha \in A$, there exists a unique $y_\alpha, z_\alpha \in U_\alpha$ with $p(y_\alpha) = y$ and $p(z_\alpha) = z$; we deduce that $p^{-1}(z) = \{z_\alpha\}_{\alpha \in A}$ and $p^{-1}(y) = \{y_\alpha\}_{\alpha \in A}$ so $p^{-1}(z) \to p^{-1}(y), z_\alpha \mapsto y_\alpha$ is a bijection.[22] This concludes the proof that $\mathcal{O}_x$ is open; the proof that $\mathcal{O}_x$ is closed, i.e. that $X \setminus \mathcal{O}_x$ is open is similar and is left to the reader

We set $G := \pi_1(X, x_0)$ and $H := p_*(\pi_1(\widetilde{X}, \widetilde{x}_0))$. It now suffices to argue that $|p^{-1}(x_0)| = [G\colon H]$. To prove this, we define a bijection $\Phi\colon \pi_1(X, x_0)/H \to p^{-1}(x_0)$. First we define the map $\varphi\colon \pi_1(X, x_0) \to p^{-1}(x_0)$ by $[\gamma] \mapsto \widetilde{\gamma}(1)$, where $\widetilde{\gamma}$ is the lift of $\widetilde{\gamma}$ starting at $\widetilde{x}_0$. We now argue that φ descends to a bijection on the quotient.

We first show that φ descends to the quotient, i.e. that $\varphi(hg) = \varphi(g)$ for every $h \in H$.[23] To see this, note that $\varphi(hg) = \widetilde{hg}(1) = (\widetilde{h} \cdot \widetilde{g})(1) = \widetilde{g}(1)$, where in the last equality we use that $\widetilde{h}$ is a loop in $\widetilde{X}$ based at $\widetilde{x}_0$. Next, we prove that φ is surjective, which implies that Φ is also surjective. Given $\widetilde{x} \in p^{-1}(x_0)$, since $\widetilde{X}$ is path-connected, there exists a path $\widetilde{\gamma}\colon I \to \widetilde{X}$ from $\widetilde{x}_0$ to $\widetilde{x}$. Set $\gamma := p \circ \widetilde{\gamma}$ so that $\varphi(\gamma) = \widetilde{\gamma}(1) = \widetilde{x}$, proving surjectivity. Finally, we prove that Φ is injective. Assume that γ_0, γ_1 are such that $\widetilde{\gamma}_0(1) = \widetilde{\gamma}_1(1)$. This implies that $\gamma_0 \cdot \overline{\gamma}_1$ lifts to a loop at $\widetilde{x}_0$. In turn this is equivalent to $[\gamma_0][\gamma_1]^{-1} \in H$ (by Proposition 2.63) and this is equivalent to $[\gamma_0]$ and $[\gamma_1]$ being equal in the set G/H. This concludes the proof of the fact that Φ is a bijection. ☐

2.4 The classification of covering spaces

The objective is to show that under some mild assumptions on a space X, there is a bijection between isomorphism classes of path-connected covering spaces of X and subgroups of $\pi_1(X, x_0)$. In Subsection 2.4.1, we introduce the terminology needed to make sense of this result, state the theorem and give some examples. Subsections 2.4.2 and 2.4.3 are then devoted to the proof. It is worth mentioning that contrarily to the proof of van Kampen's theorem, in some proofs, we will not give all details in the interest of expositional clarity.

2.4.1 The classification theorem: statement and examples

We state the classification of covering spaces and give examples; references include [Mun00, Section 79] and [Hat02, Chapter 1.3]. First however, we make precise what we mean by an isomorphism of covering spaces and discuss the assumptions that we will require of X.

Terminology 2.66. The classification theorem will apply once we make some mild restrictions on the spaces we consider.

1. A space X is *locally path-connected* if for every $x \in X$ and every neighborhood $U \ni x$, there exists a path-connected neighorhood V with $x \in V \subset U$; we already saw this definition in Active learning session 1.75. Path-connectedness does not imply local path-connectedness.[24]

[22] Here is the proof I gave in class. Since X is path-connected, we pick a path γ from y to z and consider the assignment $p^{-1}(y) \to p^{-1}(z), \widetilde{y} \mapsto \gamma_{\widetilde{y}}(1)$; here $\gamma_{\widetilde{y}}$ is the unique path lifting γ and starting at $\widetilde{y}$. This map is injective: if $\gamma_{\widetilde{y}}(1) = \gamma_{\widetilde{y}'}(1)$, then $\overline{\gamma}_{\widetilde{y}}$ and $\overline{\gamma}_{\widetilde{y}'}$ are two paths starting at the same point and lifting $\overline{\gamma}$; they must therefore must be equal and so in particular $\widetilde{y} = \widetilde{y}'$. The same argument shows that $p^{-1}(z)$ injects into $p^{-1}(y)$ and so the sets are in bijection.

[23] Recall that the set G/H is defined as $G/\sim$ where $g_1 \sim g_2$ if and only if $g_1 g_2^{-1} \in H$ (this can be defined without H being normal: being normal is used to ensure that G/H becomes a group). In particular $g_1 \sim g_2$ if and only if $g_1 = hg_2$ for some $h \in H$.

[24] The interested reader can consult https://en.wikipedia.org/wiki/Comb_space.

2. A space X is *semilocally simply-connected* if every $x \in X$ admits a neighborhood U such that the inclusion $U \to X$ induces the trivial homomorphism $\pi_1(U, x) \to \pi_1(X, x)$.

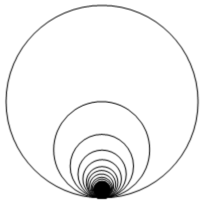


Figure 2.26: An example of a path-connected, locally path-connected space that is not semilocally simply-connected; the problematic point is the bottommost one. Source: https://wildtopology.files.wordpress.com/2012/05/hawaiian-earring-1.png

Remark 2.67. Here are some remarks concerning semilocally simply-connectedness.

1. All the spaces we have encountered in Chapter 2 are semilocally simply-connected: if X is a CW complex, then it is semilocally simply-connected [Hat02, Proposition A.4].[25]
2. An example of a non semilocally simply-connected space is illustrated in Figure 2.26, but the proof is omitted.
3. Being semilocally simply-connected is a necessary condition in Theorem 2.70: if X admits a simply-connected covering space $\widetilde{X}$, then we argue it must be semilocally simply-connected. By definition of a covering space, every $x \in X$, is contained in an open set $U \subset X$ such that there is an open set $\widetilde{U} \subset \widetilde{X}$ with $p|\colon \widetilde{U} \to U$ a homeomorphism. To prove that $\pi_1(U, x_0) \to \pi_1(X, x_0)$ is the trivial map, use $p|^{-1}$ to lift γ to a loop $\widetilde{\gamma}$ in $\widetilde{U} \subset \widetilde{X}$ based at $\widetilde{x}_0 \in \widetilde{U} \cap p^{-1}(x_0)$, use the simple-connectedness of $\widetilde{X}$ to deduce that $\widetilde{\gamma} \simeq c_{\widetilde{x}_0}$ and then project down this homotopy using p to deduce that $\gamma \simeq c_{x_0}$ in X, as required.

Next, we make precise what we mean by considering two covering spaces as being "the same".

Definition 2.68. Let X be a space and let $x_0 \in X$.

1. Two coverings $p_1\colon \widetilde{X}_1 \to X$ and $p_2\colon \widetilde{X}_2 \to X$ are *isomorphic* if there is a homeomorphism $f\colon \widetilde{X}_1 \to \widetilde{X}_2$ with $p_2 \circ f = p_1$.
2. Two coverings $p_1\colon (\widetilde{X}_1, \widetilde{x}_1) \to (X, x_0)$ and $p_2\colon (\widetilde{X}_2, \widetilde{x}_2) \to (X, x_0)$ are *basepoint preserving isomorphic* if there is a homeomorphism $f\colon \widetilde{X}_1 \to \widetilde{X}_2$ with $p_2 \circ f = p_1$ and $f(\widetilde{x}_1) = \widetilde{x}_2$.

Remark 2.69. We make some remarks concerning isomorphisms of covering spaces.

1. Isomorphisms take fibers to fibers: $f(p_1^{-1}(x)) = p_2^{-1}(x)$; this follows from $p_2 \circ f = p_1$.
2. We leave it to the reader to verify that (basepoint-preserving) isomorphism defines an equivalence relation on the set of (based) covering spaces over a fixed (based) space.
3. Example 2.72 below will describe an example of two covering spaces of $S^1 \vee S^1$ that are isomorphic but not basepoint preserving isomorphic.

We can now state the main theorem in the theory of covering spaces.

Theorem 2.70. *Let X be a path-connected, locally path-connected, semilocally simply-connected space, and let $x_0 \in X$. Mapping a covering space $p\colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ to its group $p_*(\pi_1(\widetilde{X}, \widetilde{x}_0))$ gives rise to a bijection between the two following sets:*

[25] In [Hat02, Proposition A.4], it is proved that CW complexes are locally contractible, a property that implies semilocally simply-connected.

1. *the set of basepoint preserving isomorphism classes of path-connected covers of* (X, x_0)*;*
2. *the set of subgroups of* $\pi_1(X, x_0)$.

Remark 2.71. Here are some remarks about this classification theorem.

1. Theorem 2.70 represents the best possible situation in topology: the topological objects are completely understood and classified in terms of some concrete algebraic data.

2. For a path-connected, locally path-connected, semilocally simply-connected space X, Theorem 2.70 implies that for every subgroup $H \leq \pi_1(X, x_0)$, there is a cover $p\colon (X_H, \widetilde{x}_0) \to (X, x_0)$ such that the fundamental group of the total space X_H is H:

$$\pi_1(X_H, \widetilde{x}_0) \cong p_*(\pi_1(X_H, \widetilde{x}_0)) = H.$$

   In particular X admits a simply-connected cover, unique up to isomorphism, which is called the *universal cover* of X. Additionally, Theorem 2.70 allows us to specify a cover simply by describing a subgroup of $\pi_1(X, x_0)$.

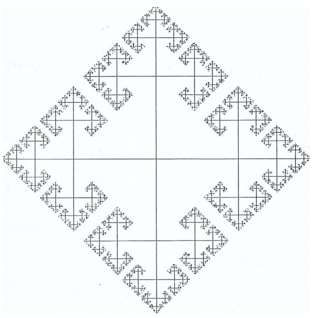


Figure 2.27: The universal cover of $S^1 \vee S^1$. Source: https://i.stack.imgur.com/O0X7E.jpg

Example 2.72. Here are some Examples of Theorem 2.70.

1. If X is simply-connected (e.g. $X = \mathbb{R}^n$ or S^n) then it admits a unique (up to isomorphism) cover; the trivial cover $\mathrm{id}_X\colon X \to X$ (which also happens to be the universal cover).

2. For $X = \mathbb{R}P^n$ with $n \geq 2$, we have $\pi_1(X) = \mathbb{Z}_2$ and so $\mathbb{R}P^n$ only admits two covers (up to isomorphism), namely the trivial cover $\mathbb{R}P^n \to \mathbb{R}P^n$ and the universal cover $S^n \to \mathbb{R}P^n$.

3. For $X = S^1$, the subgroups of $\pi_1(X) = \mathbb{Z}$ are $n\mathbb{Z}$ and the corresponding covers are the trivial cover, the cover p_n and the exponential map $\exp\colon \mathbb{R} \to S^1$.

4. For the torus $X = S^1 \times S^1$, with $\pi_1(S^1 \times S^1) = \mathbb{Z}^2$, the universal cover is $\mathbb{R}^2$, while Example 2.64 discussed the cover corresponding to $\mathbb{Z} \oplus 0$.

5. Example 2.64 included the example of the covering space of $S^1 \vee S^1$ associated to $\langle a, b^2, bab^{-1} \rangle$ while Figure 2.27 shows the universal cover of $X = S^1 \vee S^1$ and Figure 2.28 shows examples of basepoint-preserving isomorphic and non- basepoint preserving isomorphic coverings of X.

There is also a classification result without basepoints which was not discussed in class. The interested reader can consult the second part of [Hat02, Theorem 1.38] but essentially, the punchline is that there is a bijection between the set of isomorphism classes of path-connected covering spaces $p\colon \widetilde{X} \to X$ and the set of conjugacy classes of subgroups of $\pi_1(X, x_0)$. Note that if $\pi_1(X, x_0)$ is abelian, then the set of conjugacy classes of subgroups of $\pi_1(X, x_0)$ coincides with its set of subgroups of $\pi_1(X, x_0)$.

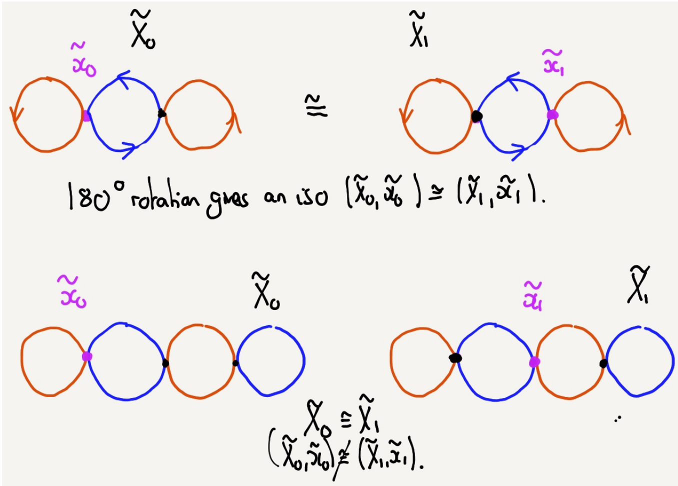


Figure 2.28: A pair of basepoint preserving isomorphic covering spaces (top) as well as a pair of isomorphic but not basepoint-preserving isomorphic covering spaces (bottom).

2.4.2 Covers that induce the same subgroup are isomorphic

The aim of this subsection is to prove the injectivity part of the classification theorem; references include [Mun00, Section 79] and [Hat02, Chapter 1.3]. More concretely, we want to show that if two covers of (X, x_0) induce the same subgroup of $\pi_1(X, x_0)$, then they are isomorphic.

Theorem 2.73. *Let X be a path-connected locally path-connected space and let $p_i\colon (\widetilde{X}_i, \widetilde{x}_i) \to (X, x_0)$ be a path-connected cover for $i = 1, 2$. The following assertions are equivalent:*

1. *the coverings p_1 and p_2 are isomorphic via a basepoint preserving isomorphism;*
2. *the coverings p_1 and p_2 induce the same subgroup: $p_{1*}(\pi_1(\widetilde{X}_1, \widetilde{x}_1)) = p_{2*}(\pi_1(\widetilde{X}_2, \widetilde{x}_2))$.*

The proof of Theorem 2.73 is just as interesting as its statement: many of the intermediate propositions are very useful in their own right. We start with the following wonderful lifting criterion (wonderful because it gives a complete algebraic answer to a topological question).

Proposition 2.74. *Let $p\colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ be a cover and let $f\colon (Y, y_0) \to (X, x_0)$ be a continuous map, where Y is a path-connected and locally path-connected space. The following assertions are equivalent:*

1. *the continuous map f lifts to a continuous map $\widetilde{f}\colon (Y, y_0) \to (\widetilde{X}, \widetilde{x}_0)$;*
2. *the following inclusion holds $f_*(\pi_1(Y, y_0)) \subset p_*(\pi_1(\widetilde{X}, \widetilde{x}_0))$.*

Proof. The implication (1) $\Rightarrow$ (2) is quickly proved: if we have a continuous map $\widetilde{f}$ with $p \circ \widetilde{f} = f$, then $f_*(\pi_1(Y, y_0)) = p_*(\widetilde{f}_*(\pi_1(Y, y_0))) \subset p_*(\pi_1(\widetilde{X}, \widetilde{x}_0))$.

We now prove the reverse direction, i.e. (2) $\Rightarrow$ (1). We assume that $f_*(\pi_1(Y, y_0)) \subset p_*(\pi_1(\widetilde{X}, \widetilde{x}_0))$ and we want to define a contimap $\widetilde{f}\colon Y \to \widetilde{X}$, that lifts f, i.e. we want to define $\widetilde{f}(y) \in \widetilde{X}$ with $p(\widetilde{f}(y)) = f(y)$ for every $y \in Y$. Since Y is path-connected, there exists a path $\gamma\colon I \to Y$ from y_0 to y and therefore $f \circ \gamma\colon I \to X$ is a path from x_0 to $f(y) = (f \circ \gamma)(1)$. Applying the path lifting property from Proposition 2.60 gives a unique lift $\widetilde{f \circ \gamma}\colon I \to \widetilde{X}$ starting at $\widetilde{x}_0$ and we set

$$\widetilde{f}(y) := \widetilde{f \circ \gamma}(1).$$

Note that $\widetilde{f}(y_0) = \widetilde{x}_0$: choose $\gamma = c_{y_0}$, the constant path at y_0 in the definition above. Additionally, observe that $\widetilde{f}$ lifts f: indeed $p \circ \widetilde{f}(y) = p(\widetilde{f \circ \gamma}(1)) = (f \circ \gamma)(1) = f(y)$ for every $y \in Y$.

We prove that $\widetilde{f}$ is well defined, i.e. that it does not depend on the choice of the path γ. Assume that γ and γ' are two paths from y_0 to y; we must show that $\widetilde{f \circ \gamma}(1) = \widetilde{f \circ \gamma'}(1)$. Observe that for the loop $h_0 := (f \circ \gamma') \cdot (\overline{f \circ \gamma})$ at x_0, we have

$$[h_0] = [(f \circ \gamma') \cdot \overline{f \circ \gamma}] = [f \circ (\gamma' \cdot \overline{\gamma})] = f_*([\gamma' \cdot \overline{\gamma}]) \in f_*(\pi_1(Y, y_0) \subset p_*(\pi_1(\widetilde{X}, \widetilde{x}_0)),$$

Proposition 2.63 implies that h_0 lifts to a loop $\widetilde{h}_0$ at $\widetilde{x}_0$. Applying the uniqueness statement in path-lifting, the first half of $\widetilde{h}_0$ is $\widetilde{f \circ \gamma'}$ and its second half is $\widetilde{\overline{f \circ \gamma}}$. Thus, we have $\widetilde{h}_0 = \widetilde{f \circ \gamma'} \cdot \widetilde{\overline{f \circ \gamma}}$, with midpoint $\widetilde{f \circ \gamma}(1) = \widetilde{\overline{f \circ \gamma}}(0) = \widetilde{f \circ \gamma'}(1)$, as required. We omit the proof that $\widetilde{f}$ is continuous, but note that it uses the fact that Y is locally path-connected. □

The next proposition gives a criterion to ensure the uniqueness of a lift.

Proposition 2.75. *Let $p\colon \widetilde{X} \to X$ be a cover and let $f\colon Y \to X$ be a continuous map with Y connected. If $\widetilde{f}_1, \widetilde{f}_2$ are two lifts of f such that $\widetilde{f}_1(y) = \widetilde{f}_2(y)$ for some $y \in Y$, then $\widetilde{f}_1 \equiv \widetilde{f}_2$.*

Proof. Consider the subset $\mathcal{O} = \{y \in Y \mid \widetilde{f}_1(y) = \widetilde{f}_2(y)\}$ of Y. We want to prove that $\mathcal{O} = Y$. Since $\mathcal{O}$ is non-empty and since Y is connected, it suffices to prove that $\mathcal{O}$ is both open and closed in Y. We prove that $\mathcal{O}$ is open in Y by showing it is a neighborhood of each of its points. Given $y \in \mathcal{O}$, we have $\widetilde{f}_1(y) = \widetilde{f}_2(y) \in X$ and by definition of a covering space, there is an open set $U \subset X$ containing $f(y)$ and an open sets $\widetilde{U} \subset \widetilde{X}$ containing $\widetilde{f}_1(y)$ and $\widetilde{f}_1(y)$ so that $p|\colon \widetilde{U} \to U$ is a homeomorphism. Since $\widetilde{f}_i$ is continuous, there is an open set $N \subset Y$ containing y with $\widetilde{f}_i(N) \subset \widetilde{U}$ and $\widetilde{f}_2(N) \subset \widetilde{U}$ (e.g. take $N := \widetilde{f}_1^{-1}(\widetilde{U}) \cap \widetilde{f}_2^{-1}(\widetilde{U})$). Since $p| \circ \widetilde{f}_1|_N = f|_N = p| \circ \widetilde{f}_2|_N$ and $p|$ is a homeomorphism, it follows that $\widetilde{f}_1|_N = \widetilde{f}_2|_N$ and therefore $N \subset \mathcal{O}$ as required. The proof that $\mathcal{O}$ is closed, i.e. that $Y \setminus \mathcal{O}$ is open is very similar and is therefore omitted. □

We now prove Theorem 2.73 which states that based coverings of (X, x_0) induce the same subgroup of $\pi_1(X, x_0)$ if and only if they are basepoint preserving isomorphic.

Proof of Theorem 2.73. If $f\colon (\widetilde{X}_1, \widetilde{x}_1) \to (\widetilde{X}_2, \widetilde{x}_2)$ is a basepoint preserving isomorphism, then

$$(p_1)_*(\pi_1(\widetilde{X}_1, \widetilde{x}_1)) = (p_2)_*(f_*(\pi_1(\widetilde{X}_1, \widetilde{x}_1))) \subset (p_2)_*(\pi_1(\widetilde{X}_2, \widetilde{x}_2)).$$

The reverse inclusion is proved similarly by writing $p_2 = p_1 \circ f^{-1}$.

We now prove the converse, namely we assume that $p_{1*}(\pi_1(\widetilde{X}_1, \widetilde{x}_1)) = p_{2*}(\pi_1(\widetilde{X}_2, \widetilde{x}_2))$ and prove the existence of a basepoint preserving isomorphism $(\widetilde{X}_1, \widetilde{x}_1) \cong (\widetilde{X}_2, \widetilde{x}_2)$. Consider the problem of lifting $p_1\colon (\widetilde{X}_1, \widetilde{x}_1) \to (X, x_0)$ to $(\widetilde{X}, \widetilde{x}_2)$, and similarly for p_2. By the criterion for the existence of a lift from Proposition 2.74[26], there are continuous maps

$$\widetilde{p}_1\colon (\widetilde{X}_1, \widetilde{x}_1) \to (\widetilde{X}_2, \widetilde{x}_2),$$
$$\widetilde{p}_2\colon (\widetilde{X}_2, \widetilde{x}_2) \to (\widetilde{X}_1, \widetilde{x}_1),$$

with $p_2 \circ \widetilde{p}_1 = p_1$ and $p_1 \circ \widetilde{p}_2 = p_2$. We show that $\widetilde{p}_1$ provides the required isomorphism (we already know it is basepoint preserving) with inverse $\widetilde{p}_2$. The equality $\widetilde{p}_2 \circ \widetilde{p}_1 = \mathrm{id}_{\widetilde{X}_1}$ follows from the uniqueness criterion of Proposition 2.75: $\widetilde{X}_1$ is path-connected and both maps lift p_1 and agree on $\widetilde{x}_1$. The proof that $\widetilde{p}_1 \circ \widetilde{p}_2 = \mathrm{id}_{\widetilde{X}_2}$ is identical. □

[26] We can apply this criterion because the $\widetilde{X}_i$ are path-connected and locally path-connected. For path-connectedness, this was assumed while for local path-connectedness, a short verification shows that a cover of a locally path-connected space is locally path-connected.

2.4.3 Covering spaces arise from subgroups

The aim of this subsection is to prove the surjectivity part of the classification theorem; references include [Mun00, Section 80 and 82] and [Hat02, Chapter 1.3]. First a quick reminder: at this stage, given a space X and $x_0 \in X$, we have associated a subgroup of $\pi_1(X, x_0)$ to every based cover of (X, x_0). Moreover we know from Theorem 2.73 that X is path-connected and locally path-connected and if two based covers of X induce the same subgroup of $\pi_1(X, x_0)$, then they are basepoint-preserving isomorphic. We now prove that if X is semilocally simply-connected, then every subgroup of $\pi_1(X, x_0)$ is induced by a based covering of (X, x_0).

Theorem 2.76. *If X is a path-connected, locally path-connected and semilocally simply-connected space and if $x_0 \in X$, then for every subgroup $H \leq \pi_1(X, x_0)$, there is a covering space $(p_H)\colon X_H \to X$ such that $(p_H)_*(\pi_1(X_H, \widetilde{x}_0)) = H$ for some basepoint $\widetilde{x}_0 \in p_H^{-1}(x_0)$.*

The main step in the proof of Theorem 2.76 consists of showing that a path-connected, locally path-connected and semilocally simply-connected space X admits a simply-connected cover; as we mentioned in Subsection 2.4.1, such a cover is called the universal cover of X.[27]

Construction 2.77. Let X be a path-connected, locally path-connected and semilocally simply-connected space, and let $x_0 \in X$. Consider the set of all homotopy classes of paths in X starting at x_0

$$\widetilde{X} = \{[\gamma] \mid \gamma\colon I \to X, \gamma(0) = x_0\}$$

and the projection $p\colon \widetilde{X} \to X, [\gamma] \mapsto \gamma(1)$. The map p is well defined because path homotopies preserve endpoints. We give an idea of how a topology is defined on $\widetilde{X}$. First of all, we consider the following collection of open subsets of X:

$$\mathcal{U} = \{U \subset X \mid U \text{ is open, path-connected and } \pi_1(U) \to \pi_1(X) \text{ is trivial}\}.$$

For an open set $U \in \mathcal{U}$ and a path $\gamma\colon I \to X$ with $\gamma(0) = x_0$ and $\gamma(1) \in \mathcal{U}$, we then consider

$$U_{[\gamma]} = \{[\gamma \cdot \eta] \mid \eta \text{ is a path in } U \text{ with } \eta(0) = \gamma(1)\}.$$

One can now check that $\{U_{[\gamma]}\}_{U \in \mathcal{U}, [\gamma]}$ forms a basis for a topology and that $p\colon \widetilde{X} \to X$ is a covering map. The proofs are omitted: they are not overly difficult, but at this point, it makes sense to avoid some details and to keep an eye on the big picture; the interested reader is referred to [Hat02, page 64] for details.

Proposition 2.78. *If X is a path-connected, locally path-connected and semilocally simply-connected space, then it admits a simply-connected covering space.*

Proof. We prove that the space $\widetilde{X}$ defined in Construction 2.77 is simply-connected. We first argue that $\widetilde{X}$ is path-connected: given $[\gamma] \in \widetilde{X}$, we consider $\gamma_t\colon I \to X, s \mapsto \gamma(st)$ and note that $t \mapsto [\gamma_t]$ is a path in $\widetilde{X}$ from $[c_{x_0}]$ to $[\gamma]$. We now prove that $\pi_1(\widetilde{X}, [c_{x_0}]) = 1$. Since p_* is injective (recall Proposition 2.63), it suffices to prove that $p_*(\pi_1(\widetilde{X}, [c_{x_0}])) = 1$. Recall from Proposition 2.63 that this subgroup consists of those (homotopy clases of) loops γ in X that lift to a loop in $\widetilde{X}$ based at $[c_{x_0}]$. By uniqueness of lifts, this loop must be of the form $t \mapsto [\gamma_t]$. But now the fact that this is a loop at $[c_{x_0}]$ precisely means that $[c_{x_0}] = [\gamma_1] = [\gamma]$, i.e. that γ is nullhomotopic in X. □

We have therefore proved that under some mild conditions, the universal cover of a space always exists. As Hatcher puts it however "In concrete cases one usually constructs a simply-connected covering space by more direct methods" [Hat02, page 65]. We already saw this principle in action in Example 2.72 and more examples of universal covers are given on the twelfth problem set. We now conclude the proof of Theorem 2.76 and therefore the proof of Theorem 2.70.

[27] We can speak of "the" universal cover thanks to Theorem 2.73.

Proof of Theorem 2.76. We construct X_H as a quotient space of the universal cover $\widetilde{X}$ of X. Declare $[\gamma], [\gamma'] \in \widetilde{X}$ to be equivalent if $\gamma(1) = \gamma'(1)$ and $[\gamma \cdot \overline{\gamma}'] \in H$. We check that this is an equivalence relation:

- reflexivity: $[\gamma] \sim [\gamma]$ because $\gamma(1) = \gamma(1)$ and $[\gamma \cdot \overline{\gamma}] = [c_{x_0}] = 1 \in H$;
- symmetry: if $[\gamma] = [\gamma']$, then $[\gamma'] = [\gamma]$ because $[\gamma' \cdot \overline{\gamma}] = [\gamma \cdot \overline{\gamma}']^{-1} \in H$;
- transitivity: if $[\gamma \cdot \overline{\gamma}'] \in H$ and $[\gamma' \cdot \overline{\gamma}''] \in H$, then $[\gamma \cdot \overline{\gamma}''] = [\gamma \cdot \overline{\gamma}'][\gamma' \cdot \overline{\gamma}''] \in H$.

We then define $X_H := \widetilde{X}/\sim$ and $p_H \colon X_H \to X$ as the map induced by the covering projection $p \colon \widetilde{X} \to X$ on X_H. We omit the proof that p_H is a covering map but refer the interested reader to [Hat02, proof of Proposition 1.36] for details.

Write $\pi \colon \widetilde{X} \to \widetilde{X}/\sim = X_H$ for the canonical projection and set $\widetilde{x}_0 := \pi([c_{x_0}])$. We check that $(p_H)_*(\pi_1(X_H, \widetilde{x}_0)) = H$. By Proposition 2.63, asking for $[\gamma] \in \pi_1(X, x_0)$ to lie in $(p_H)_*(\pi_1(X_H, \widetilde{x}_0))$ is equivalent to asking for γ to lift to a loop in X_H at $\widetilde{x}_0$. This in turn is equivalent to asking for γ to lift to a path $\widetilde{\gamma}$ in $\widetilde{X}$ starting at $[c_{x_0}]$ and with $\widetilde{\gamma}(1) \sim [c_{x_0}]$; here recall that $\widetilde{\gamma}(1)$ is a path in X (in fact, by uniqueness of path lifting, $\widetilde{\gamma}$ must be the path $I \to \widetilde{X}, t \mapsto [\gamma_t]$ already in mentioned in Proposition 2.78). Since $\widetilde{\gamma}(1) = [\gamma] \in \widetilde{X}$, this is equivalent to asking for $[\gamma] \sim [c_{x_0}]$, where $\sim$ denotes the previously defined equivalence relation on $\widetilde{X}$. By definition of this equivalence relation, this occurs if and only if $\gamma(1) = x_0$ and $[\gamma] = [\gamma \cdot c_{x_0}] \in H$, i.e. $[\gamma] \in H$ (note that $\gamma(1) = x_0$ is automatic because γ is a loop at x_0). This concludes the proof of the theorem. □

We have now proved Theorem 2.70. Together with the following active learning session, this concludes the material that we will cover in this class.

Active learning 2.79. Here is a summary of what was discussed in the tenth active learning session:[28]

- We recalled from the twelfth problem set that the *deck transformation group* $G(\widetilde{X})$ of a cover $\widetilde{X} \to X$ consists of all isomorphisms of $\widetilde{X}$; we also recalled the definition of the normaliser $N(H)$ of a subgroup $H \leq G$:

$$N(H) = \{g \in G \mid gHg^{-1} = H\}.$$

  We also recalled from the twelfth problem set that if G acts properly discontinuously by homeomorphisms on a space Y, then $G(Y) = G$. This implies that the deck transformation group of $p_n \colon S^1 \to S^1$ is $\mathbb{Z}_n$ and that the deck transformation group of $\mathbb{R} \to S^1$ is $\mathbb{Z}$.

- A covering space $p \colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ is *normal* if the subgroup $p_*(\pi_1(\widetilde{X}, \widetilde{x}_0)) \leq \pi_1(X, x_0)$ is normal (this turns out to be independent of $\widetilde{x}_0$). If a space has abelian fundamental group, then all its covers are normal. We gave an examples of normal and non-normal covers of $S^1 \vee S^1$.

- We stated and proved the following proposition (which is a subset of [Hat02, Proposition 1.39]:

  Proposition 2.80. *Let $p \colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ be a path-connected cover of a path-connected, locally path-connected space X and set $H := p_*(\pi_1(\widetilde{X}, \widetilde{x}_0))$. Then there is an isomorphism*

$$G(\widetilde{X}) \cong N(H)/H.$$

  From this proposition, we deduced that deck transformation group of the universal cover of a space X is $\pi_1(X)$.

[28]Depending on the year, this material won't be covered.

Chapter 3

Afterword: What comes next?

At this point, the interested reader might be wondering what comes next. There are in fact two distinct questions to wonder about:

1. if this class continued, what would be the next topics?

2. what is the broader picture, i.e. what are some themes that motivate researchers in topology?

Regarding the first question, the course could continue in several directions.

- One possibility would be to learn more algebraic topology, e.g. homology, cohomology and higher homotopy groups, all of which are covered in [Hat02]. Very briefly, these are other topological invariants that are sometimes more powerful than the fundamental group to prove that spaces are not homotopy equivalent (and so, in particular, not homeomorphic).

- Another possibility would be to learn about smooth manifolds (the study of topological manifolds is also a possibility, although a less common one). For instance after learning about the definitions, one usual learns about embeddings, immersions and tangent spaces as in [Tu11]. Depending on the flavor of the course, one might then either focus on the geometry of manifolds (this is called *differential geometry* or *global analysis*) or on the topology of manifolds (this is, broadly speaking, a subset of geometric topology).

- A third possibility (still mostly within geometric topology) would be to focus on low dimensional manifolds, i.e. dimensions $1, 2, 3$ and 4. To name only a few examples, 3-manifold topology and 4-manifold topology are fields in their own right, as are the study of surfaces (this is connected to dynamics and geometric group theory) and knot theory (the study of smooth simple closed curves in 3-space). A classical reference in low dimensional topology is [Rol90] while a panorama of 4-manifold topology can be found [Sco05]

In order to give a flavor of some of these topics, we attempt to answer the (difficult) second question, namely "what are some themes that motivate researchers in topology". As a first step, let us return to day one and ask again "what is topology?" On the first page of these notes, our first answer was that "*informally, topology studies the properties of shapes that are preserved under continuous deformations.*" During Chapter 1, we saw a more precise answer: topology is, broadly speaking, the study of topological spaces up to homeomorphism (or homotopy equivalence).[1] With the tools and vocabulary of Chapters 1 and 2, here is an attempt at a third answer:

Topology is the study of

- *topological spaces, considered either up to homeomorphism or homotopy equivalence,*
- *continuous maps between topological spaces;*

in both cases, particular emphasis is placed on manifolds and CW complexes.

Let's try to make this more concrete by giving some examples (a disclaimer: these examples might slightly reflect my own interests, they are by no means meant to be exhaustive).

1. What do we mean by "studying topological spaces up to homeomorphism, with a particular emphasis on CW complexes and manifolds"?

   - A guiding goal of topology is to classify manifolds up to homeomorphism and smooth manifolds up to diffeomorphism (*in what follows, for simplicity "manifold" will always mean "compact, connected oriented manifold without boundary"*). For example, every 2-manifold (i.e. every surface) is homeomorphic to S^2 or Σ_g for $g \geq 1$ [Mun00, Chapter 12].
   - In general, classifying manifolds is too ambitious a goal, so it makes sense to restrict the topology of the manifolds we wish to study. For example, one could ask for a classification of simply-connected manifolds. In dimension 2, the only example is the 2-sphere. In dimension 3, it was conjectured by Poincaré that the only such manifold is the 3-sphere, and this was proved by Perelman in 2002 (in fact, Perelman proved even more than the so-called *Poincaré conjecture*, but stating the geometrization conjecture would take us too far astray; an account of this can be found in this note by Milnor https://www.claymath.org/sites/default/files/poincare.pdf). In dimension 4, Freedman classified topological 4-manifolds up to homeomorphism; a hint of the statement is also given in Milnor's note from a few lines above.[2] Simply-connected smooth 4-manifolds are poorly understood and an active area of research.
   - Building on the previous bullet point, the smooth/topological *generalised Poincaré conjecture* posits that if a smooth/topological M is homotopy equivalent to S^n, then it is diffeomorphic/homeomorphic to S^n.[3] The topological Poincaré conjecture is known in all dimensions, while the smooth Poincaré is false in general (the smallest *exotic sphere* is 7-dimensional), it is true in some dimensions (including 1,2,3,5,6). As mentioned in https://www.claymath.org/sites/default/files/poincare.pdf,

[1] As Chapter 1 progressed, the reader might have noted that topological spaces can be very pathological and so it made sense to restrict the class of spaces under considerations and two very natural choices are the study of manifolds and CW complexes.

[2] A very brief overview of Freedman's incredible proof can be found in [Sco05] for recent developpments on Freedman's work, see https://www.quantamagazine.org/new-math-book-rescues-landmark-topology-proof-20210909/.

[3] If a compact connected oriented 3-manifold is simply-connected, then it must be homotopy equivalent to S^3, so for $n = 3$ the generalised Poincaré conjecture does indeed reduce to the "original" Poincaré conjecture.

some relevant names are Smale, Stallings, Zeeman, Newman in high dimensions, Milnor for exotic spheres and Freedman for the topological Poincaré conjecture. The smooth 4-dimensional Poincaré conjecture is still wide open and active area of research.

2. What do we mean by "the study of continuous maps between topological spaces with a particular emphasis on CW complexes and manifolds"?

   - Given based CW complexes X, Y, it is natural to wonder whether one can describe the set $[X, Y]_*$ of base-point preserving homotopy classes of maps from X to Y? For example, when $X = S^1$, we saw that the fundamental group $[S^1, X]_* = \pi_1(X)$ is well understood. In general however the situation quickly gets out of hand as even the case of spheres is unknown in general: for $X = S^n$, the study of the *higher homotopy groups* $[S^n, X]_* = \pi_n(X)$ has arguably been a driver for a lot of research in algebraic topology and homotopy theory; this table indicates that things get bad pretty quicky: https://en.wikipedia.org/wiki/Homotopy_groups_of_spheres#Table
   - Given an n-manifold M and a k-dimensional manifold K with $k \leq n$, understand embeddings $K \hookrightarrow M$ (here an embedding is a continuous map that is a homeomorphism onto its image). For example, when $K = S^1$ and $M = S^3$, this is classical knot theory[4], while $K = S^2$ and $M = S^4$ is the study of 2-knots. An even more ambitious goal which is more in the realm of algebraic topology is to study the homotopy type of the space of embeddings of K into M (with the appropriate topology). Given a manifold M, algebraic topologists are also interested in understand the homotopy type of the space $\mathrm{Homeo}(M)$ of self-homeomorphism of M or, in the case M is smooth the space $\mathrm{Diff}(M)$ of diffeomorphisms of M, e.g. the famous Smale conjecture (proved by Hatcher) asserts that $\mathrm{Diff}(S^3) \simeq O(4)$; in 2018 Watanabe disproved the 4-dimensional Smale conjecture: $\mathrm{Diff}(S^4) \not\simeq O(5)$; see https://www.quantamagazine.org/how-tadayuki-watanabe-solved-a-topological-mystery-about-spheres-20211026/

Acknowledgements

I would like to thank David Cimasoni for sharing his notes from the topology courses he taught at the university of Geneva and which inspired several parts of this class. My memories from my time in your classes and as your student have deeply shaped both these notes and my way of teaching.

Thanks also go to the students of MIT's 18.901, and especially to Ben Lou, for pointing out typos and inaccuracies in these notes.

[4]Knot theory is related to 4-dimensional topology though the study of *knot concordance*; see https://www.quantamagazine.org/graduate-student-solves-decades-old-conway-knot-problem-20200519/

Bibliography

[FNOP19] Stefan Friedl, Matthias Nagel, Patrick Orson, and Mark Powell. A survey of the foundations of four-manifold theory in the topological category. *ArXiv 1910.07372*, 2019.

[Hat02] Allen Hatcher. *Algebraic topology.* Cambridge University Press, Cambridge, 2002.

[Mun00] James R. Munkres. *Topology.* Prentice Hall, Inc., Upper Saddle River, NJ, 2000. Second edition of [MR0464128].

[Rol90] Dale Rolfsen. *Knots and links*, volume 7 of *Mathematics Lecture Series.* Publish or Perish, Inc., Houston, TX, 1990. Corrected reprint of the 1976 original.

[Sco05] Alexandru Scorpan. *The wild world of 4-manifolds.* American Mathematical Society, Providence, RI, 2005.

[Tu11] Loring W. Tu. *An introduction to manifolds.* Universitext. Springer, New York, second edition, 2011.