Thomas Krämer

Algebraic Topology

TU Chemnitz, Summer 2026

(Version of July 12, 2026)

Contents

Chapter I
Introduction

In topology we study geometric shapes from a qualitative viewpoint, with a focus on features that are invariant under continuous deformation. In this context metric aspects such as distances or angles become secondary. Instead, we use the much more flexible notion of a topological space. These are made of rubber: We may bend, stretch or compress them continuously, as long as we don't tear them apart and don't glue pieces together. More precisely, the only data remembered by a topological space are its open subsets (one could say that these give a notion of two points being near to each other without talking about distances):

Definition. A *topology* on a set X is a collection $\mathscr{T}$ of subsets of X such that

a) $\mathscr{T}$ is stable under arbitrary unions,
b) $\mathscr{T}$ is stable under finite intersections, and
c) both X and the empty set $\varnothing$ belong to $\mathscr{T}$.

We then call $(X, \mathscr{T})$ a *topological space*. The sets in $\mathscr{T}$ are called the *open subsets* of the topological space. A subset $Z \subset X$ is called *closed* if $X \smallsetminus Z$ is open.

Example. The following are topological spaces:

a) $X = \mathbb{R}^n$ with the Euclidean topology, where by definition a subset $U \subset X$ is open iff for every $p \in U$ there exists $\varepsilon > 0$ such that

$$B_\varepsilon(p) \; := \; \{q \in X \mid d(p,q) < \varepsilon\} \; \subset \; U.$$

Here $d(p,q) := |p-q|$ denotes the Euclidean distance between p and q.

b) If X is a topological space, then any subset $Y \subset X$ inherits a topology where we define a subset $V \subset Y$ to be open iff

$$V \; = \; U \cap Y \quad \text{for some open subset} \quad U \; \subset \; X.$$

We call this the *subspace topology* and $Y \subset X$ a *topological subspace* of X.

c) By combining these two examples, any subset $Y \subset \mathbb{R}^n$ has a natural topology.

d) For any topological spaces X and Y, the *product topology* on $Z = X \times Y$ is by definition the topology whose open subsets are unions

$$W = \bigcup_{i \in I} U_i \times V_i \subset Z = X \times Y$$

where $U_i \subset X$ and $V_i \subset Y$ are open and I is an arbitrary index set.

e) For any topological spaces X and Y we endow their *disjoint union* $Z = X \sqcup Y$ with the topology whose open subsets are the unions

$$W = U \sqcup V \quad \text{of open subsets} \quad U \subset X \quad \text{and} \quad V \subset Y.$$

f) Any topology on a set X sits between the following two extreme topologies:

- The *discrete topology* where every subset of X is defined to be open.
- The *indiscrete topology* whose only open sets are X and the empty set $\varnothing$.

When dealing with topological spaces, we are mainly interested in maps between them that are compatible with the respective topologies:

Definition. A map $f\colon X \to Y$ between topological spaces is *continuous* if for any open $V \subset Y$ the preimage $f^{-1}(V) \subset X$ is again open. We call f a *homeomorphism* if it is continuous, bijective and if the inverse $f^{-1}\colon Y \to X$ is also continuous. We then also write

$$X \cong Y$$

and say that X and Y are *homeomorphic*.

Example. Saying that two spaces are homeomorphic is a precise version of saying that they look 'the same' topologically. Here are some examples:

a) The unit circle is homeomorphic to a square. An explicit homeomorphism can be obtained by projecting points radially from the center:

[PICTURE]

b) The exponential function gives a continuous bijective map

$$f\colon \quad [0, 2\pi) \to S^1 := \{z \in \mathbb{C} \mid |z| = 1\}, \quad t \mapsto \exp(it),$$

but its inverse is not continuous. In fact the topological spaces $[0, 2\pi)$ and S^1 are not homeomorphic. A quick argument for this uses compactness: A topological space X is called *compact* if every cover $X = \bigcup_{i \in I} U_i$ by open sets $U_i \subset X$ admits a finite subcover, i.e. one can find a finite collection of indices $i_1, \dots, i_n \in I$ such that

$$X = U_{i_1} \cup \cdots \cup U_{i_n}$$

Since S^1 is compact while $[0, 2\pi)$ is not, they cannot be homeomorphic.

c) The unit interval $I = [0,1]$ is not homeomorphic to the square $Q = I \times I$. This is not as obvious as it seems: Both have the same cardinality and there even exist continuous surjective maps

$$f\colon \quad I \to Q = I \times I$$

known as space-filling curves (`en.wikipedia.org/wiki/Peano_curve`). A quick argument for the non-existence of homeomorphisms uses the notion of path-connectedness: A topological space X is called *path-connected* if for any points $p, q \in X$ there is a continuous map $f\colon [0,1] \to X$ with $f(0) = p$ and $f(1) = q$ as shown in the following picture:

[PICTURE]

If we remove an interior point from I, the remaining open subset is no longer path-connected. By way of contrast, in Q the complement of any point is still path-connected. Hence I and Q cannot be homeomorphic.

d) It seems obvious that the following two spaces are not homeomorphic:

[PICTURE]

Intuitively, the reason is that the former has only one hole while the latter has two holes. But how can we make this into a formal argument?

The basic idea of algebraic topology is to attach to any topological space algebraic invariants (groups, rings, vector spaces, ...) which encode properties such as 'holes' in a meaningful way. The constructions will be made so that any continuous map of topological spaces induces a homomorphism between their algebraic invariants, and so that homeomorphisms induce isomorphisms. In particular, if two spaces have non-isomorphic algebraic invariants, then the spaces cannot be homeomorphic. As a slogan:

Algebraic topology converts topology (hard) into algebra (easy)!

In this class we will see two main examples for this approach, homotopy groups and homology groups. They are not only good for showing that two topological spaces are not homeomorphic, but have plenty of other applications. For instance, we will soon be able to give very simple proofs of

- the fundamental theorem of algebra: Every non-constant polynomial over the complex numbers has a complex root,
- the Brouwer fixed point theorem: Every continuous map from a disk to itself has a fixed point,
- the Jordan curve theorem: Every simple closed curve in the plane separates the plane into an inside and an outside region,
- the Borsuk–Ulam theorem: There are always two opposite points on Earth with the same temperature and air pressure,
- ...

Chapter II
The fundamental group

1 Preliminaries about paths

Let X be a topological space. We want to study it in terms of paths:

Definition 1.1. A *path in* X is a continuous map $\gamma\colon [0,1] \to X$. To keep track of the end points $a = \gamma(0)$ and $b = \gamma(1)$, we also say γ is a path from a to b. The *inverse* of γ is defined by

$$\gamma^{-}\colon \quad [0,1] \to X, \quad s \mapsto \gamma(1-s).$$

A path $\mu\colon [0,1] \to X$ is said to be *composable with* $\gamma\colon [0,1] \to X$ if $\mu(0) = \gamma(1)$, in which case we put

$$\gamma\cdot\mu\colon \quad [0,1] \to X, \quad s \mapsto \begin{cases} \gamma(2s) & \text{if } 0 \le s < \frac{1}{2}, \\ \mu(2s-1) & \text{if } \frac{1}{2} \le s \le 1. \end{cases}$$

Note that we denote the concatenation of paths from left to right!

[PICTURE OF INVERSE AND COMPOSITE]

Lemma 1.2. *Define a binary relation ~ on X by writing $a \sim b$ if there exists a path from a to b in X. Then ~ is an equivalence relation.*

Proof. The relation is

- reflexive: For any $a \in X$ we have $a \sim a$ via the constant path $\gamma\colon [0,1] \to \{a\}$.
- symmetric: If $a \sim b$ via a path γ, then $b \sim a$ via γ^-.
- transitive: If $a \sim b$ via a path γ and $b \sim c$ via a path μ, then $a \sim c$ via $\gamma \cdot \mu$. □

The equivalence classes for the above relation $\sim$ are called the *path components* of X. We put

$$\pi_0(X) \;=\; \{\text{path-components of } X\}.$$

Note that this is just a set, with no extra structure. We say that X is *path-connected* if it has only one path-component, that is, if any two points $a, b \in X$ can be joined by a path in X. In general, any topological space X decomposes canonically as the disjoint union

$$X \;=\; \bigsqcup_{Y \in \pi_0(X)} Y$$

of its path components $Y \subset X$, and each of those is path-connected. So in discussing paths we may assume without loss of generality that X is path-connected.

2 Definition of the fundamental group

The notion of a path is too rigid to be useful. In practice we only care about paths up to continuous deformation:

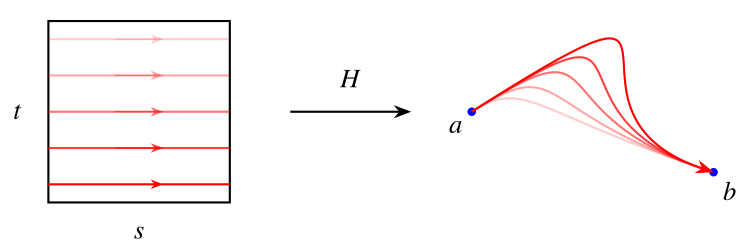
Definition 2.1. Let $\gamma_0, \gamma_1\colon [0,1] \to X$ be two paths in X. A *homotopy* between them is a continuous map

$$H\colon\; [0,1] \times [0,1] \;\to\; X \quad \text{with} \quad H(s,0) = \gamma_0(s) \quad \text{and} \quad H(s,1) = \gamma_1(s) \quad \text{for all } s.$$

If γ_0 and γ_1 are paths from a to b, we call H a *homotopy through paths from a to b* if moreover

$$H(0,t) \;=\; a \quad \text{and} \quad H(1,t) \;=\; b \quad \text{for all } t.$$

We then say that the two paths γ_0 and γ_1 are *equivalent* and write $\gamma_0 \sim \gamma_1$.



Lemma 2.2. *The relation $\sim$ is an equivalence relation on the set of paths in X.*

Proof. The relation is

- reflexive: We have $\gamma \sim \gamma$ via the constant homotopy $H(s,t) := \gamma(s)$.
- symmetric: If $\gamma_0 \sim \gamma_1$ via some homotopy H, then also $\gamma_1 \sim \gamma_0$ via the inverse homotopy

$$H^-\colon \quad [0,1]\times[0,1] \to X \quad \text{defined by} \quad H^-(s,t) := H(s,1-t).$$

- transitive: If $\gamma_0 \sim \gamma_1$ via a homotopy H' and $\gamma_1 \sim \gamma_2$ via a homotopy H'', then

$$\gamma_0 \sim \gamma_2 \quad \text{via the homotopy} \quad H(s,t) := \begin{cases} H'(s,2t) & \text{for } 0 \le t < \frac{1}{2} \\ H''(s,1-2t) & \text{for } \frac{1}{2} \le t \le 1. \end{cases}$$

Note that the map H defined by the above formula is continuous. □

We denote by $[\gamma]$ the equivalence class of a path γ for the equivalence relation in the above lemma. The inversion and composition of paths are well-defined on the level of these equivalence classes:

Lemma 2.3. *Let γ_0, γ_1 be two paths in X.*

a) If $\gamma_0 \sim \gamma_1$, then also their inverses satisfy $\gamma_0^- \sim \gamma_1^-$.
b) If μ_0, μ_1 are composable with γ_0, γ_1 and if $\mu_0 \sim \mu_1$, then also $\gamma_0 \cdot \mu_0 \sim \gamma_1 \cdot \mu_1$.

Proof. To prove $b)$, pick any homotopies $H'\colon \gamma_0 \sim \gamma_1$ and $H''\colon \mu_0 \sim \mu_1$ between the respective paths. We can then define a homotopy between their composites via the formula

$$H(s,t) := \begin{cases} H'(2s,t) & \text{for } 0 \le 2 < \frac{1}{2}, \\ H''(2s-1,t) & \text{for } \frac{1}{2} \le 2 \le 1. \end{cases}$$

The proof of $a)$ is similar and left as an exercise. □

We now specialize to the most important case of loops, i.e. paths that begin and end at a given reference point:

Definition 2.4. A *loop based at p* is path $\gamma\colon [0,1] \to X$ with $\gamma(0) = \gamma(1) = p$. The set of equivalence classes

$$\pi_1(X,p) := \{[\gamma] \mid \gamma \text{ is a loop in } X \text{ based at } p\}$$

is called the *fundamental group of X with respect to the base point p*.

Lemma 2.5. *The set $\pi_1(X,p)$ forms a group with respect to the composition of loops. Its neutral element is the class $[\varepsilon_p]$ of the constant loop $\varepsilon_p\colon [0,1] \to \{p\}$, and inverses are given by*

$$[\gamma]^{-1} = [\gamma^-] \quad \textit{for} \quad [\gamma] \in \pi_1(X,p).$$

Proof. Let $\varepsilon_p\colon [0,1] \to \{p\}$ be the constant loop. We claim that for all loops γ, μ, ν the following properties are satisfied:

- Associativity: $(\gamma\cdot\mu)\cdot\nu \sim \gamma\cdot(\mu\cdot\nu)$.
- Existence of a neutral element: $\varepsilon_p\cdot\gamma \sim \gamma \sim \gamma\cdot\varepsilon_p$.
- Existence of inverses: $\gamma\cdot\gamma^- \sim \varepsilon_p \sim \gamma^-\cdot\gamma$.

To see this one may use the homotopies that are given by the reparametrizations illustrated below:

[PICTURE OF THE ABOVE HOMOTOPIES]

□

Note that the passage to equivalence classes is essential for the statement in the above lemma: For a non-constant path $\gamma\colon [0,1] \to X$ the composite $\gamma^-\cdot\gamma$ is not constant, it is only *equivalent* to a constant path. Likewise the composition of paths is not associative, it is only so up to equivalence of paths. To simplify the notation for products of paths we will nonetheless omit brackets and make the convention that products are formed from left to right, i.e.

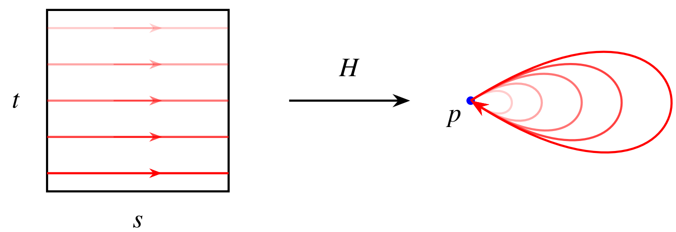
$$\gamma\cdot\mu\cdot\nu := (\gamma\cdot\mu)\cdot\nu$$

etc. Of course any other way of putting brackets will lead to an equivalent path which is in fact a reparametrization of the previous one by a piecewise linear bijection from the unit interval onto itself.

Example 2.6. We have $\pi_1(\mathbb{R}^n, 0) \simeq 0$ since any loop γ based at the origin in $\mathbb{R}^n$ is homotopic to the constant loop via the homotopy

$$H(s,t) := (1-t)\cdot\gamma(s)$$

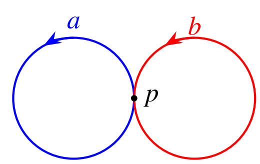
that contracts γ by rescaling it:



Example 2.7. We do not yet have tools to compute any non-trivial fundamental group, but let us already take an informal look at two intuitive examples:

a) We will prove soon that $\pi_1(S^1, 1) \simeq \mathbb{Z}$ via the isomorphism that sends $n \in \mathbb{Z}$ to the loop
$$\gamma_n\colon \quad [0,1] \to S^1, \quad s \mapsto \exp(2\pi i n s).$$

b) In general the group $\pi_1(X,p)$ can be non-abelian. For instance, let $X = S^1 \vee S^1$ be the topological space obtained by glueing two circles together at a point p, then the two loops a and b around the two circles do not commute:

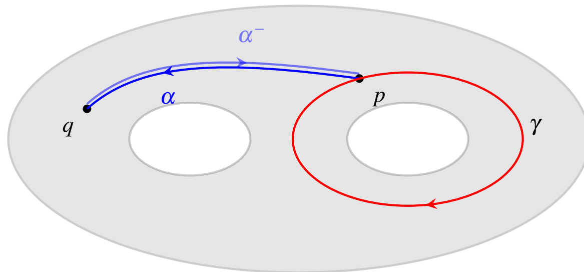


We will soon develop computational tools to prove this formally.

Before proceeding further towards the computation of fundamental groups, let us clarify how it depends on the choice of the base point. First, it is clear from the definitions that $\pi_1(X,p) = \pi_1(Y,p)$ where $Y \subset X$ denotes the path component of p in X, so we can only hope to compare fundamental groups for base points in the same path component. For those we have:

Lemma 2.8. *Let $p, q \in X$. Then any path $\alpha\colon [0,1] \to X$ from p to q gives rise to an isomorphism*
$$[-]^\alpha\colon \quad \pi_1(X,q) \xrightarrow{\sim} \pi_1(X,p), \quad [\gamma] \mapsto [\gamma]^\alpha := [\alpha^- \cdot \gamma \cdot \alpha]$$
of groups, and this isomorphism only depends on the equivalence class $[\alpha]$.

Proof. If $\gamma\colon [0,1] \to X$ is a loop based at p, then $\alpha^- \cdot \gamma \cdot \alpha$ is a loop based at $q = \alpha(1)$ as illustrated in the following picture:



By lemma 2.3 the equivalence class of this loop factors as

$$[\alpha^- \cdot \gamma \cdot \alpha] = [\alpha^-] \cdot [\gamma] \cdot [\alpha],$$

hence it only depends on the equivalence classes $[\alpha]$ and $[\gamma]$. Therefore we obtain a well-defined map

$$[-]^\alpha\colon \quad \pi_1(X, q) \to \pi_1(X, p), [\gamma] \mapsto [\alpha^- \cdot \gamma \cdot \alpha]$$

which depends only on $[\alpha]$, and an immediate computation shows that this map is a group homomorphism and has an inverse given by $[\gamma] \mapsto [\alpha \cdot \gamma \cdot \alpha^-]$. □

In particular, for a path-connected topological space X the isomorphism type of the group $\pi_1(X, p)$ does not depend on the choice of the point $p \in X$, so the following definition makes sense:

Definition 2.9. A topological space X is *simply connected* if it is path-connected and has trivial fundamental group.

This following alternative characterization of simply connected spaces explains their name. It also illustrates why in complex analysis simply connected subsets of the complex plane play an important role for the construction of primitives of holomorphic functions by path integration:

Lemma 2.10. *A topological space X is simply connected if and only if for any two points $a, b \in X$ there is a* unique *equivalence class of paths from a to b in X.*

Proof. If all points $a, b \in X$ can be connected by a path in X, then by definition X is pathwise-connected. If moreover this path is unique up to equivalence of paths, then taking $b = a$ we get that $\pi_1(X, a) \simeq \{0\}$, hence X is simply connected.

Conversely, if X is simply connected, then it is in particular path-connected, so for all $a, b \in X$ there exists a path in X from a to b. To show uniqueness of this connecting path up to equivalence, suppose we are given two paths $\gamma_0, \gamma_1\colon [0,1] \to X$ from a to b. Then

$$\gamma = \gamma_1^- \cdot \gamma_0 \quad \text{is a loop based at } a.$$

By assumption $\pi_1(X, a) \simeq \{0\}$, hence it follows that γ is equivalent to a constant loop, which is equivalent to saying that $\gamma_1 \sim \gamma_0$ as desired. □

3 Functoriality and homotopy invariance

Let $f\colon X \to Y$ be a continuous map of topological spaces. Since the composite of continuous maps is again continuous, the composite of f with any path $\gamma\colon [0,1] \to X$ gives rise to a path

$$f \circ \gamma\colon \quad [0,1] \to Y, \quad t \mapsto f(\gamma(t)).$$

In the case of loops this gives a homomorphism between fundamental groups:

Lemma 3.1. *Let $p \in X$. Then any continuous map $f\colon X \to Y$ gives rise to a group homomorphism*

$$f_*\colon \quad \pi_1(X,p) \to \pi_1(Y, f(p)), \quad [\gamma] \mapsto [f \circ \gamma].$$

Proof. It follows from the definitions that for any two paths $\gamma_1, \gamma_2\colon [0,1] \to X$ we have the implication

$$\gamma_1 \sim \gamma_2 \quad \Longrightarrow \quad f \circ \gamma_1 \sim f \circ \gamma_2.$$

Hence we get a well-defined map $f_*\colon \pi_1(X,p) \to \pi_1(Y,f(p)), [\gamma] \mapsto [f \circ \gamma]$. This is a group homomorphism:

$$\begin{aligned}((f\circ\gamma)\cdot(f\circ\mu))(s) &= \begin{cases}(f\circ\gamma)(2s) & \text{for } 0 \le s < \frac{1}{2} \\ (f\circ\mu)(2s-1) & \text{for } \frac{1}{2} \le s \le 1\end{cases} \\ &= (f\circ(\gamma\cdot\mu))(s) \qquad \text{for all } s \in [0,1]\end{aligned}$$

implies $f_*(\alpha)\cdot f_*(\beta) = f_*(\alpha\cdot\beta)$ for $\alpha = [\gamma], \beta = [\mu] \in \pi_1(X,p)$. □

Since the fundamental group only sees loops up to continuous deformation, the above homomorphism f_* should remain unchanged if we deform the map $f\colon X \to Y$ continuously. This leads us to the following notion:

Definition 3.2. Let $f_0, f_1\colon X \to Y$ be two continuous maps. A *homotopy* between them is a continuous map

$$H\colon\ X \times [0,1] \to Y \ \text{ with } H(x,0) = f_0(x) \text{ and } H(x,1) = f_1(x) \text{ for all } x \in X.$$

We say that f_0 and f_1 are *homotopic* if there exists a homotopy between them.

Homotopic map induce ‘the same’ maps on fundamental groups, provided that we keep track of what happens to the respective base points:

Lemma 3.3. *Let $f_0, f_1\colon X \to Y$ be two continuous maps, and let $H\colon X \times [0,1] \to Y$ be a homotopy between them. Then for any point $p \in X$ we have a commutative diagram*

$$\begin{tikzcd}[column sep=small] & \pi_1(X,p) \arrow[dl, "f_{0,*}"'] \arrow[dr, "f_{1,*}"] & \\ \pi_1(Y,f_0(p)) \arrow[rr, "\cong", "{[\mu]\mapsto[\mu]^\alpha}"'] & & \pi_1(Y,f_1(p)) \end{tikzcd}$$

where the vertical arrow is induced by the path $\alpha\colon [0,1]\to Y, t\mapsto H(p,1-t)$.

Proof. Let $\gamma\colon [0,1]\to X$ be a loop based at the point p. We must show that the two loops

$$f_{1,*}(\gamma) \quad \text{and} \quad \alpha^-\cdot f_{0,*}(\gamma)\cdot\alpha$$

are equivalent. For this we must find a homotopy between them through loops based at $f_1(p)$. This can be done as follows:

[PICTURE OF THE HOMOTOPY DESCRIBED BELOW]

For $t\in[0,1]$, consider the path

$$\alpha_t\colon \quad [0,1] \to Y, \quad s \mapsto \alpha(st)$$

given by the initial segment of length t in the path α, reparametrized so that its domain is again the unit interval. This path begins at the point $\alpha_t(0)=f_1(p)$ and ends at $\alpha_t(1)=\alpha(t)$. On the other hand, for each t we can define a loop based at this end point via

$$\gamma_t\colon \quad [0,1] \to Y, \quad s \mapsto H(\gamma(s),1-t).$$

The loops $\delta_t := \alpha_t^-\cdot\gamma_t\cdot\alpha_t$ are then all based at the same point $f_1(p)$, and one checks that the map

$$[0,1]\times[0,1] \to Y \quad \text{given by} \quad (s,t) \mapsto \delta_t(s)$$

is a homotopy between the loops $\delta_0 = f_{1,*}(\gamma)$ and $\delta_1 = \alpha^-\cdot f_{0,*}(\gamma)\cdot\alpha$. $\square$

In practice we often fix a point $p\in X$ and call the pair (X,p) a *pointed space*. By a map

$$f\colon \quad (X,p) \to (Y,q)$$

of pointed spaces we mean a continuous map $f\colon X\to Y$ with $f(p)=q$. For such maps the appropriate notion of deformation is homotopy through maps of pointed spaces. This is the special case $Z=\{p\}$ of the following notion:

Definition 3.4. Let $f_0, f_1 \colon X \to Y$ be continuous maps, and let $H \colon X \times [0,1] \to Y$ be a homotopy between them. Suppose that $f_0|_Z = f_1|_Z$ for some subspace $Z \subset X$. We call H a *homotopy rel(ative to) Z* if

$$H(z,t) \quad \text{is independent of } t \text{ for each } z \in Z.$$

We then say that the two maps f_0 and f_1 are *homotopic rel(ative to) Z*.

Example 3.5. With the above notation,

a) two paths $\gamma_0, \gamma_1 \colon X = [0,1] \to Y$ are equivalent in the sense of definition 2.1 if and only if they are homotopic relative to $Z = \{0,1\} \subset X = [0,1]$.
b) two maps $f_0, f_1 \colon (X,p) \to (Y,q)$ of pointed spaces are homotopic through maps of pointed spaces if and only if they are homotopic relative to $Z = \{p\} \subset X$. In this case we also omit the set brackets and call the maps homotopic relative to p.

In the category of pointed spaces we therefore obtain:

Corollary 3.6. *If two maps of pointed spaces $f_0, f_1 \colon (X,p) \to (Y,q)$ are homotopic relative to the base point $Z = \{p\}$, then they induce on fundamental groups the same homomorphism*

$$f_{0,*} = f_{1,*} \colon \quad \pi_1(X,p) \to \pi_1(Y,q).$$

Proof. Saying that f_0 and f_1 are homotopic relative to p means by definition that we can choose a homotopy between them such that in corollary 3.3 the path α is constant. But then the map $\pi_1(Y,q) \to \pi_1(Y,q), [\mu] \mapsto [\mu]^\alpha$ is the identity. □

The above suggests that the notion of homeomorphism of topological spaces still keeps too much irrelevant information: What matters in topology is not whether a continuous map is invertible, but rather whether it is invertible up to homotopy. This leads to the following notion:

Definition 3.7. A continuous map $f \colon X \to Y$ is called a *homotopy equivalence* if there is a continuous map $g \colon Y \to X$ such that the composites $g \circ f$ and $f \circ g$ are homotopic to the identity maps id_X respectively id_Y. If such a homotopy equivalence exists, we also write

$$X \simeq Y$$

and say that the topological spaces X and Y are *homotopy equivalent*.

A topological space X is called *contractible* if it is homotopy equivalent to the space $Y = \{*\}$ consisting of a single point. The terminology is explained by the following observation:

Lemma 3.8. *A topological space X is contractible if and only if there exists $p \in X$ and a continuous map*

$$H \colon \quad X \times [0,1] \to X \quad \textit{with} \quad H(x,0) = p \quad \textit{and} \quad H(x,1) = x \quad \textit{for all} \quad x \in X.$$

In particular, every contractible space X is simply connected.

Proof. For a topological space $Y = \{*\}$ consisting of a single point,

- there is a unique map $f\colon X \to Y$, and
- giving a map $g\colon Y \to X$ amounts to giving a point $p = g(*) \in X$.

These maps are continuous for trivial reasons. Moreover we have $f \circ g = id_Y$ and the map $g \circ f\colon X \to \{p\} \subset X$ is constant. Therefore X is homotopy equivalent to a point if and only if the identity map id_X is homotopic to a constant map, which happens if and only if there exists a homotopy H with the stated properties. In this case every point $a \in X$ can be connected to the point p by the path

$$\gamma_a\colon [0,1] \to X, s \mapsto H(a,s),$$

hence all $a, b \in X$ can be connected by the path $\gamma_a^- \cdot \gamma_b$. So X is path-connected, and similarly one sees

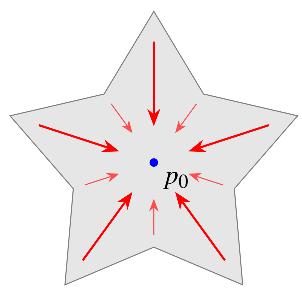
$$\pi_1(X,p) \simeq \{0\}$$

because any loop γ at p can be contracted via the homotopy $(s,t) \mapsto H(\gamma(s),t)$. □

Example 3.9. A subset $X \subset \mathbb{R}^n$ is called *star-shaped* if there is a point $p_0 \in X$ such that

$$\lambda p_0 + (1-\lambda)p \in X \quad \text{for every } p \in X \text{ and every } \lambda \in [0,1].$$

It follows from lemma 3.8 that any such subset is contractible:



Note that a star-shaped subset of $\mathbb{R}^n$ need not be convex: In fact, a subset $X \subset \mathbb{R}^n$ is convex if and only if it is star-shaped with respect to *every* point $p_0 \in X$.

Remark 3.10. As we noted at the end of lemma 3.8, any contractible space is simply connected. The converse is not true: For example, one easily checks that the sphere

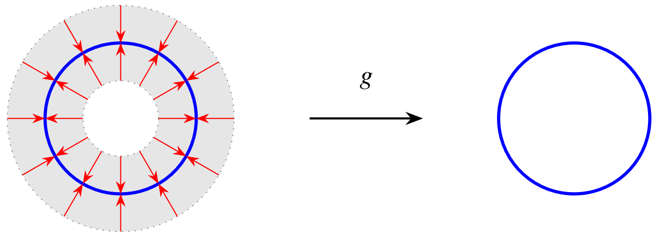
$$S^2 = \{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$$

is simply connected, but we will see later that it is not contractible.

Example 3.11. The unit circle is homotopy equivalent to an annulus, indeed the inclusion

$$f\colon \quad X = \{z \in \mathbb{C} \mid |z| = 1\} \hookrightarrow Y = \{z \in \mathbb{C} \mid \tfrac{1}{2} < |z| < \tfrac{3}{2}\}$$

is a homotopy equivalence — an inverse up to homotopy is $g\colon Y \to X, z \mapsto z/|z|$:



Note that by the same argument the unit circle is also homotopy equivalent to a Möbius strip. These two examples already give some idea about what information we lose in passing from homeomorphism to homotopy equivalence. However, all information about fundamental groups is kept:

Lemma 3.12. *If X and Y are homotopy equivalent topological spaces and both of them are path-connected, then their fundamental groups are isomorphic.*

Proof. By assumption there exist $f\colon X \to Y$ and $g\colon Y \to X$ such that $g \circ f$ and $f \circ g$ are homotopic to the identity map. Since the identity map of a topological space induces the identity map on fundamental groups, it then follows from lemma 3.3 that the two maps

$$\begin{aligned} (g \circ f)_*\colon \quad & \pi_1(X, x_0) \to \pi_1(X, x_1) \quad \text{with} \quad x_1 = g(f(x_0)) \\ (f \circ g)_*\colon \quad & \pi_1(Y, y_0) \to \pi_1(Y, y_1) \quad \text{with} \quad y_1 = f(g(y_0)) \end{aligned}$$

are given by conjugation with certain paths α and β from x_0 to x_1 and from y_0 to y_1 respectively. In particular these two maps are isomorphisms of groups. This holds for any choice of base points $x_0 \in X$ and $y_0 \in Y$. Since for path-connected spaces the fundamental groups for different base points are isomorphic, we may as well assume that $y_0 = f(x_0)$. Then also $g(y_0) = g(f(x_0)) = x_1$ and hence $f(x_1) = f(g(y_0)) = y_1$, and the two above maps factor as

$$\begin{aligned} (g \circ f)_* &= g_* \circ f_*\colon \quad \pi_1(X, x_0) \to \pi_1(Y, y_0) \to \pi_1(X, x_1), \\ (f \circ g)_* &= f_* \circ g_*\colon \quad \pi_1(Y, y_0) \to \pi_1(X, x_1) \to \pi_1(Y, y_1). \end{aligned}$$

The first row being an isomorphism implies that the map $g_*\colon \pi_1(Y, y_0) \to \pi_1(X, x_1)$ is surjective, while the second row being an isomorphism implies that the same map is injective. So $g_*\colon \pi_1(Y, y_0) \to \pi_1(X, x_1)$ is an isomorphism and we are done. □

4 The fundamental group of the circle

As a first nontrivial example we will now compute the fundamental group of the unit circle

$$S^1 = \{z \in \mathbb{C} \mid |z| = 1\}.$$

Our computation will be a special case of arguments that we will develop later in the chapter about covering spaces, but we here give an independent discussion of this special case. For each $n \in \mathbb{Z}$, the path

$$\gamma_n: \quad [0,1] \to S^1, \quad s \mapsto \exp(2\pi i n s)$$

is a loop based at the point 1, and we show:

Theorem 4.1. *We have an isomorphism $\varphi: \mathbb{Z} \xrightarrow{\sim} \pi_1(S^1, 1)$ given by $n \mapsto [\gamma_n]$.*

Proof. One easily checks the identity $[\gamma_m] \cdot [\gamma_n] = [\gamma_{m+n}]$, hence the map φ is a group homomorphism. To show that it is in fact an isomorphism, consider the continuous map

$$p: \quad \mathbb{R} \to S^1, \quad s \mapsto \exp(2\pi i s).$$

It wraps each interval of length one once around the circle counterclockwise. For each $n \in \mathbb{Z}$ one has

$$\gamma_n = f \circ \tilde{\gamma}_n \quad \text{for the path} \quad \tilde{\gamma}_n: \quad [0,1] \to \mathbb{R}, \quad s \mapsto ns.$$

We also say that $\tilde{\gamma}_n$ is a path that *lifts* γ_n. Such lifts exist more generally:

For every path $\gamma: [0,1] \to S^1$ starting at the base point $\gamma(0) = 1$, there is a unique path

$$\tilde{\gamma}: \quad [0,1] \to X \quad \textit{with} \quad \tilde{\gamma}(0) = 0$$

such that $\gamma = p \circ \tilde{\gamma}$ as shown in the following commutative diagram:

$$\begin{tikzcd} & \mathbb{R} \arrow[d, "p"] \\ {[0,1]} \arrow[ur, "\exists!\tilde{\gamma}"] \arrow[r, "\gamma"'] & S^1 \end{tikzcd} \qquad (*)$$

To see this, note that for any small enough path-connected open subset $U \subset S^1$ the preimage is a disjoint union

$$p^{-1}(U) = \bigsqcup_{n \in \mathbb{Z}} V_n$$

where each $V_n = V_0 + n \subset \mathbb{R}$ is an interval of length < 1 on which the map p restricts to a homeomorphism

$$p|_{V_n}: \quad V_n \xrightarrow{\sim} U.$$

So any path in $U \subset S^1$ lifts uniquely to a path in each of the intervals $V_n \subset \mathbb{R}$, and the choice of n is determined by a choice of the initial point of the lift. Let us call an open subset $U \subset S^1$ with the above properties *small*. Now by compactness $[0,1]$ can be covered by finitely many preimages $\gamma^{-1}(U)$ with $U \subset S^1$ small, hence γ is a composition of finitely many paths with image contained in small subsets. Each of those finitely many paths has a unique lift starting from a given lift of the starting point. The lifting property $(*)$ then follows by continuously gluing these unique lifts with given starting points for each of the finitely many appearing small subsets.

Having checked the lifting property $(*)$, let us now return to the proof of the theorem. To any loop $\gamma\colon [0,1] \to S^1$ based at 1 we may attach the end point $\tilde{\gamma}(1)$ of the unique lift $\tilde{\gamma}$ that we constructed above (starting at $\tilde{\gamma}(0) = 0$). Applying a lifting argument similar to the above, by suitably subdividing the compact set $[0,1] \times [0,1]$ which is the domain of a homotopy of loops we see that the value

$$\tilde{\gamma}(1) \in p^{-1}(1) = \mathbb{Z}$$

only depends on the equivalence class $[\gamma] \in \pi_1(S^1, 1)$. So we obtain a well-defined map

$$\deg\colon \quad S^1 \to \mathbb{Z}, \quad [\gamma] \mapsto \tilde{\gamma}(1)$$

which we call the *degree map*. One easily checks that the degree map is a group homomorphism. Moreover, by construction we have

$$\deg \circ \varphi = id_{\mathbb{Z}}\colon \quad \mathbb{Z} \xrightarrow{\varphi} \pi_1(S^1, 1) \xrightarrow{\deg} \mathbb{Z}.$$

Hence it suffices to show that deg is injective. To see this, suppose $\gamma_0, \gamma_1\colon [0,1] \to S^1$ are two loops with $\deg(\gamma_0) = \deg(\gamma_1)$. Then the composite

$$\gamma := \gamma_1^- \cdot \gamma_0 \quad \text{has degree} \quad \deg(\gamma) = \deg(\gamma_0) - \deg(\gamma_1) = 0.$$

By definition this means that the unique lift $\tilde{\gamma}\colon [0,1] \to \mathbb{R}$ starting at the origin also ends at the origin, i.e. it is a loop in $\mathbb{R}$. Since the real line is contractible, it follows that $\tilde{\gamma}$ is homotopic to a constant loop. But then $\gamma = p \circ \tilde{\gamma}$ is also homotopic to a constant loop, which implies $[\gamma_1] = [\gamma_0]$ as desired. □

As a first application of the above, we can give a simple proof of the fact that the unit disk

$$D := \{z \in \mathbb{C} \mid |z| \le 1\}$$

cannot be retracted onto its outer boundary $\partial D = S^1$ without cutting a hole into it:

Corollary 4.2. *There is no continuous map $r\colon D \to \partial D$ with $r|_{\partial D} = id$.*

Proof. Let $i\colon \partial D \hookrightarrow D$ be the inclusion of the unit circle as the boundary of the unit disk. If there exists a continuous map $r\colon D \to \partial D$ with $r \circ i = id$, then by functoriality we have

$$r_* \circ i_* = id\colon \quad \pi_1(\partial D, 1) \xrightarrow{i_*} \pi_1(D, 1) \xrightarrow{r_*} \pi_1(\partial D, 1).$$

But $\pi_1(D,1) = \{0\}$ since D is contractible, while $\pi_1(\partial D, 1) = \pi_1(S^1, 1) \simeq \mathbb{Z} \neq \{0\}$ by the previous lemma. This gives a contradiction. □

Corollary 4.3 (Brouwer fixed point theorem). *Every continuous map $f\colon D \to D$ has at least one fixed point, i.e. there exists a point $p \in D$ with $f(p) = p$.*

Proof. Suppose $f\colon D \to D$ is a continuous map with no fixed point. Then $f(p) \neq p$ for any $p \in D$, so we can define a map $r\colon \quad D \to \partial D = S^1$ by sending a point p to be the unique point of intersection with S^1 of the half-ray from p to $f(p)$. Since f was assumed to be continuous, one easily checks that r is continuous in contradiction to the previous corollary. □

Another nice application of our computation of the fundamental group of the circle is a simple proof of the following result:

Corollary 4.4 (Fundamental theorem of algebra). *Let*

$$f(x) \;=\; x^n + c_{n-1}x^{n-1} + \cdots + c_0 \;\in\; \mathbb{C}[x]$$

be a monic polynomial of degree $n > 0$ with coefficients $c_0, \dots, c_{n-1} \in \mathbb{C}$. Then we have

$$f(a) \;=\; 0 \quad \textit{for some} \quad a \;\in\; \mathbb{C}.$$

Proof. Suppose that $f(a) \neq 0$ for all $a \in \mathbb{C}$. Using this for $a \in S^1$ we may define a continuous map

$$g\colon \quad S^1 \;\to\; S^1 \quad \text{by} \quad g(a) \;:=\; \frac{f(a)}{|f(a)|}.$$

We define the degree $\deg(g) \in \mathbb{Z}$ of this map to be the image of $[g \circ \exp] \in \pi_1(S^1, 1)$ under the isomorphism

$$\deg\colon \quad \pi_1(S^1, 1) \;\xrightarrow{\sim}\; \mathbb{Z}$$

from theorem 4.1. We will compute this degree in two different ways:

a) Since $f(a) \neq 0$ for $|a| < 1$, we may define a homotopy $h\colon S^1 \times [0,1] \to S^1$ by the formula

$$h(s,t) \;:=\; \frac{f_t(s)}{|f_t(s)|} \quad \text{for} \quad f_t(s) \;:=\; f(st).$$

This is a homotopy from g to a constant map, hence we get $\deg(g) = 0$.

b) Since $f(a) \neq 0$ for $|a| > 1$, we may define a homotopy $H\colon S^1 \times [0,1] \to S^1$ by the formula

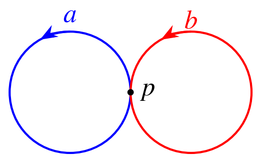
$$\begin{aligned} H(s,t) \;:=\; \frac{F_t(s)}{|F_t(s)|} \quad \text{for} \quad F_t(s) \;&:=\; t^n f(s/t) \\ &=\; s^n + t(c_1 s^{n-1} + t c_2 s^{n-2} + \cdots + t^{n-1} c_n). \end{aligned}$$

This is a homotopy from g to the loop γ_n given by $s \mapsto s^n$, hence $\deg(g) = n$.

For $n > 0$ this gives a contradiction, so we must have $f(a) = 0$ for some $a \in \mathbb{C}$. □

5 The Seifert – van Kampen theorem

We now discuss an important tool to compute fundamental groups of topological spaces in terms of open covers. For instance, we will see that the fundamental group of the figure eight is a free group on the two generators a,b that are shown in the following picture:



In other words, any element of $\pi_1(X,p)$ can be written uniquely as a finite product

$$g_1 g_2 \cdots g_k \quad \text{with} \quad k \in \mathbb{N}_0 \quad \text{and} \quad g_1,\dots,g_k \in \{a,b\}.$$

This is an instance of the following abstract construction in group theory:

Definition 5.1. The *free product* $G_1 * G_2$ of two groups G_1 and G_2 is the set of all finite formal expressions

$$g_1 g_2 \cdots g_k \quad \text{with} \quad k \in \mathbb{N} \quad \text{and} \quad g_i \in G_1 \sqcup G_2,$$

modulo the equivalence relation $\sim$ that is generated by the following simplification rules for $\alpha = 1,2$:

- Any occurence of the neutral element $g_i = e \in G_\alpha$ may be removed.
- For any two consecutive elements $g_i, g_{i+1} \in G_\alpha$ that both lie in the group G_α, the formal product of these two elements may be replaced by their actual product in that group, i.e.

$$\cdots g_i g_{i+1} \cdots \sim \cdots g \cdots \quad \text{for the element} \quad g = g_i g_{i+1} \in G_\alpha.$$

One immediately verifies that $G_1 * G_2$ is a group with respect to the product defined by

$$(g_1 g_2 \cdots g_k) \cdot (h_1 h_2 \cdots h_l) := g_1 g_2 \cdots g_k h_1 h_2 \cdots h_l.$$

Note that this group is non-abelian as soon as G_1 and G_2 are both non-trivial!

We have given an explicit construction, but the free product $G_1 * G_2$ can also be characterized more conceptually by the following universal property: For any other group G and any pair of group homomorphisms $\varphi_\alpha\colon G_\alpha \to G$, there is a unique group homomorphism

$$\varphi\colon \quad G_1 * G_2 \longrightarrow G$$

with $\varphi \circ \iota_\alpha = \varphi_\alpha$ for the two inclusions $\iota_\alpha\colon G_\alpha \hookrightarrow G_1 * G_2$. One easily sees that this property determines the free product uniquely up to canonical isomorphism. In practice we often want a variant where we only consider pairs of homomorphisms that agree on some given common subgroup of the two groups G_1 and G_2. This leads to the following notion:

Definition 5.2. Let H be a group. The *pushout* or *amalgamated product* of G_1, G_2 with respect to two group homomorphisms $f_\alpha\colon H \to G_\alpha$ ($\alpha = 1,2$) is defined to be the quotient

$$G_1 *_H G_2 \;:=\; (G_1 * G_2)/N,$$

where we denote by

$$N \;\trianglelefteq\; G_1 * G_2$$

the smallest normal subgroup containing $f_1(h)^{-1} \cdot f_2(h)$ for all $h \in H$.

Lemma 5.3. *The pushout has the following universal property: For any group G and any pair of homomorphisms $\varphi_\alpha\colon G_\alpha \to G$ with $\varphi_1 \circ f_1 = \varphi_2 \circ f_2$, there is a unique homomorphism*

$$\varphi\colon \quad G_1 *_H G_2 \longrightarrow G$$

such that $\varphi \circ \iota_\alpha = \varphi_\alpha$ for the natural homomorphisms $\iota_\alpha\colon G_\alpha \to G_1 *_H G_2$:

$$\begin{tikzcd}
H \arrow[r, "f_1"] \arrow[d, "f_2"'] & G_1 \arrow[d, "\iota_1"] \arrow[ddr, bend left, "\varphi_1"] & \\
G_2 \arrow[r, "\iota_2"'] \arrow[drr, bend right, "\varphi_2"'] & G_1 *_H G_2 \arrow[dr, dotted, "\exists!\varphi"] & \\
& & G
\end{tikzcd}$$

Proof. Left as an exercise. □

Let us now come back to topology. The goal of this section is to compute the fundamental group of a topological space in terms of an open cover:

Theorem 5.4 (Seifert – van Kampen). *Consider a topological space with an open cover*

$$X = U_1 \cup U_2$$

where $U_1, U_2 \subset X$ are path-connected open subsets whose intersection $U_{12} = U_1 \cap U_2$ is also path-connected. Then for any $p \in U_{12}$ we have

$$\pi_1(X,p) \;\simeq\; \pi_1(U_1,p) *_{\pi_1(U_{12},p)} \pi_1(U_2,p).$$

In fact, at the end of this section we will prove a generalized version of this theorem for covers involving more than two open subsets. But before coming to the proof, let us illustrate the theorem by some simple examples:

Example 5.5. For the above theorem the path-connectedness of $U_{12} = U_1 \cap U_2$ is crucial: For example, the unit circle $S^1 \subset \mathbb{C}$ can be covered by the simply connected open subsets

$$U_1 = \{z \in \mathbb{C} \mid \mathrm{Im}(z) > -\tfrac{1}{2}\} \quad \text{and} \quad U_2 = \{z \in \mathbb{C} \mid \mathrm{Im}(z) < \tfrac{1}{2}\}.$$

Here the groups $\pi_1(U_1, p)$ and $\pi_1(U_2, p)$ are trivial, hence so is their amalgamated product. But the Seifert – van Kampen theorem does not apply: Here $U_1 \cap U_2$ is not path-connected. Indeed

$$\pi_1(S^1, p) \simeq \mathbb{Z} \neq \{0\},$$

so the conclusion of the Seifert – van Kampen theorem fails in this case!

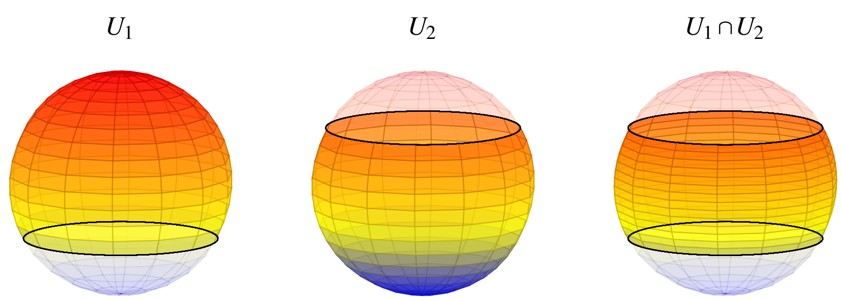
Example 5.6. For $n \geq 2$ the unit sphere $S^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1^2 + \cdots + x_n^2 = 1\}$ is simply connected: It is obviously path-connected, hence we only need to verify that

$$\pi_1(S^n, p) \simeq \{0\} \quad \text{for all} \quad n > 1.$$

This follows easily by decomposing the sphere as a union $S^n = U_1 \cup U_2$ of the two open subsets

$$\begin{aligned} U_1 &= \{(x_1, \dots, x_n) \in S^n \mid x_n > -\tfrac{1}{2}\}, \\ U_2 &= \{(x_1, \dots, x_n) \in S^n \mid x_n < +\tfrac{1}{2}\}. \end{aligned}$$

as shown in the following picture:



Each of these subsets is homeomorphic to an open ball in $\mathbb{R}^{n-1}$ via a stereographic projection, so $\pi_1(U_1, p) \simeq \pi_1(U_2, p) \simeq \{0\}$. In contrast to the previous example, their intersection

$$U_{12} = U_1 \cap U_2 = \{(x_1, \dots, x_n) \in S^n \mid -\tfrac{1}{2} < x < \tfrac{1}{2}\}$$

is path-connected. Hence the Seifert – van Kampen theorem gives

$$\pi_1(S^n, p) \simeq \pi_1(U_1, p) *_{\pi_1(U_{12}, p)} \pi_1(U_2, p) \simeq \{0\}.$$

Example 5.7. If $X = U_1 \cup U_2$ with path-connected open subsets $U_1, U_2 \subset X$ whose intersection U_{12} is simply connected, then the Seifert – van Kampen theorem says

$$\pi_1(X, p) \simeq \pi_1(U_1, p) * \pi_1(U_2, p) \quad \text{for any} \quad p \in U_1 \cap U_2.$$

For instance, the fundamental group of the figure eight is the free group

$$\pi_1(\infty, p) \simeq \mathbb{Z} * \mathbb{Z}$$

on two generators as mentioned earlier. In general, define the *wedge sum* of two pointed topological spaces (X_α, p_α) to be the topological space

$$X_1 \vee X_2 := (X_1 \sqcup X_2)/\sim$$

where $\sim$ is the equivalence relation that identifies p_1 with p_2. If each $p_\alpha \in X_\alpha$ is a deformation retract of some open neighborhood

$$V_\alpha \subset X_\alpha$$

in the sense that the inclusion $\{p_\alpha\} \hookrightarrow V_\alpha$ is a homotopy equivalence, then the same argument as above shows that

$$\pi_1(X_1 \vee X_2, p) \simeq \pi_1(X_1, p_1) * \pi_1(X_2, p_2).$$

Inductively we then also obtain for any finite number of pointed spaces (X_α, p_α) whose base points are deformation retracts of some open neighborhoods, their wedge sum has fundamental group

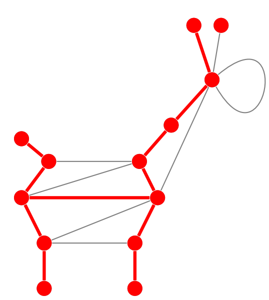
$$\pi_1(X_1 \vee \cdots \vee X_n, p) \simeq \pi_1(X_1, p_1) * \cdots * \pi_n(X_n, p_n),$$

where the iterated free product on the right hand side can be taken in any order since $*$ is associative. For instance, for a bouquet $X = S^1 \vee \cdots \vee S^1$ of n circles we get

$$\pi_1(S^1 \vee \cdots \vee S^1, p) \simeq \mathbb{Z} * \cdots * \mathbb{Z}.$$

This group is called the *free group on n generators*. It appears for instance as the fundamental group of finite connected graphs as in the following example:

Example 5.8. Let X be a *finite connected graph*, i.e. a connected topological space that is obtained from a finite set of vertices by attaching a finite number of edges such that each edge connects a vertex either to itself or to another vertex. By a *tree* in X we mean a subgraph $X_0 \subset X$ in which any two vertices are connected by a *unique* path, i.e. a subgraph which is simply connected. Any finite connected graph contains a maximal tree:



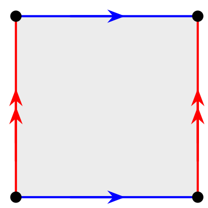
Moreover, any maximal tree $X_0 \subset X$ contains all the vertices of X. Fixing such a maximal tree, let n be the number of edges in $X \smallsetminus X_0$. Since the endpoints of each edge lie on the chosen maximal tree, we obtain by contracting this tree to a point a map

$$f\colon \quad X \to S^1 \vee \cdots \vee S^1$$

to a bouquet of n circles. One may check that f is a homotopy equivalence, so by the previous example we get that

$$\pi_1(X,p) \simeq \mathbb{Z} * \cdots * \mathbb{Z}.$$

Example 5.9. A torus X can be obtained by taking a square and identifying pairs of opposite sides in a way preserving the orientation:



Consider the open cover $X = U_1 \cup U_2$ where $U_1 \subset X$ is a small open disk and $U_2 \subset X$ is the complement of a slightly smaller closed disk. Thus the subset U_1 is contractible while U_2 is homotopy equivalent to $S^1 \vee S^1$ as one may see by radial projection to the boundary of the unit square: Indeed, with the given identifications the boundary of the unit square becomes a bouquet of two circles. The intersection $U_{12} = U_1 \cap U_2$ is homotopy equivalent to a circle. If $\gamma\colon [0,1] \to U_{12}$ is a loop going round this circle once and $a, b \in \pi_1(U_2, *)$ are the two loops which are given by the two sides of the unit square, we may write the morphism induced by the inclusion $U_{12} \hookrightarrow U_2$ as

$$\pi_1(U_{12},*) = \langle \gamma \rangle \simeq \mathbb{Z} \to \pi_1(U_2,*) = \langle a,b \rangle \simeq \mathbb{Z} * \mathbb{Z}, \quad \gamma \mapsto aba^{-1}b^{-1}$$

So the Seifert – van Kampen theorem gives

$$\begin{aligned} \pi_1(X,*) &= \pi_1(U_1,*) *_{\pi_1(U_{12},*)} \pi_1(U_2,*) \\ &= \pi_1(U_2,*)/\pi_1(U_1 \cap U_2,*) = \langle a,b \mid aba^{-1}b^{-1} = 1 \rangle \simeq \mathbb{Z}^2. \end{aligned}$$

6 Proof of the Seifert – van Kampen theorem

The theorem of Seifert – van Kampen in the version that we stated above suffers from two drawbacks:

- the intersection $U_1 \cap U_2$ had to be path-connected, and
- we could only do covers by *two* open subsets at a time.

We now explain a generalization that deals with both issues in a more conceptual way, and then directly prove this generalized version.

The issue with path-connectedness is related to the fact that we want to fix the same base point for all fundamental groups: If $U_1 \cap U_2$ is not path-connected, then in which path component should we choose our base point? Conceptually it is better to use paths between arbitrary points. Such paths cannot always be composed, but there is a notion of a group-like object with a partially defined multiplication:

Definition 6.1. A *groupoid* G consists of the following data:

a) a non-empty set X of objects,
b) for all $a, b \in X$, a set $G_{a,b}$ of morphisms (for some a, b this set may be empty),
c) for all $a, b, c \in X$, composition maps

$$G_{a,b} \times G_{b,c} \to G_{a,c} \quad (\gamma, \mu) \mapsto \gamma \cdot \mu$$

satisfying associativity $(\gamma \cdot \mu) \cdot \lambda = \gamma \cdot (\mu \cdot \lambda)$ for any three composable γ, μ, λ,
d) for each $a \in X$, a neutral element $\varepsilon_a \in G_{a,a}$ such that

$$\varepsilon_a \cdot \gamma = \gamma \quad \text{and} \quad \mu \cdot \varepsilon_a = \mu \quad \text{for all } \gamma \in G_{a,b}, \mu \in G_{b,a},$$

e) for all $a, b \in X$, an inversion map $G_{a,b} \to G_{b,a}, \gamma \mapsto \gamma^{-1}$ such that

$$\gamma \cdot \gamma^{-1} = e_a \quad \text{and} \quad \gamma^{-1} \cdot \gamma = e_b \quad \text{for all } \gamma \in G_{a,b}.$$

In categorical language, items a) – d) amount to having a *small category*, and a groupoid is simply a small category in which every morphism is invertible. Any group can be viewed as a groupoid with only a single object. Groupoids with more objects arise naturally as follows:

Definition 6.2. Let X be a topological space. For a non-empty subset $X_0 \subset X$ we define the *fundamental groupoid*

$$G = \pi_1(X, X_0)$$

with base points in X_0 to be the following groupoid:

- Objects are the points in the subset $X_0 \subset X$.
- Morphisms between points $a, b \in X$ are given by

$$\pi_1(X, X_0)_{a,b} := \{\text{paths } \gamma \colon [0,1] \to X \text{ from } a \text{ to } b\} / \sim$$

  where $\sim$ is the equivalence relation of homotopy through paths from a to b.
- Neutral elements are given by constant paths, inverses by inverse paths.

When $X_0 = X$, we drop the base points from the notation and call $\Pi(X) = \pi_1(X, X)$ the *fundamental groupoid of* X. It will be useful in our discussion of covering spaces later. At the opposite extreme, we may also take $X_0 = \{p\}$ to be a singleton set. In this case we recover the notion of fundamentl groups:

Example 6.3. If $X_0 = \{p\}$ consists of a single point, then

$$\pi_1(X, \{p\}) = \pi_1(X, p)$$

when the group on the right hand side is viewed as a groupoid with a single object.

Using fundamental groupoids, we will get a version of the Seifert – van Kampen theorem without any path-connectedness assumptions. For this we have to view the fundamental groupoid as a functor from pairs of topological spaces to groupoids:

Remark 6.4. By a *pair of topological spaces* we mean pair (X, X_0), where X is a topological space and $X_0 \subset X$ is a subset endowed with the subspace topology. By defininition a *map of pairs*

$$f\colon \quad (V, V_0) \to (X, X_0)$$

is a continuous map $f\colon V \to X$ with $f(V_0) \subset X_0$. Any such map of pairs induces a morphism of groupoids

$$f_*\colon \quad \pi_1(V, V_0) \to \pi_1(X, X_0),$$

where the notion of morphisms of groupoids is defined in the obvious way.

In particular, given a pair of topological spaces (X, X_0) and a cover $X = U_1 \cup U_2$ by open subsets $U_1, U_2 \subset X$, we can apply the above remark to all the maps of pairs in the diagram

$$\begin{tikzcd} (U_{12}, U_{12} \cap X_0) \arrow[r] \arrow[d] & (U_2, U_2 \cap X_0) \arrow[d] \\ (U_1, U_1 \cap X_0) \arrow[r] & (X, X_0) \end{tikzcd}$$

with $U_{12} = U_1 \cap U_2$ to get a corresponding diagram of groupoids:

$$\begin{tikzcd} \pi_1(U_{12}, U_{12} \cap X_0) \arrow[r] \arrow[d] & \pi_1(U_2, U_2 \cap X_0) \arrow[d] \\ \pi_1(U_1, U_1 \cap X_0) \arrow[r] & \pi_1(X, X_0) \end{tikzcd}$$

The condition $X = U_1 \cup U_2$ amounts to saying that the first diagram is a pushout in the category of pairs of topological spaces, in the sense that the universal property analogous to the one in lemma 5.3 holds. The essence of the Seifert – van Kampen theorem is that under suitable assumptions the second diagram is a pushout in the category of groupoids, or as a slogan: The fundamental groupoid commutes with pushouts. This formulation of the theorem immediately generalizes to arbitrary open covers

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

if we replace the pushout by the notion of a coequalizer:

Definition 6.5. Let $\mathcal{C}$ be a category. Suppose we are given an index set I together with

- an object $U_i \in \mathcal{C}$ for every $\alpha \in I$,
- an object $U_{ij} = U_{ji} \in \mathcal{C}$ with a morphism f_{ij}: $U_{ij} \to U_i$ for all $i, j \in I$

Then an object $X \in \mathcal{C}$ with morphisms f_i: $U_i \to X$ is a *coequalizer* of the f_{ij} if it has the following universal property:

- We have $f_i \circ f_{ij} = f_i \circ f_{ji}$ for all $i, j \in I$.
- For any other $Z \in \mathcal{C}$ and g_i: $U_i \to Z$ such that $g_i \circ f_{ij} = g_i \circ f_{ji}$ for all $i, j \in I$, there is a unique morphism

$$g: \quad X \to Z \quad \text{with} \quad g_i = g \circ f_i \quad \text{for all } i \in I.$$

If such a coequalizer exists, it is determined uniquely up to unique isomorphism and we write

$$X = \operatorname{colim}_{i,j \in I}\Big(U_{ij} \rightrightarrows U_i\Big).$$

For $I = \{1,2\}$ we also call say that X is the *pushout* of the diagram $U_1 \leftarrow U_{12} \to U_2$.

For instance, taking $\mathcal{C}$ to be the category of topological spaces, we can view any topological space as the colimit of an arbitrary open cover. The general form of the Seifert – van Kampen theorem says that in this case the fundamental groupoid of the space is the colimit of the fundamental groupoids of the covering sets:

Theorem 6.6. *Let (X, X_0) be a pair of topological spaces, and let $X = \bigcup_{i \in I} U_i$ be a cover by open subsets such that the subset X_0 meets every path component of the intersections $U_{ijk} = U_i \cap U_j \cap U_k$ for all triples of indices $i, j, k \in I$. Then*

$$\pi_1(X, X_0) = \operatorname{colim}_{i,j \in J}\Big(\pi_1(U_{ij}, U_{ij} \cap X_0) \rightrightarrows \pi_1(U_i, U_i \cap X_0)\Big).$$

Proof. We have to verify that $\pi_1(X, X_0)$ has the universal property required for the coequalizer. So let G be any groupoid with morphisms

$$g_i: \quad \pi_1(U_i, U_i \cap X_0) \to G$$

of groupoids such that for all $i, j \in I$ the following diagrams commute:

$$\begin{tikzcd}
\pi_1(U_{ij}, U_{ij} \cap X_0) \arrow[r, "f_{ji}"] \arrow[d, "f_{ij}"'] & \pi_1(U_j, U_j \cap X_0) \arrow[d, "g_j"] \\
\pi_1(U_i, U_i \cap X_0) \arrow[r, "g_i"'] & G
\end{tikzcd}$$

Here f_{ij} denotes the homomorphism induced by the inclusion $U_{ij} \hookrightarrow U_i$. We have to show that there is a unique morphism

$$g: \quad \pi_1(X, X_0) \to G$$

of groupoids with the property that

$$g_i = g \circ f_i: \quad \pi_1(U_i, U_i \cap X_0) \to \pi_1(X, X_0) \to G \quad \text{for all } i \in I,$$

where f_i denotes the homomorphism induced by the inclusion $U_i \hookrightarrow X$.

We begin by proving the uniqueness: Let $[\gamma] \in \pi_1(X, X_0)(a, b)$ be the homotopy class of a path

$$\gamma: \quad [0,1] \to X \quad \text{between two points } a, b \in X_0.$$

The preimages $\gamma^{-1}(U_i)$ with $i \in I$ form an open cover of the unit interval, hence by compactness of that interval we may find a subdivision $0 = s_0 < s_1 < \cdots < s_m = 1$ such that for each $\alpha \in \{0, 1, \dots, m-1\}$ the given path restricts on the corresponding subinterval to a path

$$\gamma_\alpha: \quad [s_\alpha, s_{\alpha+1}] \to U_{i_\alpha} \subset X \quad \text{for some} \quad i_\alpha \in I.$$

To simplify the notation, we fix such an index i_α for each α and put $V_\alpha = U_{i_\alpha}$ in what follows. Note that the end points of γ_α do not necessarily lie in X_0. But by assumption X_0 meets each path component of the intersections

$$V_\alpha \cap V_{\alpha+1} \subset X,$$

so we can find paths

$$\mu_\alpha: [0,1] \to V_\alpha \cap V_{\alpha+1} \quad \text{with} \quad \mu_\alpha(0) = \gamma(s_\alpha) \quad \text{and} \quad \mu_\alpha(1) \in X_0.$$

Moreover, we can assume μ_0 and μ_m are the constant paths at a resp. b. Then

$$\gamma \sim \gamma_0 \cdot \gamma_1 \cdots \gamma_{m-1} \sim \beta_0 \cdot \beta_1 \cdots \beta_{m-1} \quad \text{for the paths} \quad \beta_\alpha := \mu_\alpha^- \cdot \gamma_\alpha \cdot \mu_{\alpha+1}.$$

Now each β_α is a path inside $V_\alpha = U_{i_\alpha}$ and its two endpoints lie in X_0. Hence any morphism

$$g: \quad \pi_1(X, X_0) \to G \quad \text{inducing the given} \quad g_{i_\alpha} = g \circ f_{i_\alpha}$$

which induces the given morphisms $g_{i_\alpha} = g \circ f_{i_\alpha}$ must satisfy

$$g[\gamma] = g_{i_0}[\beta_0] \cdot g_{i_1}[\beta_1] \cdots g_{i_{m-1}}[\beta_{m-1}] \quad \text{for the images} \quad g_{i_\alpha}[\beta_\alpha] \in G.$$

Since γ was an arbitrary path, this shows the uniqueness of the morphism g.

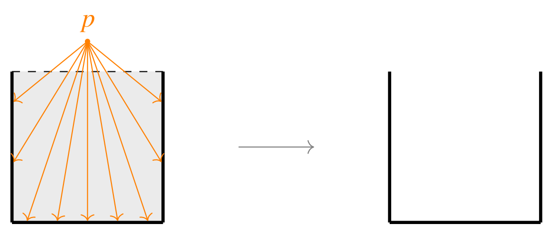
To show existence, we have to check that the expression that we obtained for $g[\gamma]$ only depends on the homotopy class $[\gamma]$ but not on the choice of the factorization of the path. As a preparation, let us draw our homotopy $\gamma \sim \beta_0\beta_1\cdots\beta_{m-1}$ as follows:

$$\begin{tikzcd}
\bullet \arrow[r, dashed, "\beta_0"] \arrow[d, no head, "\mu_0"'] & \bullet \arrow[r, dashed, "\beta_1"] & \bullet \arrow[r, dotted, no head] & \bullet \arrow[r, dashed, "\beta_{m-1}"] & \bullet \arrow[d, no head, "\mu_m"'] \\
\bullet \arrow[r, "\gamma_0"'] & \circ \arrow[u, "\mu_1"'] \arrow[r, "\gamma_1"'] & \circ \arrow[u, "\mu_2"'] \arrow[r, dotted, no head] & \circ \arrow[u, "\mu_{m-1}"'] \arrow[r, "\gamma_{m-1}"'] & \bullet
\end{tikzcd}$$

Here $\bullet$ indicates points whose image lies in X_0, while $\circ$ denotes arbitrary points. The homotopy is given on the solid boundary part

$$\Delta_\alpha := [0,1]\times\{0\}\cup\{0,1\}\times[0,1] \subset Q_\alpha = [0,1]\times[0,1]$$

of each of the small squares by the paths μ_α, γ_α, $\mu_{\alpha+1}$ as shown (with the thick lines on the very left and on the very right denoting the constant paths μ_0, μ_m), and we extend the homotopy continuously from this boundary part to the full square by precomposing with the retraction $r_\alpha\colon\ Q_\alpha \to \Delta_\alpha$ that is given by the projection from some point above the square:



Such filling arguments will be used repeatedly in what follows. Returning to the independence of choices, suppose we are given paths $\gamma,\gamma'\colon [0,1]\to X$ from a to b in the same homotopy class

$$[\gamma] = [\gamma'] \in \pi_1(X,X_0)(a,b)$$

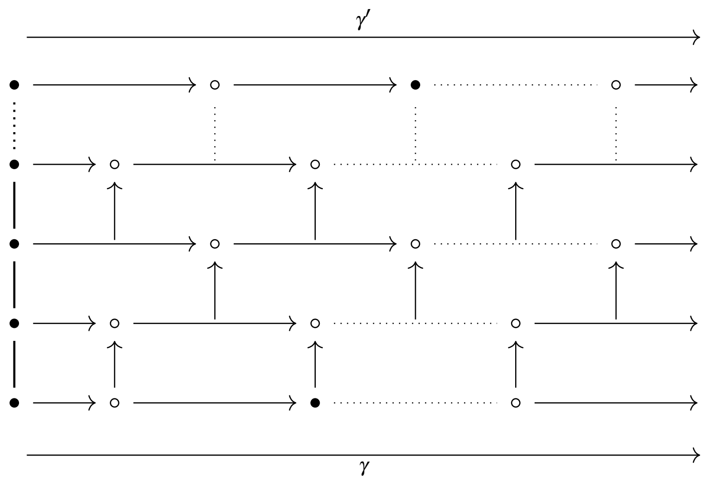
and two decompositions

$$\gamma = \beta_0\beta_1\cdots\beta_m \quad\text{and}\quad \gamma' = \beta_0'\beta_1'\cdots\beta_n'$$

where each $\beta_\alpha,\beta_\alpha'$ lies in one of the open subsets U_i and has end points in X_0. Pick a homotopy

$$H\colon\quad Q = [0,1]\times[0,1] \to X \quad\text{from}\quad \gamma \quad\text{to}\quad \gamma'.$$

By compactness we may subdivide the unit square into smaller rectangles such that each of the smaller rectangles is mapped into one of the open subsets U_i. We may assume that our chosen subdivision refines the given subdivision into a product of paths on $[0,1]\times\{0,1\}$. By slightly perturbing all except for the top and bottom rows to obtain a 'brick wall decomposition' as shown below, we may assume any point of the domain of our homotopy lies on at most *three* of the rectangles:



Here the left and the right vertical boundary are constant paths. The bottom and top row are given by γ and γ' respectively. On those top and bottom rows, all nodes corresponding to end points of $\beta_0,\dots,\beta_m$ and $\beta'_0,\dots,\beta'_n$ are labelled $\bullet$, but all other nodes on those rows and in the interior of the square can only be labelled $\circ$.

The idea is now to deform the homotopy H to a new homotopy via a continuous map

$$\tilde{H}\colon\quad Q\times[0,1]\to X\quad\text{with}\quad \tilde{H}|_{Q\times\{0\}}=H$$

such that the original subdivision of the square is kept but all the labels $\circ$ can be replaced by $\bullet$. More precisely, we impose the following properties:

- If $R\subset Q$ is a rectangle in our subdivision with $H(R)\subset U_i$ for some $i=i_\alpha\in I$, then

$$\tilde{H}(R,t)\subset U_i\quad\text{for all}\quad t\in[0,1].$$

- For any node $p\in Q$ of our subdivision we have $\tilde{H}(p,1)\in X_0$.
- If the node already satisfied $H(p)\in X_0$, then $\tilde{H}(p,t)=H(p)$ for all t.

Such a homotopy $\tilde{H}$ can be constructed as follows: For each rectangle $R\subset Q$ in our subdivision of the square, fix an index $i=i(R)\in I$ with $H(R)\subset U_i$. Given any node p in our subdivision, we have ensured by the brick wall decomposition that the node lies on at most three different rectangles R_1,R_2,R_3. Put $i_\nu=i(R_\nu)\in I$. By our assumption X_0 meets each path component of

$$U = U_{i_1} \cap U_{i_2} \cap U_{i_3},$$

so we may choose a path μ_p: $[0,1] \to U$ joining $H(p)$ to a point in X_0. If $H(p) \in X_0$, we take this path to be constant. Using these paths, we can now define the desired homotopy by proceeding on each brick $R \times [0,1]$ separately:

- On the bottom face $R \times \{0\}$ it is given by $\tilde{H}(x,0) = H(x)$.
- For each vertex p of the bottom face, put the path μ_p on the corresponding vertical edge of the brick, and extend this to a continuous map on each of the lateral faces using the retraction argment from before.
- Finally, fill in the entire brick via the retraction given again by projection from a point slightly above the brick, as we did previously for the retraction from a square but now in three dimensions.

If we apply this construction in each of the rectangles of our subdivision, then on the top face $Q \times \{1\}$ of the cube $Q \times [0,1]$ we obtain a homotopy

$$\tilde{Q}|_{Q\times\{1\}}: \quad Q \to X$$

which is a homotopy between two paths in the same classes as γ and γ' relative to the end points, and with the property that this homotopy sends all the nodes in our given subdivision to points of A_0. This new homotopy therefore sends each rectangle $R \subset Q$ in our subdivision to a commutative square in the groupoid $\pi_1(U_i, U_i \cap X_0)$ for $i = i(R) \in I$. Its image under the given morphism g_i: $\pi_1(U_i, U_i \cap X_0) \to G$ is then a commutative square in G. By combining all these commutative squares in the groupoid G we get

$$g_{i_0}[\beta_0] \cdots g_{i_m}[\beta_m] = g_{i'_0}[\beta'_0] \cdots g_{i'_n}[\beta' n'],$$

which proves that g: $\pi_1(X, X_0) \to G$ is well-defined. □

Example 6.7. From the above version of the Seifert – van Kampen theorem, we may give another computation of the fundamental group of the circle using the open cover $S^1 = U_1 \cup U_2$ with

$$U_1 = \{(x,y) \in S^1 \mid y > -1/2\} \quad \text{and} \quad U_2 = \{(x,y) \in S^1 \mid y < 1/2\}.$$

As a set of base points, take $X_0 = \{p,q\}$ with $p = (1,0)$ and $q = (-1,0)$. In this case the fundamental groupoids of the open subsets have the following shape where the nodes p,q stand for objects (= points) and the arrows for morphisms (= paths):

$$\pi_1(U_1,\{p,q\}): \qquad \begin{tikzcd} p \arrow[loop left] \arrow[r, bend left] & q \arrow[l, bend right] \arrow[loop right] \end{tikzcd}$$

$$\pi_1(U_2,\{p,q\}): \qquad \begin{tikzcd} p \arrow[loop left] \arrow[r, bend right] & q \arrow[l, bend left] \arrow[loop right] \end{tikzcd}$$

$$\pi_1(U_1 \cap U_2, \{p,q\}): \qquad \begin{tikzcd} p \arrow[loop left] & & & q \arrow[loop right] \end{tikzcd}$$

The coequalizer is then the following groupoid:

$$\pi_1(S^1, \{p,q\}): \qquad \begin{tikzcd} p \arrow[loop left] \arrow[rrr, bend left=10] \arrow[rrr, bend right=10] & & & q \arrow[loop right] \arrow[lll, bend right=20] \arrow[lll, bend left=20] \end{tikzcd}$$

If we restrict to morphisms starting and ending at p, we see that $\pi_1(S^1, p) \simeq \mathbb{Z}$ is generated by a single loop which is the product *ab* in the above picture.

For reference, let us also state the fundamental group version of the Seifert – van Kampen theorem in the case of covers by more than two open sets. Given a family of groups $(G_i)_{i \in I}$ indexed by an arbitrary set I, we define their *free product*

$$\mathop{\ast}_{i \in I} G_i$$

to be the set of all formal expressions

$$g_1 g_2 \cdots g_k \quad \text{with} \quad k \in \mathbb{N} \quad \text{and} \quad g_1, \dots, g_k \in \bigsqcup_{i \in I} G_i,$$

modulo the same relations as in definition 5.1. With this notation we have:

Corollary 6.8 (Seifert – van Kampen). *Let X be a topological space that admits a cover*

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

by open subsets $U_i \subset X$ such that for all $i, j, k \in I$, the intersections $U_i \cap U_j \cap U_k$ are path-connected. Suppose that there is a point $p \in \bigcap_{i \in I} U_i$, and consider the group homomorphism

$$\Phi\colon \quad G = \mathop{\ast}_{i \in I} \pi_1(U_i, p) \longrightarrow \pi_1(X, p)$$

induced by the inclusions $U_i \hookrightarrow X$. Then Φ induces an isomorphism

$$\overline{\Phi}\colon \quad G/N \xrightarrow{\sim} \pi_1(X, p)$$

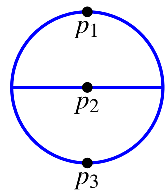
where $N \trianglelefteq G$ is the smallest normal subgroup containing for all $\gamma \in \pi_1(U_i \cap U_j, p)$ the elements

$$\iota_{ij*}(\gamma)^{-1} \cdot \iota_{ji*}(\gamma) \quad \textit{for the inclusions} \quad \iota_{ij}\colon U_i \cap U_j \hookrightarrow U_i.$$

The example of $X = S^1$ has already shown that in general the path-connectedness of the double intersections $U_i \cap U_j$ is needed for the above statement. To see that the path-connectedness of the triple intersections $U_i \cap U_j \cap U_k$ is also essential, consider the topological space

$$X = S^1 \cup [-1, 1] \subset \mathbb{R}^2$$

shown below:



For $i \in \{1, 2, 3\}$, let $U_i = X \smallsetminus \{p_i\}$ be the complement of the point p_i shown in the picture. Then we have

$$X = U_{-1} \cup U_0 \cup U_1 \quad \text{but} \quad \pi_1(X, *) \simeq \mathbb{Z} * \mathbb{Z} \not\simeq \mathbb{Z} * \mathbb{Z} * \mathbb{Z}.$$

Chapter III
Covering spaces

1 Covering spaces and lifting

When we computed the fundamental group of the circle in theorem 4.1, the key point was to consider lifts of paths under the map

$$p\colon \quad \mathbb{R} \to S^1 = \{z \in \mathbb{C} \mid |z| = 1\}, \quad x \mapsto \exp(2\pi i x).$$

The basic feature which guaranteed the existence and uniqueness of lifts was the observation that this map is a covering map in the following sense:

Definition 1.1. A continuous map $p\colon Y \to X$ between topological spaces is called a *covering map* or a topological *cover* if it is surjective and every $x \in X$ has an open neighborhood $U \subset X$ such that $p^{-1}(U) \subset Y$ is a disjoint union of open subsets each of which is mapped by p homeomorphically onto the open subset $U \subset X$. We call any path-connected open $U \subset B$ with this property a *fundamental neighborhood* and call $Y_x = p^{-1}(x)$ the *fiber* over the point $x \in X$.

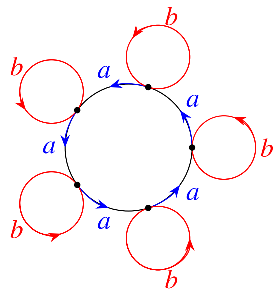
For any set F endowed with the discrete topology, the projection $p\colon X \times F \to X$ is a covering map. We also refer to this as a *trivial cover*. Any covering map can be identified locally on the target with a trivial cover, but globally there are usually many nontrivial covers:

Example 1.2. For any $n \in \mathbb{N}$ the map $p_n\colon S^1 \to S^1, z \mapsto z^n$ is a covering map. We have a commutative diagram

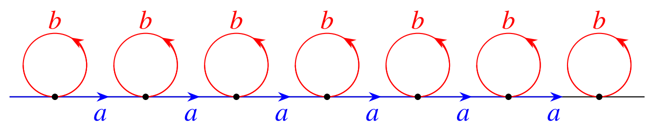
$$\begin{tikzcd} \mathbb{R} \arrow[r, "\exists q_n"] \arrow[dr, "p"'] & S^1 \arrow[d, "p_n"] \\ & S^1 \end{tikzcd}$$

where $q_n\colon \mathbb{R} \to S^1$ denotes the unique covering map with $q_n(0) = 1$, so all these covers can be understood in terms of the exponential cover considered above.

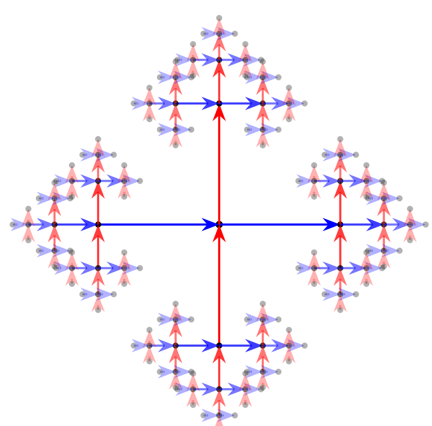
Example 1.3. The figure eight graph $X = S^1 \wedge S^1$ has many interesting covers. For instance, we can put the n-fold cover from the previous example over one of the two circles and attach a copy of the other circle at each of the n nodes, as illustrated in the following picture for $n = 5$:



We could also take the universal cover of one circle and attach a copy of the other circle over each of the countably many nodes:



We could also mix these constructions: For instance, we could replace each loop b in the previous picture by an n-fold cover and accordingly attach $n-1$ blue circles at each of those loops, etc. The 'largest' path-connected cover of the figure eight is obtained by gluing together infinitely many copies of the universal cover of each of the two unit circles. The result will be the infinite tree sketched below, where we omitted the labels a, b and rescaled the segments at each iteration:



Again each of the previous covers of $X = S^1 \vee S^1$ can be obtained from this last cover by suitable identifications between the infinitely many edges and vertices.

The crucial ingredient in our computation of $\pi_1(S^1,*)$ in the proof of theorem 4.1 was that every path in the unit circle admits a unique lift under the exponential map, once a lift of the starting point is fixed. The same holds for arbitrary covers:

Theorem 1.4 (Path Lifting Property). *Let $p\colon Y \to X$ be a covering map. Fix $x \in X$ and a point*

$$y \in p^{-1}(x).$$

in the fiber above it. Then the following properties hold:

a) For any path $\gamma\colon [0,1] \to X$ with $\gamma(0) = x$, there is a unique path $\tilde{\gamma}\colon [0,1] \to Y$ with

$$p \circ \tilde{\gamma} = \gamma \quad \textit{and} \quad \tilde{\gamma}(0) = y.$$

b) Every homotopy $\gamma \sim \mu$ lifts under $p\colon Y \to X$ to a unique homotopy $\tilde{\gamma} \sim \tilde{\mu}$.

Proof. For part $a)$, note that by compactness of the unit interval we may find for any path γ a subdivision $0 = s_0 < s_1 < \cdots < s_m = 1$ of the unit interval such that γ maps each interval $[s_{i-1}, s_i]$ into some fundamental neighborhood $U_i \subset X$. Thus for each i the preimage $p^{-1}(U_i) = \sqcup_j V_{ij}$ is a disjoint union of open subsets $V_{ij} \subset Y$ such that each

$$p_{ij} \;=\; p|_{V_{ij}}\colon \quad V_{ij} \;\longrightarrow\; U_i$$

is a homeomorphism. We can then lift the path inductively as follows:

- Put $\tilde{\gamma}(0) = y$.
- Inductively, if $\tilde{\gamma}$ has already been defined on the interval $[s_0, s_i]$ for some i, let j be the unique index with $\tilde{\gamma}(s_i) \in V_{ij}$, and define

$$\tilde{\gamma}|_{[s_i,s_{i+1}]} \;:=\; p_{ij}^{-1} \circ \gamma|_{[s_i,s_{i+1}]}\colon \quad [s_i, s_{i+1}] \to V_{ij}.$$

Part $b)$ follows in the same way: Given a homotopy $H\colon Q = [0,1] \times [0,1] \to X$, we subdivide Q into smaller squares that map into fundamental open neighborhoods and then lift the homotopy on each of those smaller squares as above. Note that both in $a)$ and $b)$ the uniqueness of the lift is also clear by the same argument. □

We will classify all covers of a given topological space in terms of subgroups of its fundamental group, starting from the following observation:

Corollary 1.5. *Let $p\colon Y \to X$ be a cover. Then for any $x \in X$ and $y \in f^{-1}(x)$, the induced map*

$$p_*\colon \pi_1(Y,y) \to \pi_1(X,x) \quad \textit{is injective}.$$

Proof. If $\tilde{\gamma}\colon [0,1] \to Y$ is a loop based at y whose image $\gamma = p \circ \tilde{\gamma}$ is homotopic to a constant loop through a homotopy keeping the base point fixed, then by the previous lemma the homotopy can be lifted to a homotopy from $\tilde{\gamma}$ to the constant loop at $y = \tilde{\gamma}(0)$, and the lifted homotopy again keeps the base points fixed. □

Example 1.6. The covers $p\colon \mathbb{R} \to S^1$ and $p_n\colon S^1 \to S^1$ of the unit circle give rise to the subgroups

$$\operatorname{im}(p_*) = \{0\} \subset \operatorname{im}(p_{n,*}) = n\mathbb{Z} \subset \pi_1(S^1, 1) = \mathbb{Z}.$$

Example 1.7. The covers of $X = S^1 \vee S^1$ that we considered in example 1.3 give rise to the subgroups

$$\begin{aligned}\{1\} &\subset \{a^{e_1}b^{f_1}\cdots a^{e_k}b^{f_k} \mid e_1 + \cdots + e_k = 0\} \\ &\subset \{a^{e_1}b^{f_1}\cdots a^{e_k}b^{f_k} \mid e_1 + \cdots + e_k \equiv 0 \mod n\} \subset \langle a, b\rangle = \pi_1(X, *).\end{aligned}$$

Note that unlike in the previous example, there are many more subgroups of the free group $\langle a, b\rangle = \mathbb{Z} * \mathbb{Z}$, and indeed $X = S^1 \vee S^1$ has many more covers!

2 Abstract covering theory via groupoids

The correspondence between covers of a topological space and subgroups of its fundamental group can be phrased purely algebraically in terms of groupoids. The key property of coverings is the lifting property for paths with given starting point but variable end point. These paths form a star in the fundamental groupoid:

Definition 2.1. Let G be a groupoid and X its set of objects. For any $a \in X$, we define its *star* to be the set

$$G_{a,\star} := \{(b, \gamma) \mid b \in X,\ \gamma \in G_{a,b}\}$$

of all morphisms $\gamma\colon a \to b$ from the given point a to arbitrary other points $b \in X$.

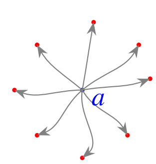
Example 2.2. For a topological space X, let $\Pi(X) = \Pi(X, X)$ be its fundamental groupoid of homotopy classes of paths with arbitrary starting and end points. The star of $a \in X$ is

$$\Pi(X)_{a,\star} = \big\{ \text{paths } \gamma\colon [0,1] \to X \mid \gamma(0) = a \big\}/\sim$$

where $\sim$ denotes homotopy rel $\{0, 1\}$. If X is simply connected, then the evaluation map

$$\Pi(X)_{a,\star} \to X, \quad [\gamma] \mapsto \gamma(1)$$

is a bijection. Fixing a path in each homotopy class, we get the following picture of the star:



Note that if X is not simply connected, then the evaluation map will not be injective.

In the abstract framework of groupoids, the notion of a covering map has the following counterpart:

Definition 2.3. A *cover* of a groupoid G is a morphism $p\colon H \to G$ from another groupoid such that

- p induces a surjection between the underlying sets of objects, and
- p induces a bijection of stars

$$p|_{H_{y,\star}}\colon \quad H_{y,\star} \to G_{x,\star} \quad \text{for each object } y \text{ of } H \text{ with image } x = p(y).$$

Note that in this case, in terms of the fiber $p^{-1}(x) = \{\text{objects } y \text{ of } H \mid p(y) = x\}$ we have

$$p^{-1}(G_{x,\star}) \;=\; \bigsqcup_{y \in p^{-1}(x)} H_{y,\star}.$$

The main example for us will be the following:

Lemma 2.4. *For any covering map $p\colon Y \to X$ of topological spaces, the induced morphism of fundamental groupoids*

$$\Pi(Y) \to \Pi(X) \quad \textit{is a covering of groupoids}.$$

Proof. Immediate by the unique path lifting property in theorem 1.4. □

We have seen that any cover of topological spaces induces an injective map of fundamental groups. We can do the same for covers of groupoids: For an object x in a groupoid G denote its automorphism group by

$$G_a \;:=\; G_{a,a}.$$

If $G = \Pi(X)$ is the fundamental groupoid of a topological space, then $G_a = \pi_1(X,a)$ is the fundamental group. Our earlier result now generalizes as follows:

Proposition 2.5. *Let $p\colon H \to G$ be a cover of groupoids and x an object of G.*

a) For any $y \in p^{-1}(x)$, the induced morphism $p_*\colon H_y \to G_x$ is injective.

b) For any $z \in p^{-1}(x)$ with $H_{y,z} \neq \varnothing$, the subgroup

$$p_*(H_z) \le G_x \quad \textit{is conjugate to} \quad p_*(H_y) \le G_x.$$

c) Conversely, all conjugates of the subgroup $p_*(H_y) \le G_x$ arise like this.

Proof. Part $a)$ follows immediately from the injectivity of the map $H_{y,\star} \to G_{x,\star}$ between stars. For part $b)$ pick any $\alpha \in H_{y,z}$. Then conjugation by α gives a group isomorphism

$$H_y \to H_z, \quad \gamma \mapsto \alpha^{-1}\gamma\alpha$$

that maps under p to the conjugation by the image $\beta \in G_{x,x} = G_x$ of α. For part $c)$ note that by the surjectivity of the map $H_{y,\star} \to G_{x,\star}$ between stars, any $\beta \in G_x = G_{x,x}$ is the image of some $\alpha \in H_{y,\star}$, hence the computation from part $b)$ applies. □

A groupoid G is called *connected* if $G_{a,b} \neq \varnothing$ for all objects a, b of G. So the fundamental groupoid of a topological space is connected iff the topological space is path-connected. Covers of groupoids have the following lifting property:

Theorem 2.6. *Let $p\colon H \to G$ be a cover of groupoids and $g\colon A \to G$ an arbitrary morphism from a connected groupoid A. Let x be an object of A, put $z = g(x)$, and pick any $y \in p^{-1}(z)$. Then the following two properties are equivalent:*

a) There is a morphism $h\colon A \to H$ of groupoids with $h(x) = y$ such that $g = p \circ h$:

$$\begin{tikzcd} & H \arrow[d, "p"] \\ A \arrow[r, "g"'] \arrow[ur, dashed, "\exists h"] & G \end{tikzcd}$$

b) We have $g_*(A_x) \subset p_*(H_y)$ as subgroups of the group G_z.

Moreover, if these two equivalent conditions hold, there is a unique such h.

Proof. Clearly $a)$ implies $b)$. For the converse, suppose $g_*(A_x) \subset p_*(H_y)$. We can then define a morphism of groupoids $h\colon A \to H$ as follows:

- Definition on objects:

  For any object x' in A, there exists a path $\gamma \in A_{x,x'}$ from x to x' because A is connected. By applying the morphism g we get a path $\mu = g(\gamma) \in G_{z,\star}$. Since the morphism $p\colon H \to G$ is a covering, this last path admits a unique lift $\tilde{\mu} \in H_{y,\star}$, and we define
  $$h(x') := \tilde{\mu}(1)$$
  to be the end point of this lift. We need to check this is well-defined: Given any two paths $\gamma_i \in A_{x,x'}$ for $i = 1, 2$, the paths $\mu_i = g(\gamma_i) \in G_{z,\star}$ both end at $z' = g(x')$ and we have
  $$\mu_1 \mu_2^{-1} = g(\gamma_1 \gamma_2^{-1}) \in g_*(A_x) \subset p_*(H_y)$$
  where the last inclusion holds by our assumption $b)$. Hence $\mu_1 \mu_2^{-1} = p(\delta)$ for some loop $\delta \in H_y$. This implies
  $$\mu_1 = p(\delta) \cdot \mu_2.$$
  Since p is a covering map, the loop $p(\delta) \in G_z$ lifts uniquely to a path in $H_{y,\star}$ and this unique lift has to coincide with the obvious lift δ which happens to be a loop. Therefore the unique lifts to $H_{y,\star}$ of the two paths $\mu_1 = p(\delta) \cdot \mu_2$ and μ_2 are related by
  $$\tilde{\mu}_1 = \delta \cdot \tilde{\mu}_2,$$
  which implies for the end points of the lifts that $\tilde{\mu}_1(1) = \tilde{\mu}_2(1)$ as desired.

- Definition on morphisms:

  For any morphism γ in A, denote by $x_0 = \gamma(0)$ its source and by $x_1 = \gamma(1)$ its target. Because A is connected, we may pick morphisms $\gamma_i \in A_{x,x_i}$ from x to x_i

for $i = 1,2$. Consider the paths $\mu_i := g(\gamma_i) \in G_{z,\star}$ and their unique lifts $\tilde{\mu}_i \in H_{y,\star}$ starting at y. By the construction in the previous step, the end points of these lifts are the objects

$$y_i := \tilde{\mu}_i(1) = h(x_i) \quad \text{in} \quad p^{-1}(z_i)$$

where $z_i = g(x_i)$. Since p is a covering map, the path $g(\gamma) \in G_{z_1,z_2} \subset G_{z_1,\star}$ admits a unique lift to a path in the star $H_{y_1,\star}$ starting at y_1, and the same argument as in the previous step shows that this unique lift ends at the point y_2, i.e. it gives a path

$$h(\gamma) := \widetilde{g(\gamma)} \in H_{y_1,y_2}$$

between the points $y_i = h(x_i)$ as desired. Again one checks that the path $h(\gamma)$ only depends on γ but not on the choice of μ_1, μ_2.

In each of the above steps, it follows from the uniqueness of lifts to a cover that the above construction is the only possible choice if we want $h(x) = y$ and $g = p \circ h$. □

The above is particularly useful when g is also a cover of groupoids. In this case the lift h will be a morphisms of covers:

Definition 2.7. Let G be a groupoid with covers $p_i\colon H_i \to G$ for $i = 1,2$. A *morphism* between the covers is a morphism $h\colon H_1 \to H_2$ of groupoids such that the following diagram commutes:

$$\begin{tikzcd} H_1 \arrow[rr, "h"] \arrow[dr, "p_1"'] & & H_2 \arrow[dl, "p_2"] \\ & G & \end{tikzcd}$$

We denote by $\mathrm{Hom}_G(H_1, H_2)$ the set of all such morphisms of covers.

Remark 2.8. Any morphism of covers with connected target is itself a cover.

Proof. Let $h \in \mathrm{Hom}_G(H_1, H_2)$ be a morphism of covers, with H_2 connected. We have to show:

a) Every object y_2 of H_2 has the form $y_2 = h(y_1)$ for some object y_1 of H_1.
b) The induced map $h\colon (H_1)_{y_1,\star} \to (H_2)_{y_2,\star}$ between the stars is a bijection.

For $a)$ consider $x = p_2(y_2)$. Since $p_1\colon H_1 \to G$ is a cover, there is an object z_1 of H_1 with $p_1(z_1) = x$. The image $z_2 = h(z_1)$ may differ from y_2, but since H_2 is connected, there exists a path $\gamma_2 \in (H_2)_{z_2,y_2}$. The image of this path is a loop

$$\alpha = p_2(\gamma) \in G_{x,x} = G_x.$$

Since $p_1\colon H_1 \to G$ is a cover of groupoids, this loop lifts to a path $\gamma_1 \in (H_1)_{z_1,\star}$. By the uniqueness of lifts for the cover $p_2\colon H_2 \to G$ then

$$h(\gamma_1) = \gamma_2 \in (H_2)_{z_2,\star}$$

since both sides are lifts of α. Thus the end point $y_2 = \gamma_2(1)$ coincides with the end point $h(y_1)$ of $h(\gamma_1)$, where we have put $y_1 = \gamma_1(1)$. Hence $y_2 = h(y_1)$ as desired.

Part $b)$ follows by considering the commutative diagram of morphisms between the stars

$$\begin{tikzcd} (H_1)_{y_1,\star} \arrow[rr, "h"] \arrow[dr, "p_1"'] & & (H_2)y_2,\star \arrow[dl, "p_1"] \\ & G_{x,\star} & \end{tikzcd}$$

Since p_1 and p_2 are covers of groupoids, both downward arrows in this diagram are bijections, hence so is the horizontal arrow h in the top row of the diagram. □

In the situation of remark 4 we also say that the cover $p_1 \colon H_1 \to G$ *dominates* the cover $p_2 \colon H_2 \to G$, and since we like to draw covers vertically, we depict the situation in a diagram of the following form:

$$\begin{tikzcd} H_1 \arrow[dd, "p_1"'] \arrow[dr, "h"] & \\ & H_2 \arrow[dl, "p_2"] \\ G & \end{tikzcd}$$

This is reminiscent of a Hasse diagram in group theory, and in fact theorem 2.6 allows us to describe the dominance relation between covers of a given groupoid in terms of subgroups. For this we need to keep track of base points:

Definition 2.9. A *pointed groupoid* is a pair (G,x) consisting of a groupoid G and an object x of G. Pointed groupoids form a category in a natural way, where we define a morphism of pointed groupoids

$$p \colon (H,y) \;\to\; (G,x)$$

to be a morphism $p \colon H \to G$ of the underlying groupoids with $p(y) = x$. Such a morphism is said to be a cover of pointed groupoids iff the underlying morphism of groupoids is a cover. For any two such pointed covers $p_i \colon (H_i, y_i) \to (G,x)$, we denote by

$$\mathrm{Hom}_{(G,x)}((H_1,y_1),(H_2,y_2)) \;=\; \{h \in \mathrm{Hom}_G(H_1,H_2) \mid h(y_1) = y_2\}$$

the set of all morphisms between them.

Theorem 2.10. *Let $p_i \colon (H_i, y_i) \to (G,x)$ for $i = 1,2$ be two connected pointed covers.*

a) There exists a morphism $h \in \mathrm{Hom}_{(G,x)}((H_1,y_1),(H_2,y_2))$ of pointed covers iff we have an inclusion

$$p_1((H_1)_{y_1}) \subset p_2((H_2)_{y_2}) \quad \textit{inside} \quad G_x.$$

b) More generally, there exists a morphism $h \in \mathrm{Hom}_G(H_1, H_2)$ *iff for some* $g \in G_x$ *we have an inclusion*

$$p_1((H_1)_{y_1}) \subset g \cdot p_2((H_2)_{y_2}) \cdot g^{-1} \quad \textit{inside} \quad G_x.$$

c) In both cases h is an isomorphism iff the inclusion $\subset$ *is an equality.*

Proof. Part $a)$ follows immediately from the lifting criterion in theorem 2.6, and $b)$ is a consequence of part $a)$ since in a connected groupoid any two different choices of base points can be connected by a path g. Part $c)$ follows by symmetry. □

Corollary 2.11. *We have a fully faithful functor*

$$\begin{aligned} \varphi\colon \quad \{\textit{connected pointed covers of } (G,x)\} &\hookrightarrow \{\textit{subgroups of } G_x\} \\ (p\colon (H,y) \to (G,x)) &\mapsto p_*(H_y) \end{aligned}$$

where the morphisms in the target category are inclusions between subgroups.

Proof. By theorem 2.10, any morphism $(H_1, y_1) \to (H_2, y_2)$ of pointed covers gives an inclusion of subgroups

$$p_{1,*}(H_{1,y_1}) \subset p_{2,*}(H_{2,y_2}),$$

thus φ indeed defines a functor. Fullness in this case amounts to saying that there is a morphism between two connected pointed covers whenever there is an inclusion between the associated subgroups, which again holds by theorem 2.10. Faithfulness amounts to saying that in this case there is a unique such morphism, which is true by the uniqueness of lifts in theorem 2.6. □

We will see soon that the above functor is also essentially surjective, i.e. every subgroup of G_x arises from some connected pointed cover. So φ induces a bijection between

- isomorphism classes of connected pointed covers of (G,x), and
- subgroups of the group G_x.

To prove this we need to construct a connected cover for each subgroup of G_x. For this we will use yet another description of connected covers which is based on the monodromy action on fibers.

3 Monodromy and the Galois correspondence

A version of corollary 2.11 that avoids the choice of base point in the covering groupoids can be given as follows. Recall that a *right action* of a group Γ on a set S is given by a map

$$S \times \Gamma \to S, \quad (s, \gamma) \mapsto s \cdot \gamma$$

such that for all $s \in S$ the following two properties hold:

- $s \cdot (\gamma \cdot \mu) = (s \cdot \gamma) \cdot \mu$ for all $\gamma, \mu \in \Gamma$,
- $s \cdot \varepsilon = s$ for the neutral element $\varepsilon \in \Gamma$.

A set S together with a right action of Γ will also be called a Γ-set. These form a category in a natural way where the morphisms are taken to be the maps $f \colon S \to T$ between Γ-sets with

$$f(s \cdot \gamma) \;=\; f(s) \cdot \gamma \quad \text{for all } s \in S,\, \gamma \in \Gamma.$$

For any Γ-set S, the set of its automorphisms forms a group $\mathrm{Aut}_\Gamma(S)$.

Example 3.1. We say that a group Γ acts *transitively* on a set S if for some $s_0 \in S$ the map

$$\Gamma \;\to\; S, \quad \gamma \mapsto s_0 \cdot \gamma$$

is surjective. If this is the case, then the same property holds for any $s_0 \in S$, and for any fixed choice of the point s_0 the above surjection factors through an isomorphism of Γ-sets

$$\Gamma_0 \backslash \Gamma \;\xrightarrow{\sim}\; S \quad \text{for the stabilizer subgroup} \quad \Gamma_0 \;=\; \{\gamma \in \Gamma \mid s_0 \cdot \gamma = s_0\} \;\le\; \Gamma.$$

Here we endow the set

$$\Gamma_0 \backslash \Gamma \;=\; \{\text{cosets } \Gamma_0 \cdot \gamma \subset \Gamma \mid \gamma \in \Gamma\}$$

with the right action of Γ via multiplication. Note that this set also comes with a natural *left* action of the normalizer $N(\Gamma_0) := \{\gamma \in \Gamma \mid \gamma^{-1}\Gamma_0\gamma = \Gamma_0\}$, where $n \in N(\Gamma_0)$ acts via

$$\Gamma_0 \cdot \gamma \;\mapsto\; n \cdot \Gamma_0 \cdot \gamma \;=\; \Gamma_0 \cdot n \cdot \gamma.$$

This left action commutes with the right action of the group Γ and thus defines an automorphism of the Γ-set $\Gamma_0 \backslash \Gamma$. In fact one easily checks that every automorphism of this Γ-set arises like this, so the automorphism group of the Γ-set $\Gamma_0 \backslash \Gamma$ is given by the quotient group

$$\mathrm{Aut}_\Gamma(\Gamma_0 \backslash \Gamma) \;=\; N(\Gamma_0)/\Gamma_0 \;=:\; W(\Gamma, \Gamma_0).$$

The notation on the right hand side is chosen to be reminiscent of *Weyl groups* that appear as quotients of certain normalizers in the theory of Lie groups.

We will use the above example to describe the following types of Γ-sets arising from connected covers of groupoids:

Definition 3.2. For any connected cover $p \colon H \to G$ of groupoids, the group $\Gamma = G_x$ acts naturally on the fiber

$$p^{-1}(x) \;=\; \{\text{points } z \text{ of } H \mid p(z) = x\}$$

as follows: Let $z \in p^{-1}(x)$ be a point in the fiber. Since p is a cover, any loop $\gamma \in G_x$ has a unique lift

$$\tilde{\gamma}_z \in H_{z,\star}$$

that starts at the chosen point in the fiber. We then define $z \cdot \gamma := \tilde{\gamma}_z(1)$ to be the end point of this unique lift. It follows from the definitions that this indeed gives a right action of G_x on the fiber $p^{-1}(x)$. We call this action the *monodromy action*.

Example 3.3. Let $p \colon H = \Pi(Y) \to G = \Pi(X)$ be the cover of fundamental groupoids induced by the map

$$Y = \mathbb{R} \longrightarrow X = S^1, \quad t \mapsto \exp(2\pi i t).$$

Put $x = 1$. Then the fundamental group

$$G_x = \pi_1(S^1, x) \simeq \mathbb{Z} \quad \text{acts by translations on the fiber} \quad p^{-1}(x) = \mathbb{Z} \subset \mathbb{R}.$$

Proof. Consider the isomorphism $\mathbb{Z} \xrightarrow{\sim} \pi_1(S^1, x)$ sending an element $n \in \mathbb{Z}$ to the loop

$$\gamma_n \colon [0,1] \to S^1, \quad t \mapsto \exp(2\pi i n t).$$

The unique lift of this loop starting at $y \in p^{-1}(x) = \mathbb{Z}$ is $\tilde{\gamma}_n \colon [0,1] \to \mathbb{R}, t \mapsto y + nt$, and it ends at $\tilde{\gamma}_n(1) = y + n$. □

In general, for any connected cover of a pointed groupoid the monodromy action on the fiber can be described explicitly as an action on cosets for a suitable subgroup as in example 3.1:

Lemma 3.4. *For any connected cover of pointed groupoids $p \colon (H, y) \to (G, x)$, the monodromy action of the group $\Gamma = G_x$ on the fiber is transitive and hence induces an isomorphism of Γ-sets*

$$\Gamma_0 \backslash \Gamma \xrightarrow{\sim} p^{-1}(x), \quad \Gamma_0 \cdot \gamma \mapsto y \cdot \gamma \quad \textit{where} \quad \Gamma_0 = \{\gamma \in \Gamma \mid y \cdot \gamma = y\} = p_*(H_y).$$

Proof. The connectedness of H implies that the monodromy action is transitive: For any $z \in p^{-1}(x)$ there exists a path $\gamma \in H_{y,z}$, and then we have $z = y \cdot \gamma$ by definition of the monodromy action. From the transitivity of the monodromy action we get by example 3.1 an isomorphism of Γ-sets

$$\Gamma_0 \backslash \Gamma \xrightarrow{\sim} p^{-1}(x) \quad \text{for the subgroup} \quad \Gamma_0 = \{\gamma \in \Gamma \mid y \cdot \gamma = y\}.$$

It remains to show that this subgroup is given by $\Gamma_0 = p_*(H_y)$. This follows again from the definitions: For $\gamma \in \Gamma = G_x$, let $\tilde{\gamma} \in H_{y,\star}$ be the unique lift starting at y. Then we have $y \cdot \gamma = \tilde{\gamma}(1)$, hence

$$\gamma \in \Gamma_0 \iff \tilde{\gamma} \in H_{y,y} = H_y \iff \gamma \in p_*(H_y)$$

and the claim follows. □

The above can be reversed, so we obtain a construction of a connected cover for each subgroup:

Theorem 3.5. *Let (G,x) be a pointed groupoid. Then for any subgroup $\Gamma_0 \leq \Gamma = G_x$ there exists a connected pointed cover $p\colon (H,y) \to (G,x)$ with $p_*(H_y) = \Gamma_0$.*

Proof. To motivate the construction, we do reverse engineering: If such a cover exists, then after fixing a reference point $y \in p^{-1}(x)$ we can by connectedness obtain any other point of H as the end point of some path $\tilde{\mu} \in H_{y,\star}$. Since the map between stars

$$p\colon H_{y,\star} \ \to \ G_{x,\star}$$

is bijective, it follows that any point z of H is the end point of the unique lift $\tilde{\mu} \in H_{y,\star}$ of some path $\mu \in G_{x,\star}$. The path μ is not determined by the point z, but for any two paths $\mu_1, \mu_2 \in G_{x,\star}$ whose unique lifts $\tilde{\mu}_1, \tilde{\mu}_2 \in H_{y,\star}$ end at the same point z, we get by composition a loop

$$\tilde{\gamma} \ = \ \tilde{\mu}_1 \cdot \tilde{\mu}_2^{-1} \ \in \ H_{y,y} \ = \ H_y,$$

and then $\tilde{\mu}_1 = \tilde{\gamma} \cdot \tilde{\mu}_2$. In other words, any two paths $\mu_1, \mu_2 \in G_{x,\star}$ whose unique lifts have the same end points are related by

$$\mu_1 \ = \ \gamma \cdot \mu_2 \quad \text{for the image} \quad \gamma \ = \ p_*(\tilde{\gamma}) \ \in \ \Gamma_0.$$

After this motivating discussion, it should be clear how to proceed in the opposite direction. Given a subgroup $\Gamma_0 \leq \Gamma = G_x$, we *define* a pointed groupoid (H,y) as follows:

- The set of objects (= points) of H is the quotient set

  $$\Gamma_0 \backslash G_{x,\star} \ := \ \{\text{orbits } \Gamma_0 \cdot \mu \subset G_{x,\star} \mid \mu \in G_{x,\star}\},$$

  with base point given by the orbit $y = \Gamma_0 \cdot \varepsilon_x$ of the neutral element $\varepsilon_x \in G_x \subset G_{x,\star}$.
- The set of morphisms between two objects $y_1 = \Gamma_0 \cdot \mu_1$ and $y_2 = \Gamma_0 \cdot \mu_2$ is defined by

  $$H_{y_1,y_2} \ := \ \{\gamma \in G_{x_1,x_2} \mid \mu_1 \cdot \gamma \cdot \mu_2^{-1} \in \Gamma_0\} \quad \text{where} \quad x_i \ := \ \mu_i(1),$$

  as illustrated in the following picture:

  [PICTURE]

It follows immediately from the definition that we have a morphism of connected pointed groupoids

$$p\colon (H,y) \to (G,x) \quad \begin{cases} \text{given on objects by} \quad \Gamma_0 \cdot \mu \mapsto \mu(1) \\ \text{and on morphisms by} \quad \gamma \mapsto \gamma. \end{cases}$$

Moreover p is a cover of groupoids: It is surjective on objects since any point of G is the end point of some path $\mu \in G_{x,\star}$, and the map on stars $H_{y,\star} \to G_{x,\star}$ is bijective since it can be identified with the identity map by the way in which we defined morphisms in H. Finally, it is clear from the construction that the given subgroup is the stabilizer

$$\Gamma_0 = \{\gamma \in G_x \mid y \cdot \gamma = y\}$$

of y under the monodromy action, so example 3.1 shows that $\Gamma_0 = p_*(H_y)$. □

This gives the essential surjectivity that was missing in corollary 2.11 and leads to the following Galois correspondence for connected covers of groupoids, where by a *pointed transitive G_x-set* we mean a pair (S,p) of a transitive G_x-set S and a point $p \in S$:

Corollary 3.6. *Let (G,x) be a pointed groupoid.*

a) We have equivalences of categories

$$\begin{tikzcd} \{\text{pointed connected covers of } (G,x)\} \arrow[r, "\simeq", "\varphi"'] \arrow[dr, "\eta"'] & \{\text{subgroups of } G_x\} \arrow[d, "\psi"', "\simeq"] \\ & \{\text{pointed transitive } G_x\text{-sets}\} \end{tikzcd}$$

where

- *φ sends $p\colon (H,y) \to (G,x)$ to the subgroup $p_*(H_y)$,*
- *ψ sends $p\colon (H,y) \to (G,x)$ to the pointed G_x-set $(p^{-1}(x),y)$,*
- *η sends $\Gamma_0 \le G_x$ to the quotient set $\Gamma_0 \backslash G_x$ with base point the coset $\Gamma_0 \cdot 1$.*

b) By forgetting base points, the functor η induces an equivalence of categories

$$\bar{\eta}\colon \quad \{\text{connected covers of } G\} \xrightarrow{\sim} \{\text{transitive } G_x\text{-sets}\}$$

$$(p\colon H \to G) \mapsto p^{-1}(x).$$

In particular, the automorphism group of any connected cover $p\colon H \to G$ is given by

$$\mathrm{Aut}_G(H) \simeq W(G_x, H_y) \quad \text{for any} \quad y \in p^{-1}(x).$$

Proof. The functor φ is fully faithful as we have seen in corollary 2.11, and it is essentially surjective by theorem 3.5. The functor ψ is an equivalence of categories by example 3.1. Hence claim $a)$ follows, and by varying the base point $y \in p^{-1}(x)$ we then also obtain $b)$. □

Definition 3.7. A connected cover of groupoids $p\colon H \to G$ is a *Galois cover* if for any point $y \in p^{-1}(x)$,

$$H_y \trianglelefteq G_x \quad \text{is a normal subgroup.}$$

Note that this condition does not depend on the choice of $y \in p^{-1}(x)$. If it holds, then by the formula in corollary 3.6 $b)$ the automorphism group of the cover is the quotient group

$$\mathrm{Aut}_G(H) \simeq G_x/H_y$$

because for normal subgroups $H_y \trianglelefteq G_x$ we have $W(G_x, H_y) = G_x/H_y$.

Intuitively speaking, a connected cover of groupoids $p\colon H \to G$ is Galois iff its automorphism group $\mathrm{Aut}_G(H)$ (also known as its *deck transformation group*) is as large as possible. To make this precise, note that the fiber $p^{-1}(x)$ comes with a natural *left* action

$$\mathrm{Aut}_G(H) \times p^{-1}(x) \to p^{-1}(x), \quad (h,x) \mapsto h \cdot x := h(x)$$

of the deck transformation group. Since a deck transformation h is determined uniquely by its value at any single point, this action is *free* in the sense that it has trivial stabilizers:

$$\{h \in \mathrm{Aut}_G(H) \mid h \cdot y = y\} = \{id\} \quad \text{for all} \quad y \in p^{-1}(x).$$

Thus for fixed $y \in p^{-1}(x)$ the map $\mathrm{Aut}_G(H) \hookrightarrow p^{-1}(x), h \mapsto h(y)$ is injective, so the deck transformation group is as large as possible if this map is also surjective, i.e. if the deck transformation group acts transitively on the fiber $p^{-1}(x)$.

Lemma 3.8. *A connected cover of groupoids $p\colon H \to G$ is Galois iff the action of the group $\mathrm{Aut}_G(H)$ on any fiber $p^{-1}(x)$ is transitive.*

Proof. Fix a point $y \in p^{-1}(x)$. By the formula for the deck transformation group in corollary 3.6 $b)$ we have identifications

$$\mathrm{Aut}_G(H) = N(H_y)/H_y \quad \text{for the normalizer} \quad N(H_y) = \{\gamma \in G_x \mid \gamma^{-1} H_y \gamma = H_y\} \le G_x.$$

But the natural left action of this normalizer on the set $p^{-1}(x) = H_y \backslash G_x$ is transitive iff $N(H_y) = G_x$, which happens iff $H_y \trianglelefteq G_x$ is a normal subgroup. This last property is the definition of the cover being Galois. □

We have written the action of deck transformations as a *left* action since this fits with the conventions for the composition of maps. On the other hand, there is also a natural *right* action of another group on the fiber $p^{-1}(x)$, namely the monodromy action

$$p^{-1}(x) \times G_x \longrightarrow p^{-1}(x), \quad y \mapsto y \cdot \gamma.$$

This is a right action because we chose to write the composition of paths from left to right. The following observation explains why this is the correct choice if we want to consider both actions together:

Lemma 3.9. *For any connected cover $p\colon H \to G$, the action by deck transformations commutes with the monodromy action in the sense that*

$$h\cdot(z\cdot\gamma) \;=\; (h\cdot z)\cdot\gamma \quad \textit{for all} \quad h\in \mathrm{Aut}_G(H),\, z\in p^{-1}(x),\, \gamma\in G_x.$$

In fact the deck transformation group can be identified with the centralizer of the monodromy action:

$$\mathrm{Aut}_G(H) \;=\; \{h\colon p^{-1}\to p^{-1} \textit{ bijective} \mid h\cdot(z\cdot\gamma) = (h\cdot z)\cdot\gamma \textit{ for all } \; z\in p^{-1}(x),\, \gamma\in G_x\}.$$

Proof. In corollary 3.6 we have identified the category of connected covers of G with the category of transitive G_x-sets. Thus $\mathrm{Aut}_G(H)$ can be viewed as a group of automorphisms of G_x-sets, and by definition the automorphisms of a G_x-set are precisely the bijections from the set to itself that commute with the G_x-action. □

Note that the description of $\mathrm{Aut}_G(Z)$ as a centralizer in the above lemma holds for all connected covers. Hence we may equivalently characterize Galois covers as those connected covers of groupoids for which the centralizer of the monodromy acts transitively on a fiber.

We conclude our discussion of abstract Galois theory for covers of groupoids by showing the existence of a universal cover:

Definition 3.10. A cover $p\colon \tilde{G}\to G$ of groupoids is a *universal cover* if $\tilde{G}$ is simply connected in the sense that

- $\tilde{G}$ is connected, and
- $\tilde{G}_x = \{0\}$ for all objects x of $\tilde{G}$.

By the Galois correspondence for connected covers we have:

Corollary 3.11. *Up to isomorphism of covers, any connected groupoid G admits a unique universal cover $p\colon \tilde{G}\to G$, and this universal cover dominates any other connected cover $q\colon H\to G$:*

$$\begin{tikzcd}
\tilde{G} \arrow[dd, "p"'] \arrow[dr, dotted, "\exists h"] & \\
& H \arrow[dl, "q"] \\
G &
\end{tikzcd}$$

Proof. The existence of the universal cover follows by taking the trivial subgroup in the Galois correspondence in corollary 3.6. The uniqueness up to isomorphism follows from the link between isomorphism of covers and conjugacy of subgroups in theorem 2.10 b), and the universal property comes from the link between dominance and inclusions of subgroups in theorem 2.10 a). □

4 Continuity of lifts under covering maps

Given a continuous map $p\colon Y \to X$ of topological spaces, let $p_*\colon \Pi(Y) \to \Pi(X)$ be the induced morphism between their fundamental groupoids. If p is a covering map, then p_* is a cover of groupoids. We want to upgrade the lifting properties from the previous sections from groupoids to topological spaces. For example, for any continuous map $g\colon Z \to X$ from a *simply connected* space Z, theorem 2.6 gives a morphism φ of groupoids such that the following diagram commutes:

$$\begin{tikzcd} & \Pi(Y) \arrow[d, "p_*"] \\ \Pi(Z) \arrow[ur, dashed, "\exists\varphi"] \arrow[r, "g_*"'] & \Pi(X) \end{tikzcd}$$

On points this gives a map $h\colon Z \to Y$ with $p \circ h = g$, but this map h might not be continuous. The issue is that our construction of the lift via the path lifting property is not a local one — nearby points need not be connected by a nearby path. So we should revisit our notions of connectedness:

Definition 4.1. A topological space X is

a) *connected* if it cannot be written as a disjoint union of two non-empty open sets.

b) *path-connected* if any two points $p, q \in X$ are connected by a path $\gamma\colon [0,1] \to X$.

c) *locally path-connected* if for every point $p \in X$, every open neighborhood $V \subset X$ of p contains a path-connected open neighborhood $U \subset V$ of p: [PICTURE]

Remark 4.2. The definitions easily imply that any path-connected topological space is also connected. Conversely, one may show that if a topological space is both connected and locally path-connected, then it is path-connected. These are the only general implications between the above three notions (see example 4.4).

For locally path-connected topological spaces, our construction of lifts via the path lifting property is indeed a local one:

Theorem 4.3. *Let $p: Y \to X$ be a covering map and $g: Z \to X$ any continuous map, where all three topological spaces in question are path-connected and Z is moreover locally path-connected. Let $z \in Z$, put $x = g(z)$, and pick any $y \in p^{-1}(x)$. Then the following two properties are equivalent:*

a) There is a continuous map $h: Z \to Y$ with $h(z) = y$ such that $g = p \circ h$:

$$\begin{tikzcd} & Y \arrow[d, "p"] \\ Z \arrow[ur, dashed, "\exists h"] \arrow[r, "g"'] & X \end{tikzcd}$$

b) We have $g_*(\pi_1(Z,z)) \subset p_*(\pi_1(Y,y))$ as subgroups of $\pi_1(X,x)$.

Moreover, if these two equivalent conditions hold, there is a unique such h.

Proof. The necessity of $b)$ is clear by functoriality of fundamental groups, and the uniqueness holds by the same statement for groupoids in theorem 2.6. If $b)$ holds, then by applying theorem 2.6 on the level of points we get a map h, constructed via path lifting, such that $p \circ h = g$ and $h(z) = y$.

So far we have only used path-connectedness. What remains to be shown is that if Z is also *locally* path-connected, then the map h constructed by path lifting is continuous. Indeed, given any $a \in Z$, pick an open neighborhood $W \subset Y$ of $h(a)$ such that $p|_W: W \to X$ is a homeomorphism onto its image. Now $V = g^{-1}(p(W)) \subset Z$ is an open neighborhood of a. By local path-connectedness there exists a path-connected open neighborhood $U \subset V$ of a. Since U is path-connected with $g(U) \subset p(W)$, it then follows from the construction of h that $h(U) \subset W$. □

Before proceeding with the general theory in parallel to the groupoid case, let us illustrate the role of local path-connectedness by a pathological example:

Example 4.4. The *topologist's sine curve*

$$S := \{(x, \sin(2\pi/x) \mid 0 < x \leq 1\} \cup \{(0,0)\} \subset \mathbb{R}^2$$

is connected, but neither path-connected nor locally path-connected: The origin $(0,0) \in S$ cannot be connected to any other point of S by a continuous path. The following picture illustrates the situation:

[PICTURE]

We can make a path-connected version of the above example by removing a small segment from a circle and replacing it by the topologist's sine curve. The resulting topological space Z is called the *Warsaw circle*:

[PICTURE]

By construction the Warsaw circle Z is path-connected. In fact it is even simply connected, since by continuity it is impossible for any path $\gamma\colon [0,1] \to Z$ to cross the origin. But Z is not locally path-connected, and in fact the lifting property from theorem 4.3 fails in this case: For example, let $g\colon Z \to X = S^1$ be the continuous map given by collapsing the topologist's sine curve. The path lifting property for the covering map $p\colon Y = \mathbb{R} \to X, y \mapsto \exp(2\pi i y)$ gives a map $h\colon Z \to \mathbb{R}$ with $h(0,0) = 0$ such that $p \circ h = g$. But h cannot be continuous:

- $h(z) \to 0$ as $z \to (0,0)$ from the left.
- $h(z) \to 1$ as $z \to (0,0)$ from the right.

As in the abstract setting, given two covers $p_i\colon Y_i \to X$ of a topological space, we define a *morphism of covers* between them to be a continuous map $f\colon Y_1 \to Y_2$ such that the diagram

$$\begin{tikzcd} Y_1 \arrow[rr, "f"] \arrow[dr, "p_1"'] & & Y_2 \arrow[dl, "p_2"] \\ & X & \end{tikzcd}$$

commutes. We denote by $\mathrm{Hom}_X(Y_1, Y_2)$ the set of such morphisms of covers. As in remark one verifies that any morphism from a cover to a connected cover is again a covering map. There is also a version of these notions for pointed topological spaces, which is defined in the same way as for pointed groupoids. Theorem 2.10 now has the following topological counterpart:

Theorem 4.5. *Let $p_i\colon (Y_i, y_i) \to (X, x)$ for $i = 1, 2$ be pointed covers of topological spaces which are connected and locally path-connected.*

a) There exists a morphism $h \in \mathrm{Hom}_{(X,x)}((Y_1, y_1), (Y_2, y_2))$ of pointed covers iff we have an inclusion

$$p_{1,*}(\pi_1(Y_1, y_1)) \subset p_{2,*}(\pi_1(Y_2, y_2)) \quad \textit{inside} \quad \pi_1(X, x).$$

b) More generally, there is a morphism $h \in \mathrm{Hom}_X(Y_1, Y_2)$ iff for some $\gamma \in \pi_1(X, x)$ we have an inclusion

$$p_{1,*}(\pi_1(Y_1, y_1)) \subset \gamma \cdot p_{2,*}(\pi_1(Y_2, y_2)) \cdot \gamma^{-1} \quad \textit{inside} \quad \pi_1(X, x).$$

c) In both cases h is an isomorphism iff the inclusion $\subset$ is an equality.

Proof. Immediate from theorem 4.3. □

5 The universal cover of a topological space

We want to reformulate the above as a Galois correspondence as we did in the abstract case of groupoids. Recall that the main point of the Galois correspondence was to have a universal cover in the sense of groupoids. The corresponding notion for topological spaces is the following:

Definition 5.1. A *universal cover* of a topological space X is a cover $p\colon \tilde{X} \to X$ from a simply connected topological space $\tilde{X}$.

Note that if such a universal cover exists, then X has to be path-connected. If we moreover assume X to be locally path-connected, we can actually talk about *the* universal cover since it is determined uniquely:

Corollary 5.2. *Up to isomorphism, any connected, locally path-connected space X admits at most one universal cover $p\colon \tilde{X} \to X$. Moreover, if such a universal cover exists, then it dominates any other connected cover $q\colon Y \to X$:*

$$\begin{tikzcd} \tilde{X} \arrow[dd, "p"'] \arrow[dr, dotted, "\exists h"] & \\ & Y \arrow[dl, "q"] \\ X & \end{tikzcd}$$

Proof. If X is locally path-connected, then the definition of covering maps implies that any cover of X is again locally path-connected. The statement therefore follows immediately from theorem 4.5. □

For abstract groupoids we have seen that the universal cover always exists. The situation is somewhat different for topological spaces:

Remark 5.3. If the universal cover $p\colon \tilde{X} \to X$ exists, any point $x \in X$ has an open neighborhood $U \subset X$ such that every loop in this neighborhood is contractible in X, i.e. such that

$$\operatorname{im}(\pi_1(U,x) \to \pi_1(X,x)) \;=\; \{1\}.$$

Proof. Since $p\colon \tilde{X} \to X$ is a covering map, any point $x \in X$ has a neighborhood $U \subset X$ whose preimage is homeomorphic to a disjoint union of copies of U. In particular, any loop $\gamma\colon S^1 \to U$ lifts not only to a path but to a loop $\tilde{\gamma}\colon S^1 \to p^{-1}(U) \subset \tilde{X}$. The lifted loop is contractible in $\tilde{X}$ since $\pi_1(\tilde{X},*) = \{0\}$. But then the original loop $\gamma = p(\tilde{\gamma})$ is contractible in X, as claimed. □

Note that the loop γ need not be contractible in $U \subset X$, it is only so in X. To distinguish the two properties we introduce the following terminology:

Definition 5.4. A topological space X is

a) *locally simply connected* if for every point $x \in X$, every open neighborhood $V \subset X$ of x contains an open neighborhood $U \subset V$ of x with $\pi_1(U,x) = \{0\}$.

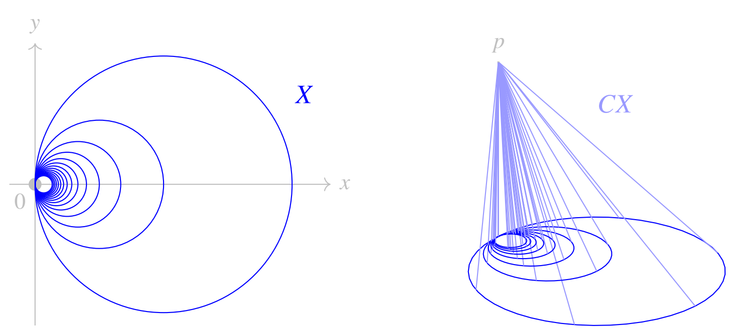
b) *semilocally simply connected* if for every $x \in X$, every open neighborhood $V \subset X$ of x contains an open neighborhood $U \subset V$ of x with

$$\operatorname{im}(\pi_1(U,x) \to \pi_1(X,x)) \;=\; \{1\}.$$

Example 5.5. By the *Hawaiian earring* we mean the union

$$X \;=\; \bigcup_{n\geq 1} \left\{ z \in \mathbb{C} \mid |z - \tfrac{1}{n}| = \tfrac{1}{n} \right\}$$

of circles of radius $1/n$ which all meet tangentially at the origin, as shown in the first of the two pictures below:



The topology on the Hawaiian earring differs from the one on an infinite wedge sum of circles, for instance X is not semilocally simply connected: Any neighborhood of the origin in X contains infinitely many circles which give rise to loops that are not contractible in X. In particular X is not locally simply connected. On the other hand, let

$$CX = (X \times [0,1])/\sim \quad \text{with } (x,0) \sim (y,0) \text{ for all } x,y \in X$$

be the cone shown in the second of the above pictures. This cone is not locally simply connected either, but it *is* semilocally simply connected because it is even contractible (it deformation retracts to the cone point $p = (*,0)$).

Remark 5.3 says that a necessary condition for the existence of a universal cover of a topological space is that the space is semilocally simply connected. In fact this condition is also sufficient:

Theorem 5.6. *A connected, locally path-connected topological space X admits a universal cover if and only if it is semilocally simply connected.*

Proof. Suppose X is semilocally simply connected, and fix a point $x \in X$. We apply the construction of covers of groupoids in the proof of theorem 3.5 to $G = \Pi(X)$ and $\Gamma_0 = \{1\} \leq G_x = \pi_1(X,x)$: As a set we define

$$\tilde{X} := G_{x,\star} = \{\text{paths } \gamma\colon [0,1] \to X \text{ with } \gamma(0) = x\}/\sim$$

where $\sim$ denotes homotopy of paths in X relative to the end points, and we consider the map

$$p\colon \quad \tilde{X} \to X, \quad [\gamma] \mapsto \gamma(1).$$

We will define on $\tilde{X}$ a simply connected topology making p a covering map.

In order to do so, recall that X is semilocally simply connected, so it admits a basis of open subsets $U \subset X$ such that the induced morphisms $\pi_1(U,u) \to \pi_1(X,u)$ are trivial. In what follows we will refer to these simply as *basic open sets*. We will construct the topology so that p restricts to a trivial cover over each basic open subset. Concretely, fix a basic open subset $U \subset X$. The fiber $p^{-1}(u)$ over a point $u \in U$ is the set of paths $\gamma\colon [0,1] \to X$ from x to u, taken modulo homotopies in X that preserve the end points of the paths. For any such γ we put

$$U_\gamma := \{\gamma \cdot \mu \mid \mu\colon [0,1] \to U \text{ is a path with } \mu(0) = u\}/\sim .$$

With this notation we have

$$p^{-1}(U) = \bigsqcup_{[\gamma]} U_\gamma \tag{$\star$}$$

where the disjoint union ranges over all homotopy classes $[\gamma] \in p^{-1}(u)$: Indeed, given any $\lambda \in p^{-1}(U)$, let $\mu\colon [0,1] \to U$ be a path from u to $\lambda(1)$. Our definition of the basic open subsets guarantees that any two different choices of such a path μ

will become homotopic in X. So the path $\gamma = \lambda\mu^{-1}$ is unique up to homotopy in X and we have

$$\lambda \in U_\gamma \quad \text{for this unique homotopy class} \quad [\gamma] \in p^{-1}(u),$$

thereby proving $(\star)$. Note that here we used that X is locally simply connected.

We next claim that the collection of all sets of the form $U_\gamma \subset \tilde{X}$ forms a basis for a topology on the set $\tilde{X}$: Indeed, given two such sets U_γ and V_λ and a point $\mu \in U_\gamma \cap V_\lambda$, we have

$$W_\mu \subset U_\gamma \cap V_\lambda$$

for any basic open subset $W \subset U \cap V$ containing the point $\mu(1)$. If we endow $\tilde{X}$ with this topology, it follows immediately from the definitions that on each of our chosen basic open subsets $U_\gamma \subset \tilde{X}$, the restriction $p|_{U_\gamma}\colon U_\gamma \to U$ is a homeomorphism. So p is a covering map in view of equation $(\star)$. Moreover, another tautological argument shows that $\tilde{X}$ is path-connected and locally path-connected. To see that $\tilde{X}$ is simply connected, it suffices to show for any $\tilde{x} \in p^{-1}(x)$ that as a subgroup of $\pi_1(X,x)$ we have

$$p_*(\pi_1(\tilde{X},\tilde{x})) = \{1\}.$$

By the dictionary between covers and their monodromy operation, this last property is equivalent to saying that $\pi_1(X,x)$ acts freely on the fiber $p^{-1}(x)$. This is again a tautology: By construction this fiber is the set of homotopy classes of loops based at x, i.e. we have a natural identification

$$p^{-1}(x) = \pi_1(X,x),$$

and via this identification the monodromy action is given by right multiplication of the group $\pi_1(X,x)$ on itself. □

6 Classification of topological covers

The existence of the universal cover means that the trivial subgroup of $\pi_1(X,x)$ is realized as $p_*(\pi_1(Y,y))$ for some connected cover $p\colon (Y,y) \to (X,x)$. Once we have this, we can in fact realize *any* subgroup of $\pi_1(X,x)$ in the same way, so we obtain a full Galois correspondence as in the case of abstract groupoids. The idea is that any cover will be a quotient of the universal one:

Definition 6.1. Let Γ be an abstract group, acting on a topological space Y from the left via homeomorphisms. We define the *quotient space* $\Gamma\backslash Y$ to be the topological space given by the set of orbits

$$\Gamma\backslash Y := \{\Gamma y \mid y \in Y\},$$

endowed with the coarsest topology making the quotient map $Y \to \Gamma\backslash Y$ continuous.

Example 6.2. Quotient maps need not be covering maps:

a) Let $\Gamma = \{\pm 1\}$ act on $Y = \mathbb{C}$ via multiplication. The quotient map can be identified with the squaring map

$$p\colon \quad \mathbb{C} \to \Gamma\backslash\mathbb{C} \simeq \mathbb{C}, \quad z \mapsto z^2$$

which is not a covering map (the fiber over the origin consists of only one point but the fibers over all other points consist of two distinct points).

b) The problem in the above example was that the origin was a fixed point of the action. But even free actions can pose problems: For instance, let $\alpha \in \mathbb{R} \smallsetminus \mathbb{Q}$, and let $\Gamma = \mathbb{Z}$ act on $Y = S^1$ via rotations by multiples of $2\pi\alpha$:

$$S^1 \times \mathbb{Z} \to S^1, \quad (z,n) \mapsto e^{2\pi i n\alpha} z.$$

Since α is irrational, this action is free. But the quotient map $p\colon S^1 \to \Gamma\backslash S^1$ is not a covering map: For instance, the quotient space $\Gamma\backslash S^1$ is not Hausdorff, but for any covering map the target is Hausdorff iff the source is.

To avoid the above type of problems, we make the following definition:

Definition 6.3. Let Y be a topological space. A group Γ of homeomorphisms from Y to itself is said to act *properly discontinuously* if every point $y \in Y$ admits an open neighborhood $U \subset Y$ with

$$\gamma(U) \cap U = \varnothing \quad \text{for all } \gamma \in \Gamma \smallsetminus \{id\}.$$

Lemma 6.4. *Let Y be connected, and let Γ be a group of homeomorphisms from Y to itself. Then the following are equivalent:*

a) The quotient map $p\colon Y \to X = \Gamma\backslash Y$ is a cover.

b) The action of Γ on Y is properly discontinuous.

Proof. Let $U \subset X$ be an open subset. The image $V = p(U)$ is open in $X = \Gamma\backslash Y$ by definition of the quotient topology, since its preimage

$$p^{-1}(V) = \bigcup_{\gamma \in \Gamma} \gamma(U) \subset Y$$

is open. If Γ acts properly discontinuously on Y, then any $x = p(y) \in X$ has an open neighborhood $U \subset Y$ of y with the property in definition 6.3. This property implies that the union displayed above is disjoint: For $\gamma_1, \gamma_2 \in \Gamma$ with $\gamma_1(U) \cap \gamma_2(U) \neq \varnothing$ the element $\gamma = \gamma_2^{-1}\gamma_1$ satisfies $\gamma(U) \cap U \neq \varnothing$, forcing $\gamma = id$ and therefore $\gamma_1 = \gamma_2$ as desired. Moreover, it is clear by construction that the quotient map p restricts to a bijective map

$$p|_{\gamma(U)}\colon \quad \gamma(U) \xrightarrow{\sim} U$$

for each fixed $\gamma \in \Gamma$ since for any two points $u_1, u_2 \in U$ with $p(\gamma(u_1)) = p(\gamma(u_2))$ there exists $\mu \in \Gamma$ with $\gamma(u_1) = \mu(\gamma(u_2))$, which by the pairwise disjointness from the previous step forces that $\gamma = \mu\gamma$ and hence that $\mu = id$ as required. Finally, it follows from the definition of the quotient topology that the set-theoretic disjoint union

$$p^{-1}(V) = \bigsqcup_{\gamma\in\Gamma} \gamma(U)$$

is in fact a topological one and the bijection $p|_{\gamma(U)}$ is a homeomorphism. Thus p is indeed a covering map.

Conversely, suppose the quotient map $p\colon Y \to X = \Gamma\backslash Y$ is a covering map. Then by definition any point $x = p(y)$ has an open neighborhood $V \subset X$ such that $p^{-1}(V)$ is topologically a disjoint union of copies of V. Let $U \subset Y$ be one of those disjoint copies. We claim that U has the property in definition 6.3: Indeed, if $\gamma(U) \cap U \neq \varnothing$, then there exist $u_1, u_2 \in U$ with $\gamma(u_1) = u_2$. Thus $p(u_1) = u_2$. But since by our choice of U the restriction $p|_U$ is injective, this forces $u_1 = u_2$. Thus we have

$$\gamma(u_1) = u_1.$$

Since $\gamma \in \mathrm{Aut}_X(Y)$, the uniqueness of lifts under covering maps then implies $\gamma = id$ and hence the property in definition 6.3 holds. □

We will use the above quotient construction to set up a Galois correspondence for covering maps of topological spaces that adds a topological layer to the one for groupoids in corollary 3.6. For this we introduce the following notion:

Definition 6.5. A cover $p\colon Y \to X$ of connected topological spaces is a *Galois cover* if the induced cover of fundamental groupoids $\Pi(Y) \to \Pi(X)$ is Galois, i.e. if for any $x \in X$ and $y \in p^{-1}(x)$,

$$\pi_1(Y,y) \trianglelefteq \pi_1(X,x) \quad \text{is a normal subgroup.}$$

Example 6.6. Being Galois is a symmetry condition:

a) With the same drawing conventions as in example 1.3, let $Z \to X = S^1 \vee S^1$ be the following cover:

Put oOo with b: small loops, a: large half-loop (in the middle)

This cover is Galois, it corresponds to the normal subgroup

$$\pi_1(Y,y) = \langle a^{e_1} b^{f_1} \cdots a^{e_k} b^{f_k} \mid e_1 + \cdots + e_k \equiv 0 \bmod 2 \rangle \trianglelefteq \pi_1(X,x) = \langle a,b \rangle.$$

b) On the other hand, let $Y \to X = S^1 \vee S^1$ be the following cover of the figure eight graph:

Put oOoO here

This cover is not Galois: The subgroup $\pi_1(Y,y) \leq \pi_1(X,x) = \langle a,b \rangle$ is not normal, for instance

$$a \in \pi_1(Y,y) \quad \text{but} \quad bab^{-1} \notin \pi_1(Y,y).$$

In general, Galois covers can be understood as quotient maps:

Remark 6.7. For connected locally path-connected spaces, theorem 4.5 says that any automorphism of covers of fundamental groupoids is induced by a continuous deck transformation. In this case, a cover $p\colon Y \to X$ is Galois iff the group $\mathrm{Aut}_X(Y)$ of such deck transformations acts transitively on any fiber $p^{-1}(x)$. In the Galois case we then get for any $y \in p^{-1}(x)$ a natural identification

$$\mathrm{Aut}_X(Y) \simeq \pi_1(X,x)/\pi_1(Y,y)$$

of the group of deck transformations with a quotient of the fundamental group, and in these terms

$$X \simeq \Gamma \backslash Y$$

is the quotient of Y by the group $\Gamma = \mathrm{Aut}_X(Y)$ of deck transformations. This allows us to construct intermediate covers: For any subgroup $\Gamma_0 \leq \Gamma$ we can consider the partial quotient

$$Y_0 := \Gamma_0 \backslash Y$$

and get a diagram as shown below where all three maps are covering maps:

$$\begin{tikzcd} Y \arrow[dd, "p"'] \arrow[dr, "\tilde{q}"] & \\ & Y_0 \arrow[dl, "q"] \\ X & \end{tikzcd}$$

One easily checks that $\tilde{q}$ is again a Galois cover with

$$\mathrm{Aut}_{Y_0}(Y) \simeq \Gamma_0,$$

while q is Galois iff the subgroup $\Gamma_0 \trianglelefteq \Gamma$ is normal, in which case $\mathrm{Aut}_X(Y_0) \simeq \Gamma/\Gamma_0$.

By applying this to the universal cover $Y \to X$ we get a topological version of the Galois correspondence of corollary 3.6. Recall that as we have seen in theorem 5.6, a topological space X has a universal cover iff it is connected, locally path-connected and semilocally simply connected. In this case we have:

Corollary 6.8. *Let X be a topological space that has a universal cover $p\colon \tilde{X} \to X$, and fix arbitrary base points $x \in X$ and $\tilde{x} \in p^{-1}(x)$. Then we have an equivalence of categories*

$$\varphi\colon \quad \{\textit{connected pointed covers of } (X,x)\} \xrightarrow{\sim} \{\textit{subgroups of } \pi_1(X,x)\}$$

$$(q\colon (Y,y) \to (X,x)) \mapsto q_*(\pi_1(Y,y))$$

where morphisms of covers correspond to inclusions of subgroups. Explicitly, any subgroup $\Gamma_0 \le \pi_1(X,x)$ is realized by the cover

$$q\colon \quad \Gamma_0 \backslash (\tilde{X},\tilde{x}) \to (X,x)$$

Proof. As in the Galois correspondence for abstract groupoids, only the essential surjectivity remains to be shown. By the above discussion we may identify the group of deck transformations of the universal cover with the fundamental group of the base space:

$$\mathrm{Aut}_X(\tilde{X}) \simeq \pi_1(X,x).$$

So for any $\Gamma_0 \le \Gamma = \pi_1(X,x)$ we may form the quotient space $Y_0 = \Gamma_0 \backslash \tilde{X}$. We have a factorization

$$p = q \circ \tilde{q}\colon \quad \tilde{X} \xrightarrow{\tilde{q}} Y_0 = \Gamma_0 \backslash \tilde{X} \xrightarrow{q} X = \Gamma \backslash \tilde{X},$$

hence q is again a covering map. Moreover, from the identification $p^{-1}(x) \simeq \Gamma$ we get an identification

$$q^{-1}(x) \simeq \Gamma_0 \backslash \Gamma$$

which is compatible with the monodromy operation, hence $q_*(\pi_1(Y_0,y)) = \Gamma_0 \le \Gamma$ by the classification of covers of groupoids via their monodromy operation. □

Example 6.9. Let $\Gamma \subset \mathbb{R}^2$ be a lattice, i.e. an abelian subgroup of rank two generated by two $\mathbb{R}$-linearly independent vectors. Then $X = \mathbb{R}^2/\Gamma$ is a torus, and the quotient map

$$p\colon \quad \tilde{X} = \mathbb{R}^2 \longrightarrow X = \mathbb{R}^2/\Gamma$$

is a cover by a simply connected space, hence coincides with the universal cover of the torus. From the Galois correspondence we then recover that $\pi_1(X,*) \simeq \Gamma$ is a free abelian group of rank two. Moreover, every connected cover of X is isomorphic to the Galois cover

$$q\colon \quad Y = \mathbb{R}^2/\Gamma_0 \longrightarrow X = \mathbb{R}^2/\Gamma \quad \text{with} \quad \mathrm{Aut}_X(Y) = \Gamma/\Gamma_0$$

for some subgroup $\Gamma_0 \le \Gamma$. For $\mathrm{rk}(\Gamma_0) = 1$ the cover Y is homeomorphic to an infinite cylinder whereas for $\mathrm{rk}(\Gamma_0) = 2$ it is again homeomorphic to a torus.

We have used the Galois correspondence to construct all connected covers of a space with known fundamental group, but we can also go in the other direction:

Lemma 6.10. *Let Y be a simply connected topological space, and let Γ be a group of homeomorphisms acting properly discontinously on Y. Then the quotient space $X = \Gamma\backslash Y$ has fundamental group*

$$\pi_1(X,*) \simeq \Gamma.$$

Proof. Since the action of Γ is properly discontinuous, the quotient map $p\colon Y \to X$ is a cover by lemma 6.4. Since Y is simply connected, it is in fact the universal cover of X and therefore the Galois correspondence gives $\Gamma = \mathrm{Aut}_X(Y) \simeq \pi_1(X,*)$. □

Example 6.11. For $m,n \in \mathbb{N}$, fix integers $\ell_i \in \mathbb{Z}$ with $\gcd(\ell_i,m) = 1$ for $i = 1,\dots,n$, and endow the sphere

$$Y = S^{2n-1} \subset \mathbb{C}^n$$

with the action of $\Gamma = \mathbb{Z}/m\mathbb{Z}$ defined by

$$k\cdot(y_1,\dots,y_n) = (e^{2\pi ik\ell_1/m}\cdot y_1,\dots,e^{2\pi ik\ell_n/m}\cdot y_n) \quad \text{for} \quad k \in \Gamma.$$

This action is free and hence properly discontinuous since the group in question is finite. The quotient

$$X_{m,\ell_1,\dots,\ell_n} = \Gamma\backslash Y$$

is called a *lens space*. For instance, if $m = 2$, then the action of Γ on the sphere is given by the antipodal map and in this case the corresponding lens space becomes the real projective space $\mathbb{RP}^{2n-1} = S^{2n-1}/\langle\pm 1\rangle$. In general, from lemma 6.10 we get that

$$\pi_1(X_{m,\ell_1,\dots,\ell_n},*) \simeq \mathbb{Z}/m\mathbb{Z}.$$

Example 6.12. The 2-torus, the Klein bottle and the real projective plane can all be obtained by identifying opposite sides of the unit square with suitable orientations as shown below:

PICTURE

We can write them as follows as the quotient of a simply connected space Y by a subgroup $\Gamma \leq \mathrm{Aut}(Y)$ of homeomorphisms acting properly discontinuously:

a) For the torus T we take $\Gamma = \mathbb{Z}^2$ acting on $Y = \mathbb{R}^2$ via translations. Hence we recover our earlier result that the torus has fundamental group $\pi_1(T,*) \simeq \mathbb{Z}^2$.

b) For the Klein bottle K we can still take $Y = \mathbb{R}^2$ but now the subgroup Γ is generated by

- a translation τ: $(x,y) \mapsto (x+1,y)$, and
- a glide reflection σ: $(x,y) \mapsto (-x,y+1)$.

These two generators do not commute, indeed

$$(\sigma^{-1}\tau\sigma)(x,y) = \sigma^{-1}(\tau(-x,y+1)) = \sigma^{-1}(-x+1,y+1) = (x-1,y) = \tau^{-1}(x,y)$$

Thus the fundamental group of the Klein bottle is

$$\pi_1(K,*) \simeq \langle \sigma,\tau \mid \sigma^{-1}\tau\sigma = \tau^{-1}\rangle \simeq \mathbb{Z} \rtimes \mathbb{Z},$$

where the semidirect product uses the action $\mathbb{Z} \to \mathrm{Aut}(\mathbb{Z}), n \mapsto (-1)^n \cdot id$. Note that we have a double cover

$$p\colon T \to K \quad \text{given by} \quad (x,y) \mapsto (x,2y),$$

and on the level of fundamental groups this is given by the subgroup

$$\pi_1(T,*) = \mathbb{Z} \times 2\mathbb{Z} \subset \pi_1(K,*) = \mathbb{Z} \rtimes \mathbb{Z}.$$

c) The real projective plane $\mathbb{RP}^2$ cannot be written as the quotient of $\mathbb{R}^2$ by a properly discontinuous group action (in contrast to the previous examples, the tiling of the plane by copies of the unit square with the given orientations now has rotational symmetries leading to fixed points). Instead $\mathbb{RP}^2$ is the quotient of the sphere $Y = S^2$ modulo the properly discontinous action of $\Gamma = \mathbb{Z}/2\mathbb{Z}$ via the antipodal involution. We can relate this to the previous picture by identifying the upper and the lower hemisphere with copies of the unit square as shown below:

PICTURE

Thus the fundamental group of the real projective plane is $\pi_1(\mathbb{RP}^2,*) \simeq \mathbb{Z}/2\mathbb{Z}$.

Chapter IV
Homology groups

1 Simplicial homology

Fundamental groups are defined via loops (which are 1-dimensional objects), so they lose a lot of information about higher-dimensional geometry: E.g. they cannot distinguish between the spheres $S^n \subset \mathbb{R}^{n+1}$ for $n \geq 2$, even though we will see that two spheres of different dimension are never homotopy equivalent. The natural next step would be to introduce higher homotopy groups, where instead of loops one considers homotopy classes of maps

$$\gamma\colon \quad S^n \to X.$$

However, these higher homotopy groups are notoriously hard to compute, even for spheres $X = S^m$ they are in general unknown when $n \gg m$. In this chapter we will introduce a family of much simpler homotopy invariants, the *homology groups* of a topological space. These will be very easy to compute and yet capture a lot of information about higher-dimensional features of a topological space.

There are various versions of homology groups available. The most intuitive one, simplicial homology, will be defined on a class of topological spaces that come with a triangulation as in the following picture:

[PICTURE]

Example 1.1. The torus can be obtained by gluing together two triangles in the way shown below:

[PICTURE]

In other words, it can be constructed by taking a vertex, attaching to this vertex three intervals, and then attaching to the resulting space two triangles. In general, the basic building blocks for triangulations are simplices:

Definition 1.2. For $n \in \mathbb{N}_0$ we define the *n-dimensional standard simplex* to be the convex hull

$$\Delta^n = \left\{ (t_0, \dots, t_n) \in \mathbb{R}^{n+1} \mid t_0 + \dots + t_n = 1 \text{ and } t_i \geq 0 \text{ for all } i \right\}$$

For $i \in \{0, 1, \dots, n\}$, we define the i-th *face* of the standard simplex to be the image of the *face map*

$$\varphi_i\colon \quad \Delta^{n-1} \hookrightarrow \Delta^n, \quad (t_0, \dots, t_{i-1}, t_{i+1}, \dots, t_n) \mapsto (t_0, \dots, t_{i-1}, 0, t_{i+1}, \dots, t_n).$$

The union of all these faces is called the *boundary* $\partial\Delta^n \subset \Delta^n$ of the simplex, and its complement will be denoted

$$\mathring{\Delta}^n = \Delta^n \setminus \partial\Delta^n$$

[PICTURE]

There are various notions of triangulations of topological spaces, which differ in the way in which simplices are allowed to be glued together. We will here adopt the following simple but very flexible notion:

Definition 1.3. A *Δ-complex* is a pair $\mathbb{X} = (X, (\sigma_\alpha)_{\alpha \in I})$ of a topological space X and a family of continuous maps

$$\sigma_\alpha\colon \quad \Delta^{n_\alpha} \to X \quad \text{for suitable} \quad n_\alpha \in \mathbb{N}_0,$$

indexed by some set I, such that the following properties are satisfied:

a) The restriction of each σ_α to the corresponding open simplex is an injective map
$$\mathring{\sigma}_\alpha\colon \quad \mathring{\Delta}^{n_\alpha} \hookrightarrow X,$$
and if $\mathring{\Delta}_\alpha \subset X$ denotes the image of this injective map, then as a set X is the disjoint union
$$X = \bigsqcup_{\alpha\in I} \mathring{\Delta}_\alpha$$

b) A subset $U \subset X$ is open iff its preimage $\sigma_\alpha^{-1}(U) \subset \Delta^{n_\alpha}$ is open for all α.

c) For each α and each $i \in \{0,1,\ldots,n_\alpha\}$, the composite of σ_α with the i-th face map φ_i is again of the form
$$\sigma_\alpha \circ \varphi_i = \sigma_\beta \quad \text{for some} \quad \beta \in I.$$

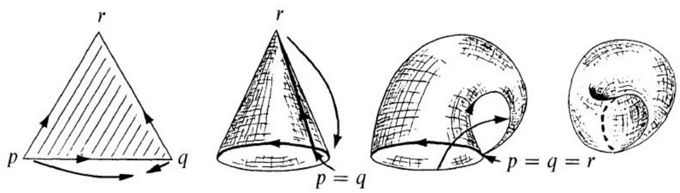
Remark 1.4. The literature often uses the notion of a *simplicial complex*, which are topological spaces obtained by gluing together simplices in a way so that no two points on the same simplex are identified and any two simplices are either disjoint or their intersection is another simplex. This requires more simplices, e.g. the smallest structure of a simplicial complex on a torus requires 7 vertices, 21 edges and 14 triangles as discovered by Möbius in 1861:

[PICTURE] Csaszar polyhedron

Thus Δ-complexes are easier to handle than simplicial complexes, although one may show that any Δ-complex can be subdivided to get a simplicial complex. For the rest of this section we will stick with Δ-complexes for simplicity.

Note that the vertices of each standard simplex come in a natural order, and the face maps ι_i respect this ordering. This imposes a restriction on identifications between faces of simplices:

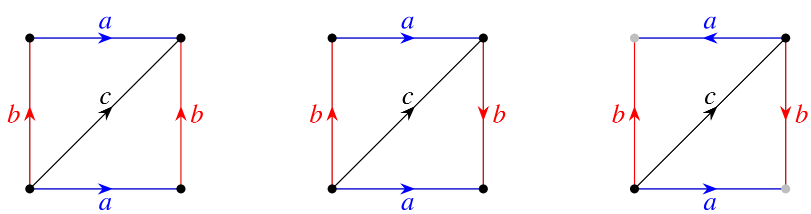
Example 1.5. If we identify the three edges of a 2-simplex with the orientation given by the labelling of their vertices, then we obtain a Δ-complex which is known as the *Dunce hat* (the picture below is from Bredon's book *Topology and Geometry*):



Identifying the three edges with a different orientation as in the first picture below does not give us a Δ-complex. However, we can subdivide it to a Δ-complex on the same underlying topological space as shown in the second picture below:

[PICTURE]

Example 1.6. We still have orientation choices in the way in which simplices are glued together. For instance, the torus, real projective space and the Klein bottle can all be described by a Δ-complex that is glued from two triangles:



Note that the Δ-complex for the real projective plane has *two* 0-simplices, unlike those for the torus and for the Klein bottle which have only one 0-simplex.

In what follows, we will attach to any Δ-complex $\mathbb{X} = (X, (\sigma_\alpha)_{\alpha \in I})$ a series of natural invariants, its homology groups. In contrast to the fundamental group, these will always be abelian groups but see information in dimension > 1. As a motivation let us start with the one-dimensional picture:

Example 1.7. Suppose that we want to compute the abelianization

$$\pi_1^{ab}(X,x) \ = \ \pi_1(X,x)/[\pi_1(X.x), \pi_1(X,x)],$$

i.e. the maximal abelian quotient of the fundamental group. This abelianization has a combinatorial description in terms of the simplices in our Δ-complex: If we choose x to be one of the 0-simplices, then clearly any loop based at x will be homotopic to a loop passing only through 1-simplices. In the abelianization $\pi_1^{ab}(X,x)$ it does not matter in which order the 1-simplices are passed, all we need no know is how often the loop passes along each 1-simplex. Thus if for $k \in \mathbb{N}_0$ we denote by

$$I(k) = \{\alpha \in I \mid n_\alpha = k\}$$

the set of indices parametrizing the k-simplices in our given Δ-complex, we can record elements of $\pi_1^{ab}(X,x)$ as a formal linear combinations

$$\sum_{\alpha \in I(1)} m_\alpha \cdot \sigma_\alpha$$

where $m_\alpha \in \mathbb{Z}$ says how often the loop passes along the 1-simplex σ_α. Here we put negative signs for passages through a 1-simplex in the opposite orientation, recalling that all our simplices come with a given orientation. However, this does not yet give a complete description of $\pi_1^{ab}(X,x)$:

a) The linear combinations coming from elements of $\pi_1^{ab}(X,x)$ have the property that for any given 0-simplex p, the total number of times it occurs as a starting point of 1-simplices in the loop agrees with the total number of times it occurs as an end point:
$$\sum_{\sigma_\alpha \circ \varphi_1 = p} m_\alpha \;-\; \sum_{\sigma_\beta \circ \varphi_0 = p} m_\beta \;=\; 0.$$

b) Any loop which is the boundary of a 2-simplex is contractible. Taking into account the orientation of the faces, we thus have for each index $\alpha \in I(2)$ the relation
$$\sigma_\alpha \circ \varphi_0 - \sigma_\alpha \circ \varphi_1 + \sigma_\alpha \circ \varphi_2 \sim 0 \quad \text{in} \quad \pi_1^{ab}(X,x).$$

Both conditions have a homological interpretation. To explain this, consider for $k \in \mathbb{N}_0$ the free abelian group

$$C_k(\mathbb{X}) = \bigoplus_{\alpha \in I(k)} \mathbb{Z} \cdot [\sigma_\alpha]$$

on the family of k-simplices in our given Δ-complex. Here $[\sigma_\alpha] \in C_k(\mathbb{X})$ denotes a formal basis vector corresponding to the index $\alpha \in I(k)$. In these terms, condition a) is equivalent to saying that

$$\gamma := \sum_{\alpha \in I(1)} m_\alpha \cdot [\sigma_\alpha] \in C_1(\mathbb{X})$$

lies in the kernel of the linear map

$$\delta_1\colon \quad C_1(\mathbb{X}) \to C_0(\mathbb{X}), \quad [\sigma_\alpha] \mapsto [\sigma_\alpha \circ \varphi_1] - [\sigma_\alpha \circ \varphi_0],$$

while b) amounts to saying that γ only matters modulo addition of elements in the image of the map

$$\delta_2\colon \quad C_2(\mathbb{X}) \to C_1(\mathbb{X}), \quad [\sigma_\alpha] \mapsto [\sigma_\alpha \circ \varphi_0] - [\sigma_\alpha \circ \varphi_1] + [\sigma_\alpha \circ \varphi_2].$$

An immediate computation shows $\delta_1 \circ \delta_2 = 0$, so we have $\mathrm{im}(\delta_2) \subset \ker(\delta_1)$. It is now a simple exercise to show that if X is path-connected, then the above construction gives a natural isomorphism of groups

$$\pi_1^{ab}(X,x) \simeq \ker(\delta_1)/\mathrm{im}(\delta_2).$$

The quotient on the right hand side is the first homology group of the given Δ-complex:

Definition 1.8. Let $\mathbb{X} = (X,(\sigma_\alpha)_{\alpha\in I})$ be a Δ-complex, and for $k \in \mathbb{N}_0$ let $C_k(\mathbb{X})$ be the free abelian group generated by the set of its k-simplices. We define the *k-th boundary map* to be the linear map

$$\delta_k\colon \quad C_k(\mathbb{X}) \to C_{k-1}(\mathbb{X}), \quad [\sigma_\alpha] \mapsto \sum_{i=0}^{k} (-1)^i \cdot [\sigma_\alpha \circ \varphi_i].$$

An immediate computation shows that $\delta_k \circ \delta_{k+1} = 0$ and hence $\mathrm{im}(\delta_{k+1}) \subset \ker(\delta_k)$ for all $k \in \mathbb{N}_0$ (for $k = 0$ we put $\delta_0 = 0$). In other words

$$C_\bullet(\mathbb{X}) := \big[\cdots \to C_k(\mathbb{X}) \to C_{k-1}(\mathbb{X}) \to \cdots \to C_1(\mathbb{X}) \to C_0(\mathbb{X}) \big]$$

is a complex of abelian groups. The homology groups

$$H_k(\mathbb{X}) := H_k(C_\bullet(\mathbb{X})) = \ker(\delta_k)/\mathrm{im}(\delta_{k+1})$$

of this complex are called the *simplicial homology* of the Δ-complex $\mathbb{X}$.

Example 1.9. Consider the 2-torus, the Klein bottle and the real projective plane with the structure of Δ-complexes from example 1.6: Let σ,τ be the two triangles in the picture with their boundary oriented counter-clockwise, and let a,b,c be the three edges with the shown orientations. Then in all three cases the corresponding chain complex takes the form

$$C_2(\mathbb{X}) = \mathbb{Z}\sigma \oplus \mathbb{Z}\tau \xrightarrow{\delta_2} C_1(\mathbb{X}) = \mathbb{Z}a \oplus \mathbb{Z}b \oplus \mathbb{Z}c \xrightarrow{\delta_1} C_0(\mathbb{X}) = \mathbb{Z}^n$$

where for simplicity we have omitted the brackets around simplices. Here $n = 1$ for the 2-torus and for the Klein bottle, but $n = 2$ for the real projective plane.

a) For the torus T we have $\delta_0 = 0$ since there is only one vertex, so the boundary of every edge vanishes. Hence

$$H_0(\mathbb{X}) = \mathrm{cok}(\delta_1) \simeq \mathbb{Z}.$$

Moreover $\delta_2(\sigma) = a+b-c$ and $\delta_2(\tau) = -a-b+c$, so the linear map $\delta_2\colon \mathbb{Z}^3 \to \mathbb{Z}^2$ is given by the matrix

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Thus $H_2(\mathbb{X}) = \ker(\delta) \simeq \mathbb{Z}$ is spanned by the so-called *fundamental class* $\sigma + \tau$ which contains each triangle in the fundamental square once. For the remaining homology group we get

$$H_1(\mathbb{X}) \simeq \mathbb{Z}^3/\operatorname{im}(A) = \mathbb{Z}^3/\mathbb{Z}\cdot(1,-1,1)^t \simeq \mathbb{Z}^2$$

which is in line with our expectation, since $\pi_1(T,*) \simeq \mathbb{Z}^2$ by example 6.12.

b) For the Klein bottle $\mathbb{X} = K$ again $\delta_2 = 0$ because there is only one 0-simplex, hence $H_0(\mathbb{X}) \simeq \mathbb{Z}$ as before. However, now $\delta_2(\sigma) = a - b - c$ and $\delta_2(\tau) = a - b + c$ and hence the linear map δ_2 is given by the matrix

$$A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Now $H_0(\mathbb{X}) \simeq \ker(A) = \{0\}$. Intuitively, there is no fundamental class because the orientations of the two triangles in the plane do not fit together: The Klein bottle is not orientable. To compute the remaining homology group $H_1(\mathbb{X})$, we bring the matrix A in Smith normal form via elementary row and column operations that are invertible over the integers:

$$A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \\ -1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix} =: A'$$

It follows that

$$H_1(\mathbb{K}) \simeq \mathbb{Z}^3/\operatorname{im}(A) \simeq \mathbb{Z}^3/\operatorname{im}(A') \simeq \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}.$$

This group is indeed the abelianization of $\pi_1(K,*) \simeq \mathbb{Z} \ltimes \mathbb{Z}$ from example 6.12.

c) For $\mathbb{X} = \mathbb{RP}^2$ we have two 0-simplices, let's call them p, q. Up to interchanging these two simplices we have $\delta_1(a) = \delta_1(b) = p - q$ and $\delta_1(c) = 0$, so δ_1 is given by the matrix

$$B = \begin{pmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix}$$

from which one immediately sees

$$H_0(\mathbb{RP}^2) \simeq \mathbb{Z}^2/\operatorname{im}(B) \simeq \mathbb{Z}.$$

Moreover we have $\delta_2(\sigma) = a - b - c$ and $\delta_2(\tau) = a - b + c$, hence δ_2 is given by the matrix

$$A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Again $H_2(\mathbb{RP}^2) \simeq \ker(A) = \{0\}$. For the remaining homology group we again compute the Smith normal form of the matrix:

$$A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \\ -1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix} =: A'$$

By applying lemma 1.10 below with $m = 3$, $r = 2$, $s = 1$ and $a_1 = 1$, $a_2 = 2$, we obtain that

$$H_1(\mathbb{RP}^2) \simeq \ker(B)/\mathrm{im}(A) \simeq \mathbb{Z}/2\mathbb{Z}.$$

Again this is in line with the fact that $\pi_1(\mathbb{RP}^2, *) \simeq \mathbb{Z}/2\mathbb{Z}$ by example 6.12.

Let us summarize the results of the above computations in the following small table of homology groups:

$\mathbb{X}$	$H_0(\mathbb{X})$	$H_1(\mathbb{X})$	$H_2(\mathbb{X})$
T	$\mathbb{Z}$	$\mathbb{Z}^2$	$\mathbb{Z}$
K	$\mathbb{Z}$	$\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$	0
$\mathbb{RP}^2$	$\mathbb{Z}$	$\mathbb{Z}/2\mathbb{Z}$	0

For completeness we also add a proof of the following lemma that was used above and allows to compute the cohomology groups of any complex of finitely generated abelian groups using only elementary linear algebra:

Lemma 1.10. *For any complex of abelian groups* $[\cdots \to \mathbb{Z}^l \xrightarrow{A} \mathbb{Z}^m \xrightarrow{B} \mathbb{Z}^n \to \cdots]$ *we have*

$$\ker(B)/\mathrm{im}(A) \simeq \mathbb{Z}^{m-r-s} \oplus \bigoplus_{i=1}^{r} \mathbb{Z}/a_i\mathbb{Z},$$

where $r = \mathrm{rk}(A)$, $s = \mathrm{rk}(B)$ *and* $a_1, \dots, a_r \in \mathbb{Z} \smallsetminus \{0\}$ *are the elementary divisors of* A.

Proof. By transforming the matrix A in Smith normal form via elementary row and column operations, we see that

$$\mathbb{Z}^m/\mathrm{im}(A) \simeq \mathbb{Z}^{m-r} \oplus T \quad \text{for the torsion part} \quad T = \bigoplus_{i=1}^{r} \mathbb{Z}/a_i\mathbb{Z},$$

where $a_1, \dots, a_r \neq 0$ are the elementary divisors of A. Since $B \cdot A = 0$, we know B factors over a map

$$\overline{B}\colon \quad \mathbb{Z}^m/\mathrm{im}(A) \simeq \mathbb{Z}^{m-r} \oplus T \longrightarrow \mathbb{Z}^n$$

and $T \subset \ker(\overline{B})$ since the target $\mathbb{Z}^n$ is torsion-free. Hence it only remains to determine the subgroup

$$\ker(\overline{B}|_{\mathbb{Z}^{n-r}}) \subset \mathbb{Z}^{n-r}$$

Note that this is a subgroup of a free abelian group and hence again free, so we only need to compute its rank. This can be done by the dimension formula from linear algebra:

$$\mathrm{rk}(\ker(\overline{B})|_{\mathbb{Z}^{n-r}}) = m-r-\mathrm{rk}(\overline{B}|_{\mathbb{Z}^{n-r}}) = m-r-s$$

where in the last step we have used that $\mathrm{im}(\overline{B}|_{\mathbb{Z}^{n-r}}) = \mathrm{im}(B)$ by construction. □

2 Singular homology

We have seen that the simplicial homology of a Δ-complex is very easy to compute, but as it stands it is not very useful as an invariant of topological spaces: A priori it depends on the choice of a triangulation! We will now discuss a more intrinsic way to define the homology of topological spaces that does not require the choice or even the existence of a triangulation.

The idea is simple – rather than choosing preferred simplices for a triangulation, we consider all possible simplices at once, allowing even for degenerate ones:

Definition 2.1. Let X be a topological space. By a *singular n-simplex* in X we mean a continuous map

$$\sigma\colon \quad \Delta^n \to X.$$

We denote by

$$C_n(X) = \bigoplus_{\sigma} \mathbb{Z}\cdot[\sigma]$$

the free abelian group generated by the set of all singular n-simplices. The elements of this group are called *singular n-chains*. For $[\sigma] \in C_n(X)$ we define its boundary by

$$\partial[\sigma] := \sum_{i=0}^{n}(-1)^i[\sigma\cdot\varphi_i] \in C_{n-1}(X)$$

where $\varphi_i\colon \Delta^{n-1} \to \Delta^n$ are the face maps for standard simplices. By linear extension we get group homomorphisms $\partial\colon C_n(X) \to C_{n-1}(X)$. We call

$$Z_n(X) := \ker(\partial) \subset C_n(X) \quad \text{and} \quad B_n(X) := \mathrm{im}(\partial) \subset C_n(X)$$

the group of *singular n-cycles* and the group of *singular n-boundaries*, respectively.

Lemma 2.2. *We have $B_n(X) \subseteq Z_n(X)$, in other words $\partial\circ\partial = 0$.*

Proof. It suffices to check that $\partial\circ\partial\colon C_n(X) \to C_{n-2}(X)$ maps every n-simplex to zero. Let $\sigma\colon \Delta^n \to X$ be a singular simplex, then

$$\begin{aligned}\partial(\partial(\sigma)) &= \partial\sum_{i=0}^{n}(-1)^i\cdot[\sigma\circ\varphi_i]\\ &= \sum_{j=0}^{n-1}\sum_{i=0}^{n}(-1)^{i+j}\cdot[\sigma\circ\varphi_{i,j}] \quad \text{for the composite} \quad \varphi_{i,j} = \varphi_i\circ\varphi_j.\end{aligned}$$

Now recall that j-th face $\varphi_i\colon \Delta^{n-1} \hookrightarrow \Delta^n$ is obtained by increasing all indices $\geq j$ of variables by one and setting the new j-th variable to be zero. So for $j \geq i$ it makes no difference whether we

a) first insert a zero in position i and then another zero in position $j+1$, or
b) first insert a zero in position j and then insert another zero in position i.

Indeed, in case b) the first inserted zero is moved up to position $j+1$ by the second insertion. In other words

$$\varphi_{i,j} = \varphi_{j+1,i} \quad \text{for} \quad j \geq i.$$

Therefore our previous alternating sum can be written as

$$\begin{aligned}
\partial(\partial(\sigma)) &= \sum_{j<i} (-1)^{i+j} [\sigma \circ \varphi_{i,j}] + \sum_{j \geq i} (-1)^{i+j} [\sigma \circ \varphi_{i,j}] \\
&= \sum_{j<i} (-1)^{i+j} [\sigma \circ \varphi_{i,j}] + \sum_{j \geq i} (-1)^{i+j} [\sigma \circ \varphi_{j+1,i}] \\
&= \sum_{j<i} (-1)^{i+j} [\sigma \circ \varphi_{i,j}] + \sum_{j'<i'} (-1)^{i'+j'+1} [\sigma \circ \varphi_{i',j'}] \\
&= 0
\end{aligned}$$

where for third equality we have used the relabelling $i' = j+1$ and $j' = i$. □

Definition 2.3. The *singular chain complex* of X is the complex

$$C_\bullet(X) := \left[\cdots \xrightarrow{\partial} C_n(X) \xrightarrow{\partial} C_{n-1}(X) \xrightarrow{\partial} \cdots \xrightarrow{\partial} C_0(X) \right]$$

The homology groups of this complex are called the *singular homology groups* of X and are denoted by

$$H_n(X) := H_n(C_*(X)) = Z_n(X)/B_n(X).$$

In contrast to simplicial homology, the above definition is not very suitable for direct computations since the groups $C_n(X)$ are huge. However, it does not require the choice of a triangulation and has the advantage of being functorial:

Lemma 2.4. *Any continuous map $f\colon X \to Y$ of topological spaces induces group homomorphisms*

$$f_*\colon \quad H_n(X) \to H_n(Y) \quad \textit{for all} \quad n \in \mathbb{N}.$$

Proof. For any singular simplex $\sigma\colon \Delta^n \to X$ the composite $f \circ \sigma\colon \Delta^n \to Y$ is a singular simplex, so we get a group homomorphism

$$f_*\colon \quad C_n(X) \to C_n(Y), \quad \sum_\sigma n_\sigma \cdot [\sigma] \mapsto \sum_\sigma n_\sigma \cdot [f \circ \sigma]$$

for each $n \in \mathbb{N}_0$. These homomorphisms commute with the differentials ∂ of the respective chain complexes:

$$f_*(\partial[\sigma]) = f_*\left(\sum_{i=0}^{n}(-1)^i[\sigma \circ \varphi_i]\right) = \sum_{i=0}^{n}(-1)^i\,(f \circ \sigma \circ \varphi_i) = \partial(f \circ \sigma) = \partial(f_*[\sigma]).$$

Thus we obtain a morphism of complexes $f_*\colon C_*(X) \to C_*(Y)$ and hence an induced morphism between homology, see lemma 2.2 in the appendix to this chapter. □

In what follows we will gather some simple properties of singular homology groups that allow to compute them in many situations. There is one case where one can actually compute singular homology directly from the definition:

Proposition 2.5. *For a singleton space $X = \{p\}$ we have*

$$H_n(X) \simeq \begin{cases} \mathbb{Z} & \text{for } n = 0, \\ 0 & \text{for } n > 0. \end{cases}$$

Proof. Since X consists of a single point, there exists a unique map $\sigma_n\colon \Delta^n \to X$ for each n. Thus we have a canonical identification $C_n(X) = \mathbb{Z}$ for each $n \in \mathbb{Z}$. Moreover we trivially have $\sigma_n \circ \varphi_i = \sigma_{n-1}$ for $i = 0, 1, \ldots, n$, hence the boundary maps are given by

$$\partial\colon \quad C_n(X) \to C_{n-1}(X), \quad [\sigma_n] \mapsto \sum_{i=0}^{n}(-1)^i \cdot [\sigma_n \circ \varphi_i] = \begin{cases} [\sigma_{n-1}] & \text{for } n \text{ even}, \\ 0 & \text{for } n \text{ odd}. \end{cases}$$

Therefore the singular chain complex takes the form

$$\begin{array}{ccccccccccc} \cdots & \longrightarrow & C_3(X) & \xrightarrow{\partial=0} & C_2(X) & \xrightarrow{\partial=id} & C_1(X) & \xrightarrow{\partial=0} & C_0(X) & \longrightarrow & 0 \\ & & \| & & \| & & \| & & \| & & \\ & & \mathbb{Z} & & \mathbb{Z} & & \mathbb{Z} & & \mathbb{Z} & & \end{array}$$

and the claim follows by taking homology of this complex. □

Similar to the situation for fundamental groups, it suffices to compute homology for path-connected topological spaces:

Proposition 2.6. *Let $X = \bigsqcup_\alpha X_\alpha$ be the decomposition into path-components. Then we have a canonical isomorphism*

$$H_n(X) \simeq \bigoplus_\alpha H_n(X_\alpha).$$

Proof. Continuous images of path-connected spaces are path-connected, so each singular simplex $\sigma\colon \Delta^n \to X$ has its image contained in one of the path-components $X_\alpha \subset X$. Therefore we have a canonical decomposition $C_n(X) = \oplus_\alpha C_n(X_\alpha)$ for all

$n \in \mathbb{N}$, and these decompositions are compatible with the differentials in the sense that $\delta(C_n(X_\alpha)) \subset C_{n-1}(X_\alpha)$. Thus we obtain direct sum decompositions

$$B_n(X) = \bigoplus_\alpha B_n(X_\alpha) \subset Z_n(X) = \bigoplus_\alpha Z_n(X_\alpha)$$

and the claim follows. □

In particular, the degree zero homology group $H_0(X)$ of a topological space X counts the number of its path-components:

Proposition 2.7. *If X is path-connected and non-empty, then $H_0(X) \simeq \mathbb{Z}$.*

Proof. By definition $H_0(X) = C_0(X)/\partial(C_1(X))$. Since we assumed that $X \neq \varnothing$, the homomorphism

$$\deg\colon \quad C_0(X) \to \mathbb{Z}, \quad \sum_\sigma n_\sigma \cdot [\sigma] \mapsto \sum_\sigma n_\sigma$$

is surjective. So it only remains to show $\ker(\deg) = \partial(C_1(X))$. The inclusion $\supset$ is clear since the boundary of every 1-simplex is a difference of two 0-simplices. For the reverse inclusion, consider a 0-chain

$$\tau = \sum_i n_i \cdot [p_i] \in \ker(\deg) \subset C_0(X)$$

with $n_i \in \mathbb{Z}$ and points (i.e. singular 0-simplices) $p_i \in X$. We have

$$\sum_i n_i = \deg(\tau) = 0.$$

Now fix $p \in X$. Since X is path-connected, we can find for each i a path γ_i from p to p_i. We may view this path as a 1-simplex $[\gamma_i] \in C_1(X)$ with $\partial[\gamma_i] = [p_i] - [p]$, and the 1-chain

$$\gamma = \sum_i n_i \cdot [\gamma_i]$$

satisfies

$$\partial\gamma = \sum_i n_i \cdot \partial[\gamma_i] = \sum_i n_i \cdot ([p_i] - [p]) = \sum_i n_i \cdot [p_i] - \deg(\tau) \cdot [p] = \tau.$$

Thus $\tau \in \partial(C_1(X))$ as required. □

It is sometimes convenient to work with a version of singular homology where a singleton space has no homology in any degrees (including degree zero). This can be done as follows:

Remark 2.8. For $n \in \mathbb{N}$, we define the n-th *reduced homology* of a topological space X to be the kernel

$$\tilde{H}_n(X) := \ker(p_*\colon H_n(X) \to H_n(*))$$

where $p\colon X \to *$ is the unique map to a point. As $H_n(*) = 0$ for $n > 0$ but $H_n(*) = \mathbb{Z}$, we have

$$\tilde{H}_n(X) \;=\; \begin{cases} H_n(X) & \text{for } n > 0, \\ \mathbb{Z}^{r-1} & \text{for } n = 0, \end{cases}$$

where r is the number of path-components of X. In particular $\tilde{H}_n(*) = 0$ for all n. In general, we have a natural splitting

$$H_0(X) \;\simeq\; \tilde{H}_0(X) \oplus \mathbb{Z}$$

using the map $i_*\colon \mathbb{Z} \to H_0(X)$ induced by the inclusion $i\colon * \hookrightarrow X$.

The first homology of a path-connected space also has a very clear geometric interpretation, in line with the motivation that we gave for the definition of simplicial homology:

Theorem 2.9 (Hurewicz). *Let X be path-connected. Then for any $p \in X$, viewing loops as 1-simplices gives a natural isomorphism from the abelianized fundamental group to the first homology:*

$$\pi_1^{ab}(X,p) \;\xrightarrow{\sim}\; H_1(X).$$

Proof. A path $\gamma\colon [0,1] \to X$ is the same thing as a singular 1-simplex. If γ is a loop based at p, then

$$\partial[\gamma] = [\gamma(1)] - [\gamma(0)] = [p] - [p] = 0.$$

We claim that if $\gamma_0 \sim \gamma_1$ are two such loops defining the same element of $\pi_1(X,p)$, then as 1-simplices they differ only by the boundary of a 2-simplex. Indeed, pick a homotopy

$$H\colon \quad [0,1] \to [0,1] \;\to\; X \quad \text{with} \quad \begin{cases} H(s,0) = \gamma_0(s) & \text{for all } s, \\ H(s,1) = \gamma_1(s) & \text{for all } s, \\ H(0,t) = p & \text{for all } t. \end{cases}$$

By cutting the domain of the homotopy in two triangles, we obtain two singular 2-simplices $\sigma_0, \sigma_1\colon \Delta^2 \to X$ with the ordering of vertices as shown in the following picture:

$$\begin{tikzcd} 2 \arrow[r, "\gamma_1"] & 3 \\ 0 \arrow[u, "\varepsilon_p"] \arrow[r, "\gamma_0"'] \arrow[ur, "\delta"] & 1 \arrow[u, "\varepsilon_p"'] \end{tikzcd}$$

Let $\delta\colon [0,1] \to X, t \mapsto H(t,t)$ be the diagonal path and $\varepsilon_p\colon [0,1] \to X$ the constant path on the vertical sides of the above square, then we have

$$\begin{aligned}\partial[\sigma_0] &= [\varepsilon_p]-[\delta]+[\gamma_0],\\ \partial[\sigma_1] &= [\gamma_1]-[\delta]+[\varepsilon_p].\end{aligned}$$

Hence it follows that

$$[\gamma_1]-[\gamma_0] \;=\; \partial\sigma \quad \text{for the 2-simplex} \quad \sigma \;=\; [\sigma_1]-[\sigma_0] \;\in\; C_2(X),$$

which means that $[\gamma_1]$ and $[\gamma_0]$ have the same image in $H_1(X)$. So the image of a path in homology depends only on the homotopy class of the path, i.e. we obtain a well-defined map

$$\tilde{h}\colon \pi_1(X,p)\to H_1(X)$$

sending path to its homology class. This a group homomorphism: For $\gamma,\mu\in\pi_1(X,p)$ we have

$$\tilde{h}(\gamma\cdot\mu)=\tilde{h}(\gamma)+\tilde{h}(\mu)$$

since the 1-chain $[\gamma\cdot\mu]-[\gamma]-[\mu]\in C_1(X)$ is the boundary of the singular 2-simplex which is given by first projecting orthogonally from Δ^2 to one of its sides and then putting $\gamma\cdot\mu$ on that side:

[PICTURE]

Since the homomorphism $\tilde{h}$ has as its target an abelian group, it factors uniquely over the abelianization as shown below:

$$\begin{tikzcd}
\pi_1(X,p) \arrow[rr, "\tilde{h}"] \arrow[dr] & & H_1(X) \\
& \pi_1^{ab}(X,p) \arrow[ur, dotted, "\exists!h"'] &
\end{tikzcd}$$

To show that h is an isomorphism, we pick for every point $x\in X$ a path $\mu_x\colon [0,1]\to X$ from p to x. This is possible since X is path-connected. We may moreover assume that for the base point $x=p$ the path $w_p=\varepsilon_p$ is constant. We then argue as follows:

- For surjectivity, take an arbitrary 1-cycle $c=\sum_i n_i[\sigma_i]\in Z_1(X)\subset C_1(X)$. Denote by $y_i=\sigma_i(0)$ and $z_i=\sigma_i(1)$ the starting and end points of the simplicices that appear in the 1-cycle. Then we have the relation

$$\sum_i n_i\cdot([z_i]-[y_i]) \;=\; \partial c \;=\; 0$$

in the free abelian group $C_0(X)$. It follows that in the free abelian group $C_1(X)$ we have the relation

$$\sum_i n_i \cdot ([w_{z_i}] - [w_{y_i}]) = 0.$$

But then

$$\begin{aligned}
c &= c - 0 \\
&= c - \sum_i n_i \cdot ([w_{z_i}] - w_{y_i}]) \\
&= \sum_i m_i \cdot ([w_{y_i}] - [w_{z_i}] + [\sigma_i]) \\
&= \sum_i m_i \cdot [\gamma_i] \quad \text{for the loop} \quad \gamma_i = w_{y_i} \cdot \sigma_i \cdot w_{z_i}^- \in \pi_1(X,p) \\
&= h(\gamma) \quad \text{for the loop} \quad \gamma = \prod_i \gamma_i^{n_i} \in \pi_1^{ab}(X,p).
\end{aligned}$$

- For injectivity we will construct a group homomorphism $g\colon H_1(X) \to \pi_1^{ab}(X,p)$ with $g \circ h = id$, which will then be the inverse of h. This can be done by the same type of construction: Consider the homomorphism

$$\tilde{g}\colon C_1(X) \to \pi_1^{ab}(X,p) \quad \text{given on generators by} \quad [\sigma] \mapsto [w_{\sigma(0)} \cdot \sigma \cdot w_{\sigma(1)}^-].$$

We claim that this homomorphism vanishes on the subgroup $B_1(X) = \partial(C_2(X))$ of boundaries: Indeed, take any singular 2-simplex $\tau\colon \Delta^2 \to X$, and let $\tau_i = \tau \circ \varphi_i$ for $i = 0,1,2$. Then

$$\begin{aligned}
\tilde{g}(\partial\tau) &= \tilde{g}(\tau_0) \cdot \tilde{g}(\tau_1)^{-1} \cdot \tilde{g}(\tau_2) \\
&= [w_{\tau_0(0)} \cdot \tau_0 \cdot w_{\tau_0(1)}^-] \cdot [w_{\tau_1(1)} \cdot \tau_1^- \cdot w_{\tau_1(0)}^-] \cdot [w_{\tau_2(0)} \cdot \tau_2 \cdot w_{\tau_2(1)}^-] \\
&= [w_{\tau_0(0)} \cdot \tau_0 \cdot \tau_1^- \cdot \tau_1 \cdot w_{\tau_2(1)}] \\
&= [w_{\tau_0(0)} \cdot w_{\tau_2(1)}^-] \\
&= 1,
\end{aligned}$$

where the third equality uses that $\tau_0(1) = \tau_1(1)$ and $\tau_1(0) = \tau_2(0)$, and the fourth equality applies a homotopy that contracts the boundary of the 2-simplex to the single vertex $\tau_0(0) = \tau_2(1)$. Thus $\tilde{g}\colon C_1(X) \to \pi_1^{ab}(X,p)$ vanishes on boundaries and hence induces a homomorphism

$$g\colon \quad H_1(X) \to \pi_1^{ab}(X,p),$$

and by construction we have $\bar{g} \circ h = id$. This finishes the proof. □

3 Homotopy invariance

Recall that two continuous maps $f, g\colon X \to Y$ between topological spaces are called *homotopic* if there exists a continuous map $H\colon X \times [0,1] \to Y$ with $H(x,0) = f(x)$ and $H(x,1) = g(x)$ for all $x \in X$. We want to show:

Theorem 3.1. *If two maps $f, g\colon X \to Y$ are homotopic, then they induce the same maps on homology, i.e.*

$$f_* = g_*\colon \quad H_n(X) \to H_n(Y) \quad \textit{for all } n \in \mathbb{N}_0.$$

Proof. It suffices to show $f_*[\sigma] = g_*[\sigma]$ for each singular simplex $\sigma\colon \Delta^n \to X$. For this let $H\colon X \times [0,1] \to Y$ be a homotopy from f to g, and consider the composite map

$$H \circ (\sigma \times id_{[0,1]})\colon \quad \Delta^n \times [0,1] \to Y$$

on the prism $\Delta^n \times [0,1]$. We subdivide the prism into $n+1$ different $(n+1)$-simplices as shown below in the case $n = 2$:

[PICTURE]

Explicitly, let $p_0, \ldots, p_n$ be the vertices of Δ^n in their natural order. The bottom face of the prism has vertices $a_i = (p_i, 0)$, and the top face has vertices $b_i = (p_i, 1)$. In this notation, the $(n+1)$-simplices in our subdivision of the prism are given by the affine linear maps

$$f_i\colon \quad \Delta^{n+1} \hookrightarrow \Delta^n \times [0,1] \quad (i = 0, 1, \ldots, n)$$

that send the vertices of Δ^{n+1} to the points $a_0, \ldots, a_i, b_i, \ldots, b_{n+1}$. We then consider the singular chain

$$h(\sigma) := \sum_{i=0}^{n} (-1)^i \cdot H \circ (\sigma \times id_{[0,1]}) \circ f_i \in C_{n+1}(Y).$$

By linear extension this defines group homomorphisms $h\colon C_n(X) \to C_{n+1}(Y)$ as indicated in the following diagram:

$$\begin{tikzcd}
\cdots & C_{n-1}(X) \arrow[r] \arrow[dl, "h"'] & C_n(X) \arrow[r] \arrow[d, "f_*-g_*"] \arrow[dl, "h"'] & C_{n+1}(X) \arrow[r] \arrow[d, "f_*-g_*"] \arrow[dl, "h"'] & \cdots \\
{} & C_{n-1}(Y) \arrow[r] & C_n(Y) \arrow[r] & C_{n+1}(Y) \arrow[r] & \cdots
\end{tikzcd}$$

A computation analogous to the one for $\partial \circ \partial = 0$ shows that

$$\partial \circ h + h \circ \partial \;=\; f_* - g_*\colon \quad C_n(X) \to C_n(Y).$$

In other words, the maps $h\colon C_n(X) \to C_{n+1}(Y)$ define a *chain homotopy* between the two maps $f_*, g_*\colon C_\bullet(X) \to C_\bullet(Y)$. The claim now follows from the general fact that if there is a chain homotopy between two maps of complexes, then the two maps induce the same maps on homology (see lemma 2.3 in the appendix). □

Corollary 3.2. *If two topological spaces X and Y are homotopy equivalent, then we have an isomorphism*

$$H_n(X) \xrightarrow{\sim} H_n(Y) \quad \textit{for each } n \in \mathbb{N}_0.$$

Proof. By definition X and Y are homotopy equivalent if and only if there exist continuous maps $a\colon X \to Y$ and $b\colon Y \to X$ such that

- $a \circ b$ is homotopic to id_Y,
- $b \circ a$ is homotopic to id_X.

By the above theorem then the composite maps

$$H_n(Y) \xrightarrow{b_*} H_n(X) \xrightarrow{a_*} H_n(Y) \quad \text{and} \quad H_n(X) \xrightarrow{a_*} H_n(Y) \xrightarrow{b_*} H_n(X)$$

are the identity, and the claim follows. □

Example 3.3. If X is contractible, then

$$H_n(X) \simeq \begin{cases} \mathbb{Z} & \text{for } n = 0, \\ 0 & \text{for } n \neq 0. \end{cases}$$

In the next section we will see how to compute the homology of more interesting spaces, which will show in particular that spheres S^n are not contractible.

4 Relative homology and the excision theorem

Often one wants to relate the homology of a topological space X to the homology of a given subspace $A \subset X$. Note that while the inclusion of the subspace always induces a natural group homomorphism $H_n(A) \to H_n(X)$, this homomorphism is usually not injective:

Example 4.1. Every topological space A can be embedded as a subspace in the cone

$$X = CA = (A \times [0,1])/(A \times \{0\}).$$

Here $H_\bullet(X) = \{0\}$ since the cone is contractible. But $H_\bullet(A)$ can be non-trivial!

However, it *is* true that the group of singular chains of the subspace $A \subset X$ is a subgroup

$$C_n(A) \subset C_n(X),$$

namely the subgroup generated by simplices that factor as $\sigma\colon \Delta^n \to A \subset X$. This leads to the following definition:

Definition 4.2. The group of *relative n-chains* with respect to a subspace $A \subset X$ is the quotient group

$$C_n(X,A) := C_n(X)/C_n(A).$$

Since the differentials $\partial\colon C_n(X) \to C_{n-1}(X)$ satisfy $\partial(C_n(A)) \subset C_{n-1}(A)$, they give rise to differentials between the groups of relative n-chains which, so we obtain a complex

$$C_\bullet(X,A) = \Big[\dots \xrightarrow{\partial} C_n(X,A) \xrightarrow{\partial} C_{n-1}(X,A) \xrightarrow{\partial} \cdots \xrightarrow{\partial} C_0(X,A)\Big]$$

of relative chains. Its homology groups

$$H_n(X,A) := H_n(C_*(X,A))$$

are called the *relative homology groups* of the pair (X,A).

Example 4.3. For $A = \varnothing$ we recover the singular homology $H_n(X,\varnothing) = H_n(X)$.

Example 4.4. For a point $A = *$ we may identify the relative homology with the reduced homology

$$H_n(X,*) \simeq \tilde{H}_n(X)$$

as defined in remark 2.8. One way to see this is to use the long exact sequence:

Proposition 4.5. *For any pair $A \subset X$ of topological spaces, we have a long exact sequence of homology groups*

$$\cdots \to H_n(A) \to H_n(X) \to H_n(X,A) \to H_{n-1}(A) \to \cdots$$

Proof. By design we have an exact sequence $0 \to C_\bullet(A) \to C_\bullet(X) \to C_\bullet(X,A) \to 0$ of complexes, hence the result follows by taking the associated long exact sequence from theorem 3.2. ◻

We have defined relative homology via quotients $C_n(X,A)$ of $C_n(X)$, but there is also a slightly more convenient description via *subgroups*. To formulate it, consider inside $C_n(X)$ the subgroups of *relative cycles* and *relative boundaries* which are defined by

$$\begin{aligned} Z_n(X,A) &:= \{\sigma \in C_n(X) \mid \partial\sigma \in C_{n-1}(A)\}, \\ B_n(X,A) &:= \{\sigma \in C_n(X) \mid \exists \sigma' \in C_n(A)\colon \sigma - \sigma' \in B_n(X)\} = B_n(X) + C_n(A). \end{aligned}$$

Lemma 4.6. *We have $H_n(X,A) \simeq Z_n(X,A)/B_n(X,A)$.*

Proof. For $\sigma \in C_n(X)$, let us denote by $\overline{\sigma} \in C_n(X,A) = C_n(X)/C_n(A)$ its image. Then by definition

$$\overline{\partial}\overline{\sigma} := \overline{\partial\sigma} \quad \text{in} \quad C_{n-1}(X,A) = C_{n-1}(X)/C_{n-1}(A).$$

where $\overline{\partial}$ denotes the differential of the relative singular chain complex $C_\bullet(X,A)$. Hence we have

$$\begin{aligned} \ker(\overline{\partial}) &= \{\overline{\sigma} \in C_n(X,A) \mid \partial\sigma \in C_n(A)\} = Z_n(X,A)/C_n(A), \\ \operatorname{im}(\overline{\partial}) &= \{\overline{\sigma} \in C_n(X,A) \mid \sigma \in B_n(X)\} = B_n(X,A)/C_n(A). \end{aligned}$$

and therefore

$$\begin{aligned} H_n(X,A) &= \ker(\overline{\partial})/\operatorname{im}(\overline{\partial}) \\ &= (Z_n(X,A)/C_n(A))/(B_n(X,A)/C_n(A)) \\ &\simeq Z_n(X,A)/B_n(X,A) \end{aligned}$$

as claimed. □

To make use of the long exact cohomology sequence we need an efficient way to compute relative cohomology in geometrically interesting examples. The most important method for this is the following result known as the excision theorem:

Theorem 4.7. *Let (X,A) be a pair of topological spaces and $U \subset X$ a subpace whose closure is contained in the interior of A, i.e. $\overline{U} \subset \operatorname{int}(A) = A \setminus \partial A$. Then we have a natural isomorphism*

$$H_n(X,A) \simeq H_n(X \setminus U, A \setminus U).$$

We will prove this in the next section via an argument using subdivisions of simplices. But before doing so, let us take a look at a few examples. We begin with an equivalent reformulation of the theorem:

Corollary 4.8. *For any two subspaces $A,B \subset X$ with $X = \operatorname{int}(A) \cup \operatorname{int}(B)$, we have a natural isomorphism*

$$H_n(X,A) \simeq H_n(B, A \cap B)$$

Proof. Take $U = X \setminus B$ in theorem 4.7, and use that $\overline{U} = X \setminus \operatorname{int}(B) \subset \operatorname{int}(A)$. □

An important consequence of the excision theorem is the following result which is known as the *Mayer-Vietoris sequence*. It allows to compute the homology of a space from an open cover in a way similar to what we did for fundamental groups in the Seifert – van Kampen theorem:

Theorem 4.9 (Mayer-Vietoris). *For any $X_1, X_2 \subset X$ with $X = \operatorname{int}(X_1) \cup \operatorname{int}(X_2)$, we have a long exact sequence*

$$\cdots \to H_n(X_1 \cap X_2) \to H_n(X_1) \oplus H_n(X_2) \to H_n(X) \to H_{n-1}(X_1 \cap X_2) \to \cdots$$

Proof. The long exact homology sequences for the pairs $X_1 \cap X_2 \subset X_1$ and $X_2 \subset X$ fit in the following commutative diagram with exact rows:

$$\begin{tikzcd}
H_n(X_1 \cap X_2) \arrow[r] \arrow[d] & H_n(X_1) \arrow[r] \arrow[d] & H_n(X_1, X_1 \cap X_2) \arrow[r] \arrow[d, "h"] & H_{n-1}(X_1 \cap H_2) \arrow[r] \arrow[d] & \cdots \\
H_n(X_2) \arrow[r] & H_n(X) \arrow[r] & H_n(X, X_2) \arrow[r] & H_{n-1}(X) \arrow[r] & \cdots
\end{tikzcd}$$

By corollary 4.8 the maps h are isomorphisms for each n. A simple diagram chase therefore gives the claimed long exact sequence, see lemma 3.3 in the appendix. $\square$

Example 4.10. For $n > 1$ the sphere $S^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1^2 + \cdots + x_n^2 = 1\}$ has the homology groups

$$H_k(S^n) \simeq \begin{cases} \mathbb{Z} & \text{for } k = 0, n, \\ 0 & \text{for } k \neq 0, n. \end{cases}$$

Proof. We may assume $k > 0$. As in our computation of the fundamental group of spheres, we decompose the sphere $X = S^n$ as a union $X = X_1 \cup X_2$ of the two open subsets

$$\begin{aligned} X_1 &= \{(x_1, \dots, x_n) \in S^n \mid x_n > -\tfrac{1}{2}\}, \\ X_2 &= \{(x_1, \dots, x_n) \in S^n \mid x_n < +\tfrac{1}{2}\}. \end{aligned}$$

These are contractible, hence $H_k(X_1) = H_k(X_2) = 0$ for $k > 0$. The Mayer-Vietoris sequence gives

$$0 \cdots \to \underbrace{H_k(X_1) \oplus H_k(X_2)}_{=0} \to H_k(S^n) \to H_{k-1}(X_1 \cap X_2) \to \underbrace{H_{k-1}(X_1) \oplus H_{k_1}(X_2)}_{=0} \to \cdots$$

for $k > 1$. Since $X_1 \cap X_2$ is homotopy equivalent to the sphere S^{n-1} and since singular homology is homotopy invariant, the claim now follows by induction on k. The base of the induction is the case $k = 1$, in which the last part

$$\underbrace{H_1(X_1) \oplus H_1(X_2)}_{=0} \to H_1(S^n) \to \underbrace{H_0(X_1 \cap X_2)}_{=\mathbb{Z}} \to \underbrace{H_0(X_1) \oplus H_0(X_2)}_{=\mathbb{Z} \oplus \mathbb{Z}} \to \underbrace{H_0(X)}_{=\mathbb{Z}} \to 0.$$

of the Mayer-Vietoris sequence gives the desired result. $\square$

Recall that in corollary 4.3 we have proven the Brouwer fixed point theorem for the disk $D \subset \mathbb{R}^2$ using that $\pi_1(\partial D, *) \simeq \mathbb{Z}$. At the time we could not replace the disk by higher-dimensional balls

$$D^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1^2 + \cdots + x_n^2 \leq 1\} \subset \mathbb{R}^n$$

because $\pi_1(\partial D^n, *) = 0$ for $n > 2$. However, we can now run the same argument using homology instead of fundamental groups:

Corollary 4.11 (Brouwer fixed point theorem). *For $n \in \mathbb{N}$, let*

$$B^n = \{(x_1,\dots,x_n) \in \mathbb{R}^n \mid x_1^2 + \cdots + x_n^2 \leq 1\}$$

be the n-dimensional ball. Then every continuous map

$$f\colon \quad B^n \to B^n \quad \textit{has a fixed point}.$$

Proof. If $f\colon B^n \to B^n$ has no fixed point, then for any $p \in B^n$ the half-ray from $f(p)$ to p is well-defined. This half-ray has a unique point of intersection with the boundary $\partial B^n \simeq S^{n-1}$ and this point of intersection depends continuously on p. Hence we obtain a continuous map $r\colon B^n \to S^{n-1}$ such that the composite

$$S^{n-1} = \partial B^n \hookrightarrow B_n \to S^{n-1}$$

is the identity map. For $n > 2$ this gives a contradiction by applying H_{n-1}, indeed the composite map

$$f_*\colon \quad \mathbb{Z} = H_{n-1}(S^{n-1}) \longrightarrow \underbrace{H_{n-1}(B^n)}_{=0} \longrightarrow H_{n-1}(S^{n-1}) = \mathbb{Z}$$

cannot be the identity: The middle term vanishes since B^n is contractible. $\square$

Another nice application of example 4.10 is the intuitively obvious statement that for $m \neq n$ the spaces $\mathbb{R}^m$ and $\mathbb{R}^n$ are not homeomorphic: Indeed, if there was a homeomorphism between them, then by restriction to the complement of a point we would get a homeomorphism between $\mathbb{R}^m \setminus \{*\}$ and $\mathbb{R}^n \setminus \{*\}$, which is impossible since these two spaces have different homology: By homotopy equivalence and by example 4.10 we have

$$H_k(\mathbb{R}^n \setminus \{*\}) \simeq H_k(S^{n-1}) = \begin{cases} \mathbb{Z} & \text{for } k \in \{0,n\}, \\ 0 & \text{otherwise.} \end{cases}$$

Remark 4.12. We can also compute the relative homology of the pair $\mathbb{R}^n \setminus \{*\} \subset \mathbb{R}^n$ easily: The long exact homology sequence has the form

$$\cdots \to \underbrace{H_k(\mathbb{R}^n)}_{=0 \text{ for } k>0} \to H_k(\mathbb{R}^n, \mathbb{R}^n \setminus \{*\}) \to H_{k-1}(\mathbb{R}^n \setminus \{*\}) \to \underbrace{H_{k-1}(\mathbb{R}^n)}_{=0 \text{ for } k>1} \to \cdots$$

Inserting the homology groups $H_k(\mathbb{R}^n \setminus \{*\})$ from above, we therefore get

$$H_k(\mathbb{R}^n, \mathbb{R}^n \setminus \{*\}) \simeq \begin{cases} \mathbb{Z} & \text{for } k = n, \\ 0 & \text{otherwise.} \end{cases}$$

This leads to a simple proof of the following result:

Corollary 4.13 (Invariance of dimension). *If open subsets $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$ are homeomorphic, then the ambient spaces have the same dimension $m = n$.*

Proof. For $n = 0$ the result is clear, and for $n = 1$ it holds because removing a point makes V disconnected. So in what follows we may assume that $n \geq 2$. If there exists a homeomorphism

$$f: \quad U \xrightarrow{\sim} V$$

then by removing a point $u \in U$ and its image $y = f(u)$, we get a homeomorphism of pairs

$$(U, U \setminus \{u\}) \simeq (V, V \setminus \{v\}).$$

The idea is to obtain from this a contradiction by taking relative homology. For this we will replace the unknown open subset $U \subset \mathbb{R}^n$ by the entire space $X = \mathbb{R}^n$ via excision: Write

$$X = A \cup B \quad \text{with} \quad A = \mathbb{R}^n \setminus \{u\}. \quad \text{and} \quad B = U.$$

Since $A, B \subset X$ are open, we have $X = \mathrm{int}(A) \cup \mathrm{int}(B)$. Hence the version of the excision theorem in corollary 4.8 gives

$$H_k(U, U \setminus \{*\}) = H_k(B, A \cap B) \simeq H_k(X, A) = H_k(\mathbb{R}^n, \mathbb{R}^n \setminus \{*\}).$$

By the computation preceding this proof, these relative homology groups determine the value of n uniquely. Hence for $m \neq n$ the pairs $(U, U \setminus \{u\})$ and $(V, V \setminus \{v\})$ cannot be homeomorphic and our claim follows. $\square$

Remark 4.14. The above computation also shows why in the excision theorem for a pair $A \subset X$ we may only remove subspaces whose closure lies in $\mathrm{int}(A)$: Indeed, for

$$U = A = \mathbb{R}^n \setminus \{*\} \subset X = \mathbb{R}^n$$

we have

$$H_n(X, A) = H_n(\mathbb{R}^n, \mathbb{R}^n \setminus \{*\}) \simeq \mathbb{Z} \neq 0,$$

but

$$H_n(X \setminus U, A \setminus U) = H_n(\{*\}, \varnothing) = H_n(\{*\}) = 0 \quad \text{for} \quad n > 0.$$

The proof of excision in the next section will explain why the boundary ∂A matters.

5 Proof of the excision theorem: Barycentric subdivision

Let us now turn to the proof of the excision theorem: We want to show that for any topological space X with subspaces $U \subset A \subset X$ such that the closure $\overline{U}$ is contained in the interior $\mathrm{int}(A) = A \setminus \partial A$, the inclusion $X \setminus U \hookrightarrow X$ induces on relative homology isomorphisms

$$H_n(X \setminus U, A \setminus U) \xrightarrow{\sim} H_n(X, A)$$

for all $n \in \mathbb{N}_0$. The idea of the proof is simple:

- Simplices $\sigma\colon \Delta^n \to A$ do not contribute to $H_n(X,A)$.
- Simplices $\sigma\colon \Delta^n \to X \setminus U$ come from $H_n(X \setminus U, A \setminus U)$.
- So try to write any class in $H_n(X,A)$ as a sum of simplices of the above types.

To realize the last goal, we will subdivide arbitrary simplices into pieces which are small in the following sense:

Definition 5.1. A simplex $\sigma\colon \Delta^n \to X$ is called *small* (relative to the given $U \subset A \subset X$) if

$$\sigma(\Delta^n) \subset A \quad \text{or} \quad \sigma(\Delta^n) \subset X \setminus U.$$

Let $C'_n(X) \subset C_n(X)$ be the subgroup which is generated by the small simplices, and let

$$C'_n(X,A) \;=\; C'_n(X)/C_n(A) \;\subset\; C_n(X,A) \;=\; C_n(X)/C_n(A)$$

denote its quotient modulo the singular chains in A. Since the boundary of a small chain is again small, the small chains form a subcomplex which fits in a short exact sequence

$$0 \to C_\bullet(A) \to C'_\bullet(X) \to C'_\bullet(X,A) \to 0.$$

The complex on the right hand side computes the groups $H_n(X \setminus U, A \setminus U)$:

Lemma 5.2. *The inclusion $i\colon X \setminus U \hookrightarrow X$ induces an isomorphism of complexes*

$$i_*\colon \quad C_\bullet(X \setminus U, A \setminus U) \xrightarrow{\sim} C'_\bullet(X,A).$$

In particular, for every n we have an isomorphism $H_n(X \setminus U, A \setminus U) \simeq H_n(C'_\bullet(X,A))$.

Proof. The existence of the induced map i_* follows from the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & C_n(A \setminus U) \arrow[r] \arrow[d] & C_n(X \setminus U) \arrow[r] \arrow[d] & C_n(X \setminus U, A \setminus U) \arrow[r] \arrow[d, dashed, "\exists! i_*"] & 0 \\
0 \arrow[r] & C_n(A) \arrow[r] & C'_n(X) \arrow[r] & C'_n(X,A) \arrow[r] & 0
\end{tikzcd}$$

with exact rows. The injectivity of i_* also follows from the diagram and from the definition of the relative chain groups. The surjectivity comes from the fact that by definition any element of $C'_n(X,A)$ can be represented by a sum $\alpha + \beta$ with $\alpha \in C_n(A)$ and $\beta \in C_n(X \setminus U)$. □

So for the proof of the excision theorem it suffices to show that if $\overline{U} \subset \operatorname{int}(A)$, then the two chain complexes $C'_\bullet(X,A)$ and $C_\bullet(X,A)$ have the same homology. Note that in any case we have a natural morphism

$$C'_\bullet(X,A) \longrightarrow C_\bullet(X,A)$$

between these two chain complexes, induced by the inclusion of the group of small chains in the one of all chains via the commutative diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & C_\bullet(A) & \longrightarrow & C'_\bullet(X) & \longrightarrow & C'_\bullet(X,A) & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow \exists! & & \\
0 & \longrightarrow & C_\bullet(A) & \longrightarrow & C_\bullet(X) & \longrightarrow & C_\bullet(X,A) & \longrightarrow & 0
\end{array}$$

with exact rows. By theorem 3.2 each of the two rows gives rise to a long exact homology sequence, and by naturality of this sequence we then get a commutative diagram

$$\begin{array}{ccccccccc}
H_n(A) & \longrightarrow & H'_n(X) & \longrightarrow & H_n(X\smallsetminus U, A\smallsetminus U) & \longrightarrow & H_{n-1}(A) & \longrightarrow & H'_{n-1}(X) \\
\| & & \downarrow & & \downarrow & & \| & & \downarrow \\
H_n(A) & \longrightarrow & H_n(X) & \longrightarrow & H_n(X,A) & \longrightarrow & H_{n-1}(A) & \longrightarrow & H_{n-1}(X)
\end{array}$$

where for simplicity we have denoted by

$$H'_n(X) = H_n(C'_\bullet(X))$$

the homology of the complex of small chains. To prove the excision theorem 4.7 we have to show that the middle vertical arrow in the above diagram is an isomorphism; by the 5-lemma (appendix, lemma 1.1) this follows from the next result, which says that homology can be computed via small chains:

Proposition 5.3. *The inclusion of complexes $C'_\bullet(X) \hookrightarrow C_\bullet(X)$ induces for each n an isomorphism*

$$H'_n(X) \xrightarrow{\sim} H_n(X).$$

Proof. The idea is to subdivide simplices into small simplices. To achieve this we will define a natural morphism of complexes

$$bs\colon \quad C_\bullet(X) \to C_\bullet(X)$$

via *barycentric subdivision*, and a chain homotopy $h\colon C_\bullet(X) \to C_{\bullet+1}(X)$ between bs and the identity $id_{C_\bullet(X)}$, such that the following properties hold:

a) Any chain $\sigma \in C_n(X)$ becomes small after subdividing sufficiently many times, i.e.

$$bs^k(\sigma) := (\underbrace{bs\circ\cdots\circ bs}_{k\text{ times}})(\sigma) \in C'_n(X) \quad \text{for} \quad k \gg 0.$$

b) The morphism bs and the chain homotopy h send small chains to small chains, hence they restrict to

- a morphism $bs\colon C'_\bullet(X) \to C'_\bullet(X)$ of complexes, and
- a chain homotopy $h\colon C'_\bullet(X) \to C'_{\bullet+1}(X)$ between bs and the identity.

The construction of bs and h with properties $a)$ and $b)$ will take the rest of this section. Assuming these properties for the moment, let us first explain how to finish the proof: The existence of the chain homotopy $h\colon C_\bullet(X) \to C_{\bullet+1}(X)$ between bs and id implies by lemma 2.3 that

$$bs_* = id\colon \quad H_n(X) \to H_n(X)$$

is the identity map on homology groups. On the other hand, property $a)$ says that any given $\sigma \in C_n(X)$ becomes small after sufficiently many barycentric subdivisions, i.e.

$$bs^k(\sigma) := (\underbrace{bs \circ \cdots \circ bs}_{k \text{ times}})(\sigma) \in C'_n(X) \quad \text{for} \quad k \gg 0.$$

Since bs_* is the identity on homology, it follows that the homology class of σ can be represented by a small chain:

$$[\sigma] = [bs^k(\sigma)] \in \operatorname{im}(H'_n(X) \to H_n(X)).$$

Thus the map $H'_n(X) \to H_n(X)$ is surjective. It is also injective:

Indeed, by $b)$ our given subdivision restricts to a morphism $bs\colon C'_\bullet(X) \to C'_\bullet(X)$ on the complex of small chains, and h restricts to a chain homotopy between this morphism and the identity. Hence as above, for any small cycle $\sigma \in C'_n(X)$ we know that

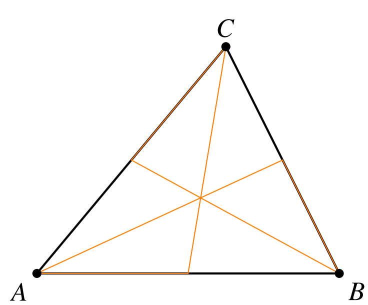
$$[\sigma] = [bs(\sigma)] \quad \text{also inside} \quad H'_n(X).$$

If σ maps to zero in $H_n(X)$, then there exists a chain $\tau \in C_{n+1}(X)$ with $\sigma = \partial\tau$. The chain τ need not be small, but assumption $a)$ says that for $k \gg 0$ the chain $bs^k(\tau)$ is small. Hence from

$$bs^k(\sigma) = bs^k(\partial\tau) = \partial(bs^k(\tau)) \in \partial(C'_{n-1}(X))$$

we see that $[\sigma] = [bs^k(\sigma)] = 0$ holds also inside $H'_n(X) = H_n(C'_\bullet(X))$. □

To fill in the missing piece in the above proof, let us now see how to define the morphism bs and the chain homotopy h with the above properties. The idea is to replace simplices by their barycentric subdivision as indicated in the following picture:



In general, we define the *barycenter* of the standard simplex $\Delta^n \subset \mathbb{R}^{n+1}$ to be the point
$$p_n = \tfrac{1}{n+1} \cdot (1,1,\dots,1) \in \Delta^n = \{(t_0,\dots,t_n) \in \mathbb{R}^{n+1}_{\geq 0} \mid \sum_i t_i = 1\}$$

As indicated in the picture, the barycentric subdivision of the standard simplex will be obtained inductively by first taking the barycentric subdivision of each face of the simplex, and then taking the cone over it with vertex at the barycenter:

Definition 5.4. Let $K \subset \mathbb{R}^m$ be a convex subset and $p \in K$. The *cone with vertex p* over a singular n-simplex
$$\sigma\colon \quad \Delta^n \to K$$
is the singular $(n+1)$-simplex

$$\mathrm{Cone}_p(\sigma)\colon \quad \Delta^{n+1} \to K, \quad t = (t_0,\dots,t_{n+1}) \mapsto \begin{cases} p & \text{for } t_0 = 1 \\ t_0 \cdot p + (1-t_0)\cdot \sigma(t^*) & \text{for } t_0 < 1 \end{cases}$$

where for $t_0 < 1$ we put
$$t^* = \tfrac{1}{1-t_0} \cdot (t_1,\dots,t_n) \in \Delta^n$$

Thus $\mathrm{Cone}_p(\sigma)\colon \Delta^{n-1} \to K$ is the unique map which

- takes the value p on the vertex $(1,0,\dots,0)$,
- restricts to σ on the face opposite to that vertex, and
- is affine-linear on each ray from $(1,0,\dots,0)$ to a point of the opposite face.

By linear extension we get a homomorphism

$$\mathrm{Cone}_p\colon \quad C_n(K) \to C_{n+1}(K), \quad \sigma \mapsto \mathrm{Cone}_p(\sigma).$$

An immediate computation shows:

Remark 5.5 (Cone formula). We have

$$\partial \mathrm{Cone}_p(\sigma) = \sigma - \mathrm{Cone}_p(\partial\sigma) \quad \text{for all} \quad \sigma \in C_n(K).$$

Returning to simplices in an arbitrary topological space X, we want to construct natural homomorphisms

$$bs_{X,n}\colon \quad C_n(X) \to C_n(X)$$

which subdivide each simplex into smaller ones. Here *natural* means that for any continuous map $f\colon Y \to X$ of topological spaces we want that our morphisms fit into a commutative diagram

$$\begin{tikzcd} C_n(Y) \arrow[r, "bs_{Y,n}"] \arrow[d, "f_*"'] & C_n(Y) \arrow[d, "f_*"] \\ C_n(X) \arrow[r, "bs_{X,n}"] & C_n(X) \end{tikzcd}$$

Any such natural subdivision is determined uniquely by its value on the identity simplex

$$\varepsilon_n = id\colon \quad \Delta^n \to Y = \Delta^n.$$

Indeed, for any other topological space X and for any n-simplex $\sigma\colon \Delta^n \to X$, the naturality forces

$$bs_{X,n}(\sigma) = bs_{X,n}(\sigma_*(\varepsilon_n)) = \sigma_*(bs_{\Delta^n,n}(\varepsilon_n)) = \sigma_*(\beta_n)$$

for the *universal subdivided simplex*

$$\beta_n := bs_{\Delta^n,n}(\varepsilon_n) \in C_n(\Delta^n).$$

This can be turned into an inductive definition of the barycentric subdivision:

Definition 5.6. Define $bs_{X,n}\colon C_n(X) \to C_n(X)$ by induction on n as follows:

- For $n = 0$ put $bs_{X,0} = id$.
- For $n \geq 0$ such that $bs_{X,n}$ has already been defined for all X, take $X = \Delta^{n+1}$ and define
$$\beta_{n+1} := \mathrm{Cone}_p(bs_{\Delta^{n+1},n}(\partial\varepsilon_{n+1})).$$
where p denotes the barycenter of Δ^{n+1} (compare this to our previous picture!).
- For simplices $\sigma\colon \Delta^{n+1} \to X$ in an arbitrary topological space X, put
$$bs_{X,n+1}(\sigma) := \sigma_*(\beta_{n+1})$$
and extend this additively to a homomorphism $bs_{X,n+1}\colon C_{n+1}(X) \to C_{n+1}(X)$.

Lemma 5.7. *The above homomorphisms $bs_{X,n}$ define a natural morphism of chain complexes*

$$bs_X\colon \quad C_\bullet(X) \to C_\bullet(X).$$

Proof. We show $\partial \circ bs_{X,n} = bs_{X,n-1} \circ \partial$ by induction on n: For $n = 0$ the statement is trivial. Suppose now that for some fixed $n \geq 0$, the statement has already been shown for all X. Then for any $(n+1)$-simplex σ one computes:

$$\begin{aligned}
\partial\,(bs_{X,n+1}(\alpha)) &= \partial\,(\sigma_*(\beta_{n+1})) \\
&= \sigma_*\,(\partial(\beta_{n+1})) \\
&= \sigma_*\left(\partial(\mathrm{Cone}_p(\tau))\right) \qquad \text{for } \tau := bs_{\Delta^{n+1},n}(\partial\varepsilon_{n+1}) \\
&= \sigma_*\left(\tau - \mathrm{Cone}_p(\partial\tau)\right)
\end{aligned}$$

where in the last step we have used the cone formular 5.5. By induction we know that $\partial \circ bs_{\Delta^{n+1},n} = bs_{\Delta^{n+1},n} \circ \partial$ and hence

$$\partial\tau = \partial(bs_{\Delta^{n+1},n}(\partial\varepsilon_{n+1})) = bs_{\Delta^{n+1},n}(\partial^2\varepsilon_{n+1}) = 0.$$

Therefore the previous chain of equalities gives

$$\begin{aligned}\partial\left(bs_{X,n+1}(\alpha)\right) &= \sigma_*(\tau)\\ &= \sigma_*\left(bs_{\Delta^{n+1},n}(\partial\varepsilon_{n+1})\right)\\ &= bs_{X,n}\left(\sigma_*(\partial\varepsilon_{n+1})\right)\\ &= bs_{X,n}(\partial\sigma).\end{aligned}$$

Thus $\partial\circ bs_{X,n+1} = bs_{X,n}\circ\partial$ as required. □

Having constructed the barycentric subdivision operator, we will now define the desired homotopy between these operators and the identity:

Proposition 5.8. *There exist natural homomorphisms $h_{X,n}\colon C_n(X)\to C_{n+1}(X)$ giving a chain homotopy*

$$h_X\colon\quad C_\bullet(X)\ \to\ C_{\bullet+1}(X)\quad \textit{between}\quad bs_X\quad \textit{and}\quad id.$$

Proof. We define $h_{X,n}$ by induction on n as follows:

- For $n=0$ and $X=\Delta^0=\{*\}$ there is no choice for the map $h\colon C_0(\Delta^0)\to C_1(\Delta^0)$.
- For $n=0$ and X arbitrary, we then define $h_{X,0}(\sigma)=\sigma_*(\varepsilon_0)$ for $\sigma\colon\Delta^0\to X$.
- Now let $n\geq 0$ be arbitrary. Suppose we have already defined natural maps $h_{X,n}$ for all X in such a way that the relation

  $$\partial\circ h_{X,n}+h_{X,n-1}\circ\partial \;=\; bs_{X,n}-id$$

  holds for all X. As before, to define a natural homomorphism $h_{X,n+1}$ for all X, we only need to consider the universal case $X=\Delta^{n+1}$. In other words, we only need to define the element

  $$\gamma \;=\; h_{\Delta^{n+1},n+1}(\varepsilon_{n+1})\;\in\; C_{n+2}(\Delta^{n+1}).$$

  The only requirement we have for this element γ is that it should satisfy the relation

  $$\partial\gamma+(h_{\Delta^{n+1},n}\circ\partial)(\varepsilon_{n+1}) \;=\; (bs_{\Delta^{n+1},n+1}-id)(\varepsilon_{n+1})$$

  which is required for having a chain homotopy. Equivalently, we want $\partial\gamma=\tau$ for the chain

  $$\tau \;=\; (-h_{\Delta^{n+1},n}\circ\partial+bs_{\Delta^{n+1},n+1}-id)(\varepsilon_{n+1})\;\in\; C_{n+1}(\Delta^{n+1}).$$

  To show that there exists a chain γ with $\partial\gamma=\tau$, we use an argument which is known as the method of *acyclic models*: The simplex Δ^{n+1} is contractible and hence acyclic, i.e. its reduced homology vanishes. In particular $H_{n+1}(\Delta^{n+1})=0$, so we have the equivalence

  $$(\exists\gamma\colon\partial\gamma=\tau)\quad\Longleftrightarrow\quad\partial\tau\;=\;0.$$

The property on the right hand side can be checked by a straightforward computation since it no longer involves the unknown chain γ. Indeed the vanishing $\partial\tau = 0$ follows from

$$\begin{aligned}\partial \circ \left(-h_{\Delta^{n+1},n} \circ \partial + bs_{\Delta^{n+1},n+1} - id\right) &= h_{\Delta^{n+1},n-1} \circ \partial^2 - bs_{\Delta^{n+1},n} \circ \partial + \partial \\ &\qquad + \partial \circ bs_{\Delta^{n+1},n+1} - \partial \\ &= -bs_{\Delta^{n+1},n} \circ \partial + \partial \circ bs_{\Delta^{n+1},n+1} \\ &= 0\end{aligned}$$

where in the first step we have used the homotopy relation in degree n from our induction assumption, in the second step we have used $\partial^2 = 0$, and in the last step we have used that $bs_{\Delta^{n+1}}$ is a morphism of chain complexes. □

Let us now return to the setup of the excision theorem: Let $U \subset A \subset X$ be a triple of topological spaces such that the closure of U lies in $\mathrm{int}(A)$. As before we say that a chain $\sigma \in C_n(X)$ is *small* with respect to the given triple if it can be written as a linear combination of simplices such that each of the simplices has its image either in A or in $X \setminus U$. It is clear from the above constructions that both $bs_{X,n}$ and h_n send small chains to small chains. Hence to complete the proof of the excision theorem it only remains to check:

Lemma 5.9. *For any $\sigma \in C_n(X)$, the chain $bs^k_{X,n}(\sigma)$ is small for all $k \gg 0$.*

Proof. By linearity it suffices to prove this for simplices $\sigma\colon \Delta^n \to X$. In this case we may replace the spaces $U \subset A \subset X$ by their preimages under the continuous map σ so that $X = \Delta^n$ becomes the standard simplex. Then in particular X is a compact metric space with respect to the Euclidean metric, so we can talk about the diameter of closed subsets of it. Now for $k \to \infty$ the diameter of the images of the simplices in the k-fold barycentric subdivision $bs^k(\sigma)$ go to zero, hence the claim follows from the Lebesgue lemma applied to the open cover

$$X = U_1 \cup U_2 \quad \text{with} \quad U_1 = \mathrm{int}(A) \quad \text{and} \quad U_2 = X \setminus \overline{U}.$$

Indeed, the Lebesgue lemma says that for any compact metric space X with an open cover $X = \bigcup_{i \in I} U_i$, there exists an $\varepsilon > 0$ such that every subset of diameter $d < \varepsilon$ is completely contained in one of the open subsets U_i. □

Note that it is here that the assumption $\overline{U} \subset \mathrm{int}(A)$ is used, since otherwise the two open subsets in the proof of the above lemma will not cover X. Indeed, if the closure of U intersects the boundary of A, then a simplex at a point of the intersection $\overline{U} \cap \partial A$ will not become small, no matter how often we subdivide it:

[PICTURE]

6 Comparison with simplicial homology and the ES axioms

We now show that for topological spaces with a triangulation as a Δ-complex, the singular homology coincides with the simplicial homology as defined earlier. Recall that a Δ-complex is a pair $\mathbb{X} = (X, (\sigma_\alpha)_{\alpha \in I})$ of a topological space X and a family of continuous maps

$$\sigma_\alpha\colon \quad \Delta^{n_\alpha} \to X \quad \text{for suitable} \quad n_\alpha \in \mathbb{N}_0,$$

indexed by some set I, such that the following properties are satisfied:

a) The restriction of each σ_α to the open simplex is injective, and as a set X is the disjoint union of these open simplices.
b) A subset $U \subset X$ is open iff its preimage $\sigma_\alpha^{-1}(U) \subset \Delta^{n_\alpha}$ is open for all α.
c) Each face of a simplex σ_α is a simplex σ_β for some $\beta \in I$.

We have then considered the simplicial chain complex $C_\bullet(\mathbb{X})$ whose terms are given by

$$C_n(\mathbb{X}) = \bigoplus_{\alpha \in I(n)} \mathbb{Z} \cdot [\sigma_\alpha] \quad \text{with} \quad I(n) := \{\alpha \in I \mid n_\alpha = n\},$$

and we have defined simplicial homology by $H_n(\mathbb{X}) = H_n(C_\bullet(\mathbb{X}))$. Since every σ_α is in particular a singular simplex in X, there exists a natural embedding of chain complexes $C_\bullet(\mathbb{X}) \hookrightarrow C_\bullet(X)$, hence a natural morphism

$$H_n(\mathbb{X}) \to H_n(X)$$

on homology for each n. To prove that this is an isomorphism, it will be convenient to consider also relative simplicial homology:

Definition 6.1. A *Δ-subcomplex* of a Δ-complex $\mathbb{X} = (X, (\sigma_\alpha)_{\alpha \in I})$ is a Δ-complex of the form

$$\mathbb{A} = (A, \{\sigma_\beta\}_{\beta \in J}) \quad \text{where} \quad A = \bigcup_{\beta \in J} \operatorname{im}(\sigma_\beta) \subset X \quad \text{for a subset} \quad J \subset I.$$

We then say that $\mathbb{A} \subset \mathbb{X}$ is a *pair of Δ-complexes*. For any such pair let $C_\bullet(\mathbb{X},\mathbb{A})$ denote the corresponding relative simplicial chain complex, defined by the short exact sequence of complexes

$$0 \to C_\bullet(\mathbb{A}) \to C_\bullet(\mathbb{X}) \to C_\bullet(\mathbb{X},\mathbb{A}) \to 0.$$

and consider the relative simplicial homology groups $H_n(\mathbb{X},\mathbb{A}) = H_n(C_\bullet(\mathbb{X},\mathbb{A}))$.

As for singular homology, the relative homology groups of a pair fit in a long exact sequence, and for $\mathbb{A} = \varnothing$ we recover the ordinary simplicial homology. The comparison between simplicial and singular homology works for arbitrary pairs:

Theorem 6.2. *Let $\mathbb{A} \subset \mathbb{X}$ be a pair of Δ-complexes. Then for all $n \in \mathbb{N}_0$ we have a natural isomorphism*

$$H_n(\mathbb{X},\mathbb{A}) \xrightarrow{\sim} H_n(X,A).$$

Proof. $a)$ Let us first assume that the Δ-complex $\mathbb{X}$ is *finite* in the sense that it has only finitely many simplices. In particular the dimension of the simplices in $\mathbb{X}$ is then bounded above. Fix an index $\alpha \in I$ such that the dimension $n = n_\alpha$ is maximal, and let $\sigma := \sigma_\alpha \colon \Delta^n \to X$ be the corresponding simplex. Put $\Delta = \sigma(\Delta^n) \subset X$. We first prove the statement in the special case where

$$\mathbb{A} \subset \mathbb{X} \quad \text{is the } \Delta\text{-subcomplex with underlying space} \quad A = X \smallsetminus \mathrm{int}(\Delta).$$

In this case the group $C_i(\mathbb{X},\mathbb{A})$ vanishes for $i \neq n$ and is free abelian generated by $[\sigma]$ for $i = n$. Hence

$$H_i(\mathbb{X},\mathbb{A}) = \begin{cases} \mathbb{Z}\cdot[\sigma] & \text{for } i = n, \\ 0 & \text{otherwise.} \end{cases}$$

We need to show that this agrees with the singular homology of the pair $A \subset X$. To compute this singular homology, let $B \subset \mathrm{int}(\Delta)$ be an open ball such that the identity map id_X is homotopic to a map which retracts the complement $X \smallsetminus B$ onto A:

[PICTURE]

Then the inclusion of pairs $(X,A) \hookrightarrow (X, X \smallsetminus B)$ is a homotopy equivalence, so one computes

$$\begin{aligned} H_i(X,A) &\simeq H_i(X, X \smallsetminus B) && \text{by homotopy invariance} \\ &\simeq H_i(\mathrm{int}(\Delta), \mathrm{int}(\Delta) \smallsetminus B) && \text{by excision} \\ &\simeq H_i(\mathbb{R}^n, \mathbb{R}^n \smallsetminus \{*\}) && \text{by homotopy invariance} \\ &\simeq H_i(\mathbb{X},\mathbb{A}) && \text{by comparison with remark 4.12.} \end{aligned}$$

b) We now prove that for any finite Δ-complex $\mathbb{X}$, the maps $H_i(\mathbb{X}) \to H_i(X)$ from simplicial to singular homology as defined above are isomorphisms. This follows from the previous step by induction on the number of simplices: Let $\mathbb{A} \subset \mathbb{X}$ be the subcomplex obtained by removing one simplex of maximal dimension. The long exact sequences for simplicial and relative homology fit into a commutative diagram of the following form:

$$\begin{tikzcd}
H_{i+1}(\mathbb{X},\mathbb{A}) \arrow[r] \arrow[d] & H_i(\mathbb{A}) \arrow[r] \arrow[d] & H_i(\mathbb{X}) \arrow[r] \arrow[d] & H_i(\mathbb{X},\mathbb{A}) \arrow[r] \arrow[d] & H_{i-1}(\mathbb{A}) \arrow[d] \\
H_{i+1}(X,A) \arrow[r] & H_i(A) \arrow[r] & H_i(X) \arrow[r] & H_i(X,A) \arrow[r] & H_{i-1}(A)
\end{tikzcd}$$

By induction the second and fifth vertical maps are isomorphisms, and by step *a*) the first and the fourth vertical maps are isomorphisms as well. Hence by the 5-lemma the middle vertical map is also an isomorphism.

c) Next we show that for any finite Δ-complex $\mathbb{X}$ and any $\mathbb{A} \subset \mathbb{X}$, the induced maps $H_i(\mathbb{X},\mathbb{A}) \to H_i(X,A)$ from simplicial to singular relative homology are isomorphisms. In fact this follows by the same argument as above by applying the 5-lemma to the commutative diagram

$$\begin{tikzcd}
H_i(\mathbb{A}) \arrow[r] \arrow[d] & H_i(\mathbb{X}) \arrow[r] \arrow[d] & H_i(\mathbb{X},\mathbb{A}) \arrow[r] \arrow[d] & H_{i-1}(\mathbb{A}) \arrow[r] \arrow[d] & H_{i-1}(\mathbb{X}) \\
H_i(A) \arrow[r] & H_i(X) \arrow[r] & H_i(X,A) \arrow[r] & H_{i-1}(A) \arrow[r] & H_{i-1}(X)
\end{tikzcd}$$

where all arrows except for the middle one are isomorphisms by step *b*).

d) The result for infinite Δ-complexes can be deduced from the result for finite Δ-complexes by using that by compactness, every fixed singular chain in X meets only finitely many simplices of the Δ-complex. The details are left to the reader as an exercise (or see Hatcher's book). □

Note that in the above argument we have only used that the singular homology groups form a family of functors

$$H_i\colon \quad (\text{pairs of topological spaces}) \to (\text{abelian groups}), \quad (X,A) \mapsto H_i(X,A)$$

indexed by integers $i \geq 0$ such that the following properties hold:

(ES1) Long exact sequence: For any pair (X,A), putting $H_i(X) := H_i(X,\varnothing)$ we have a long exact sequence

$$\cdots \to H_i(A) \to H_i(X) \to H_i(X,A) \to H_{i-1}(A) \to \cdots$$

(ES2) Homotopy invariance: Any homotopy equivalence $(X,A) \to (X',A')$ induces an isomorphism $H_i(X,A) \xrightarrow{\sim} H_i(X',A')$ for all i.

(ES3) Excision: For any subspace $U \subset X$ whose closure is contained in $\mathrm{int}(A)$, we have an isomorphism $H_i(X,A) \simeq H_i(X \setminus U, A \setminus U)$.

(ES4) Dimension axiom: For a point we have

$$H_i(\mathrm{pt}) = \begin{cases} \mathbb{Z} & \text{for } i = 0, \\ 0 & \text{for } i \neq 0. \end{cases}$$

Note that we do not include the Mayer-Vietoris sequence as an axiom since we have deduced it formally from the excision theorem. The above four axioms are known as the *Eilenberg-Steenrod axioms*. The same argument as in the previous proof show that any family of functors satisfying these axioms is isomorphic to singular homology.

Chapter V
Appendix: Some homological algebra

1 Exact sequences

Let A, B, C be abelian groups. A sequence of homomorphisms $A \xrightarrow{f} B \xrightarrow{g} C$ is called *exact* if

$$\operatorname{im}(f) = \ker(g).$$

A sequence of more than two homomorphisms of abelian groups is called exact if any two successive morphisms in it form an exact sequence in the previous sense. By a *short exact sequence* we mean an exact sequence of the form

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0.$$

More explicitly,

- exactness at A means that f is injective,
- exactness at B means that $\operatorname{im}(f) = \ker(g)$, and
- exactness at C means that g is surjective.

Homological algebra is the study of relations that can be deduced from combinations of exact sequences. A typical example is the following observation:

Lemma 1.1 (5-lemma). *Let*

$$\begin{tikzcd}
A \arrow[r, "\alpha"] \arrow[d, two heads, "a"] & B \arrow[r, "\beta"] \arrow[d, "b", "\simeq"'] & C \arrow[r, "\gamma"] \arrow[d, "c"] & D \arrow[r, "\delta"] \arrow[d, "d", "\simeq"'] & E \arrow[d, hook, "e"] \\
A' \arrow[r, "\alpha'"] & B' \arrow[r, "\beta'"] & C' \arrow[r, "\gamma'"] & D' \arrow[r, "\delta'"] & E'
\end{tikzcd}$$

be a commutative diagram with exact rows where b and d are isomorphisms, a is surjective and e is injective. Then c is an isomorphism as well.

Proof. ... $\square$

The technique that we applied in the above proof is known as *diagram chasing*. Another instance is the following frequently used fact:

Lemma 1.2 (Snake lemma). *Let*

$$\begin{array}{ccccccc} & & A & \longrightarrow & B & \longrightarrow & C & \longrightarrow & 0 \\ & & \big\downarrow f & & \big\downarrow g & & \big\downarrow h & & \\ 0 & \longrightarrow & A' & \longrightarrow & B' & \longrightarrow & C' & & \end{array}$$

be a commutative diagram whose rows are exact sequences. Then we have a short exact sequence

$$\ker(f) \to \ker(g) \to \ker(h) \to \operatorname{coker}(f) \to \operatorname{coker}(g) \to \operatorname{coker}(g).$$

Proof. ... □

2 Complexes and their homology

A *complex* of abelian groups is a sequence

$$A_* = \left[\cdots \longrightarrow A_{n+1} \xrightarrow{f_{n+1}} A_n \xrightarrow{f_n} A_{n-1} \xrightarrow{f_{n-1}} \cdots \right]$$

of homomorphisms of abelian groups such that $f_n \circ f_{n-1} = 0$ for all n. Note that this is weaker than saying that the sequence is exact, and this is precisely what makes complexes interesting since it gives rise to new invariants which measure the failure of exactness:

Definition 2.1. For a complex A_* we call

- $Z_n(A_*) := \ker(f_n) \subset A_n$ the subgroup of n-cycles, and
- $B_n(A_*) := \operatorname{im}(f_{n-1}) \subset A_n$ the subgroup of n-boundaries.

The condition $f_n \circ f_{n-1} = 0$ means $B_n(A_*) \subset Z_n(A_*)$. The *homology groups* of the complex are defined by

$$H_n(A_*) := Z_n(A_*)/B_n(A_*).$$

TODO: Define morphisms of complexes.

Lemma 2.2. *Any morphism of complexes $f\colon A_* \to B_*$ induces morphisms between their homology groups*

$$f_*\colon \quad H_n(A_*) \to H_n(B_*) \quad \textit{for all } n \in \mathbb{N}_0.$$

Proof. ... □

TODO: Define chain homotopies between morphisms of complexes. S

Lemma 2.3. *If $f, g\colon A_* \to B_*$ are chain homotopic, then on homology they induce the same map*

$$f_* = g_*\colon \quad H_n(A_*) \to H_n(B_*) \quad \text{for all } n \in \mathbb{N}_0.$$

Proof. ... □

3 The long exact homology sequence

The notion of short exact sequences has an obvious extension to maps between complexes:

Definition 3.1. A short exact sequence of complexes is a sequence $0 \to A_* \to B_* \to C_* \to 0$ of morphisms of complexes such that in each degree $n \geq 0$ the sequence

$$0 \to A_n \to B_n \to C_n \to 0 \quad \text{is exact.}$$

In general, for a short exact sequence of complexes as above the induced sequence on homology groups in a given degree n is not exact. Instead, all these sequence fit together into a long exact sequence:

Theorem 3.2. *Any short exact sequence $0 \to A_* \to B_* \to C_* \to 0$ of complexes gives rise to a long exact sequence*

$$\cdots \to H_n(A_*) \to H_n(B_*) \to H_n(C_*) \to H_{n-1}(A_*) \to \cdots$$

Proof. Consider for each $n \geq 0$ the differentials $\partial_n\colon A_n \to A_{n-1}$. Since $\partial_n \circ \partial_{n+1} = 0$, the differential ∂_n factors through a map

$$\overline{\partial}_n\colon \quad \operatorname{coker}(\partial_{n+1}) \to A_{n-1}.$$

Since $\partial_{n-1} \circ \partial_n = 0$ and therefore $\partial_{n-1} \circ \overline{\partial}_n = 0$, we then have an induced map

$$\overline{\partial}_n\colon \quad \operatorname{coker}(\partial_{n+1}) \to \ker(\partial_n).$$

Putting together these maps and their couterparts for the complexes B_* and C_*, we obtain a commutative diagram

$$\begin{tikzcd}
 & \operatorname{coker}(\partial^A_{n+1}) \arrow[r] \arrow[d] & \operatorname{coker}(\partial^B_{n+1}) \arrow[r] \arrow[d] & \operatorname{coker}(\partial^C_{n+1}) \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \ker(\partial^A_{n-1}) \arrow[r] & \ker(\partial^B_{n-1}) \arrow[r] & \ker(\partial^C_{n-1}) &
\end{tikzcd}$$

The claim now follows from the snake lemma. □

In various situations we want to compare two long exact sequences of the above form. The following lemma (whose proof could have been given in section 1) achieves this when every third morphism between the terms of the sequences is an isomorphism:

Lemma 3.3. *Suppose we are given a commutative diagram*

$$\begin{tikzcd}
\cdots \arrow[r] & A_n \arrow[r, "\alpha_n"] \arrow[d, "f_n"] & B_n \arrow[r, "\beta_n"] \arrow[d, "g_n"] & C_n \arrow[r, "\gamma_n"] \arrow[d, "h_n"] & A_{n-1} \arrow[r, "\alpha_{n-1}"] \arrow[d, "f_{n-1}"] & \cdots \\
\cdots & A'_n \arrow[r, "\alpha'_n"] & B'_n \arrow[r, "\beta'_n"] & C'_n \arrow[r, "\gamma'_n"] & A'_{n-1} \arrow[r, "\alpha'_{n-1}"] & \cdots
\end{tikzcd}$$

with exact rows. If the maps h_n are isomorphisms for all n, then we have a long exact sequence

$$\cdots \longrightarrow A_n \xrightarrow{(\alpha_n, f_n)} B_n \oplus A'_n \xrightarrow{d} B'_n \xrightarrow{\delta} A_{n-1} \longrightarrow \cdots$$

Proof. ... $\square$