Lecture Notes in Elementary Algebraic Topology 1.1

Binbin Xu

December 30, 2025

Contents

This lecture note is based on the course "Elementary Algebraic Topology" which I taught in the fall 2023. For the moment, there are still some parts to be completed.

This course is an elementary introduction to Algebraic Topology for those who first meets this subject.

I would like to thank all students taking this course for their participation and all the valuable discussions.

Chapter 1

Introduction

In this chapter, we review briefly some history on general topology and algebraic topology. More details can be found in [2] and the following two websites

mathshistory.st-andrews.ac.uk
analysis-situs.math.cnrs.fr

1.1 Some interesting problems

We start by discussing some interesting problems to have an idea of what a topological problem looks like.

The Seven Bridges of Königsberg

The starting point of topology is Euler's study on the famous "The Seven Bridges Problem". Here (See Figure 1.1.1) is a map of an old town called Köningsberg which was the capital of the east Prussia in east Europe. It is now called Kaliningrad, a city of Russia.

Figure 1.1.1: Map of Köningsberg (from Wikipedia)

The city was divided into four parts by the Pregel River. The question asked by Euler was:

Is it possible to visit all parts of the city by passing each bridge exactly once?

The actual shape of each part is not essential in this problem. Figure 1.1.2 is a sketch map of the town to simplify the situation.

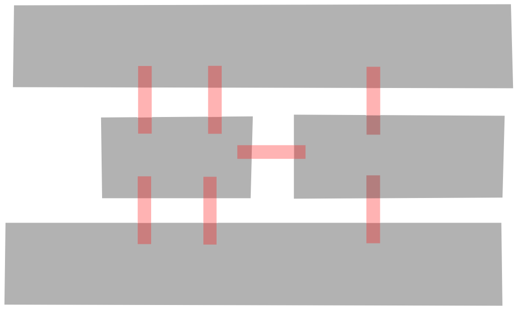


Figure 1.1.2: Sketch map of Köningsberg

Notice that when we try to solve this question by walking in the city, staying in one part will not change the result. Therefore, we can simplify the map by shrinking each part to a point and obtain the following graph (See Figure 1.1.3).

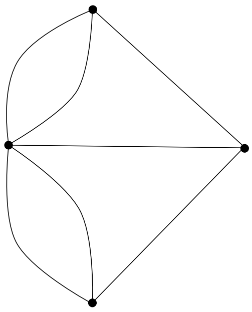


Figure 1.1.3: Graph for The Seven Bridges Problem

Euler's question is then equivalent to the following one:

Is it possible to find a circle path in the graph passing each edge of the graph exactly once?

One observation is that if we pass a vertex in the middle of the circle path, there should be one edge arriving at this vertex and another one leaving it. Hence, if we call the number of half edges adjacent to one vertex the *degree* of this vertex, then all vertices should have even degree unless it is the starting or the ending vertex.

In 1736, Euler published a paper on the solution of this question, not only showing that this is impossible, but also providing a solution for the general question which can be stated in today's language as follows.

Theorem 1.1.1 (Euler)

A finite connected graph has a circle path passing each edge exactly once if and only if there is no vertex with odd degree.

It has a path with distinct starting and ending points passing each edge exactly once if and only if there are exactly two vertices with odd degree.

Notice that the graph for "The Seven Bridges Problem" has 4 vertices where one has degree 5 and the other three have degree 3. Hence there is no way that we can visit all parts of the city by passing each bridge exactly once.

Polyhedron

Another famous work done by Euler is about convex polyhedrons in the Euclidean space. Given a convex polyhedron, Euler gives the following formula

$$v - e + f = 2$$

where v is the number of vertices of the polyhedron, e is the number of edges of the polyhedron and f is the number of faces of the polyhedron.

There are three observations which one can make from this identity. Firstly all data involved in this identity has nothing to do with the geometry of the convex polyhedron. For example, any convex polyhedron with 5 vertices with the same adjacency relation among vertices, edges and faces has a same identity (See Figure 1.1.4).

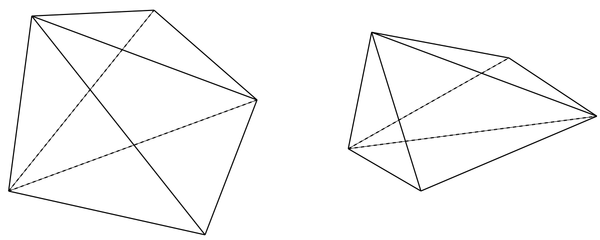


Figure 1.1.4: Identity for different geometric information: $5 - 9 + 6 = 2$

Secondly the value on the right hand side is a constant independent of the values of v, e and f. In the other words, if we consider another convex polyhedron with maybe 6 or 10 vertices, this constant is still the same (See Figure 1.1.5).

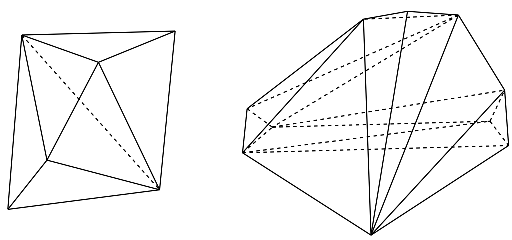


Figure 1.1.5: Identity for the left: $6 - 10 + 6 = 2$;
Identity for the right: $7 - 11 + 6 = 2$

Thirdly this identity also holds for non-convex polyhedron (See Figure 1.1.6).

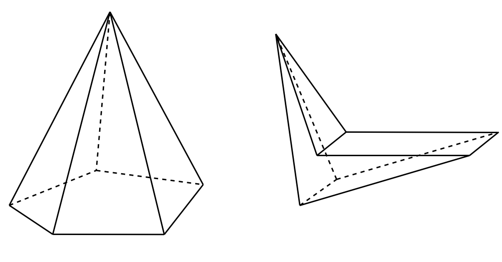


Figure 1.1.6: Identity for the left: $6 - 12 + 8 = 2$;
Identity for the right: $10 - 24 + 16 = 2$

Later, this identity is generalized by Antoine-Jean Lhuilier. He notice that Euler's formula is wrong when there are "holes" in the polyhedrons. If there are g holes, then we have

$$v - e + f = 2 - 2g.$$

For example, Figure 1.1.7 is a polyhedron with 1 hole.

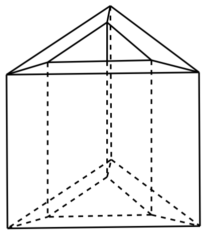


Figure 1.1.7: Identity for genus 1: $12 - 24 + 12 = 0 = 2 - 2 \times 1$

Remark 1.1.2.
Given a polyhedron with g holes, if we do not distinguish vertices, points on edges and points on faces, the surface of this polyhedron is a genus g surface. From this point of view, each polyhedron with g holes can be considered as a genus g surface marked by some distinguished points as vertices and distinguished lines as edges. The value $2 - 2g$ on the right hand side of the identity only depends on the genus g of the surface. This is called the *Euler characteristic* of the genus g surface. In particular, Euler's work consider the case $g = 0$ where the surface is a sphere.

Intersection number between two closed planar curves

Consider two closed curves

$$\begin{aligned} \alpha : S^1 &\to \mathbb{R}^2 \\ s &\mapsto (\alpha_1(s), \alpha_2(s)) \end{aligned}$$

and

$$\begin{aligned} \beta : S^1 &\to \mathbb{R}^2 \\ t &\mapsto (\beta_1(t), \beta_2(t)) \end{aligned}$$

For simplicity, we assume that both α and β are $\mathcal{C}^1$, and for any $s, t \in S^1$, we have $\dot{\alpha}(s) \neq (0,0)$ and $\dot{\beta}(t) \neq (0,0)$.

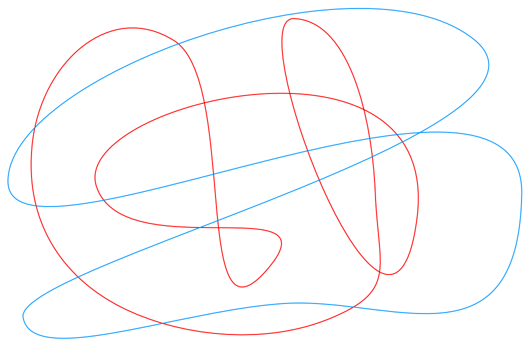


Figure 1.1.8: Two curves with 14 intersection points with each other.

Assume that $\alpha \cap \beta$ is finite. We are interested in the parity of the number of intersection points between α and β. To make this question more clear, we assume that all intersections between α and β are transversal, and α and β pass each intersection point only once. In the other words, at each intersection point in $\alpha \cap \beta$, we assume that the tangent vector of α and the one of β are linearly independent, and for each $s \in S^1$ (resp. $t \in S^1$), there is at most one $t \in S^1$ (resp. $s \in S^1$), such that $\alpha(s) = \beta(t)$.

Poincaré showed the following result.

Theorem 1.1.3 (Poincaré)

The number of intersection points between α and β is always even.

Notice that in the statement of this result, there is no condition on geometric information of α and β, although the whole problem lies in a context of Euclidean geometry.

Intuitively this is not hard to understand. If α is a round circle, it separate the plane into two parts. We call the compact part the inside of α, and the infinite part the outside of α. When we walk along β with the starting point outside of α. The intersection happens when we meet α. Since the intersection is transversal, each time when we meet α, we go from inside to outside or from outside to inside. Since the starting point is outside, we have to meet α even number of times to be outside of α, which suggests the number of intersection points between α and β should be even. Of course, one needs to consider general cases and make a rigorous proof to get a theorem.

1.2 Poincaré's analysis situs

Since 1895, Poincaré published the famous paper "Analysis Situs" and its five supplements, introducing "analysis situs" which he considered as a third geometry after the metric geometry and the projective geometry. Its key feature different from the previous ones is that there is no more notion of quantities of geometric measurements. The properties considered are all qualitative. For example, two figures are considered as the same if we can change one to the other by a continuous deformation.

In Analysis situs, Poincaré introduced the notion of manifold, Betti number, homology and cohomology and their duality, fundamental group, the Euler-Poincaré formula, etc. Notice that from the work of Euler and its generalization, we may associate numbers to topological objects as topological invariants. Poincaré extended the notion of topological invariant, so that a topological invariant could be an algebraic object, such as fundamental group, homology group, etc, instead of just a number.

For more details about "Analysis Situs", see the following CNRS website:

analysis-situs.math.cnrs.fr

Chapter 2

General Topology

2.1 Topological Space

As mentioned in the introduction, Poincaré considered topology as "third geometry" following "metric geometry" (about distance) and "projective geometry" (about lines). The key feature differentiating "topology" from the other two is that there is no distance, angle or any other quantitative measurements. For example, in the "Seven Bridges Problem", when we walk in the city, all we care about is which part of the city we are in, instead of the exact location. In other words, we still care about geometry but in a large sense. Instead of saying the exact location, we will consider neighborhoods. This will be described by so called topological structures.

Topological structures

Definition 2.1.1

Let X be a non-empty set. A ***topological structure*** on X is a collection $\mathcal{T}$ of subsets of X, satisfying the following properties:

1) $X \in \mathcal{T}$ and $\emptyset \in \mathcal{T}$;

2) for any $U, V \in \mathcal{T}$, we have $U \cap V \in \mathcal{T}$;

3) for any non-empty subset $\mathcal{A} \subset \mathcal{T}$, we have

$$\bigcup \{U \mid U \in \mathcal{A}\}.$$

Definition 2.1.2

A ***topological space*** is a couple $(X, \mathcal{T})$ where X is a non-empty set and $\mathcal{T}$ is a topological structure on X.

Given any topological structure $\mathcal{T}$ on X, we will call it a *topology* on X for short.

Remark 2.1.3 (a remark on the word "space").
Mathematically there is no essential difference between "space" and "set". Usually when we see the word "space", we should expect a set with certain structure (for example topological structure, differential structure, metric structure, symplectic structure, etc.) depending on the context, most of the time relating to geometry. In this course, by a space we usually mean a topological space.

To simplify the notation, given a topological space $(X, \mathcal{T})$, when the topological structure $\mathcal{T}$ is clear, we may simply denote it by X. Sometimes, we also say a space X without mentioning $\mathcal{T}$. This means that we already have a topological structure chosen once and for all.

Definition 2.1.4

Let $(X, \mathcal{T})$ be a topological space. Any subset $U \in \mathcal{T}$ is called an ***open subset*** of X for the topological structure $\mathcal{T}$. A subset $K \subset X$ is said to be ***closed*** if its complement is open.

Remark 2.1.5.
It is possible (not necessary) that a subset is open and closed at the same time. Two trivial examples are $\emptyset$ and X for any topology on the space X.

Remark 2.1.6.
Roughly speaking, the subsets in the collection $\mathcal{T}$ tells "neighborhood"s of each point in X.

Consider the distance on the real line $\mathbb{R}$ given by the absolute value of the difference between two points. To get to one point $x \in \mathbb{R}$ from anther point $y \in \mathbb{R}$, we may walk along $\mathbb{R}$ from y and check the distance to x from our position. When the distance becomes 0, we know that we arrive at x (See Figure 2.1.1).

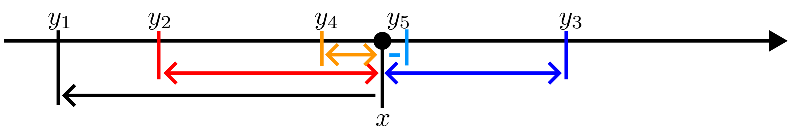


Figure 2.1.1: A sequence converges to x

When we consider this in the context of topology, there is no notion of distance. Instead, we may consider get into all neighborhoods containing the point x to say we get to the point x (See Figure 2.1.2).

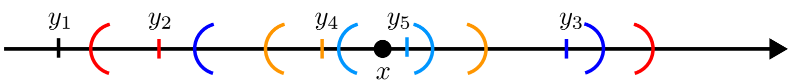


Figure 2.1.2: A sequence converges to x

On the other hand, if we are not close to x, there must be one neighborhood where we are not in. Therefore, to make this more rigorous, we have to make precise the meaning of neighborhood here, and this is the notion related to open sets.

This example suggests that it is possible to use topologies on a space to distinguish points in the Euclidean space. This may not be the case when considering more general topological structures. A trivial example is the space X equipped with the topology $\mathcal{T} = \{\emptyset, X\}$ discussed later.

Example 2.1.7 (Euclidean space $\mathbb{R}^3$ / Metric topology).
We start by an example familiar to us the most. Consider the Euclidean space $\mathbb{R}^2$, and denote by

$d_{\mathbb{E}}$ the Euclidean metric. For any point $p \in \mathbb{R}^2$ and any positive real number r, the *open ball* in $\mathbb{R}^2$ centered at p with radius r is defined to be

$$B_p(r) := \{q \in \mathbb{R}^2 \mid d_{\mathbb{E}}(p, q) < r\}.$$

We define that a subset $U \subset \mathbb{R}^2$ is said to be *open* if

$$\forall p \in U, \exists r > 0, B_p(r) \subset U.$$

We can verify that these open sets form a topology on $\mathbb{R}^2$ (See Figure 2.1.3).

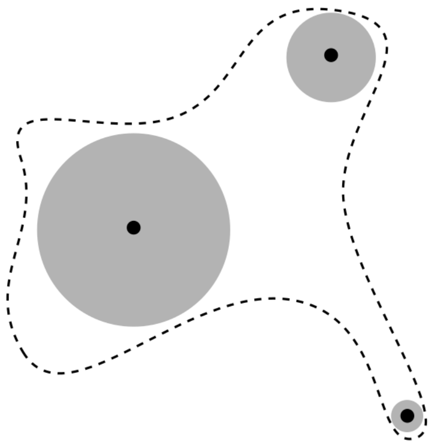


Figure 2.1.3: An open set in $\mathbb{R}^2$

This construction can be done for any metric space. We call such a topology the one induced by the metric or simply the *metric topology*.

Question 2.1.8
Show that with respect to this topology, an open ball $B_p(r)$ is open.

Example 2.1.9 (Trivial topology).
If we forget the intuition from the metric geometry and simply play with the definition, we may find the following two trivial examples. One is the following subset of $\mathcal{P}(X)$ denoted by

$$\mathcal{O}_X := \{X, \emptyset\},$$

while the other one is the whole power set $\mathcal{P}(X)$. They both satisfy all conditions in Definition 2.1.1 trivially, hence both induce topologies on X. The topology $\mathcal{O}_X$ is called the *trivial topology* on X, and the topology $\mathcal{P}(X)$ is called the *discrete topology* on X. In particular, every point $x \in X$ forms an open set in the discrete topology.

Previously, we discussed how to talk about convergences without using distance. Under the discrete topology, if a sequence of points x_n in X converges to x, since the idea is to get into every neighborhood of x, it is eventually a constant sequence with all $x_n = x$ for n bigger than some $N \in \mathbb{N}$.

Example 2.1.10 (Initial topology).
Later we will introduce the notion of continuity, which is a map between two spaces relating the

open sets in two spaces in certain way. Sometimes, we may be in a situation where we have a map (or several maps) first and try to find and study a topology with respect to which the map (or the maps) is continuous. One such example would be the projection from a Cartesian product to one of its factors. See Subsection 2.3 for details on product spaces.

We consider the construction for one map to have an idea. Let X be a non-empty set, and $(Y, \mathcal{T}_Y)$ be a topological space. Let f be any map from X to Y. Then

$$\mathcal{T} := \{f^{-1}(U) \mid U \subset \mathcal{T}_Y\}$$

gives a topological structure on X. Under this topology the map f is continuous.

Example 2.1.11 (Zariski topology).
Let F be any field. Given a natural number $n > 0$, we consider the set F^n. We say a subset $U \subset F^n$ is *Zariski closed* if it is the solution set of a family of polynomials on F with n variables. By considering the complement of a Zariski closed subset of F^n as an open set, we get the *Zariski topology* on F^n. This is used a lot in the study of Lie theory and algebraic geometry.

Given any non-empty set X, as we have seen above, the space X could be equipped with different topological structures. Hence whether a subset of X is open depends on the choice of topological structures.

From its definition, it is possible that a topology on a non-empty set could be quite arbitrary and artificial. By definition, to give a topology on a non-empty set, it is enough to describe all open sets in this topology, which is also equivalent to describe all closed sets for this topology.

Comparison between topologies

Roughly speaking, a topology on X is a subset of $\mathcal{P}(X)$ satisfying certain properties. Unless X contains only one element, such a subset in $\mathcal{P}(X)$ is not unique, i.e. the topology on X is not unique when X contains more than one element. The partial order on the set $\mathcal{P}(X)$ given by inclusion induces a partial order among all topologies on X.

Definition 2.1.12

Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be two topologies on X. We say that $\mathcal{T}_1$ is ***finer*** than $\mathcal{T}_2$ if we have

$$\mathcal{T}_2 \subset \mathcal{T}_1.$$

In this case, we also say that $\mathcal{T}_2$ is ***coarser*** than $\mathcal{T}_1$.

This is equivalent to say that an open set in $\mathcal{T}_2$ is also an open set in $\mathcal{T}_1$. This may remind us of the comparison between different partitions of a given set. We will discuss this after introducing the notion of subbasis and basis. Before that let us check two trivial examples.

Example 2.1.13.
By definition, all topologies of X contains X and $\emptyset$. Therefore, the trivial topology $\mathcal{O}_X$ is the coarsest topology of X. On the other hand, any topology if coarser than the discrete topology $\mathcal{P}(X)$. Therefore $\mathcal{P}(X)$ is the finest topology of X.

Example 2.1.14 (Topologies on $\mathbb{R}^2$ induced by different metrics).
There are different ways to define metrics on $\mathbb{R}^2$, each of which can induce a topology on $\mathbb{R}^2$. Here we consider the four metrics whose distance functions are given by the following formulas: let O denote the origin of $\mathbb{R}^2$, for any point $p = (x_1, y_1)$ and $q = (x_2, y_2)$,

1. $\mathrm{d}_{\mathbb{E}}(p,q) = \sqrt{(x_1-x_2)^2+(y_1-y_2)^2}$;
2. $\mathrm{d}_{\infty}(p,q) = \max\{|x_1-x_2|+|y_1-y_2|\}$;
3. $\mathrm{d}_{SNCF}(p,q) = \begin{cases} \mathrm{d}_{\mathbb{E}}(p,q), & \text{if } O,\ p \text{ and } q \text{ are colinear} \\ \mathrm{d}_{\mathbb{E}}(p,O)+\mathrm{d}_{\mathbb{E}}(q,O), & \text{otherwise} \end{cases}$;
4. $\widetilde{\mathrm{d}}(p,q) = 1$.

Respectively, we denote their corresponding metric topologies $\mathcal{T}_{\mathbb{E}}$, $\mathcal{T}_{\infty}$, $\mathcal{T}_{SNCF}$ and $\widetilde{\mathcal{T}}$. Then by considering balls for each metric, we can verify the following relations

$$\mathcal{T}_{\mathbb{E}} = \mathcal{T}_{\infty} \subset \mathcal{T}_{SNCF} \subset \widetilde{\mathcal{T}}.$$

Remark 2.1.15.
We remark here that not every pair of topologies can be compared.

Subbases and bases

To give a topology on a set X, we may describe all its open sets. Alternatively, we can begin with some subsets of X and try to get a topology of X by considering their intersections and unions. This is what we call "generating a topology on X from a collection of its subsets".

More precisely, we denote by $\mathbb{T} \subset \mathcal{P}(\mathcal{P}(X))$ the collection of all topologies on X. Let $\mathcal{A}$ denote a collection of non-empty subsets of X. We give the following definition.

Definition 2.1.16

The ***topology generated by*** $\mathcal{A}$ is defined to be the following one

$$\mathcal{T}_{\mathcal{A}} := \bigcap\{\mathcal{T} \in \mathbb{T} \mid \mathcal{A} \subset \mathcal{T}\}.$$

Proposition 2.1.17

The set $\mathcal{T}_{\mathcal{A}}$ is a topology on X.

Proof. For any $\mathcal{T} \in \mathbb{T}$, we have

$$\emptyset, X \in \mathcal{T},$$

from which we have

$$\emptyset, X \in \bigcap \mathbb{T} \subset \bigcap\{\mathcal{T} \in \mathbb{T} \mid \mathcal{A} \subset \mathcal{T}\} = \mathcal{T}_{\mathcal{A}}.$$

By the definition of $\mathcal{T}_{\mathcal{A}}$, for any $U, V \in \mathcal{T}_{\mathcal{A}}$, for any $\mathcal{T} \in \mathbb{T}$ containing $\mathcal{A}$, we have

$$U, V \in \mathcal{T}.$$

Since $\mathcal{T}$ is a topology on X, we have

$$U \cap V \in \mathcal{T}.$$

Since This holds for any $\mathcal{T} \in \mathbb{T}$ containing $\mathcal{A}$, we have

$$U \cap V \in \bigcap\{\mathcal{T} \in \mathbb{T} \mid \mathcal{A} \subset \mathcal{T}\} = \mathcal{T}_{\mathcal{A}}.$$

Similarly, by the definition of $\mathcal{T}_{\mathcal{A}}$, for any collection $\{U_\alpha\}_{\alpha\in\Omega}$ of sets in $\mathcal{T}_{\mathcal{A}}$, for any $\mathcal{T} \in \mathbb{T}$ containing $\mathcal{A}$, we have

$$\{U_\alpha\}_{\alpha\in\Omega} \subset \mathcal{T}.$$

Since $\mathcal{T}$ is a topology, we have

$$\bigcup_{\alpha\in\Omega} U_\alpha \in \mathcal{T}.$$

Since this holds for any $\mathcal{T} \in \mathbb{T}$ contianig $\mathcal{A}$, we have

$$\bigcup_{\alpha\in\Omega} U_\alpha \in \bigcap\{\mathcal{T} \in \mathbb{T} \mid \mathcal{A} \subset \mathcal{T}\} = \mathcal{T}_{\mathcal{A}}.$$

As a conclusion of the above discussion, we show that $\mathcal{T}_{\mathcal{A}}$ is a topology on X. □

We can also describe this topology in a constructive way.

Proposition 2.1.18

The topology $\mathcal{T}_{\mathcal{A}}$ consists of subsets in X which can be written as an arbitrary union of finite intersections of subsets in $\mathcal{A} \cup \{X\}$.

Proof. We consider the following subset of $\mathcal{P}(X)$:

$$\left\{ \bigcup_{\alpha\in\Omega} \left(\bigcap_{i=1}^{n_\alpha} U_{\alpha i} \right) \middle| \Omega \text{ arbitrary index set, } n_\alpha \in \mathbb{N}^*,\, U_{\alpha i} \in \mathcal{A} \right\}$$

By consider the distribution, we have

$$\left(\bigcup_{\alpha\in\Omega} \left(\bigcap_{i=1}^{n_\alpha} U_{\alpha i} \right) \right) \cap \left(\bigcup_{\beta\in\Theta} \left(\bigcap_{j=1}^{m_\beta} U_{\beta j} \right) \right) = \bigcup_{\alpha\in\Omega,\, \beta\in\Theta} \left(\left(\bigcap_{i=1}^{n_\alpha} U_{\alpha i} \right) \cap \left(\bigcap_{i=1}^{m_\beta} U_{\beta j} \right) \right)$$

The above set satisfies the Condition 2) and 3) in Definition 2.1.1. By taking Ω to be empty set, we can see that this set also contains $\emptyset$. The only thing not necessarily true is that it contains X.

To get over this problem, we may consider the same construction for $\mathcal{A} \cup \{X\}$. □

Definition 2.1.19

A subset $\mathcal{A}$ of $\mathcal{T}$ is called a ***subbasis*** of $\mathcal{T}$ if it generates $\mathcal{T}$.

Remark 2.1.20.

The above definition is essential the same as the one in the course "Point Set Topology". On the other hand, if $\mathcal{A}$ is required to have the property that

$$\bigcup \mathcal{A} = X.$$

Then by considering all arbitrary union of finite intersections of subsets in $\mathcal{A}$, we can have a topology on X. There is no need to consider $\mathcal{A} \cup X$. This is used as a definition for subbases in some references.

Definition 2.1.21

A subset $\mathcal{B}$ of $\mathcal{T}$ is called a ***basis*** of $\mathcal{T}$ if any $U \in \mathcal{T}$ can be written as a union of open sets in $\mathcal{B}$.

Remark 2.1.22.
Both subbases and bases of $\mathcal{T}$ can be used to generate $\mathcal{T}$. The difference is that when we generate $\mathcal{T}$ using a subbasis, we have to consider both "arbitrary unions" and "finite intersections", while when using a basis, we only need to consider "arbitrary unions".

Any non empty collection of subsets of X can be a subbasis of some topology on X, but this is not true for bases.

To check if a subbasis is actually a basis, we may consider the definition, as well as the following equivalent condition.

Proposition 2.1.23

Assume that $\mathcal{B}$ is a subbasis of $\mathcal{T}$ satisfying $\bigcup \mathcal{B} = X$. Then $\mathcal{B}$ is a basis of $\mathcal{T}$ if and only if it satisfies the following property (see Figure 2.1.4 for an illustration):

- $\forall U, V \in \mathcal{B}$, $\forall x \in U \cap V$, $\exists W \in \mathcal{B}$, $x \in W \subset U \cap V$.

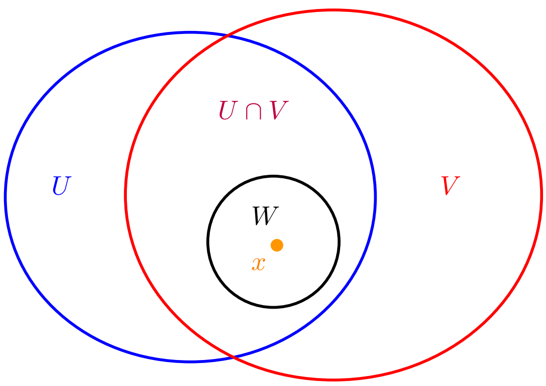


Figure 2.1.4: Condition in Proposition 2.1.23

Proof. One direction is trivial. If $\mathcal{B}$ is a basis, then for any U and V in $\mathcal{B}$, the intersection $U \cap V$ is in $\mathcal{T}$, hence is a union of sets in $\mathcal{B}$. Therefore, for any $x \in U \cap V$, there is a set $W \in \mathcal{B}$, such that

$$x \in W \subset U \cap V.$$

Now we turn to the other direction. Assume that $\mathcal{B}$ is a subbasis, satisfying the condition

- $\forall U, V \in \mathcal{B}$, $\forall x \in U \cap V$, $\exists W \in \mathcal{B}$, $x \in W \subset U \cap V$.

If we can show that any intersection between two sets U and V in $\mathcal{B}$ is still in $\mathcal{B}$, by induction, any finite intersections of sets in $\mathcal{B}$ is in $\mathcal{B}$, hence we can conclude that $\mathcal{B}$ is a basis.

In fact the condition in the proposition is weaker than this. Let U and V be any sets in $\mathcal{B}$. For any $x \in U \cap V$, by hypothesis, there is a set $W_x \in \mathcal{B}$, such that

$$x \in W_x \subset U \cap V,$$

therefore we have

$$U \cap V = \bigcup_{x \in U \cap V} W_x.$$

By induction, we may show that given any finitely many sets $U_1, ..., U_n$ in $\mathcal{B}$, there is a collection of sets $\{W_\alpha\}_{\alpha \in \Omega}$ in $\mathcal{B}$, such that

$$U_1 \cap \cdots \cap U_n = \bigcup_{\alpha \in \Omega} W_\alpha.$$

Since $\mathcal{B}$ is a subbasis of $\mathcal{T}$ with $\bigcup \mathcal{B} = X$, given any $W \in \mathcal{T}$, it can be written as a union of finite intersections among sets in $\mathcal{B}$:

$$W = \bigcup_{\alpha \in \Omega} \left(\bigcap_{i=1}^{n_\alpha} U_{\alpha i} \right).$$

By the previous discussion, for any α, there is a collection of sets $\{W_{\alpha\beta}\}_{\beta \in \Theta}$ in $\mathcal{B}$, such that

$$\bigcap_{i=1}^{n_\alpha} U_{\alpha i} = \bigcup_{\beta \in \Theta} W_{\alpha\beta},$$

therefore, we have

$$W = \bigcup_{\alpha \in \Omega} \left(\bigcup_{\beta \in \Theta} W_{\alpha\beta} \right) = \bigcup_{\alpha \in \Omega} \bigcup_{\beta \in \Theta} W_{\alpha\beta}.$$

Hence $\mathcal{B}$ is a basis of $\mathcal{T}$. □

Example 2.1.24 (Euclidean spaces).
We consider the Euclidean space $\mathbb{R}^3$. Using the same notation as before, the collection of all open balls in $\mathbb{R}^3$ form a basis of the topology of $\mathbb{R}^3$ induced by the Euclidean metric $\mathrm{d}_{\mathbb{E}}$. In fact we can choose a even smaller basis by considering only open balls with rational radius. Moreover, we can check that the intersection between any two balls is not a ball. To fill in one such intersection, we have to use infinitely many balls.

Example 2.1.25 (Partition of a set).
Let X be any non-empty set. Let $\mathcal{B}$ be a partition of X. We denote by $\mathcal{T}$ the topology on X generated by $\mathcal{B}$. Hence $\mathcal{B}$ is a subbasis of $\mathcal{T}$. Moreover, if we consider the definition of a partition of a set, the condition in Proposition 2.1.23 is satisfied vacuously, since any two subsets in $\mathcal{B}$ have empty intersection. Hence $\mathcal{B}$ is also a basis of $\mathcal{T}$, and any open set in $\mathcal{T}$ is a union of subsets in $\mathcal{B}$.

Neighborhoods and neighborhood bases

In order to study local properties near a point in a space, we introduce the notion of neighborhood and neighborhood basis.

Consider a topological space X.

Definition 2.1.26

For any non-empty subset A of X, a ***neighborhood*** of A is a subset B of X, such that there is an open subset U of X containing A and contained in B:

$$A \subset U \subset B.$$

In particular, for any point $x \in X$, we call a neighborhood of $\{x\}$ a ***neighborhood*** of x.

Remark 2.1.27.
By its definition, a neighborhood of A is not necessary to be open (See Figure 2.1.5). If a neighborhood of A is open (resp. closed) we will call it an *open neighborhood* (resp. *closed neighborhood*).

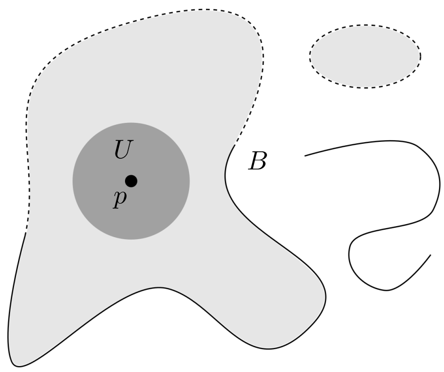


Figure 2.1.5: One neighborhood B of a point $p \in \mathbb{R}^2$, where B is the union of the three parts.

We obtain two immediate properties of neighborhoods from the above definition.

Proposition 2.1.28

Let p be a point in X.

1) If B is an open set in X containing p, then B is a neighborhood of p.

2) If B_1 and B_2 are two neighborhoods of p, then so is their intersection $B_1 \cap B_2$.

Proof. 1) Consider the definition of a neighborhood of p, since B is open, then we have

$$p \in B \subset B.$$

Therefore B is a neighborhood of p.

2) This comes from the fact that the intersection between two open sets is open. By the definition of a neighborhood of p, we have open set U_1 and U_2, such that

$$p \in U_1 \subset B_1, \; p \in U_2 \subset B_2.$$

Hence

$$p \in U_1 \cap U_2 \subset B_1 \cap B_2.$$

Since $U_1 \cap U_2$ is open, the intersection $B_1 \cap B_2$ is a neighborhood of p. □

Remark 2.1.29.
The second property relies on the fact that openness is preserved by finite intersections. Since openness is not necessary preserved by arbitrary intersection, we do not have 2) for the intersection of arbitrarily many neighborhoods of p.

Using the notion of neighborhood, we have a criteria for a subset to be open in X.

Proposition 2.1.30

A subset A of X is open if and only if for any $x \in A$, the set A is a neighborhood of x.

Proof. By the previous proposition, if A is open, then it is a neighborhood of any of its points.

Conversely, if A is a neighborhood of any $x \in A$, then by the definition of the neighborhood, for any $x \in A$, there is a open set U_x, such that

$$x \in U_x \subset A.$$

Hence we have

$$A = \bigcup_{x \in A} U_x.$$

Therefore, the set A is open. □

Remark 2.1.31.
If we recall the discussion on the open sets in an Euclidean space (See Example 2.1.7), by that definition, open balls are also open. Then similar to the above proposition, we have a subset of the Euclidean space is open if and only if any point admits a ball neighborhood contained in this subset.

Let $\mathcal{N}(x)$ denote the collection of all neighborhoods of x in X.

Definition 2.1.32

A subset $\mathcal{B} \subset \mathcal{N}(x)$ is called a ***neighborhood basis*** of x if it satisfies the following property

$$\forall U \in \mathcal{N}(x),\ \exists B \in \mathcal{B},\ B \subset U.$$

Example 2.1.33.
We consider the Euclidean space $\mathbb{R}^2$. Consider a point $p \in \mathbb{R}^2$. Let $(r_n)_{n \geq 0}$ be a sequence of positive real numbers such that

$$\lim_{n \to \infty} r_n = 0.$$

Then

$$\{B_n = B_p(r_n) \mid n \in \mathbb{N}\},$$

is a neighborhood basis of p (See Figure 2.1.6).

Example 2.1.34.
We consider another example which may look strange at the first glance. Let X be a non-empty

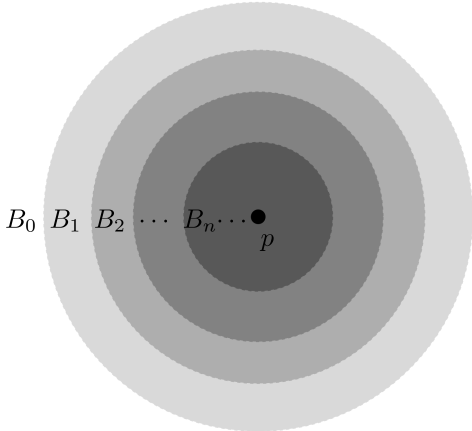


Figure 2.1.6: One neighborhood basis of $p \in \mathbb{R}^2$

set. We consider the discrete topology. Then for any $x \in X$, the single point set $\{x\}$ itself can form a neighborhood basis of x.

Limit points and limit values

Let us first recall what is the limit of a sequence in $\mathbb{R}$ that we saw in the Analysis course. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ converging to a:

$$\lim_{n\to\infty} x_n = a,$$

i.e. for any $\epsilon > 0$, there is $N \in \mathbb{N}$, such that for any natural number $n > N$, we have

$$|x_n - a| < \epsilon.$$

From the topological point of view, we consider the usual topology on $\mathbb{R}$ induced by the Euclidean metric, the following subsets of $\mathbb{R}$

$$\{(x-\epsilon, x+\epsilon) \mid \epsilon \in \mathbb{R}_{>0}\},$$

form a neighborhood basis of x. The limiting condition above is then written as for any neighborhood $(x-\epsilon, x+\epsilon)$ of x, there is a natural number N, such that for any $n > N$, we have

$$x_n \in (x-\epsilon, x+\epsilon).$$

See Figure 2.1.2 for an illustration.

Using the limit of a sequence, we can define the continuous function from $\mathbb{R}$ to itself. Let

$$f : \mathbb{R} \to \mathbb{R},$$

be a function. Let a be a point in $\mathbb{R}$. In Analysis course, we say that f is continuous at a if

$$\lim_{x\to a} f(x) = f(a),$$

which means that whatever the sequence (x_n) converging to a is, we have

$$\lim_{n\to\infty} f(x_n) = f(a).$$

Another way to say it is that

$$\forall \epsilon > 0,\, \exists \delta > 0,\, \forall x \in (a-\delta, a+\delta),\, f(x) \in (f(a)-\epsilon, f(a)+\epsilon).$$

From the topological point of view, here we consider a neighborhood basis of a

$$\mathcal{B} = \{(a-\delta, a+\delta) \mid \delta \in \mathbb{R}_{>0}\},$$

and a neighborhood basis of $f(a)$

$$\mathcal{C} = \{(f(a)-\epsilon, f(a)+\epsilon) \mid \epsilon \in \mathbb{R}_{>0}\},$$

such that

$$\forall V \in \mathcal{C},\, \exists U \in \mathcal{B},\, f(U) \subset V.$$

Figure 2.1.7 is an illustration

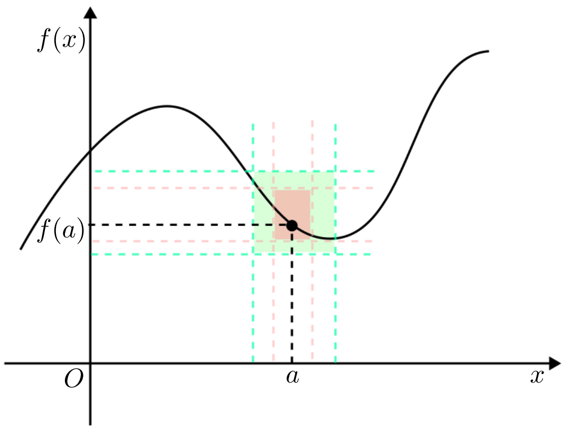


Figure 2.1.7: The function $f(x)$ continuous at a.

Following this idea, we have the following topological definition of the convergence of a sequence.

Definition 2.1.35

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a topological space X. We say that the sequence $(x_n)_{n\in\mathbb{N}}$ ***converges to a point*** $a \in X$, if for any neighborhood U of a, there exists $N \in \mathbb{N}$, such that for any $n > N$, we have $x_n \in U$.

Similar, we have a topological definition of a map continuous at a point as follows.

Definition 2.1.36

Let f be a map from a topological space X to a topological space Y. We say that f is ***continuous at*** $a \in X$, if for any neighborhood V of $f(a)$, there exists a neighborhood U of a, such that

$$f(U) \subset V.$$

Remark 2.1.37.
Notice that in the ϵ-δ language, the quantities ϵ and δ are used to find these neighborhoods in the above definitions.

These definition can also be restricted to some subset of a topological space.

Definition 2.1.38

Let A be a subset of a topological space X. A point $x \in X$ is said to be an ***adherent point*** of A if for any neighborhood U of x, we have

$$U \cap A \neq \emptyset.$$

Let f be a map from a topological space X to a topological space Y. Let A be a subset of X and a be an adherent point of A. We say that f admits a ***limit value*** $y \in Y$ ***when*** x ***tends to*** a ***in*** A, if for any neighborhood V of y and any neighborhood U of a, we have

$$f(U \cap A) \cap V \neq \emptyset.$$

We say that $f(x)$ ***admits a limit when*** x ***tends to*** a ***in*** A, if there exists $y \in Y$, such that for any neighborhood V of y, we have a neighborhood U of a,

$$f(U \cap A) \subset V.$$

Remark 2.1.39.
We will discuss the uniqueness of a limit in the next part.

Hausdorff condition

In the study in the Euclidean plane $\mathbb{R}^2$, the convergence of a sequence in $\mathbb{R}^2$ is one thing which we discuss a lot. One fact used a lot is that the limit is unique for any convergence sequence in $\mathbb{R}^2$. This relates to the property of $\mathbb{R}^2$ used here is that any two distinct points admit disjoint neighborhoods. For example, if two points in $\mathbb{R}^2$ have distance $r > 0$, then we may choose a ball neighborhood of radius $r' < r/2$ for each one of them, which are disjoint.

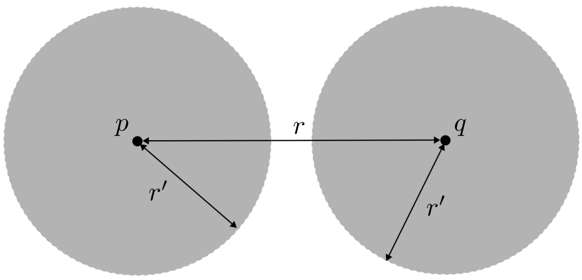


Figure 2.1.8: Disjoint neighborhoods of two points p and q in $\mathbb{R}^2$.

This can guarantee that there is no ambiguity of the limit point. However, such a property does not hold for any topological space.

Definition 2.1.40

We say that a topological space X satisfies the ***Hausdorff condition*** if any distinct points x and y in X admit disjoint neighborhoods, i.e. there exist a neighborhood U of x and a neighborhood V of y such that

$$U \cap V = \emptyset.$$

If the space X satisfies the Hausdorff condition, we say that X is ***Hausdorff***.

By this definition, the Euclidean plane, and more generally all metric spaces are Hausdorff. Now let us check some non-Hausdorff spaces.

Example 2.1.41.
Consider the union

$$\{0_-\} \cup \{0_+\} \cup (0, 1).$$

For any $x \in]0, 1]$, it has a neighborhood basis

$$\mathcal{B}(x) = \left\{ \left(x - \frac{1}{n}, x + \frac{1}{n} \right) \cap (0, 1) \,\middle|\, n \in \mathbb{N}^* \right\}.$$

A neighborhood basis of 0_- can be given by

$$\mathcal{B}(0_-) = \left\{ \{0_-\} \cup \left(0, \frac{1}{n}\right) \,\middle|\, n \in \mathbb{N}^* \right\},$$

and similarly a neighborhood basis of 0_+ can be given by

$$\mathcal{B}(0_+) = \left\{ \{0_+\} \cup \left(0, \frac{1}{n}\right) \,\middle|\, n \in \mathbb{N}^* \right\}.$$

Then we may see that it is impossible to separate 0_- and 0_+ with disjoint neighborhood.

The above example may be a little bit artificial. The following one appears a lot in the study of character varieties.

Example 2.1.42.
Consider the special linear group $\mathrm{SL}(2, \mathbb{R})$. Let χ denote the set of conjugacy classes of elements in $\mathrm{SL}(2, \mathbb{R})$. The group $\mathrm{SL}(2, \mathbb{R})$ can be considered as part of $\mathbb{R}^4$. We consider the restriction of the Euclidean metric $\mathrm{d}_{\mathbb{E}}$ on $\mathbb{R}^4$ to $\mathrm{SL}(2, \mathbb{R})$ and give a topology on $\mathrm{SL}(2, \mathbb{R})$.

Given two subsets A and B of $\mathrm{SL}(2, \mathbb{R})$, we can define the distance between them as

$$\mathrm{d}_{\mathbb{E}}(A, B) := \inf\{\mathrm{d}_{\mathbb{E}}(x, y) \mid x \in A, y \in B\}.$$

There is a natural projection from $\mathrm{SL}(2, \mathbb{R})$ to χ. By defining a subset in χ is open if it is the image of an open set of $\mathrm{SL}(2, \mathbb{R})$, we have a topology on χ (we will talk about this construction later in details).

The non-Hausdorff phenomenon appears when A and B are distinct elements in χ with $\mathrm{d}_{\mathbb{E}}(A, B) = 0$, then there is no way that we can separate A and B by disjoint neighborhoods in χ.

Such A and B do exist in χ. Consider the following two matrices

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Notice that in the conjugacy class of x, we have the elements of following form

$$\begin{bmatrix} 1 & t^{-2} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} t^{-1} & 0 \\ 0 & t \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} t & 0 \\ 0 & t^{-1} \end{bmatrix}$$

for any $t \neq 0$. As t goes to $+\infty$, we have

$$\begin{bmatrix} 1 & t^{-2} \\ 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Now we turn to the relation between the uniqueness of convergence limit and the Hausdorff condition.

Proposition 2.1.43

If the topological space X is Hausdorff, then given any convergent sequence $(x_n)_{n\in\mathbb{N}}$ in X, its limit is unique.

Proof. We prove it by contradiction. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. Assume that a and b be two distinct limits of $(x_n)_{n\in\mathbb{N}}$ in X. Since X is Hausdorff, there are U and V neighborhoods of a and b respectively, such that

$$U \cap V = \emptyset.$$

On the other hand, since a is a limit of $(x_n)_{n\in\mathbb{N}}$, there exists $N_x \in \mathbb{N}$, such that for any $n > N_x$, we have

$$x_n \in U.$$

Similarly, since b is a limit of $(x_n)_{n\in\mathbb{N}}$, there exists $N_y \in \mathbb{N}$, such that for any $n > N_y$, we have

$$x_n \in U.$$

Let $N = \max\{N_x, N_y\}$, then for any $n > N$, we have

$$x_n \in U \cap V,$$

which contradicts to the fact that $U \cap V$ is empty. □

Interior, closure and boundary

Let X be a topological space and A be one of its subset.

Definition 2.1.44

The ***interior*** of A is defined as:

$$\mathring{A} := \{x \in A \mid \exists U \text{ neighborhood of } x, U \subset A\}.$$

The ***closure*** of A denoted by $\overline{A}$ is the set of all adherent points of A, i.e.

$$\overline{A} := \{x \in X \mid \forall \text{ neighborhood } U \text{ of } x,\ U \cap A \neq \emptyset\}.$$

The ***boundary*** of A is defined to be the subset

$$\partial A := \overline{A} \setminus \mathring{A}.$$

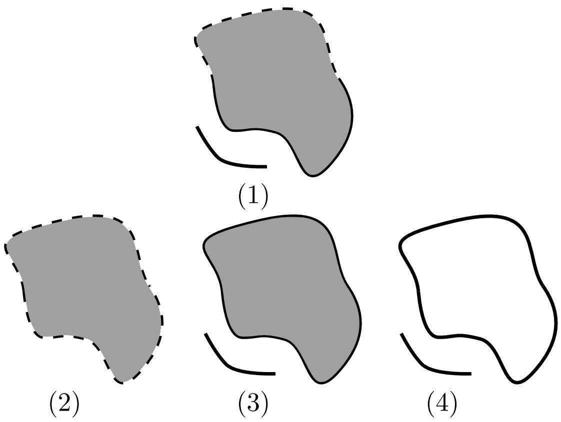


Figure 2.1.9: (1) subset A of $\mathbb{R}^2$; (2) the interior $\mathring{A}$ of A; (3) the closure $\overline{A}$ of A; (4) the boundary ∂A of A.

Example 2.1.45.
See Figure 2.1.9 for an illustration of the above definition for a subset A in $\mathbb{R}^2$. Here the dash line means those boundary points are not in A.

There are certain facts which can be verified directly using the definitions.

Proposition 2.1.46

For any subset A of X, we have

1) the interior of A is the union of all open subsets of X which are contained in A;
2) the closure of A is the intersection of all closed subsets of X which contain A;
3) the boundary of A is the intersection between $\overline{A}$ and $\overline{A^c}$.

Proof. To show the first point, it is enough to prove the following two facts:

- $\mathring{A}$ is open;
- any open subset of A is contained in $\mathring{A}$.

For any $x \in \mathring{A}$, by the definition of interior and the definition of neighborhood, there is an open neighborhood U of x contained in A. By Proposition 2.1.30, since U is open, it is also a neighborhood of any point $y \in U$. Hence every $y \in U$ is also a point in $\mathring{A}$, which implies

$$U \subset \mathring{A}.$$

Hence $\mathring{A}$ is a neighborhood of x. Since this holds for any $x \in \mathring{A}$, we have $\mathring{A}$ open.

Let B be an open subset of X contained in A. By Proposition 2.1.30, for any $x \in B$, the open set B is a neighborhood of x which is contained in A by hypothesis. Therefore, we have

$$B \subset \mathring{A},$$

and it holds for any open subset of A, including $\mathring{A}$. Hence $\mathring{A}$ is the union of all open subsets of A (in another word, $\mathring{A}$ is the largest open subset of A).

To show the second point, it is enough to prove the following two facts:

- the complement of $\overline{A}$ is open;
- any open subset of X disjoint from A is contained in $(\overline{A})^c$.

If $x \in X$ is not a limit point of A, there is an open neighborhood U of x with

$$U \cap A = \emptyset.$$

Moreover for any $y \in U$, the set U is a neighborhood of y, hence

$$y \notin \overline{A}.$$

Hence

$$U \subset (\overline{A})^c.$$

Therefore $(\overline{A})^c$ is open and $\overline{A}$ is closed.

Let B is a closed subset in X containing A, then B^c is open and disjoint from A. For any point $x \in B^c$, since B^c is open, hence it is a neighborhood of x. Hence x is not a limit point of A, therefore $x \notin \overline{A}$. Hence we have

$$B^c \cap \overline{A} = \emptyset.$$

This implies that

$$\overline{A} \subset B.$$

From the above discussion, we conclude that $\overline{A}$ is the intersection of all closed set in X containing A (in another word, $\overline{A}$ is the smallest closed subset containing A).

To show the third point, let $x \in \partial A$, by the definition of the boundary, we have

$$x \in \overline{A}, \quad x \notin \mathring{A}.$$

By the definition of $\mathring{A}$, given any neighborhood U of x, we have

$$U \cap A^c \neq \emptyset,$$

for otherwise, if $U \cap A^c = \emptyset$, we have $U \subset A$, which means that $x \in \mathring{A}$. This is a contradiction. Hence we have

$$x \in \overline{A^c},$$

and then

$$\partial A \subset \overline{A} \cap \overline{A^c}.$$

Conversely, let $x \in \overline{A} \cap \overline{A^c}$, for any neighborhood U of x, we have

$$U \cap A \neq \emptyset, \quad U \cap A^c \neq \emptyset.$$

Hence U is not a subset of A, hence $x \notin \mathring{A}$, and we have

$$x \in \overline{A} \setminus \mathring{A} = \partial A.$$

Hence we have

$$\overline{A} \cap \overline{A^c} \subset \partial A.$$

As a conclusion, we have

$$\overline{A} \cap \overline{A^c} = \partial A.$$

□

If we consider taking the interior or the closure of a map as maps from $\mathcal{P}(X)$ to $\mathcal{P}(X)$, their relations with taking the union or the intersection of two subsets, or taking the compliment of a subset are as follows.

Proposition 2.1.47

If A and B are both subsets of X, we have

1) $\mathring{A} \cup \mathring{B} \subset \widehat{A \cup B}^{\circ}$, $\quad \overline{A \cup B} = \overline{A} \cup \overline{B}$;

2) $\widehat{A \cap B}^{\circ} = \mathring{A} \cap \mathring{B}$, $\quad \overline{A \cap B} \subset \overline{A} \cap \overline{B}$;

3) $\widehat{A^c}^{\circ} = \overline{A}^c$, $\quad \overline{A^c} = \left(\mathring{A}\right)^c$, $\quad \partial A = \partial A^c$.

Proof. 1) For any $p \in \mathring{A} \cup \mathring{B}$, we have $p \in \mathring{A}$ or $p \in \mathring{B}$. Without loss of generality, we assume $p \in \mathring{A}$. Then there is a neighborhood U of p, such that

$$p \in U \subset A \subset A \cup B.$$

Hence we have

$$p \in \widehat{A \cup B}^{\circ}$$

The other direction is not correct. For example, we consider $A = (0,1) \cup \{5\}$ and $B = (4,5) \cup (5,6)$. Consider the Euclidean metric topology on $\mathbb{R}$, then 5 is an interior point for $A \cup B$, but

$$\mathring{A} = (0,1), \ \ \mathring{B} = B,$$

which shows that $5 \notin \mathring{A} \cup \mathring{B}$.

Let $p \in \overline{A \cup B}$. For any neighborhood U of p, we have

$$U \cap (A \cup B) \neq \emptyset.$$

Hence we have $U \cap A \neq \emptyset$ or $U \cap B \neq \emptyset$.

If all neighborhoods of p have non empty intersection with A, we have

$$p \in \overline{A}.$$

Otherwise, there is a neighborhood U of p disjoint from A. We claim that all neighborhoods of p have non empty intersection with B. For otherwise, there is a neighborhood V of p disjoint from B. Notice that $U \cap V$ is also a neighborhood of p, and we have

$$(U \cap V) \cap A = \emptyset = (U \cap V) \cap B.$$

This contradicts to the fact that $p \in \overline{A \cup B}$. Therefore, we have

$$\overline{A \cup B} \subset \overline{A} \cup \overline{B}.$$

Conversely, if $p \in \overline{A} \cup \overline{B}$, without loss of generality we may assume that $p \in \overline{A}$, then given any neighborhood U of p, we have

$$U \cap A \neq \emptyset.$$

this implies that

$$U \cap (A \cup B) \neq \emptyset.$$

Therefore we have $p \in \overline{A \cup B}$, and moreover

$$\overline{A \cup B} \supset \overline{A} \cup \overline{B}.$$

As a conclusion, we have

$$\overline{A \cup B} = \overline{A} \cup \overline{B}.$$

For any point in $p \in \overset{\circ}{\widehat{A \cap B}}$, there is a neighborhood U of p, such that

$$U \subset A \cap B,$$

hence

$$U \subset A \text{ and } U \subset B.$$

This implies that

$$p \in \mathring{A} \text{ and } p \in \mathring{B}.$$

from which we have

$$p \in \mathring{A} \cap \mathring{B}.$$

Hence

$$\overset{\circ}{\widehat{A \cap B}} \subset \mathring{A} \cap \mathring{B}.$$

Conversely, if $p \in \mathring{A} \cap \mathring{B}$, there are two neighborhoods U_A and U_B of p, such that

$$U_A \subset A \text{ and } U_B \subset B.$$

Hence

$$U_a \cap U_B \in A \cap B.$$

By Proposition 2.1.28, the intersection $U_A \cap U_B$ is again a neighborhood of p, hence

$$p \in \overset{\circ}{\widehat{A \cap B}}.$$

Therefore we have

$$\overset{\circ}{\widehat{A \cap B}} \supset \mathring{A} \cap \mathring{B}.$$

We conclude now that

$$\overset{\circ}{\widehat{A \cap B}} = \mathring{A} \cap \mathring{B}.$$

Now we consider the closure of A, B and $A \cap B$. If $p \in \overline{A \cap B}$, then given any neighborhood U of p, we have

$$U \cap (A \cap B) \neq \emptyset.$$

Equivalently, we have

$$U \cap A \neq \emptyset \text{ and } U \cap B \neq \emptyset$$

Hence

$$p \in \overline{A} \text{ and } p \in \overline{B}.$$

Therefore, we have

$$\overline{A \cap B} \subset \overline{A} \cap \overline{B}.$$

The other direction of inclusion is not correct. For example, we consider $A = (0, 1)$ and $B = (1, 2)$. Then we have

$$\overline{A} = [0, 1] \text{ and } \overline{B} = [1, 2].$$

Hence

$$\overline{A} \cap \overline{B} = \{1\},$$

while $A \cap B = \emptyset$ which implies that $\overline{A \cap B} = \emptyset$.

3) A point x is in $\widehat{\mathring{A^c}}$, if and only if there is a neighborhood U of x such that

$$U \subset A^c.$$

Hence $U \cap A = \emptyset$ which is equivalent to $x \notin \overline{A}$. This shows that

$$\widehat{\mathring{A^c}} = \overline{A}^c.$$

A point x is in $\overline{A^c}$, if and only if for any neighborhood U of x, we have

$$U \cap A^c \neq \emptyset.$$

This is equivalent to the fact that given any neighborhood U of x, we have

$$U \not\subseteq A,$$

which is equivalent to

$$x \notin \mathring{A}.$$

Hence we have

$$\overline{A^c} = \left(\mathring{A}\right)^c.$$

By the point 3) in Proposition 2.1.46, we have

$$\partial A = \overline{A} \cap \overline{A^c}.$$

Replacing A by A^c, we have

$$\partial A^c = \overline{A^c} \cap \overline{(A^c)^c} = \overline{A^c} \cap \overline{A} = \partial A.$$

□

2.2 Continuity

Let X and Y be two topological spaces. To build the connection between the two spaces in a topological way, we use so called continuous maps.

Definition 2.2.1

A map

$$f : X \to Y,$$

is said to be ***continuous*** if it satisfies one of the following equivalent conditions:

1) for any open set $V \subset Y$, its preimage $f^{-1}(V)$ is an open set in X;
2) for any closed set $V \subset Y$, its preimage $f^{-1}(V)$ is a closed set in X;
3) for any set $A \subset X$, we have $f(\overline{A}) \subset \overline{f(A)}$;
4) for any point $x \in X$, for any neighborhood V of $f(x) \in Y$, there is a neighborhood U of x, such that $f(U) \subset V$.

Remark 2.2.2.
From Definition 2.1.36, the last condition means that the map f is continuous at every point $x \in X$

Consider Definition 2.1.32, we may replace the last condition by an equivalent one

5) for any point $x \in X$, there is a neighborhood basis $\mathcal{B}$ of x and a neighborhood basis $\mathcal{C}$ of $f(x)$, for any neighborhood $V \in \mathcal{C}$, there is $U \in \mathcal{B}$, such that $f(U) \subset V$.

Proposition 2.2.3

All conditions listed in the definitions are equivalent to each other.

Proof. 1) $\iff$ 2) This comes from the fact that for any subset V in Y

$$(f^{-1}(V))^c = f^{-1}(V^c).$$

4) $\Longrightarrow$ 3) Let A be a subset of X. Assume that for any $x \in A$, any neighborhood V_x of $f(x)$, we have a neighborhood U_x of x, such that

$$f(U_x) \subset V_x.$$

Then for any neighborhood V_x of $f(x)$, we have

$$V_x \cap f(A) \supset f(U_x) \cap f(A) \neq \emptyset.$$

Hence $f(x) \in \overline{f(A)}$. We have

$$f(\overline{A}) \subset \overline{f(A)}.$$

3) $\Longrightarrow$ 1) Assume that for any set A, we have

$$f(\overline{A}) \subset \overline{f(A)}.$$

Let V be an open set in Y. Then its complement is closed, hence

$$\overline{V^c} = V^c.$$

Let $K = f^{-1}(V^c)$. We have

$$f(\overline{K}) \subset \overline{f(K)} = \overline{V^c} = V^c.$$

Let U denote $f^{-1}(V)$. Notice that $K = U^c$.

For any $x \in U$, if any neighborhood W of x satisfies

$$W \cap U^c \neq \emptyset,$$

we have

$$x \in \overline{K}.$$

By the previous discussion, we have

$$f(x) \in \overline{f(K)} = V^c.$$

Hence

$$x \notin f^{-1}(V) = U,$$

which is a contradiction. Therefore for any $x \in U$, there is a neighborhood W of x contained in U, hence U is also a neighborhood of x. This means U is open.

4) $\Longrightarrow$ 1) Let V be a open set of Y. For any $x \in X$, such that $f(x) \in V$, the set V is a neighborhood of $f(x)$, hence there is a neighborhood U_x of x, such that

$$f(U_x) \subset V,$$

Hence we have
$$U_x \subset f^{-1}(V).$$
By the definition of neighborhood, there is an open neighborhood U'_x of x, such that
$$U'_x \subset U_x \subset f^{-1}(V).$$
Hence we have
$$f^{-1}(V) = \bigcup_{x \in X, f(x) \in V} U'_x.$$
Therefore $f^{-1}(V)$ is open in X.

1) $\Longrightarrow$ 4) Assume that given any $x \in X$ for any neighborhood V of $f(x)$, there is an open neighborhood V_x of $f(x)$ contained in V. By 1), $f^{-1}(V_x)$ is open, hence a neighborhood of x. Let $U_x = f^{-1}(V_x)$, we then have
$$f(U_x) \subset V_x \subset V.$$
□

As mentioned in the previous in Example 2.1.10, any map from X to Y can induce a topology on X by considering the preimages of open sets in Y.

Consider a map
$$f : X \to Y.$$
We denote by $\mathcal{T}_X$ the topology on X and by $\mathcal{T}_f$ the topology induced by f.

Proposition 2.2.4

The map f is continuous if and only if the topology $\mathcal{T}_X$ is finer than $\mathcal{T}_f$.

Proof. Let $\mathcal{T}_Y$ denote the topology on Y. By the definition of $\mathcal{T}_f$, for any $U \in \mathcal{T}_f$, there is an open set $V \in \mathcal{T}_Y$, such that
$$U = f^{-1}(V).$$

By definition, the map f is continuous with respect to the topology $\mathcal{T}_X$ on X, if and only if for any $V \in \mathcal{T}_Y$, we have
$$f^{-1}(V) \in \mathcal{T}_X,$$
this is equivalent to
$$\mathcal{T}_f \subset \mathcal{T}_X,$$
i.e. $\mathcal{T}_X$ is finer than $\mathcal{T}_f$. □

Definition 2.2.5

A map
$$f : X \to Y,$$
is called a ***homeomorphism*** if it is continuous and bijective, and its inverse
$$f^{-1} : Y \to X,$$
is also continuous.

Two topological space are said to be ***homeomorphic*** if there is a homeomorphism between them.

Now assume that the map

$$f : X \to Y,$$

is bijective.

If f is continuous, by taking preimage and considering Proposition2.2.4, it induces a map

$$f^* : \mathcal{T}_Y \to \mathcal{T}_X,$$

whose image is $\mathcal{T}_f$. If f^{-1} is continuous, then by the same reason, we have a map

$$(f^{-1})^* : \mathcal{T}_X \to \mathcal{T}_Y,$$

whose image is $\mathcal{T}_{f^{-1}}$.

Assume that f is a homeomorphism. Since $f \circ f^{-1}$ is identity map, for any $V \in \mathcal{T}_Y$, we have

$$V = (f \circ f^{-1})(V) = f(f^*(V)) = (f^{-1})^*(f^*(V)),$$

hence it induces the identity map

$$(f^{-1})^* \circ f^* : \mathcal{T}_Y \to \mathcal{T}_Y.$$

Similarly, the composition

$$f^* \circ (f^{-1})^* : \mathcal{T}_X \to \mathcal{T}_X$$

is also the identity map. Therefore, both map f^* and $(f^{-1})^*$ are bijective, and we have

$$\mathcal{T}_X = \mathcal{T}_f, \quad \mathcal{T}_Y = \mathcal{T}_{f^{-1}}$$

In fact, we could have the following proposition.

Proposition 2.2.6

Let f be a bijective map from a topological space $(X, \mathcal{T}_X)$ to a topological space $(Y, \mathcal{T}_Y)$. Then f is a homeomorphism if and only if $\mathcal{T}_X = \mathcal{T}_f$.

Proof. If f is a homeomorphism, the above discussion shows that $\mathcal{T}_X = \mathcal{T}_f$.

Now let us assume that $\mathcal{T}_X = \mathcal{T}_f$. We would like to show that the bijective map f and its inverse are both continuous.

By Proposition 2.2.4, we have f continuous. To study f^{-1}, notice that the fact that f is bijective induces a bijective map

$$\begin{aligned} f : \mathcal{P}(X) &\to \mathcal{P}(Y), \\ A &\mapsto f(A). \end{aligned}$$

Its has an inverse

$$\begin{aligned} f^{-1} : \mathcal{P}(Y) &\to \mathcal{P}(X), \\ B &\mapsto f^{-1}(B). \end{aligned}$$

which is also bijective.

Let $U \in \mathcal{T}_X$ be any open set of X. Since $\mathcal{T}_X = \mathcal{T}_f$, there is an open set $V \in \mathcal{T}_Y$, such that

$$f^{-1}(V) = U.$$

Since f^{-1} is bijective and $f(U)$ satisfies

$$f^{-1}(f(U)) = U,$$

we have $V = f(U)$. Therefore, we have

$$(f^{-1})^{-1}(U) = f(U) = V$$

open for any $U \in \mathcal{T}_X$. The map f^{-1} is continuous.

As a conclusion, the map f is a homeomorphism. □

Remark 2.2.7.
All these discussions shows that if two topological spaces are homeomorphic to each other, from the topological point of view (only playing with open sets), we cannot distinguish them. We will see this phenomenon in a more concrete way later.

Remark 2.2.8.
The discussion here may reminds us the example to show that a continuous bijection is not necessary a homeomorphism. Let us check what happens here and compare it with the above proposition.

Let X be the interval $[0, 1)$ and Y be the unit circle S^1 in the Euclidean plane $\mathbb{R}^2$. We define the map

$$\begin{aligned} f : [0, 1) &\to S^1, \\ t &\mapsto (\cos 2t\pi, \sin 2t\pi). \end{aligned}$$

(See Figure 2.2.1).

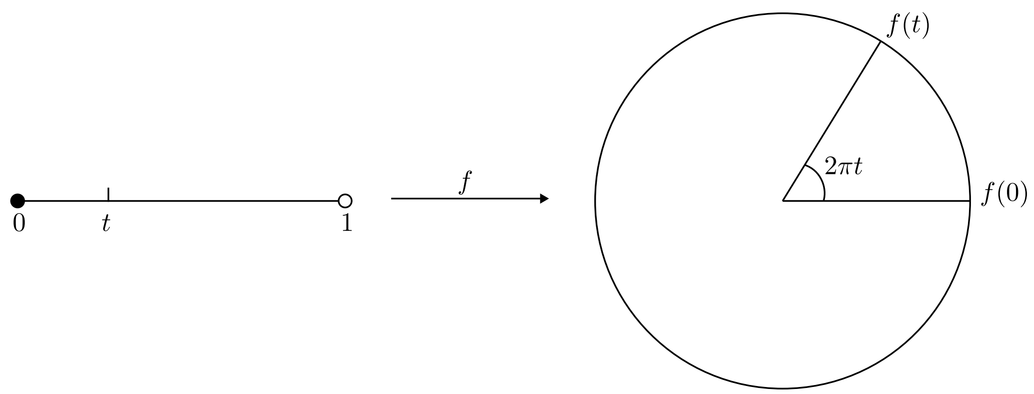


Figure 2.2.1: The map f

We consider the subspace topology (which will be talk about later). The topology in $[0, 1)$ considered here is generated by its intersections open intervals in $\mathbb{R}$, while the topology in S^1 considered here is generated its intersection with Euclidean open balls in $\mathbb{R}$.

Notice that this is a continuous map and bijective, but f^{-1} is not continuous. The problem appears when we consider the neighborhoods of 0. The interval $[0, 1/2)$ is open in $[0, 1)$, but its preimage is not in S^1.

If we would like to compare $\mathcal{T}_X$ and $\mathcal{T}_f$, we will see that a neighborhood basis of 0 in $\mathcal{T}_X$ can be given by

$$\left\{ \left[0, \frac{1}{n}\right) \middle| n \in \mathbb{N} \setminus \{0, 1\} \right\},$$

while a neighborhood basis of 0 in $\mathcal{T}_f$ can be given by

$$\left\{ \left[0, \frac{1}{n}\right) \cup \left(\frac{n-1}{n}, 1\right) \middle| n \in \mathbb{N} \setminus \{0, 1\} \right\}.$$

(See Figure 2.2.2)

Hence for any $n \in \mathbb{N} \setminus \{0, 1\}$, the subset

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

Figure 2.2.2: A neighborhood basis of 0 for $\mathcal{T}_X$ (Left); a neighborhood basis of 0 for $\mathcal{T}_f$ (Right)

is never a neighborhood of 0 for $\mathcal{T}_f$. We can verify that

$$\mathcal{T} \supsetneq \mathcal{T}_f.$$

2.3 Constructions of topologies

There are several ways which we usually use to construct topological spaces from a set possibly equipped with some (geometric, algebraic, topological, etc.) structures.

Subspace topology

Let X be a topological space. Given any non-empty subset $A \subset X$, we can define a topology on A by considering the topology on X in the following way:

- a subset of A is open if and only if it can be written as an intersection $A \cap U$, where U is open in X.

Definition 2.3.1

This topology on A is called the ***subspace topology***, and the subset A equipped with the subspace topology is called a ***topological subspace*** of X (or simply a ***subspace*** of X).

The following is an immediate consequence of the definition.

Proposition 2.3.2

Let A be a subset of X and consider the subspace topology on A.

1) Let $\mathcal{B}$ be a basis (resp. subbasis) of the topology on X, then

$$\{A \cap U \mid U \in \mathcal{B}\}$$

form a basis (resp. subbasis) of the topology on A.

2) Let p be a point in A and $\mathcal{B}_p$ be a neighborhood basis of p in X, then

$$\{A \cap U \mid U \in \mathcal{B}_p\}$$

is a neighborhood basis of p in A equipped with the sub

Example 2.3.3 (Circles in $\mathbb{R}^2$).
Denote by C a circle in the Euclidean plane $\mathbb{R}^2$. Consider the Euclidean metric topology on $\mathbb{R}^2$, and the induced subspace topology on C. Since open balls form a basis of the topology on $\mathbb{R}^2$, their intersections with C, which are open circular arcs, form a basis of the topology on C (See Figure 2.3.1 for an illustration).

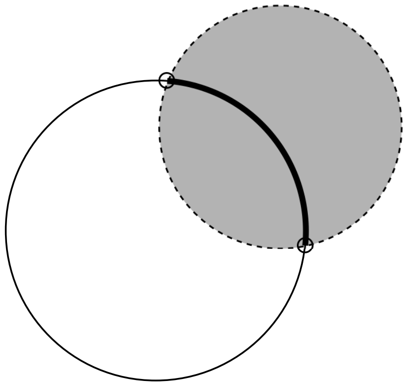


Figure 2.3.1: An open circular arc on a circle

Example 2.3.4 (Tangential circles in $\mathbb{R}^2$).
For any $n \in \mathbb{N}^*$, we consider the circle C_n in $\mathbb{R}^2$ centered at $(1/n, 0)$ of radius $1/n$. Let X be the union

$$X = \bigcup_{n \in \mathbb{N}^*} C_n.$$

This is usually called the Hawaii earring.

We consider the subspace topology on X induced by the Euclidean metric topology on $\mathbb{R}^2$. Let $p \in X$ be a point different from 0. There is a unique circle C_n, such that $p \in C_n$. The local picture around p in X would be the same as that around p in C_n.

The difference appears when we consider the neighborhood of the origin O. A neighborhood basis of O in $\mathbb{R}^2$ can be given by open disks. Let D be an open disk centered at O of radius $r > 0$. Then for any $n \in \mathbb{N}^*$ such that $n^{-1} < r$, we have

$$C_n \subset D.$$

(See Figure 2.3.2 for an illustration.)

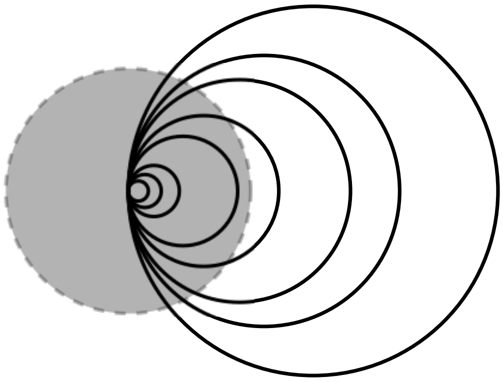


Figure 2.3.2: Any neighborhood of O contains C_n from some $N \in \mathbb{N}^*$

Hence, by taking a neighborhood of O in each circle C_n and taking the union of them, we do not necessary get a neighborhood of O (See Example 2.3.24 "Wedge Sum").

Remark 2.3.5.

If A and B are both subsets of X, with

$$A \subset B \subset X,$$

Then A has a subspace topology $\mathcal{T}_{A,X}$ induced by the topology of X. At the same time, the set A also has a subspace topology $\mathcal{T}_{A,B}$ induced by the subspace topology of B which is a subspace of X. Since $A \subset B$, given any open set U of X, we have

$$A \cap U = (A \cap B) \cap U = A \cap (B \cap U).$$

Hence $\mathcal{T}_{A,X} = \mathcal{T}_{A,B}$.

Initial topology

Let X be a non-empty set, and Y be a topological space. As in Example 2.1.10, for any map

$$f : X \to Y,$$

we can associate to X a topology defined as follows:

- a subset of X is open if and only if it is the preimage of an open set of Y.

By Proposition 2.2.4, this topology is the coarsest topology on X with respect to which f is continuous.

This construction can be done in a more general setting. We consider the following set

$$\Pi := \{((Y_\alpha, \mathcal{T}_\alpha), f_\alpha) \mid \alpha \in \Omega\},$$

where Ω is the index set, and for each $\alpha \in \Omega$, $(Y_\alpha, \mathcal{T}_\alpha)$ is a topological space and f_α is a map from X to Y_α. We then consider the following set

$$\mathcal{A} = \{U \in \mathcal{P}(X) \mid \exists \alpha \in \Omega,\, \exists V \in \mathcal{T}_\alpha,\, U = f_\alpha^{-1}(V)\},$$

as a subbasis and denote by $\mathcal{I}$ the topology on X generated by $\mathcal{A}$.

Definition 2.3.6

The topology $\mathcal{I}$ constructed above is called the ***initial topology*** on X induced by $(f_\alpha)_{\alpha \in \Omega}$.

Proposition 2.3.7

If $\mathcal{T}$ is a topology on X, and for any $\alpha \in \Omega$, f_α is continuous with respect to $\mathcal{T}$, then $\mathcal{T}$ is finer than $\mathcal{I}$.

Proof. For any $\alpha \in \Omega$, since f_α is continuous with respect to $\mathcal{T}$, for any V open set in Y_α, we have the preimage

$$U = f_\alpha^{-1}(V) \in \mathcal{T}.$$

Hence we have $\mathcal{A} \subset \mathcal{T}$, which implies $\mathcal{I} \subset \mathcal{T}$. □

Example 2.3.8 (Subspace topology).

The subspace topology is a special case of the initial topology. Let A be a non-empty subspace of a topological space X. We consider the embedding map

$$\iota : A \hookrightarrow X,$$

then the subspace topology on A is the initial topology induced by ι.

Example 2.3.9 (Weak topology and Weak-* topology).
Let $(E, \|\cdot\|)$ be a real normed vector space. Notice that $\|\cdot\|$ induces a distance on E, which moreover induces a topology on E which we usually called the *strong topology*. A map from E to a topological space is said to be *strongly continuous* if it is continuous with respect to the strong topology on E.

The dual space E^* of E is then defined to be the space of linearly strongly continuous maps from E to $\mathbb{R}$ (or linear functionals, linear forms). The initial topology induced by

$$\{l \mid l \in E^*\},$$

is called the *weak topology* on E.

Reciprocally, the space E can be considered as part of the dual space of E^*. The initial topology induced by

$$\{x \mid x \in E\},$$

is called the *weak-$*$ topology* on E^*.

Remark 2.3.10.
The term "weak topology" may be used in a more general sense, sometimes considered as the same as initial topology in some references.

How to describe an initial topology

With same notation as above, in practical, we can take a basis $\mathcal{B}_\alpha$ of each $\mathcal{T}_\alpha$, then consider the f_α-preimages in X

$$\{f_\alpha^{-1}(U) \mid U \in \mathcal{B}_\alpha\}.$$

Then

$$\bigcup_{\alpha\in\Omega} \{f_\alpha^{-1}(U) \mid U \in \mathcal{B}_\alpha\}$$

is a subbasis of $\mathcal{I}$.

Product topology

Let $(X_\alpha)_{\alpha\in\Omega}$ be a family of topological spaces. We consider their Cartesian product

$$\prod_{\alpha\in\Omega} X_\alpha = \left\{ (x_\alpha)_{\alpha\in\Omega} \in \left(\bigcup_{\alpha\in\Omega} X_\alpha\right)^\Omega \middle| \forall\, \alpha \in \Omega,\, x_\alpha \in X_\alpha \right\}.$$

For each $\alpha \in \Omega$, there is a canonical projection map

$$\begin{aligned} pr_\alpha : \prod_{\alpha\in\Omega} X_\alpha &\to X_\alpha, \\ (x_\alpha)_{\alpha\in\Omega} &\mapsto x_\alpha. \end{aligned}$$

Definition 2.3.11

The initial topology on $\prod_{\alpha\in\Omega} X_\alpha$ induced by $(pr_\alpha)_{\alpha\in\Omega}$ is called the ***product topology*** on $\prod_{\alpha\in\Omega} X_\alpha$.

Final topology

In a similar fashion, we can define the final topology on a non-empty set X by setting X as the image space instead of the domain space of a map. More precisely, consider a map

$$f : Y \to X,$$

where Y is a topological space. We can associate to X a topology defined as follows

- a subset of X is open if and only if its preimage is open in Y.

Remark 2.3.12.
Notice that the map f is not necessarily injective, hence the preimage of U being open in Y is not the same as U being the image of an open set in Y.

We denote this topology by $\mathcal{F}$. Similar to Proposition 2.2.4, we have the following proposition about the relation between this topology and the continuity of f.

Proposition 2.3.13

Given a topology $\mathcal{T}$ on X, the map f is continuous with respect to $\mathcal{T}$ if and only if $\mathcal{T}$ is coarser than $\mathcal{F}$.

Proof. If f is continuous, then for any $U \in \mathcal{T}$, the preimage $f^{-1}(U)$ is open in Y, hence by the definition of $\mathcal{F}$, we have

$$U \in \mathcal{F}.$$

This implies that

$$\mathcal{T} \subset \mathcal{F}.$$

Conversely, if $\mathcal{T} \subset \mathcal{F}$, for any $U \in \mathcal{T}$, we have $U \in \mathcal{F}$. By the definition of $\mathcal{F}$, the preimage $f^{-1}(U)$ is open in Y. Hence f is continuous. □

Similar to the initial topology, we can generalize the above discussion and consider a collection of maps from topological spaces to X. Let $(Y_\alpha, \mathcal{T}_\alpha)_{\alpha\in\Omega}$ be a collection of topological spaces. For each $\alpha \in \Omega$, we denote by

$$f_\alpha : Y_\alpha \to X,$$

a map from Y_α to X. Then we consider the following set

$$\mathcal{B} = \{U \in \mathcal{P}(X) \mid \forall\, \alpha \in \Omega,\ f_\alpha^{-1}(U) \in \mathcal{T}_\alpha\},$$

as a subbasis and denote by $\mathcal{F}$ the topology on X generated by $\mathcal{B}$.

Definition 2.3.14

The topology $\mathcal{F}$ constructed above is called the ***final topology*** on X induced by $(f_\alpha)_{\alpha\in\Omega}$.

A generalization of Proposition 2.3.13 can be stated as follows, which can also be used as the definition of the final topology on X with respect to $(f_\alpha)_{\alpha\in\Omega}$.

Proposition 2.3.15

Given any topology $\mathcal{T}$ on X, all maps f_α's are continuous with respect to $\mathcal{T}$ if and only if $\mathcal{T}$ is coarser than $\mathcal{F}$.

Example 2.3.16 (Coherent topology).
Let X be a non-empty set and $(X_\alpha)_{\alpha\in\Omega}$ be a family of subsets of X whose union is X. Denote by ι_α the inclusion of X_α into X. Assume that for each α, the subset X_α is equipped with a topology $\mathcal{T}_\alpha$. Then the final topology on X induced by $(\iota_\alpha)_{\alpha\in\Omega}$ is called the *coherent topology* on X induced by $(X_\alpha)_{\alpha\in\Omega}$, also called the weak topology on X.

If X is a topological space and $(X_\alpha)_{\alpha\in\Omega}$ is an open cover of X, then for each α, the topology on X_α is the subspace topology. Then the coherent topology coincides with the given topology on X.

How to describe a final topology

It is slightly more complicated to describe a final topology. By the definition of the final topology, any open set U in X must have the f_α-preimage open in Y_α. Hence U is necessarily the f_α-image of an open set of Y_α. Hence with the same notation as above, we start by considering the f_α-images of open sets in $\mathcal{B}_\alpha$

$$\{f_\alpha(V) \mid V \in \mathcal{T}_\alpha\}.$$

Notice that since f_α is not necessarily injective, we only have

$$V \subset f_\alpha^{-1}(f_\alpha(V)),$$

instead of equality in general. Hence we do not necessarily have every $f_\alpha(V)$ open in X with respect to the final topology. We need those $V \in \mathcal{T}_\alpha$, such that $f_\alpha^{-1}(f_\alpha(V))$ is still open in Y_α. Hence we consider the following set

$$\mathcal{C} = \bigcup_{\alpha\in\Omega} \{f_\alpha(V) \mid V \in \mathcal{T}_\alpha,\ \forall \beta \in \Omega, f_\beta^{-1}(f_\alpha(V)) \in \mathcal{T}_\beta\}.$$

which is a subbasis of the final topology.

Quotient topology

Let X be a non-empty set and $\mathcal{R} \subset X \times X$ be an equivalence relation on X. We denote by $X/\mathcal{R}$ the set of $\mathcal{R}$-equivalence classes and

$$\pi : X \to X/\mathcal{R},$$

the canonical projection.

Assume that X admits a topology, then the final topology on $X/\mathcal{R}$ induced by π is called the quotient topology on $X/\mathcal{R}$.

Example I: Quotient by a group action

We have learned groups actions on a set. By considering the orbits decomposition, each group action on a set X induces a partition of X, hence an equivalent relation on it. If moreover X admits a topological structure, then it induces a quotient topology on $X/\mathcal{R}$. In this following, we give some explicit examples.

Example 2.3.17 (Circle as a quotient space).
When talking about the circle S^1, we usually consider the unit circle in the complex plane and define it as the following set

$$S^1 = \{e^{2\pi i\theta} \mid \theta \in \mathbb{R}\}.$$

Alternative, we can also consider S^1 as a quotient space of $\mathbb{R}$ by an action. We consider the real line $\mathbb{R}$ and the group $\mathbb{Z}$ of integers acts on it by

$$\begin{aligned} f : \mathbb{Z} \times \mathbb{R} &\to \mathbb{R} \\ (n, x) &\mapsto x + n \end{aligned}$$

For any $x \in \mathbb{R}$, its orbit is

$$O_x = \{x + n \mid n \in \mathbb{Z}\}.$$

We consider

$$\mathbb{R}/\mathbb{Z} := \{O_x \mid x \in \mathbb{R}\}.$$

To describe the final topology, it is enough to describe a collection of subsets of $\mathbb{R}/\mathbb{Z}$ whose preimages form a basis of the topology of $\mathbb{R}$.

We first describe a basis of the topology on $\mathbb{R}$. By considering the Euclidean metric on $\mathbb{R}$, a basis can be given by considering all open intervals. We denote this basis by $\mathcal{B}$.

For any open interval $I \in \mathcal{B}$ of $\mathbb{R}$, we consider its image under π. Notice that the image of any open interval of length greater than 1 is $\mathbb{R}/\mathbb{Z}$. We only need to consider the open interval with length smaller or equal to 1. To give an explicit description, we use the following fact: there is a bijective map

$$\varphi : \mathbb{R}/\mathbb{Z} \to [0, 1),$$

by sending an orbit to its unique representative in $[0, 1)$.

We consider the composition

$$\varphi \circ \pi : \mathbb{R} \to [0, 1),$$

then the image of an open interval in $\mathbb{R}$ with length smaller or equal to 1 is one of the following two types:

- either it is an interval $(a, b) \subset (0, 1)$;
- or there are a and b in $(0, 1)$ with $a \leq b$, such that the image is $[0, a) \cup (b, 1)$.

Moreover their preimages in $\mathbb{R}$ are all open. Hence they can generate the final topology on $[0, 1)$.

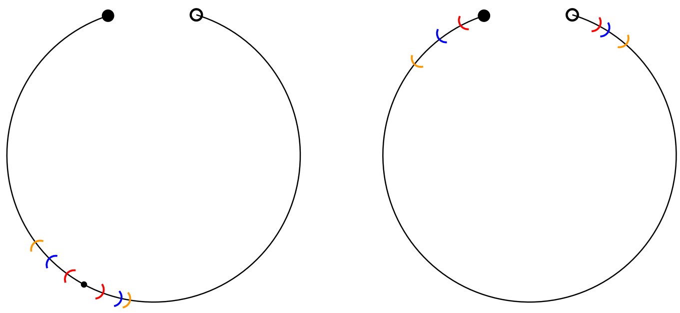


Figure 2.3.3: A neighborhood basis of a point in $(0, 1)$ (left); a neighborhood basis of 0 (right)

Consider the topology on $[0, 1)$ generated by these two kinds of open sets, we may find the resulting topological space is homeomorphic to S^1, and the following map

$$\begin{aligned} \psi : [0, 1) &\to S^1, \\ x &\mapsto e^{2\pi i x}. \end{aligned}$$

is an homeomorphism.

In this way, we found that the space $\mathbb{R}/\mathbb{Z}$ with quotient topology is topologically the same as S^1.

Remark 2.3.18.

In fact, we can even going further to talk about metric geometry in this example. Since $\mathbb{Z}$ acts isometrically on $\mathbb{R}$ "properly discontinuously", the metric on $\mathbb{R}$ induces a metric on $\mathbb{R}/\mathbb{Z}$ whose length is 1.

Remark 2.3.19.
We will come back to this example when we talk about "fundamental group" and "universal cover".

Example II: Gluing spaces by identifying points

The third way to obtain a circle topologically is familiar to everyone the most in some way. In our daily life, we can tie two ends of a rope together to get a circle. We will give a mathematical description of this process.

Example 2.3.20.
We consider the unit interval $[0,1]$ on $\mathbb{R}$ as a subspace. The goal is to "glue" 0 and 1 together. We consider the following equivalence relation:

$$\mathcal{R} := \{(x,x) \mid x \in [0,1]\} \cup \{(0,1),(1,0)\} \subset [0,1]^2.$$

Then in the quotient space

$$[0,1]/\mathcal{R},$$

the point 0 and 1 are identified together.

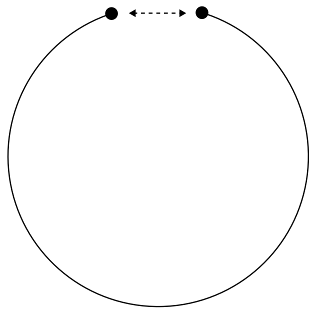


Figure 2.3.4: Identifying 0 with 1.

We can show that the set $[0,1]/\mathcal{R}$ with the quotient topology is homeomorphic to the circle. To be more precise, we notice that the topology on $[0,1]$ is the subspace topology, hence a basis can be given by considering the following three types intervals

- open intervals in $[0,1]$,
- intervals $[0,x)$ for any $x \in (0,1]$,
- intervals $(x,1]$ for any $x \in [0,1)$.

Let π denote the quotient map. Notice that the singletons 0 and 1 are not open, hence the set $\pi([0,x))$ should not be open in the quotient topology, since

$$\pi^{-1}(\pi([0,x))) = [0,x) \cup \{1\},$$

which is not open in $[0,1]$. A neighborhood basis of $\pi(0)$ can be given by

$$\{\pi([0,x) \cup (y,1]) \mid x \in (0,1],\, y \in [0,1)\}.$$

Under this topology, notice that not only 0 and 1 are glued together, so are their neighborhoods. Hence, the topological space $[0,1]/\mathcal{R}$ is homeomorphic to S^1.

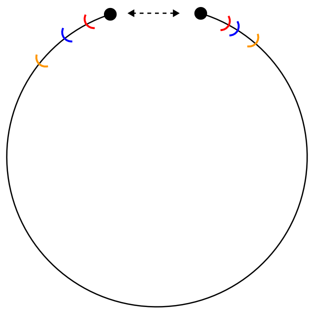


Figure 2.3.5: Gluing neighborhoods of 0 with neighborhoods of 1 in $[0, 1]$.

The above example shows how we describe "glue two point together" in a mathematical way using equivalence relation. A slightly more complicated example is the following way to get torus.

Example 2.3.21.
We consider the unit square D in $\mathbb{R}^2$ with vertices

$$v_1 = (0, 0),\ v_2 = (1, 0),\ v_3 = (1, 1),\ v_4 = (0, 1).$$

For any pair of points p and q in $\mathbb{R}^2$, the segment connecting them can be parametrized by $[0, 1]$ in the following way

$$p(t) = (1 - t)p + tq,\ t \in [0, 1].$$

We then define the following equivalent relation

$$\begin{aligned}\mathcal{R} :=&\{(u(t), v(t)) \mid \forall\, t \in [0, 1],\ u(t) = (1 - t)v_1 + tv_2,\ v(t) = (1 - t)v_4 + tv_3\}\cup \\ &\cup \{(u(t), v(t)) \mid \forall\, t \in [0, 1],\ u(t) = (1 - t)v_1 + tv_4,\ v(t) = (1 - t)v_2 + tv_3\}\cup \\ &\cup \{(v(t), u(t)) \mid \forall\, t \in [0, 1],\ u(t) = (1 - t)v_1 + tv_2,\ v(t) = (1 - t)v_4 + tv_3\}\cup \\ &\cup \{(v(t), u(t)) \mid \forall\, t \in [0, 1],\ u(t) = (1 - t)v_1 + tv_4,\ v(t) = (1 - t)v_2 + tv_3\}\cup \\ &\cup \{(v, v) \mid v \in D\} \subset D^2.\end{aligned}$$

The quotient space $D/\mathcal{R}$ is then obtained by identifying the sides v_1v_2 (resp. v_1v_4) and v_4v_3 (resp. v_2v_3). The resulting surface is a torus which we denote by T. Notice that the four vertices of D are identified together and we denote it by p.

Similar as in the previous example, roughly speaking, when we glue two points on the sides of D, we also glue their neighborhoods together to get a neighborhoods of the resulting point in T, and the neighborhoods of four vertices are glued together to get neighborhoods of p in T (See Figure 2.3.6 for an illustration).

The following example is more like what we do when making Baozi.

Example 2.3.22.
When we make a baozi, if we forget those pleats, roughly speaking we change a disk into a sphere by identifying the boundary of the disk to a point. Mathematically, we consider the unit disk $\mathbb{D}$ in $\mathbb{C}$ given by

$$\mathbb{D} = \{z \in \mathbb{C} \mid |z| \leq 1\}.$$

We consider the following equivalence relation

$$\mathcal{R} := \{(z, z) \mid z \in \mathbb{D}\} \cup (S^1)^2 \subset \mathbb{D}^2$$

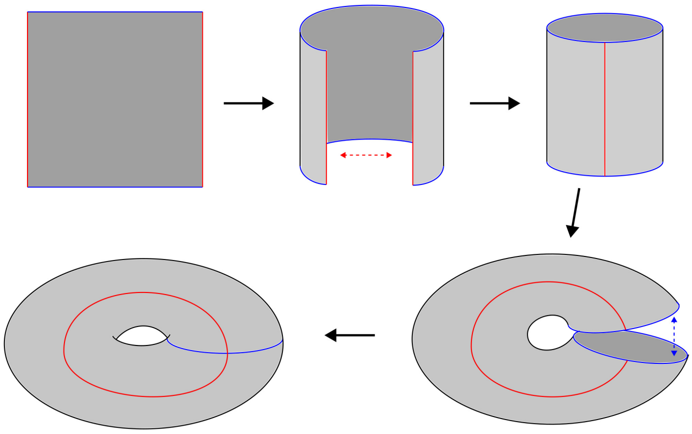


Figure 2.3.6: Gluing opposite sides of a square step by step.

Then the quotient space $\mathbb{D}/\mathcal{R}$ is a sphere (See Figure 2.3.7 for an illustration). In particular, we denote by p the resulting point by identifying points in S^1 together, and a neighborhood of S^1 is then sent to a neighborhood of p.

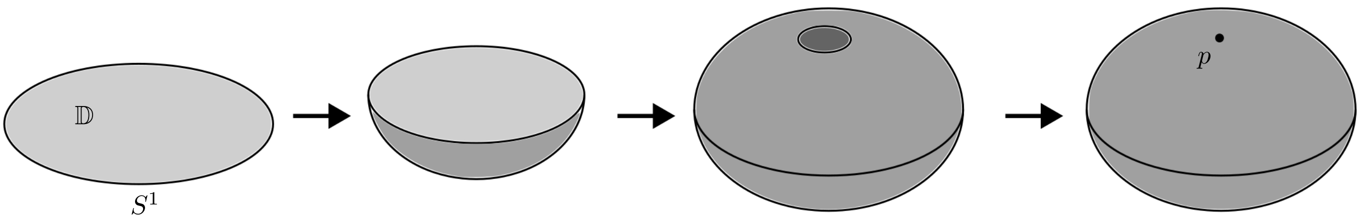


Figure 2.3.7: Identifying all points in S^1 together.

From above examples, we may conclude that by taking quotient, we may identify certain points together to get a point, and at the same time we also identifying their neighborhoods to get the neighborhood of the resulting point for the quotient topology.

Question 2.3.23
Consider the unit circle in the Euclidean plane

$$S^1 := \{e^{2\pi i\theta} \mid \theta \in \mathbb{R}\}.$$

We consider the subgroup $\langle r_\alpha \rangle$ of the isometry group of the Euclidean plane generated by the rotation r_α which rotates the plane around the origin for an angle $\alpha \notin 2\pi\mathbb{Q}$ counterclockwise. Describe the quotient topology on $S^1/\langle r_\alpha \rangle$.

In the above examples, what we did was to modify some part of the space to get a new space. Next we would like to introduce two ways to construct new spaces by connecting several given spaces together.

Example 2.3.24 (Wedge sum of spaces).
Let $(X_\alpha, x_\alpha)_{\alpha\in\Omega}$ be a family of topological spaces X_α marked by a point x_α. We consider the disjoint union of these spaces

$$\mathcal{X} = \bigsqcup_{\alpha\in\Omega} X_\alpha.$$

and the coherent topology on $\mathcal{X}$ is generated by the topologies on X_α's (See Example 2.3.16). We define the following equivalence relation

$$\mathcal{R} := \{(x, x) \in \mathcal{X}^2 \mid x \in \mathcal{X}\} \cup \{(x_\alpha, x_\beta) \in \mathcal{X}^2 \mid \alpha, \beta \in \Omega\}.$$

The quotient space $\mathcal{X}/\mathcal{R}$ is called the *wedge sum* of $(X_\alpha, x_\alpha)_{\alpha\in\Omega}$ and is denoted by

$$\bigvee_{\alpha\in\Omega} (X_\alpha, x_\alpha).$$

Intuitively, what we have done is identifying all x_α's together. Let us denote this resulting point by $y \in \mathcal{X}$. Roughly speaking, the neighborhood of y can be obtained by two steps: first taking one neighborhood for each x_α, then identifying all x_α's together.

For example, we consider the wedge sum of two circles, and the resulting space is the figure eight (See Figure 2.3.8).

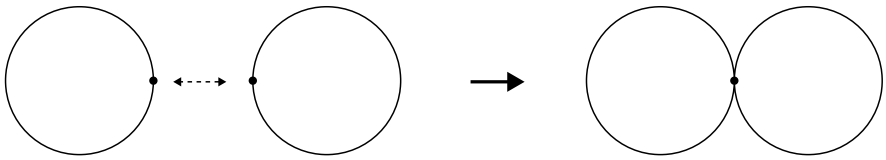


Figure 2.3.8: Wedge sum of two circles.

Example 2.3.25 (Connected sum).
Consider two n-manifolds M_1 and M_2. Let B_1 and B_2 be open n-balls in M_1 and M_2 respectively, and denote their boundary in M_1 and M_2 by A_1 and A_2. Since both A_1 and A_2 are $(n-1)$-spheres, there is an orientation reversing homeomorphism

$$f : A_1 \to A_2.$$

We consider the disjoint union

$$N = (M_1 \setminus B_1) \sqcup (M_2 \setminus B_2),$$

with the coherent topology, and construct the following equivalence relation

$$\mathcal{R} := \{(x, x) \in N^2 \mid x \in N\} \cup \{(x, f(x)) \in N^2 \mid x \in A_1\} \cup \{(f(x), x) \in N^2 \mid x \in A_1\}.$$

The quotient space $N/\mathcal{R}$ is called the connected sum between M_1 and M_2, and we denote it by

$$M_1 \# M_2.$$

Figure 2.3.9 is a connected sum of two copies of torus.

Remark 2.3.26.
Notice that the construction of a connected sum can guarrentee that the resulting space is still an n-manifold.

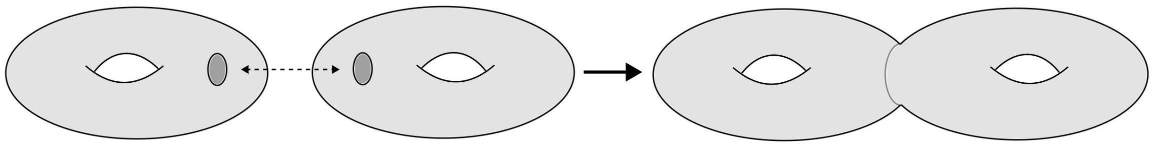


Figure 2.3.9: Connected sum of two torus.

We end this part by one example which is a technique usually called "cone off" part of a space.

Example 2.3.27 (Cone).
Let X be a topological space. We consider the following product space

$$X \times [0,1].$$

A *cone* based on X is then defined as the following quotient space

$$\mathrm{Cone}(X) := X \times [0,1]/\mathcal{R},$$

where

$$\mathcal{R} := \{((x,t),(x,t)) \in (X\times[0,1])^2 \mid x \in X\times[0,1]\} \cup \{((x,1),(y,1)) \in (X\times[0,1])^2 \mid x,y \in X\times[0,1]\}.$$

If $X = S^1$, the above construction gives exactly the cone that we are used to know (See Figure 2.3.10 for an illustration).

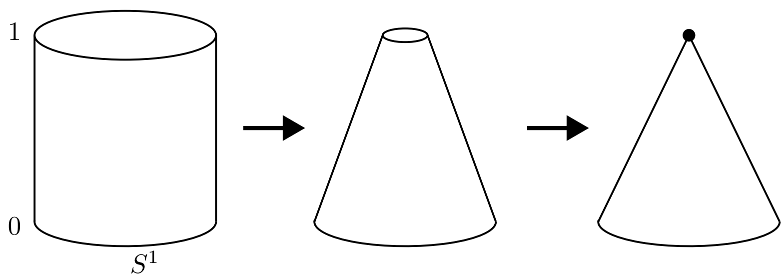


Figure 2.3.10: A cone.

The above construction can be done for a part of the space X. Let A be a non empty subset of X. We consider the following disjoint union

$$Y = X \sqcup \mathrm{Cone}(A).$$

We define an equivalent relation on it by

$$\mathcal{R} := \left\{ (y,y) \in Y^2 \middle| \, y \in Y \right\} \cup \left\{ (x,(x,0)) \in Y^2 \middle| \, x \in A \right\}.$$

Roughly speaking, the space $Y/\mathcal{R}$ can be understood as gluing $\mathrm{Cone}(A)$ to X along A. The process of changing X to $Y/\mathcal{R}$ is called "*coning off A in X*" (See Figure 2.3.11 for an illustration).

2.4 Connectivity

We recall here three basic properties of topological spaces which we usually discuss. The intuition of these properties has something to do with the study of Euclidean space. We will discuss them one by one in this part.

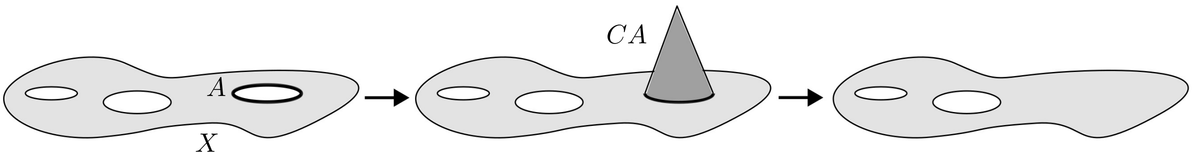


Figure 2.3.11: Coning off A in X.

Connected spaces

When we study the continuity in Euclidean space, it is intuitively related to the notion of connectedness. For example, let

$$f : [a, b] \to \mathbb{R},$$

be a continuous increasing function defined on a closed interval $[a, b] \subset \mathbb{R}$. The intermediate value theorem tells us for any $y \in [f(a), f(b)]$, there exists a point $c \in [a, b]$, such that $f(c) = y$. In other words, the image $[f(a), f(b)]$ is a interval of $\mathbb{R}$ with no gap in the middle which consists with our impression on connectedness.

When we follow this observation and study the connectedness for a general topological space, we find that the connectedness can be understood in different ways which are no longer equivalent when consider a general space. One way to say that something is connected is that it cannot be described as a union of two disjoint components. More rigorously, we have the following definition of being connected.

Definition 2.4.1

We say that a topological space X is ***connected*** if X cannot be written as a disjoint union of two non-empty open subset, i.e. there is NO pair of open sets in X denoted by U and V respectively, which satisfy the following properties:

1) U and V are non-empty;
2) $U \cap V = \emptyset$;
3) $U \cup V = X$.

A subset A of X is ***connected*** if by considering its subspace topology it is a connected topological space.

Remark 2.4.2.
Alternatively, the above definition can be rewritten as: The space X is connected if and only if the only subsets of X both open and closed are X and $\emptyset$.

Proposition 2.4.3

If A is a connected subset of X, so is $\overline{A}$.

Proof. We prove this proposition by contradiction.

Let A be a connected subset of X. If $\overline{A}$ is not connected, there are open sets U and V in X, such that

1) $U \cap \overline{A}$ and $V \cap \overline{A}$ are not empty;

2) $U \cap V \cap \overline{A}$ is empty;

3) $(U \cup V) \cap \overline{A}$ is $\overline{A}$.

Let $x \in U \cap \overline{A}$. Notice that U is a neighborhood of x. Since x is also a limit point of A, we have

$$U \cap A \neq \emptyset.$$

Similarly, we have

$$V \cap A \neq \emptyset.$$

Since $A \subset \overline{A}$, by 2) we have

$$U \cap V \cap A \subset U \cap V \cap \overline{A} = \emptyset,$$

and by 3) we have

$$(U \cup V) \cap A = ((U \cup V) \cap \overline{A}) \cap A = A.$$

The above discussion shows that A is not connected which is a contradiction. □

This result can be enhanced to the following one.

Corollary 2.4.4

Let A be a connected subset of X. If B is a subset of X with

$$A \subset B \subset \overline{A},$$

then B is connected.

Remark 2.4.5.
The proof is the same. In the previous proof, we only use the fact that points in $\overline{A}$ are limit points of A, and $A \subset \overline{A}$. These are still true, when we replace $\overline{A}$ by B.

Given a point x of X, there may be many connected subsets of X containing x.

Lemma 2.4.6

Let x be a point in X, and U and V be two connected subsets containing x, then $U \cup V$ is still connected.

Proof. We prove this lemma by contradiction. Assume that $U \cup V$ is not connected. Let A and B be two non empty open sets of $U \cup V$, such that

$$A \cap B = \emptyset, \quad A \cup B = U \cup V.$$

Since A is non empty and $A \subset U \cup V$, one of $A \cap U$ and $A \cap V$ must be non empty. Without loss of generality, we may assume that

$$A \cap U \neq \emptyset.$$

If $A \cap V$ is empty, since $V \subset A \cup B$, we have $V \subset B$, from which we have

$$x \in V \cap U \subset B \cap U.$$

Therefore, U can be written as a disjoint union of two non empty open sets $A \cap U$ and $B \cap U$ which contradicts to the fact that U is connected.

If $A \cap V$ is also non empty, then since one of $B \cap U$ and $B \cap V$ must be non empty, there is one of U can V having non empty intersections with both A and B. Without loss of generality, we may assume that it is U. Therefore, U can be written as a disjoint union of two non empty open sets $A \cap U$ and $B \cap U$ which contradicts to the fact that U is connected. □

Remark 2.4.7.
Here we consider connected sets containing x which may not be neighborhoods of x. In the proof, the existence of x can guarantee that $U \cap V$ is not empty, which is the condition needed essentially.

Using this lemma, we may define the following equivalence relation in X: given any pair of points x and y in X,

- $x \sim y$ if and only if there is a connected set U containing both x and y.

Definition 2.4.8

An equivalence class of the equivalence relation $\sim$ in X is called a ***connected component*** in X.

Here are several facts about connected components of X.

Proposition 2.4.9

The connected components of X have the following properties.

1) The space X is a disjoint union of its connected components.

2) A subset of X is a connected component if and only if it is a maximal connected subset of X.

3) A connected component is closed.

Proof. 1) This comes from the fact that each connected component is an equivalent class for an equivalence relation on X. By the properties of equivalence classes, we have the statement.

2) Let C be a connected component of X and x be any point in C. We first show that C is connected. Otherwise, there are two open sets U and V of X, such that

- $U \cap C \neq \emptyset, V \cap C \neq \emptyset$;
- $U \cap V = \emptyset$;
- $(U \cup V) \cap C = C$.

Let $x \in U \cap C$ and $y \in V \cap C$. Since $x \sim y$, we have a connected subset W of X containing x and y. Notice that all points of W are equivalent to x, hence

$$W \subset C.$$

By the relation among U, V and C, we have the following facts

- $U \cap W \neq \emptyset, V \cap W \neq \emptyset$;
- $U \cap V \cap W \subset U \cap V \cap C = \emptyset$;
- $(U \cup V) \cap W = (U \cup V) \cap (C \cap W) = C \cap W = W$.

Therefore W is not connected which is a contradiction. Hence C is connected.

Now we show the maximality of C. Let D be any connected subset of X having non empty intersection with C. Let x be in this intersection. For any $y \in D$, we have $x \sim y$, therefore $D \subset C$. Hence C is a maximal connected subset of X (no other connected subset of X containing C).

Conversely, assume that C is a maximal connected subset of X. Let x be a point in C. Since C is connected, we have

$$C \subset [x],$$

where $[x]$ is the equivalence class of x.

Let y be any point in X with $y \sim x$. By the definition of the equivalence relation, we have a connected subset U containing both x and y. Then the union $U \cup C$ is again a connected subset containing x. Since C is maximal, we have

$$U \cup C = C.$$

Therefore $U \subset C$ which implies that $y \in C$. Hence

$$[x] \subset C.$$

As a conclusion, we have

$$C = [x],$$

which is a connected component by definition.

3) If C is a connected component of X, 2) shows that C is connected. By Proposition 2.4.3, its closure $\overline{C}$ is also connected and contains C. By the maximality of C, we have

$$C = \overline{C},$$

which is closed. □

Remark 2.4.10.
The point 2) is an equivalent way to define a connected component in X.

The above proposition told us that the closedness holds all the time. Since the openness is preserved by finite intersection, we have the following immediate corollary of the above proposition.

Corollary 2.4.11

If X has finitely many connected components, then each of these components is open.

However the openness of a connected component does not hold in general. For example, we consider $\mathbb{Q}$ as a subspace of $\mathbb{R}$ (consider the Euclidean metric topology). Let p be a rational number. Then we can show that its connected component is $\{p\}$.

In fact, we can always cut $\mathbb{R}$ at an irrational number to separate $\mathbb{Q}$ and any of its subsets into two subsets which are both open and closed, unless this subset contains only one element. Notice that for a basis of the subspace topology on $\mathbb{Q}$ can be given by the intersections between $\mathbb{Q}$ and open intervals in $\mathbb{R}$.

The openness can be guaranteed when the space satisfies the following local property, called local connectedness.

Definition 2.4.12

The space X is said to be ***locally connected*** if for any point $x \in X$, for any neighborhood U of x, there is a connected neighborhood V of x satisfying $V \subset U$.

Considering a neighborhood basis of a point in X, the above definition is equivalent to the following one.

Proposition 2.4.13

The space X is locally connected if every point $x \in X$ admits a neighborhood basis consisting of connected sets.

Remark 2.4.14.
Since this definition is about local property, there is no reason that we should expect that a locally connected space is connected. For example, the subspace $(0, 1) \cup (2, 3)$ of $\mathbb{R}$ with the Euclidean metric topology is locally connected, but not connected.

When we check the other direction of implication, it is also not true in general. Notice that in the definition of local connectedness, we do not require only one connected neighborhood at each point. Instead, we require the existence of a neighborhood basis formed by connected subsets at each point which is strictly stronger. We may consider the following example to see this.

Example 2.4.15 (Connected but not locally connected).
This example is usually called the "topologist's sine curve". We consider the map

$$\begin{aligned} f : (0, 3) &\to \mathbb{R}, \\ x &\mapsto \sin \frac{1}{x}. \end{aligned}$$

Consider the graph

$$\mathrm{Graph}(f) := \{(x, f(x)) \mid x \in (0, 3)\}.$$

Then we consider the set given by

$$X = \mathrm{Graph}(f) \cup \{(0, y) \mid y \in [-1, 1]\}.$$

(See Figure 2.4.1 for an illustration.)

Consider it as a subspace of $\mathbb{R}^2$. Then given any point

$$p \in \{(0, t) \mid t \in [-1, 1]\},$$

any of its neighborhood basis will essentially contain some disconnected pieces.

Proposition 2.4.16

A space X is locally connected if and only if the connected components of any open set in X are open.

Proof. If X is locally connected, the each point of x has a neighborhood basis in which all sets are connected. Given any open set U of X, by the previous discussion, we may write it into a

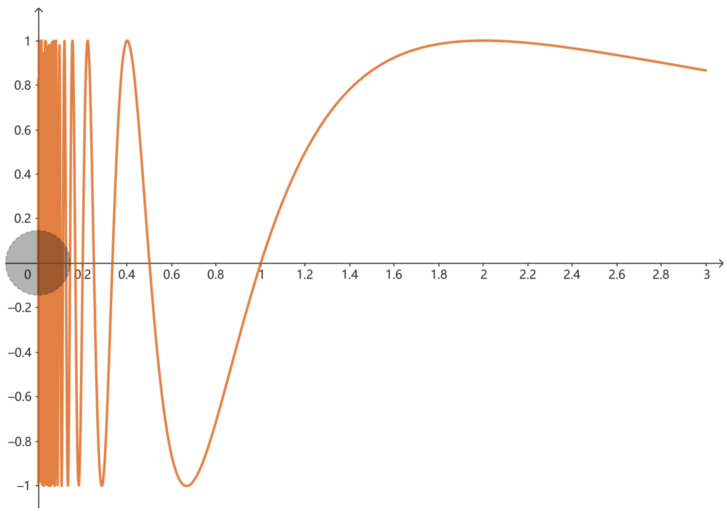


Figure 2.4.1: The set X.

disjoint union of connected components:

$$U = \bigsqcup_{\alpha \in \Omega} C_\alpha.$$

For any $\alpha \in \Omega$, and any $x \in C_\alpha$, since U is a neighborhood of x and X is locally connected, there is a connected open neighborhood V_x of x, such that

$$V_x \subset U.$$

Since V_x is connected, we have

$$V_x \subset C_\alpha.$$

Hence C_α is a neighborhood of x. Since x can be choose arbitrarily in C_α, we have C_α open. Therefore we may conclude that for any $\alpha \in \Omega$, the component C_α is open.

Conversely, for any $x \in X$, we consider a neighborhood U of x and denote U_x is an open neighborhood of x contained in U. We may write it into a disjoint union of connected components:

$$U_x = \bigsqcup_{\alpha \in \Omega} C_\alpha.$$

Choose $\beta \in \Omega$ such that x be a point of C_β. Since all connected components of U_x are open, we have C_β open in U_x. By the definition of subspace topology, there is an open subset V of X, such that

$$C_\beta = V \cap U_x.$$

Since both U_x and V are open in X, we have C_α open in X. Therefore the neighborhood U of x contains a connected neighborhood C_β of x. Hence X is locally connected. □

Corollary 2.4.17

If a space X is locally connected, then each connected component of X is both closed and open.

Remark 2.4.18.
If we work with nice topological spaces such as manifold, we can always assume that this is the case.

Another key feature of connectedness is that it is preserved by a continuous map.

Proposition 2.4.19

Let X and Y be two spaces and

$$f : X \to Y,$$

be a continuous and surjective map. If X is connected, then Y is connected.

Proof. We prove it by contradiction. If Y is not connected, then there are two nonempty open subsets of Y, denoted by V_1 and V_2, which are disjoint and whose union is Y.

Since f is continuous and surjective, the preimages

$$U_1 = f^{-1}(V_1) \neq \emptyset \quad \text{and} \quad U_2 = f^{-1}(V_2) \neq \emptyset$$

are both open in X. Moreover,

$$f^{-1}(V_1) \cap f^{-1}(V_2) = f^{-1}(V_1 \cap V_2) = \emptyset.$$
$$f^{-1}(V_1) \cup f^{-1}(V_2) = f^{-1}(V_1 \cup V_2) = X.$$

Hence X is not connected, which is a contradiction. □

Corollary 2.4.20

Let X and Y be two spaces and

$$f : X \to Y,$$

be a continuous. If X is connected, then the image $f(X)$ is a connected subset of Y.

Using this property, we may have another equivalent definition of connectedness.

Definition 2.4.21

A space X is ***connected*** if any continuous map from X to a space with discrete topology is constant.

Remark 2.4.22.
For example, we may consider the space $\{0, 1\}$ with discrete topology.

Path connected spaces

Another way of understanding connectedness is that we can go from any point to another one in a continuous way (walking along a path). This is actually the notion of being path connected.

Definition 2.4.23

A ***path*** in X is a continuous map

$$\gamma : [0, 1] \to X.$$

Definition 2.4.24

The space X is said to be ***path connected*** if for any pair of points p and q in X, there is a path
$$\gamma : [0,1] \to X,$$
such that $\gamma(0) = p$ and $\gamma(1) = q$.

Some spaces familiar to us are path connected. We give some examples.

Example 2.4.25 ($\mathbb{R}^n$).
We consider $\mathbb{R}^n$ equipped with the Euclidean metric topology. Given any pair of points p and q in $\mathbb{R}^n$, we consider the map
$$\gamma : [0,1] \to \mathbb{R}^n$$
$$t \mapsto (1-t)p + tq.$$

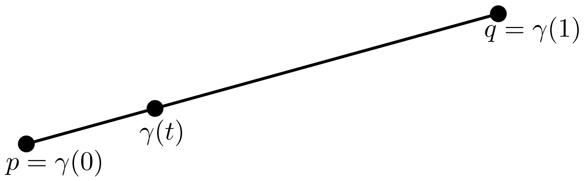


Figure 2.4.2: The segment connecting p to q

The image of γ is a segment of $\mathbb{R}^n$ connecting p to q (See Figure 2.4.2). We may verify with the definition that this is a path in $\mathbb{R}^n$. Hence $\mathbb{R}^n$ is path connected.

Example 2.4.26 (S^n).
We use the coordinates in $\mathbb{R}^{n+1}$:
$$S^n = \{(x_1, ..., x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1^2 + \cdots + x_{n+1}^2 = 1\}.$$
Let p and q be two points in $\mathbb{R}^n$, and denote by P the plane in $\mathbb{R}^{n+1}$ passing p, q and O the origin. The intersection $S^n \cap P$ is a circle of radius 1 and passing p and q. It is enough to show that a circle is path connected (See Figure 2.4.3 for an illustration for S^2).

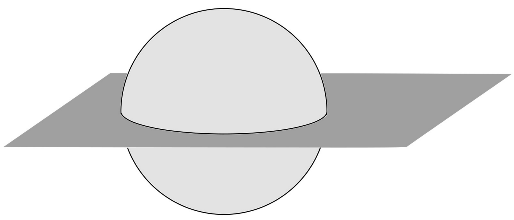


Figure 2.4.3: Cutting an 2-sphere with a 2-plane

Consider the unit circle S^1 in $\mathbb{R}^2$, which can be described as follows:
$$S^1 = \{(\cos\theta, \sin\theta) \in \mathbb{R}^2 \mid \theta \in \mathbb{R}\}.$$

Assume that

$$p = (\cos\theta_1, \sin\theta_1), \quad q = (\cos\theta_2, \sin\theta_2),$$

we can define

$$\begin{aligned} \gamma : [0,1] &\to S^1, \\ t &\mapsto (\cos((1-t)\theta_1 + t\theta_1), \sin((1-t)\theta_1 + t\theta_1)). \end{aligned}$$

This is a path in S^1 connecting p to q. Hence S^1 is path connected, so is S^n.

Another way to see this is to consider the stereographic projection. Let $N = (0, ..., 0, 1) \in \mathbb{R}^{n+1}$. We consider the map

$$\begin{aligned} \pi : S^n \setminus \{N\} &\to \{(x_1, ..., x_n, 0) \in \mathbb{R}^{n+1} \mid x_1 \in \mathbb{R}, \ldots, x_n \in \mathbb{R}\}, \\ (y_1, ..., y_{n+1}) &\mapsto (0, ..., 0, 1) + \frac{1}{1 - y_{n+1}} (y_1, ..., y_n, y_{n+1} - 1). \end{aligned}$$

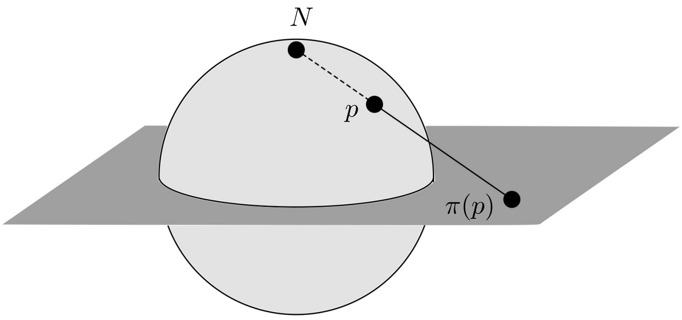


Figure 2.4.4: The stereographic projection of S^2

One may verify that this map is a homeomorphism. Since $\mathbb{R}^n$ is path connected, for any points p and q in $S^n \setminus \{N\}$, there is a path γ in $\mathbb{R}^n$ connecting $\pi(p)$ and $\pi(q)$. We consider the composition $\pi^{-1} \circ \gamma$. This is path in S^n connecting p and q. If one of p and q is N, we may use an $\mathrm{SO}(n)$ element A to move p and q away from N. The above discussion works for $A(p)$ and $A(q)$, and we have a path η connecting $A(p)$ and $A(q)$. Then $A^{-1} \circ \eta$ is a path connecting p and q.

Remark 2.4.27.
The "n" in the notation S^n stands for the dimension of S^n. Hence S^n is the unit sphere of $\mathbb{R}^{n+1}$.

Notice that $[0,1]$ is connected, hence any path in X would be connected. This indicated that the path connectivity may implies the connectivity. In fact, this is true and we state it as follows.

Proposition 2.4.28

If a space X is path connected, then it is connected.

Proof. We prove it by contradiction. Assume that X is not connected, then there are non-empty open set U and V such that

$$U \cap V = \emptyset, \quad U \cup V = X.$$

Let $x \in U$ and $y \in V$. Since X is path connected, there exists a continuous map

$$\alpha : [0,1] \to X,$$

such that $\alpha(0) = x$ and $\alpha(1) = y$.

Let

$$A = \alpha^{-1}(U), \quad B = \alpha^{-1}(V).$$

Since α is continuous, both A and B are open in $[0, 1]$. Moreover, since

$$U \cap V = \emptyset, \quad U \cup V = X,$$

we have

$$A \cap B = \emptyset, \quad A \cup B = [0, 1].$$

Notice that $0 \in A$ and $1 \in B$, both A and B are non empty. Hence $[0, 1]$ is not connected which is a contradiction. □

When we consider manifolds, it seems that there is not much difference between the connectedness and the path connectedness. However, the other direction does not hold in general. See the following example.

Example 2.4.29 (Connected not path connected).
We consider again the "topologist's sine curve" (See Example 2.4.15). Same as before, the topology on X is the subspace topology by consider X as a subset of the Euclidean space $\mathbb{R}^2$.

We use the same notation as in Example 2.4.15. Notice that the graph of f is path connected, hence is connected, and a connected component of X containing any point of $\mathrm{Graph}(f)$ must contain the entire graph. Since a connected component is also closed, therefor this connected components must contains all limit points of $\mathrm{Graph}(f)$. Notice that points in $\{(0, y) \mid y \in [-1, 1]\}$ are all limit points of $\mathrm{Graph}(f)$, hence are in its connected component. This means that there is only one connected component in X. Hence X is connected.

However, the space X is not path connected. Let $p = (0, 0)$ and $q = (1, 0)$. Given any path γ with $\gamma(0) = p$ and $\gamma(1) = p$, for any

$$p' \in \{(0, y) \mid y \in [-1, 1]\},$$

there is a sequence $(t_n)_{n\in\mathbb{N}}$ with

$$\lim_{n\to\infty} t_n = 0,$$

such that

$$\lim_{n\to\infty} \gamma(t_n) = p'.$$

This contradicts to the fact that γ is continuous.

Similar to the local connectedness, we also have a notion of local path connectedness which defined in a similar fashion.

Definition 2.4.30

The space X is said to be ***locally path connected*** if for any point $x \in X$, for any neighborhood U of x, there is a path connected neighborhood V of x satisfying $V \subset U$.

Similar to the case for the locally connected property, we have the following proposition for the locally path connected property.

Proposition 2.4.31

The space X is locally connected if every point $x \in X$ admits a neighborhood basis consisting of path connected sets.

Similar to the relation between connectedness and local connectedness, there is no implication between path connectedness and local path connectedness in either way. One direction is easy to understand. The local path connectedness only require conditions in neighborhoods at each point. Hence it cannot not implies the path connectedness. To see the other direction does not hold either, we may consider the following example.

Example 2.4.32 (Path connected not locally path connected).
Consider the following subset of $\mathbb{R}^2$:

$$X = \{(t,0) \mid t \in [0,1]\} \bigcup \{(0,s) \mid s \in [0,1]\} \bigcup \left(\bigcup_{n \in \mathbb{N}} \left\{ \left(\frac{1}{n}, r \right) \middle| r \in [0,1] \right\} \right)$$

(See Figure 2.4.5 for an illustration.)

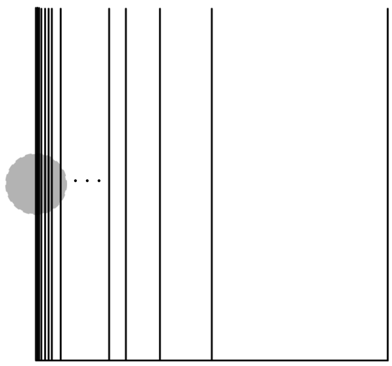


Figure 2.4.5: Being path connected but not locally path connected

We consider the Euclidean metric topology on $\mathbb{R}^2$, and consider the subspace topology on X. Notice that X is not locally path connected in any point on the vertical segment

$$\{(0,s) \mid s \in [0,1]\}$$

other than $(0,0)$.

Remark 2.4.33.
Manifolds are nice topological spaces in the sense that they are at the same time connected, path connected, locally connected and locally path connected.

2.5 Compactness

One fact familiar to us is that a continuous map from an interval $[a,b]$ to $\mathbb{R}$ is uniformly continuous, which means as along as x and y are closed enough ($< \delta$), so do their f-value ($< \epsilon$) and this constant δ is uniform, i.e. independent of choices of x and y.

In the proof of this fact, we use one properties of $[a,b]$ which is that any of its open cover contain a subcover which is finite. Here a cover is a collection of subsets whose union is the entire space. The notion related to this property is the compactness of a space.

Compact space

Definition 2.5.1

Let X be a topological space. A family $\mathcal{A} = (A_\alpha)_{\alpha\in\Omega}$ of open subsets of X is called an ***open cover*** of X if

$$X = \bigcup_{\alpha\in\Omega} A_\alpha.$$

If a subfamily $\mathcal{A}' \subset \mathcal{A}$ is also an open cover of X, we call $\mathcal{A}'$ a ***subcover*** of a cover $\mathcal{A}$.

Remark 2.5.2.
In some references, the equality

$$X = \bigcup_{\alpha\in\Omega} A_\alpha$$

is replaced by

$$X \subset \bigcup_{\alpha\in\Omega} A_\alpha,$$

which has some advantage when we discuss compact subsets.

Definition 2.5.3

A topological space X is said to be ***compact*** if every open cover of X admits a finite subcover.

A subset of a topological space X is ***compact*** if it is a compact space with respect to the subspace topology.

Remark 2.5.4.
By considering the relation between open sets and closed sets, we can also define the compactness by the following condition:

- given any collection of closed sets in X with empty intersection, it has a finite subcollection with empty intersection.

We first give some properties of a compact space.

Proposition 2.5.5

1) A closed subset of a compact space X is compact.

2) If a space X is Hausdorff, then any compact subset of X is closed.

Proof. 1) Let A be a closed subset in X. Hence its complement A^c is open. Given any open cover $\{U_\alpha\}_{\alpha\in\Omega}$ of A, for each α, there is an open set V_α of X, such that

$$U_\alpha = V_\alpha \cap A.$$

Then the following collection

$$\{V_\alpha\}_{\alpha\in\Omega} \cup \{A^c\},$$

is a open cover of X. Since X is compact, it contains a finite subcover of X denoted by

$$\{W_1, ..., W_n\} \subset \{V_\alpha\}_{\alpha\in\Omega} \cup \{A^c\}.$$

Then by removing empty intersections if exist, the collection

$$\{W_1 \cap A, ..., W_n \cap A\}$$

is an open cover on A. Notice that $A^c \cap A$ is empty set, hence we have

$$\{W_1 \cap A, ..., W_n \cap A\} \subset \{U_\alpha\}_{\alpha\in\Omega}.$$

This implies that A is compact.

2) Let K be a compact subset of X. We would like to show that its complement is open. Since X is Hausdorff, for any $x \in K^c$, for any $y \in K$, there are open neighborhood U_y of x and open neighborhood V_y of y in Y, such that

$$U \cap V = \emptyset.$$

Notice that

$$K \subset \bigcup_{y\in K} V_y.$$

By the compactness of K, there is a finite collection

$$\{V_1, ..., V_n\}$$

associated to points $y_1, ..., y_n \in K$, such that

$$K = \bigcup_{i=1}^{n} (V_i \cap K).$$

We denote by $U_1, ..., U_n$ the open neighborhoods of x associated to $y_1, ..., y_n$. Then we find a open neighborhood

$$\bigcap_{i=1}^{n} U_i$$

of x and an open neighborhood

$$\bigcup_{j=1}^{n} V_i$$

of K which are disjoint.

The above construction shows that K^c is a neighborhood of x. Since x is chosen arbitrarily, the set K^c is a neighborhood of any of its points, hence is open. Therefore K is closed. □

Compactness is also a property preserved by a continuous map.

Proposition 2.5.6

Let X and Y be two spaces and

$$f : X \to Y,$$

be a continuous and surjective map. If X is compact, then Y is compact.

Proof. Given any open cover

$$\{V_\alpha\}_{\alpha\in\Omega}$$

of Y, we consider the collection of their preimages

$$\{U_\alpha = f^{-1}(V_\alpha)\}_{\alpha\in\Omega},$$

which is an open cover of X. Since X is compact, there is a finite subcover

$$\{U_1, ..., U_n\},$$

whose images

$$V_1, ..., V_n$$

form an open cover of Y which is a subcover of

$$\{V_\alpha\}_{\alpha\in\Omega}.$$

Hence Y is also compact. □

This means that similar to the connectedness, the compactness is also preserved by continuous maps.

One thing that we have seen in analysis course is that any non constant sequence contained in a compact subset of $\mathbb{R}$ has a convergence subsequence. This holds for a general second countable space, i.e. a topological space with a countable basis.

Proposition 2.5.7

Let X be a second countable topological space. Then X is compact if and only if any sequence $(x_n)_{n\in\mathbb{N}}$ in X admits a convergent subsequence.

Proof. Notice that if $(x_n)_{n\in\mathbb{N}}$ has a constant subsequence, we may always choose this subsequence which is convergent. Hence we will assume from now on that for any $m, n \in \mathbb{N}$, we have $x_m \neq x_n$.

Assume that X is compact. If the sequence has no convergent subsequence, then for any $x \in X$, there is an open neighborhood of x whose intersection with $(x_n)_{n\in\mathbb{N}}$ has only finitely many elements. Then such open sets form an open cover of X. Since X is compact, there is a finite subcover of X denoted by

$$U_1, ..., U_n,$$

associated to points $x_1, ..., x_n$. Since

$$X = \bigcup_{i=1}^{n} U_i,$$

there must be one of $U_1, ..., U_n$ whose intersection with $(x_n)_{n\in\mathbb{N}}$ has infinitely many elements, which is a contradiction.

Conversely, if X is not compact, then there is a infinite cover $\mathcal{C}$ of X which does not admit any finite subcover of X.

Since X is second countable which admits a countable basis

$$\mathcal{B} = \{U_n \mid n \in \mathbb{N}\},$$

we consider

$$\mathcal{B}_0 = \{U \in \mathcal{B} \mid \exists V \in \mathcal{C}, U \subset V\}.$$

Then $\mathcal{B}_0$ is an infinite cover, since $\mathcal{B}$ is a basis and $\mathcal{C}$ is a cover of X. We denote

$$\mathcal{B}_0 = \{U'_n \mid n \in \mathbb{N}\}.$$

By assumption, $\mathcal{C}$ does not admit a finite subcover, hence $\mathcal{B}_0$ does not admit a finite subcover. (for otherwise, if $U_1', ..., U_n'$ in $\mathcal{B}_0$ form a cover of X, denote by $V_1, ..., V_n$ their corresponding elements in $\mathcal{C}$, then $V_1, ..., V_n$ also form a cover of X.)

Therefore for any $n \in \mathbb{N}$, we can find a point

$$x_n \notin U_1' \cup \cdots \cup U_n'.$$

Notice that since $\mathcal{B}_0$ is a cover, for each $n \in \mathbb{N}$, there must be a index $N \in \mathbb{N}$, for any $m > N$, we have

$$x_n \in U_1' \cup \cdots \cup U_m'.$$

Consider this sequence, and we would like to show that this does not have a convergent subsequence by contradiction. Assume this is the case. Then $(x_n)_{n\in\mathbb{N}}$ has not constant subsequence. By taking a subsequence of $(x_n)_{n\in\mathbb{N}}$ and taking unions of the first several U_i''s if necessary, we may assume that $(x_n)_{n\in\mathbb{N}}$ are pairwise distinct, and is convergent.

We denote by x a limit point of this sequence. Hence for any neighborhood U of x contains all but finitely many points in $(x_n)_{n\in\mathbb{N}}$. Since $x \in X$, there is an open set $U_j' \in \mathcal{B}_0$, such that $x \in U_j'$. Notice that U_j' is a neighborhood of x, hence contains all but finitely many elements in $(x_n)_{n\in\mathbb{N}}$, or more equivalently, there is an $N_0 \in \mathbb{N}$, such that for any $n > N_0$, we have

$$x_n \in U_1' \cup \cdots \cup U_j'.$$

This contradicts to the construction of x_n with $n > j$.

□

Remark 2.5.8.
The property that any sequence admits a convergent subsequence is called the *sequential compactness.* Notice that a separated metric space equipped with the metric topology is always second countable. Hence the above equivalence between compactness and the sequential compactness holds in these cases. In particular, it holds for $\mathbb{R}^n$ and any n-manifolds with $n \in \mathbb{N}^*$.

Now we consider Hausdorff spaces.

Proposition 2.5.9

Consider X and Y two Hausdorff topological spaces, and a map

$$f : X \to Y.$$

Assume that Y is compact. Consider a non-empty subset A of X and assume that $a \in \overline{A}$. Then f admits at least a limit value $y \in Y$ when x tends to a in A. If moreover such y is unique, then f has y as the limit when x tends to a in A.

Proof. See Definition 2.1.38 for precise definitions.

Let $\mathcal{V}(a)$ denote the set of all neighborhood of a, then we consider the following subset of Y:

$$\mathrm{Lim(A)} = \bigcap_{U \in \mathcal{V}(a)} \overline{f(U \cap A)}.$$

By definition, if $y \in Y$ is a limit value of f when x tends to a in A, then for any neighborhood V of y and any neighborhood U of a, we have

$$f(U \cap A) \cap V \neq \emptyset.$$

This is equivalent to $y \in \mathrm{Lim}(A)$.

The firs part of the proposition is equivalent to say that $\mathrm{Lim}(A)$ is not empty. We prove it by contradiction. Assume that it is empty. Since Y is compact, since $\overline{f(U \cap V)}$'s are closed, there are finitely many of them whose intersection is empty:

$$\overline{f(U_1 \cap A)} \cap \cdots \cap \overline{f(U_k \cap A)} = \emptyset.$$

Hence, we have

$$f(U_1 \cap \cdots \cap U_k \cap A) \subset \overline{f(U_1 \cap A)} \cap \cdots \cap \overline{f(U_k \cap A)} = \emptyset.$$

This implies that

$$U_1 \cap \cdots \cap U_k \cap A = \emptyset.$$

However, this is impossible, since $U_1 \cap \cdots \cap U_k$ is a neighborhood of a and $a \in \overline{A}$.

Assume that such y is unique. Let V be an open neighborhood of y. Hence V^c is closed. On the other hand, the above discussion shows that

$$\{y\} = \mathrm{Lim}(A) = \bigcap_{U \in \mathcal{V}(a)} \overline{f(U \cap A)}.$$

Therefore

$$V^c \cap \left(\bigcap_{U \in \mathcal{V}(a)} \overline{f(U \cap A)} \right) = \emptyset.$$

Since Y is compact, there is a finite collection of neighborhoods $U_1, ..., U_k$ of a, such that

$$V^c \cap \left(\bigcap_{i=1}^{k} \overline{f(U_i \cap A)} \right) = \emptyset,$$

which implies that

$$V^c \cap f(U_1 \cap \cdots \cap U_k \cap A) = \emptyset,$$

This is equivalent to

$$f(U_1 \cap \cdots \cap U_k \cap A) \subset V.$$

Notice that $U_1 \cap \cdots \cap U_k$ is a neighborhood of a. Hence y is the limit of f when x tends to a in A. □

Here is a corollary discuss the same problem as in Proposition 2.5.7.

Corollary 2.5.10

In a compact Hausdorff space, any sequence admits a limit point. If this point is unique, then the sequence converges to this point.

Proof. Consider $X = \{(n+1)^{-1} \mid n \in \mathbb{N}\} \cup \{0\}$ as a subspace of $\mathbb{R}$. Given any sequence $(y_n)_{n \in \mathbb{N}}$ in a compact Hausdorff space Y, define

$$f : X \to Y$$

with $f(n) = y_n$ for any $n \in \mathbb{N}$. Then the corollary can be deduced from the above proposition. □

Remark 2.5.11.
Having a limit point (need infinitely many sequence points in any neighborhood) is not exactly the same as converging to a point (need all but finitely many sequence points in any neighborhood).

Next we would like to show that the compactness is preserved when taking products among topological spaces.

Theorem 2.5.12 (Tychonov Theorem)

Any product of compact spaces is compact.

Proof. Let $\{X_\alpha\}_{\alpha\in\Omega}$ be a collection of compact topological spaces. We denote by

$$X = \prod_{\alpha\in\Omega} X_\alpha$$

their product space (equipped with the product topology). The goal is to show that this is a compact space.

By the definition of product topology, it is generated by the following subbasis

$$\mathcal{A} = \{\mathrm{pr}_\alpha^{-1}(V) \mid V \text{ open in } X_\alpha\}.$$

Lemma 2.5.13

If a topological space Y is not compact, then any subbases covering Y admits a subcover with no finite subcovers.

Remark 2.5.14.
Here the requirement on a subbasis of covering Y is due to Definition 2.1.19 used previously, where we do not require a subbasis cover the whole space.

In other words, the above tries to say that if Y is not compact, then given any subbasis $\mathcal{B}$ covering Y, it has a subset $\mathcal{C} \subset \mathcal{B}$ which covers Y and has no finite subcover.

Proof of Lemma 2.5.13. We consider the collection Θ of all open covers of Y with no finite subcover which form a subset of $\mathcal{P}(\mathcal{P}(Y))$:

$$\Theta := \{\mathcal{C} \subset \mathcal{P}(Y) \mid \mathcal{C} \text{ is a cover of } Y \text{ with no finite subcover}\}.$$

Since Y is not compact, the set Θ is no empty. Now we consider the partial order induced by inclusion in $\mathcal{P}(\mathcal{P}(Y))$. Notice that each chain has a maximal element. Hence by Zorn lemma there is a maximal element in Θ denoted by $\mathcal{C}_{max}$.

Consider any subbasis $\mathcal{B}$ which covers Y. For any $y \in Y$, there is $V \in \mathcal{C}_{max}$, such that

$$y \in V.$$

By the definition of a subbasis, there are finitely many elements

$$U_1, ..., U_k \in \mathcal{B},$$

such that

$$x \in U_1 \cap \cdots \cap U_k \subset V.$$

By the maximality, we have $U_1, ..., U_k \in \mathcal{C}_{max}$. For otherwise, without loss of generality, we may assume that $U_1 \notin \mathcal{C}_{max}$. Then since $\mathcal{C}_{max}$ is maximal, the cover

$$\{U_1\} \cup \mathcal{C}_{max},$$

has a finite cover of Y, denoted by

$$U_1, V_1, ..., V_l.$$

Since $U_1 \subset V$, the collection

$$\{V, V_1, ..., V_l\} \subset \mathcal{C}_{max}$$

is a finite cover of Y which is a contradiction.

Now we consider $\mathcal{B} \cap \mathcal{C}_{max}$, this is a subset of $\mathcal{B}$ which covers Y. Moreover it has no finite subcover. □

Now we go back to the proof of Tychonov Theorem. We consider the subbasis $\mathcal{A}$ describe above. We would like to show that this subbasis has no subcover of X with no finite subcover. More precisely, given any subcover $\mathcal{C}$ in $\mathcal{A}$, there is $\alpha\Omega$, such that there are

$$\{\mathrm{pr}_\alpha^{-1}(V) \mid V \in \mathcal{L} \subset \mathcal{T}_\alpha\} \subset \mathcal{C},$$

where $\mathcal{T}_\alpha$ is the topology on X_α, $\mathcal{L}$ is a subset of proper open sets in $\mathcal{T}_\alpha$ such that $\cup\mathcal{L} = X_\alpha$. Such an α does exists, otherwise $\mathcal{C}$ is not a covering of X.

Notice that $\mathcal{L}$ is an open cover of X_α, since X_α is compact, there is a finite subcover

$$\{V_1, ..., V_k\}$$

of X_α. This moreover implies that their preimages under pr_α form an open cover of X, which is a finite subcover of $\mathcal{C}$. Hence X is compact. □

Remark 2.5.15.

The proof of Tychonov Theorem uses the Zorn Lemma which is equivalent to the Axiom of Choice.

Locally compact spaces

There is also notion of local compactness, however the way with which we define it is different from what is used previously for local (path) connectedness.

Definition 2.5.16

A space X is ***locally compact at a point*** $x \in X$ if there is a compact neighborhood U of x. A space is ***locally compact*** if it is locally compact at every point $x \in X$.

Example 2.5.17.

If the space X is equipped with the discrete, topology, then every point x in X is a compact subset. This is not difficult to see. The subspace topology of $\{x\}$ has only two open sets: $\emptyset$ and $\{x\}$. Hence $\{x\}$ is compact. At the same time $\{x\}$ is a neighborhood of x. Hence X is locally compact.

Notice the first half argument works in any topological space. A single point subset in any topological space is compact. However such a subset is not always open in an arbitrary topological space.

Example 2.5.18.

Here is another extremal case. If X is a compact topological space, then X is locally compact, since X is a neighborhood of any of its points. Moreover, any closed subset in X is compact (See 2.5.5).

Example 2.5.19.
Let X be the n-dimensional Euclidean space with $n \in \mathbb{N}^*$, and consider its metric topology. Then it is locally compact, since any closed ball is compact in $\mathbb{R}^n$, and it contains an open ball which is open, hence it is a neighborhood of its center. We may show that every sequence in a closed ball has a converges subsequence. Then Proposition 2.5.7 shows that a closed ball in $\mathbb{R}^n$ is compact.

Proposition 2.5.20

In fact, if X is Hausdorff and locally compact at $x \in X$, then it has a compact neighborhood basis of x.

Proof. Since X is locally compact at x, there is a compact neighborhood K of x. Let U be an open subset in x contained in K, we would like to show that there is a compact neighborhood of x contained in U.

Notice that X is Hausdorff. Let y be a different from x in K, then there are open neighborhoods U_y and V_y of x and y respectively, such that

$$U_y \cap V_y = \emptyset.$$

Figure

Notice that

$$\bigcap_{y \in K \setminus \{x\}} U_y = \{x\}.$$

We denote

$$W_y = \mathring{\widehat{U_y^c}}.$$

Notice that $y \in W_y$ for any $y \in K \setminus \{x\}$. Hence

$$\{U\} \cup \{W_y \cap K \mid y \in K \setminus \{x\}\}$$

is an open cover of K. Since K is compact, hence there is a finite subcover. Hence there is a subcover of K:

$$\{U, W_{y_1} \cap K, ..., W_{y_n} \cap K\}.$$

Then

$$W = (W_{y_1} \cup \cdots \cup W_{y_n}) \cap K$$

is open in K, whose complement is closed in K. Since X is Hausdorff, so is K. Hence $K \setminus W$ is also compact. Moreover

$$W \cup U = K$$

implies that

$$K \setminus W \subset U.$$

Notice that we have

$$K \setminus W = \overline{(U_{y_1} \cup \cdots \cup U_{y_n})} \cap K$$

hence it is a neighborhood of x, since U_{y_1} is a neighborhood of x. □

Corollary 2.5.21

Any open subset of a locally compact Hausdorff space is locally compact.

Compactness and continuous maps

In this part, we discuss the relation between compactness and continuous maps.

Proposition 2.5.22

Let X and Y be compact Hausdorff topological spaces, and $f : X \to Y$ be a continuous map. If f is bijective, then f is a homeomorphism.

Proof. It is enough to show that f^{-1} is also continuous. We can show that f sends closed subsets in X to closed subsets in Y. Since both spaces are Hausdorff and compact, a closed subset in X is compact, hence $f(X)$ is also compact and hence closed in Y. (See Proposition 2.5.5) □

2.6 Topological properties/Topological invariants

When f is a homeomorphism from X to Y, its inverse f^{-1} is a homeomorphism from Y to X. Given any set Ω of topological spaces, we can verify that "being homeomorphic" satisfies the reflexivity, the symmetry and the transitivity, therefore induces an equivalence relation in Ω.

Definition 2.6.1

A property P of a topological space X is said to be a ***topological property*** if it is satisfied by any other topological space Y which is homeomorphic to X.

Remark 2.6.2.
In the study of topology, we mainly consider topological properties of a topological space. Therefore, if two spaces are homeomorphic, they cannot be distinguished from the topological point of view.

Chapter 3

Homotopy and Fundamental Groups

From this chapter, we begin to study spaces from the topological point of view. In the other words, we consider topological spaces and study what are not changed (also called topological invariants) when we modify the space globally or locally in a continuous way. The first topological invariant that we would like to introduce in this chapter is the fundamental group.

3.1 Homotopy

Continuous deformations are everywhere either in our daily life or in various areas of mathematics. For example, folding a piece of paper and blowing up a balloon are continous deformations of a piece of paper and balloon respectively. On the other hand, neither tearing a piece of paper into small prices nor puncturing a balloon is a continuous deformation.

Since we are going to considering spaces under continuous deformations, we start this chapter by making a mathematical description, which relates to the notion of homotopy.

Definition 3.1.1

Two continuous maps f and g from the space X to the space Y are said to be ***homotopic*** if there is a continuous map

$$\begin{aligned} H : X \times [0,1] &\to Y, \\ (x,t) &\mapsto H(x,t), \end{aligned}$$

such that for each $x \in X$, we have

$$H(x,0) = f(x), \quad H(x,1) = g(x).$$

We call such a map H a ***homotopy*** between f and g, and for each $t \in [0,1]$, we have the map

$$\begin{aligned} H_t : X &\to Y, \\ x &\mapsto H(x,t). \end{aligned}$$

Example 3.1.2 (Path as a homotopy).
Let $X = \{x\}$ be a single point set. Then any point $y \in Y$ can be consider as the image of the map

$$\begin{aligned} f : \{x\} &\to Y \\ x &\mapsto y \end{aligned}.$$

Notice that such maps are always continuous.

Given two points p and q in Y, we consider the map f and g from $\{x\}$ to Y, such that $f(x) = p$ and $g(x) = q$.

Consider the map

$$\begin{aligned} \phi : [0,1] &\to \{x\} \times [0,1] \\ t &\mapsto (x,t) \end{aligned}.$$

This is a homeomorphism.

If f and g are homotopic to each other, we have a continuous map

$$H : \{x\} \times [0,1] \to Y,$$

such that $H(x,0) = f(x)$ and $H(x,1) = g(x)$ (hence for any $x' \in \{x\}$).

If we define

$$\begin{aligned} \gamma : [0,1] &\to Y \\ t &\mapsto \gamma(t) := H(x,t) \end{aligned},$$

then $\gamma = H \circ \phi$ is continuous with $\gamma(0) = p$ and $\gamma(1) = q$. Hence it is a path in Y connecting p and q.

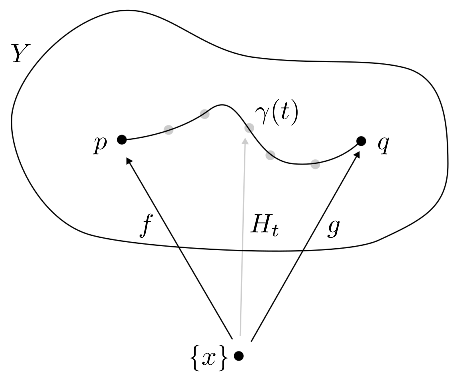


Figure 3.1.1: A path as a homotopy

Conversely, if p and q can be connected by a path

$$\gamma : [0,1] \to Y,$$

such that $\gamma(0) = p$ and $\gamma(1) = q$, then we can define a map

$$\begin{aligned} H : \{x\} \times [0,1] &\to Y \\ (x,t) &\mapsto H(x,t) := \gamma(t). \end{aligned}$$

Hence $H = \gamma \circ \phi^{-1}$ is continuous with $H(x,0) = p = f(x)$ and $H(x,1) = q = g(x)$ (hence for any $x' \in \{x\}$). Therefore H is a homotopy between f and g continuous maps from X to Y. (See Figure for an illustration of the above discussions.)

Informally speaking, a homotopy between two maps f and g from X to Y can be understood in the following way. Instead of two points (end points of a path) in Y, we consider two subsets $f(X)$ and $g(X)$ of Y "parametrized" [1] by a model space X. Instead of moving a point from one

[1] Although the maps f and g may not be injective, we borrow the terminology here.

place to another, we now move the all points in $f(X)$ to points in $g(X)$ at the same time in a continuous way.

Notice that for any point x in X, we have a continuous map

$$\begin{aligned} H_x : [0,1] &\to Y, \\ t &\mapsto H(x,t). \end{aligned}$$

In other words, the point $f(x)$ is moved in Y continuously as time passing and stop at $g(x)$ at time 1. (See Figure 3.1.2 for an illustration.)

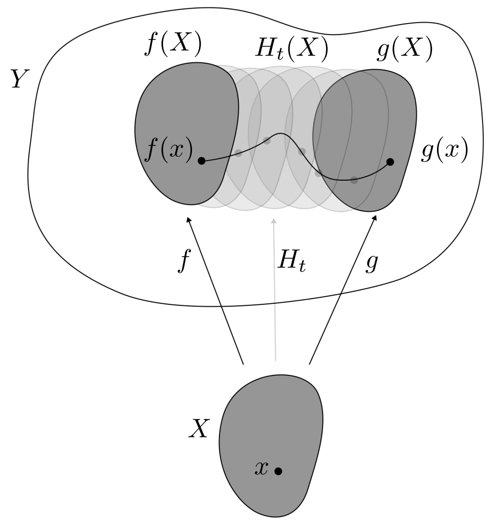


Figure 3.1.2: An illustration of a homotopy between f and g

Of course, for the map H, being continuous is a stronger condition than having all maps H_x's continuous, (or every point $f(x)$ being moved in a continuous way). Notice that the homotopy produces, at each moment t, a subspace $H_t(X)$ of Y, which is again "parametrized" by X. When we deform $f(X) = H_0(X)$ to $g(X) = H_1(X)$, there are certain properties should also be preserved. For example, if X is connected, then all $H_t(X)$ should be connected. There are other things that should be considered as well.

To be more precise, we may consider the set of all continuous maps from X to Y, associate to it a topology and talk about path in it. The topology is called the open compact topology, and its precise description will be given below. The rough idea is that we want not only to moving points continuously, but also to move any compact subset of X continuously (a single point subset is always compact).

Open-compact topology

More precisely, assume that X is locally compact, and let

$$\mathcal{C}(X,Y) := \{\text{continuous maps from } X \text{ to } Y\},$$

denote the set of all continuous maps from X to Y. Let K be any compact subset of X, and U be any open subset of Y. We consider the following subset of $\mathcal{C}(X,Y)$:

$$V(K,U) := \{f \in \mathcal{C}(X,Y) \mid f(K) \subset U\}.$$

The open compact topology on $\mathcal{C}(X,Y)$ is generated by the following subset (as a subbasis)

$$\{V(K,U) \mid K \subset X \text{ is compact and } U \subset Y \text{ is open}\}.$$

We now try to understand why

$$\begin{aligned} \Phi : [0,1] &\to \mathcal{C}(X,Y) \\ t &\mapsto H_t \end{aligned}.$$

is a path in $\mathcal{C}(X,Y)$. Let us assume that X and Y are both Hausdorff for simplicity. We consider an open set $V(K,U)$ in the subbasis for some compact subset $K \subset X$ and some open subset $U \subset Y$. Now we would like to show that $\Phi^{-1}(V(K,U))$ is open in $[0,1]$. In the other words, for any $t \in \Phi^{-1}(V(K,U))$, there is an $\epsilon > 0$, such that

$$\{s \in [0,1] \mid |s-t| < \epsilon\} \subset \Phi^{-1}(V(K,U)).$$

This is not difficult to understand if $K = \{x\}$ for some $x \in X$. Now consider the general case and we will discuss by contradiction. Assume the above fact is not true. Then for any ϵ, there is an $s \in [0,1]$, such that $|s-t| < \epsilon$ and $s \notin \Phi^{-1}(V(K,U))$. In particular, we consider $\epsilon = (n+1)^{-1}$ and denote $s_n \in [0,1]$ associated to it. Notice that

$$S = \{s_n \mid n \in \mathbb{N}\} \cup \{t\}.$$

is compact. For any s_n, we have

$$H_{s_n}(K) \cap U^c \neq \emptyset.$$

Hence we have $(x_n, s_n) \in K \times S$, such that

$$H(x_n, s_n) \in U^c.$$

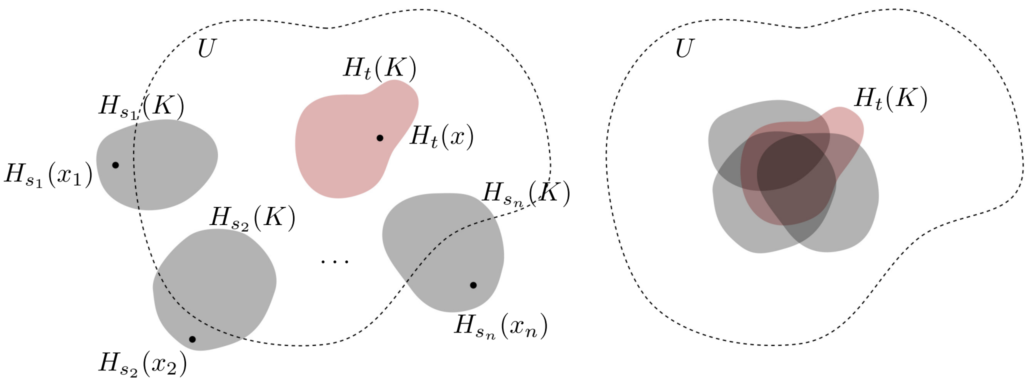


Figure 3.1.3: Not continuous (Left); Continuous (Right)

Now since K and S are compact, the sequence $(x_n, s_n)_{n\in\mathbb{N}}$ has a limit point $(x,t) \in K \times S$. On the other hand, by the continuity of H, we have

$$H(K \times S) \subset Y,$$

compact. Since H is continuous at (x,t), for any neighborhood Z of $H(x,t) \in Y$, there is a neighborhood W of (x,t) in $K \times S$, such that

$$H(W) \subset Z.$$

Since (x,t) is a limit point of $\{(x_t, s_t) \in K \times S \mid n \in \mathbb{N}\}$, we have

$$Z \cap \{H(x_n, s_n) \mid n \in \mathbb{N}\} \neq \emptyset.$$

This shows that $H(x,t)$ is a limit point of $\{H(x_n, s_n) \mid n \in \mathbb{N}\}$ which is contained in U^c. On the other hand U^c is closed, hence

$$H(x,t) = H_t(x) \in H_t(K) \cap U^c \subset U \cap U^c = \emptyset,$$

which is impossible (See Figure 3.1.3 for an illustration).

Examples

We give some elementary examples of homotopies between maps.

Example 3.1.3 (Constant homotopy).
An immediate observation from the definition of homotopy is that any continuous map is homotopic to itself. More precisely, let X and Y be two spaces and

$$f : X \to Y$$

be a continuous map. We consider the map

$$\begin{aligned} H : X \times [0,1] &\to Y \\ (x,t) &\mapsto f(x) \end{aligned}.$$

which is continuous with $H(x,0) = f(x)$ and $H(x,1) = f(x)$ for any $x \in X$. We call it a *constant homotopy* (See Figure 3.1.4 for an illustration).

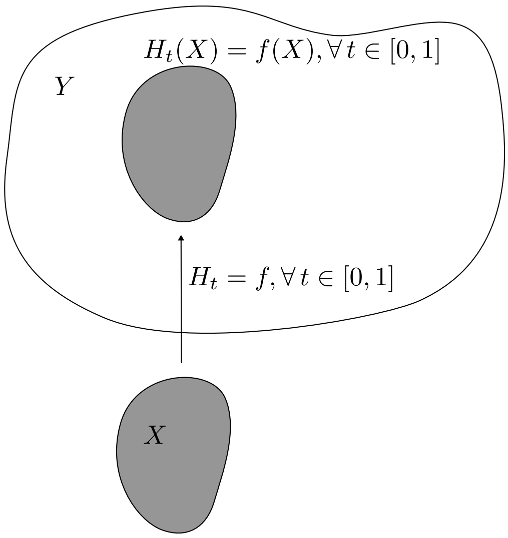


Figure 3.1.4: An illustration of a constant homotopy

Example 3.1.4 (Cylinder in $\mathbb{R}^3$).
Now let us consider an example where S^1 is not a single point set. Consider the following cylinder in $\mathbb{R}^3$ described using the coordinates of $\mathbb{R}^3$:

$$C = \{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 = 1,\ 0 \leq z \leq 1\}.$$

Consider the unit circle S^1 in $\mathbb{R}^2$:

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

We can define two continuous maps

$$\begin{aligned} f : S^1 &\to C \\ (x, y) &\mapsto (x, y, 0) \end{aligned},$$

and

$$\begin{aligned} g : S^1 &\to C \\ (x, y) &\mapsto (x, y, 1) \end{aligned}.$$

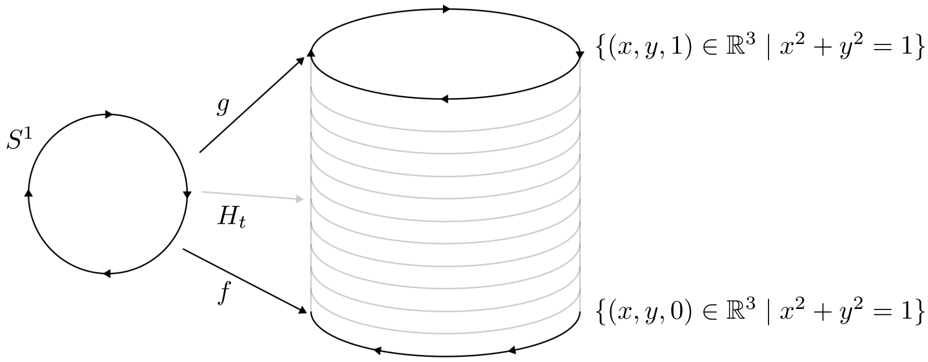


Figure 3.1.5: Moving S^1 from the bottom of the cylinder to the top.

Moving the circle from one boundary of C to another in a parallel way gives a homotopy between f and g. More precisely, we consider the map

$$\begin{aligned} H : S^1 \times [0, 1] &\to C \\ ((x, y), t) &\mapsto (x, y, t) \end{aligned}.$$

Notice that this is a continuous map (in fact a homeomorphism) with

$$H((x, y), 0) = f(x, y) \text{ and } H((x, y), 1) = g(x, y)$$

for any $(x, y) \in S^1$.

So far in all examples all maps H_t's are homeomorphisms which is not necessary for a homotopy. Let us see one simple example.

Example 3.1.5 (A disk shrinks to a point).
Consider the closed unit disk in $\mathbb{R}^2$:

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

We consider the identity map:

$$\begin{aligned} \mathrm{id} : D^2 &\to D^2 \\ (x, y) &\mapsto (x, y) \end{aligned}$$

and the constant map:

$$\begin{aligned} \mathrm{Const} : D^2 &\to D^2 \\ (x,y) &\mapsto (0,0) \end{aligned}.$$

Both of them are continuous maps, which is true for any topological spaces. We can define the following maps

$$\begin{aligned} H : D^2 \times [0,1] &\to D^2 \\ ((x,y),t) &\mapsto t(x,y) \end{aligned}.$$

This is a continuous map with $H((x,y),0) = \mathrm{id}(x,y)$ and $H((x,y),1) = \mathrm{Const}$ for any $(x,y) \in D^2$. Hence we have a homotopy between id and Const (See Figure 3.1.6).

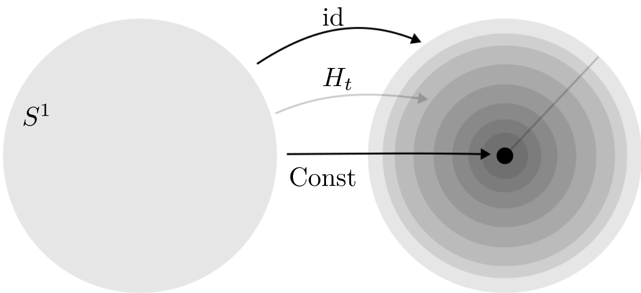


Figure 3.1.6: A disk shrinks to a point.

Example 3.1.6 (Homotopy of a path).
Take one rubber rope and fix two ends on the ground. The rope gives a path on the ground. We may move the rope without moving the two ends on the grounds. This gives homotopies between paths.

More precisely, consider γ_0 and γ_1 are two path in $\mathbb{R}^2$ such that

$$\gamma_0(0) = \gamma_1(0) = p \text{ and } \gamma_0(0) = \gamma_1(0) = q.$$

We can define a homotopy in the following way:

$$\begin{aligned} H : [0,1] \times [0,1] &\to \mathbb{R}^2 \\ (s,t) &\mapsto \gamma_0(s) + t(\gamma_1(s) - \gamma_0(s)) \end{aligned}.$$

(See Figure 3.1.7 for an illustration.)

Notice that by the definition we have in particular

$$H(0,t) = p \text{ and } H(1,t) = q,$$

for any $t \in [0,1]$, which means that the end points are fixed during the homotopy process.

Remark 3.1.7.
In the above examples, the constructions of the homotopies uses the linear structure of $\mathbb{R}^n$ which may not exist in other topological spaces.

A homotopy between two continuous maps is not unique

The key point in the definition of being homotopic is the **existence** of the homotopy map H, rather than the map H itself. Normally such a homotopy map H is never unique for several reasons.

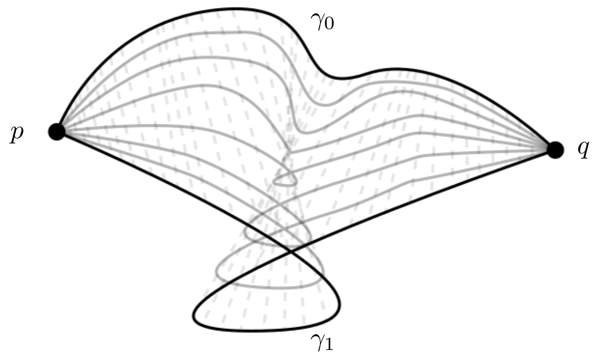


Figure 3.1.7: A homotopy between two paths with end points fixed.

-Reparametrization of a homotopy The maps H_t's changes as the time t passes. Consider the following map

$$\varphi : [0,1] \to [0,1],$$

which is continuous and increasing with $\varphi(0) = 0$ and $\varphi(1) = 1$. Then the following map

$$\begin{aligned} \Phi : X \times [0,1] &\to X \times [0,1] \\ (x,t) &\mapsto (x, \varphi(t)) \end{aligned}$$

is continuous with $\Phi(x, 0) = (x, 0)$ and $\Phi(x, 1) = (x, 1)$. We call it a *reparametrization map*.

Given f and g two homotopic continuous maps from X to Y, and denote by H a homotopy between them, the composition $\widetilde{H} = H \circ \Phi$ is again a homotopy between f and g. In this case, for any $t \in [0,1]$, we have

$$\widetilde{H}_t = H_{\varphi(t)}.$$

For example, we consider

$$\begin{aligned} \varphi : [0,1] &\to [0,1] \\ t &\mapsto \max\{0, 2t-1\} \end{aligned}.$$

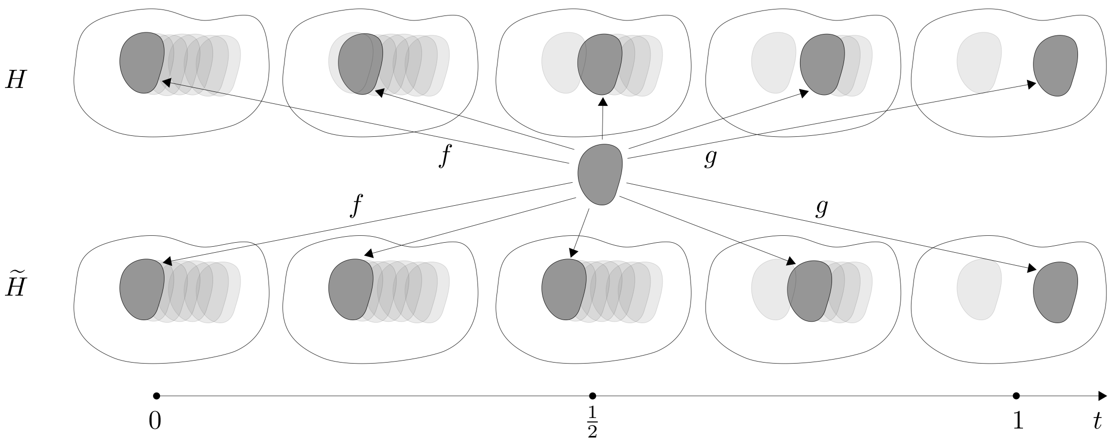


Figure 3.1.8: A reparametrization of a homotopy.

Then the new homotopy map $\widetilde{H}$ can be given as follows: for any $(x,t) \in X \times [0,1]$, we have

$$\widetilde{H}(x,t) = \begin{cases} f(x) & t \in \left[0, \frac{1}{2}\right] \\ H(x, 2t-1) & t \in \left[\frac{1}{2}, 1\right] \end{cases}$$

Roughly speaking, for $t \in [0,1/2]$, the map f is not changed, while for $t \in [1/2,1]$, the map f is changed to the map g according to the homotopy H, but with double speed (See Figure 3.1.8 for an illustration).

-Different choices of homotopy Another reason is that the collection of maps $\{H_t\}_{t\in[0,1]}$ could be different. This is not difficult to understand, if we check Example 3.1.6. The set $\{H_t\}_{t\in[0,1]}$ used to pass from γ_0 and γ_1 could be quite random and far from being unique (See Figure 3.1.9 for an illustration).

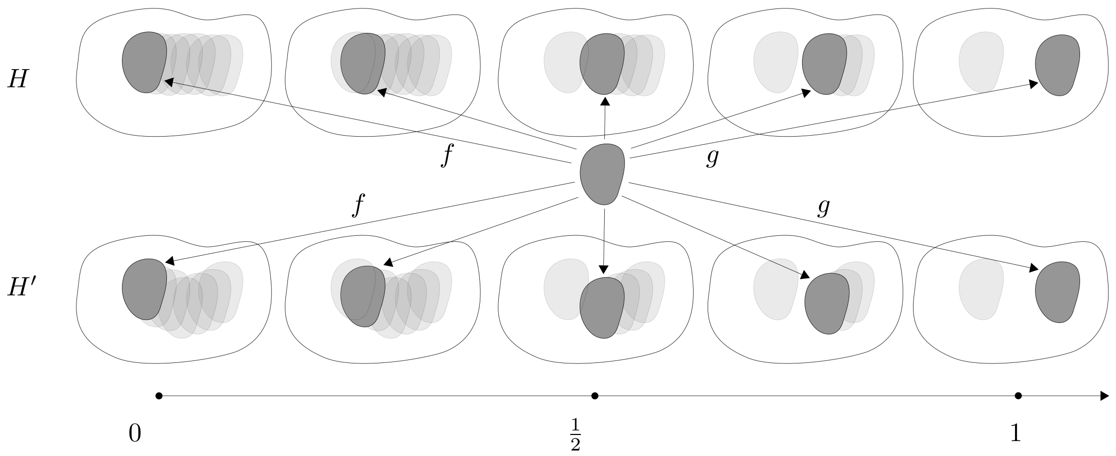


Figure 3.1.9: A different homotopy between two maps.

Inverse of a homotopy

Another thing that one may notice is that although the definition of a homotopy between maps from X to Y relies on a time parameter which seems to give an direction for the homotopy process, the notion of being homotopic is symmetric.

To be more precise, let f and g be two continuous maps from a space X to a space Y homotopic to each other, and let H be a homotopy between them as in the above definition. We can define another map

$$\overline{H} : X \times [0,1] \to Y,$$
$$(x,t) \mapsto H(x, 1-t).$$

Roughly speaking the map $\overline{H}$ gives a deformation which is given by backward playing the deformation given by H (See Figure 3.1.10 for an illustration).

Definition 3.1.8

The map $\overline{H}$ is called the ***inverse*** of H.

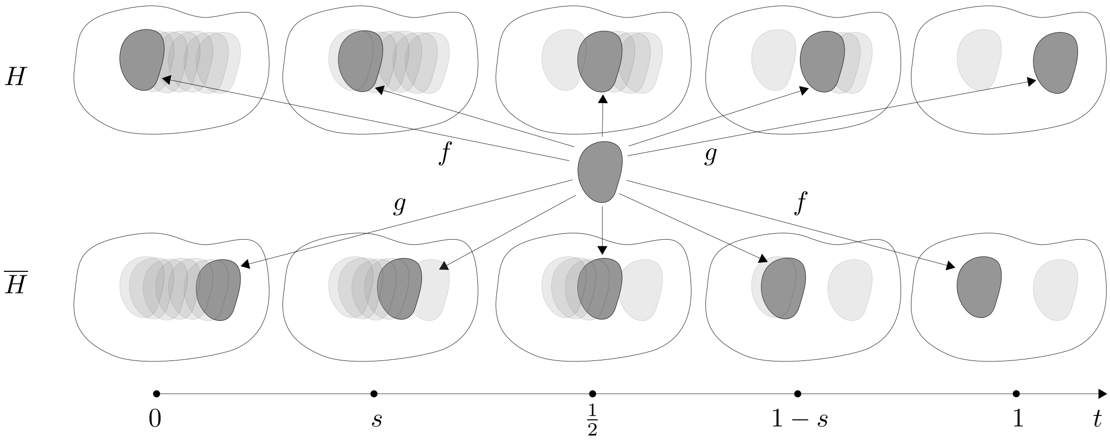


Figure 3.1.10: The inverse of a homotopy.

Induced equivalence relation on the set of continuous maps

We consider the set of all continuous maps from X to Y, denoted by

$$\mathcal{C}(X,Y) := \{\text{Continuous maps from } X \text{ to } Y\} \subset \mathcal{P}(X \times Y).$$

Proposition 3.1.9

The following relation on $\mathcal{C}(X,Y)$ is an equivalence relation: for any $f, g \in \mathcal{C}(X,Y)$,

$$f \sim g \Leftrightarrow f \text{ and } g \text{ are homotopic.}$$

Proof. We have to verify the three properties: reflexivity, symmetry and transitivity. We will omit the verification of the continuity of all homotopies appearing below.

Firstly, given any $f \in \mathcal{C}(X,Y)$, we consider the constant homotopy (see Example 3.1.3)

$$\begin{aligned} H : X \times [0,1] &\to Y, \\ (x,t) &\mapsto f(x) \end{aligned}$$

and have $f \sim f$.

Secondly, by considering the inverse of a homotopy (see Definition 3.1.8), if two maps f and g in $\mathcal{C}(X,Y)$ satisfy $f \sim g$, then $g \sim f$.

Finally, let f, g and h be three maps in $\mathcal{C}(X,Y)$. Assume that $f \sim g$ and $g \sim h$. We denote by F be the homotopy between f and g and G be the homotopy between g and h, then we can define the following map H: for any $(x,t) \in X \times [0,1]$

$$H(x,t) = \begin{cases} F(x,2t) & t \in \left[0, \dfrac{1}{2}\right] \\ G(x,2t-1) & t \in \left[\dfrac{1}{2}, 1\right] \end{cases}$$

and this is a homotopy between f and h (see Figure 3.1.11 for an illustration).

□

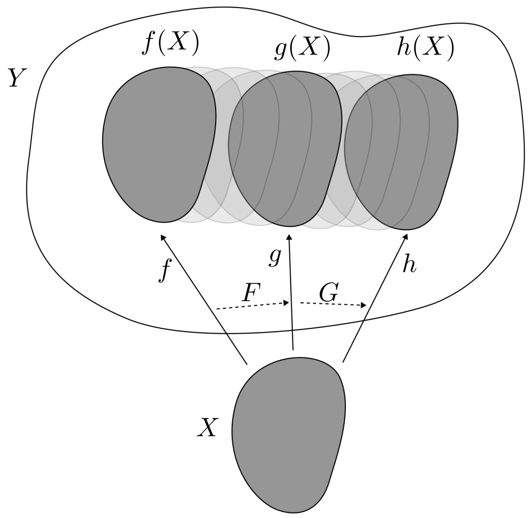


Figure 3.1.11: A composition between two homotopies.

Remark 3.1.10.
In the proof, we construct a homotopy H from the homotopy F and G. Informally speaking, if we can modify f to g in a continuous way and modify g to h in a continuous way, then we modify f to h in a continuous way through g. We denote the homotopy H constructed in the proof by

$$H = F * G,$$

(first F, then G) and call it the ***composition*** between F and G.

Definition 3.1.11

A space X is said to be ***contractible***, if the identity map

$$\begin{aligned} \mathrm{id}_X : X &\to X \\ x &\mapsto x \end{aligned}$$

is homotopic to a constant map

$$\begin{aligned} \mathrm{Const}_c : X &\to X \\ x &\mapsto c \end{aligned}$$

where $c \in X$.

Example 3.1.12 (Star-like subsets in $\mathbb{R}^2$).
A subset D in $\mathbb{R}^2$ is star-like if there is a point $c \in D$, such that for which $p \in D$, we have

$$\{c + t(p - c) \mid t \in [0, 1]\} \subset D.$$

We call such a point c a center of D (see Figure 3.1.12).

A convex subset of $\mathbb{R}^2$ is in particular is a star-like subset. In a convex subset, we may choose any point as the center.

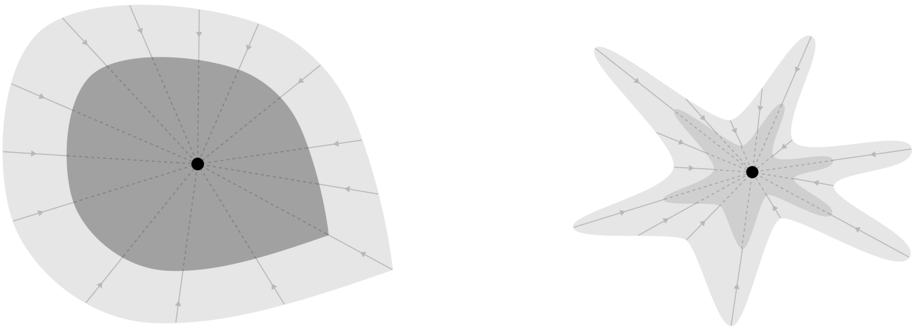


Figure 3.1.12: A convex (also star-like) set (left); a non-convex star-like set (right).

Not every star-like region is convex. For example, if we consider the union D of any three distinct rays issued from $(0,0)$. Then D is not convex, but $(0,0)$ could be the center to make D star-like.

Similar to the connectedness property, we also have a local version of this notion.

Definition 3.1.13

A space X is said to be ***locally contractible*** if every point $x \in X$ admits a neighborhood basis consisting of only contractible set.

Remark 3.1.14.
All manifolds are locally contractible.

3.2 Homotopy equivalence

We know that two homeomorphic spaces are topologically equivalent, meaning that we cannot distinguish them by any topological method. However, for some topological properties, being homeomorphic is too strong. Using homotopy we can given a weaker equivalence relation among topological spaces which are more suitable for studying certain topological properties or topological invariants.

Definition 3.2.1

Two spaces X and Y are said to be ***homotopy equivalent*** if there exists continuous maps

$$f : X \to Y, \quad g : Y \to X,$$

such that $f \circ g \sim \mathrm{id}_Y$ and $g \circ f \sim \mathrm{id}_X$.

Example 3.2.2.
If X and Y are homeomorphic to each other, we have a homeomorphism

$$f : X \to Y.$$

This means that there is a continuous map

$$g : Y \to X,$$

such that

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

Considering the constant homotopy (See Example 3.1.3), given any continuous map h from a topological space W to itself, we have h homotopic to h. Therefore, two homeomorphic topological spaces are homotopy equivalent to each other.

The two examples below show that being homotopy equivalent is a condition strictly weaker than being homeomorphic.

Example 3.2.3.
Let X be the closed unit disk in $\mathbb{R}^2$ and Y be $\{O\}$ with O the origin. Since X is uncountable and Y is finite, they cannot be homeomorphic to each other. Consider the following two maps

$$\begin{aligned} f : X &\to Y \\ p &\mapsto O \end{aligned} \quad \text{and} \quad \begin{aligned} g : Y &\to X \\ O &\mapsto O \end{aligned}$$

Then

$$f \circ g = \mathrm{id}_Y,$$

and

$$g \circ f = f$$

which is homotopic to id_X as shown in Example 3.1.5 (see Figure 3.2.1). Hence X and Y are homotopy equivalent.

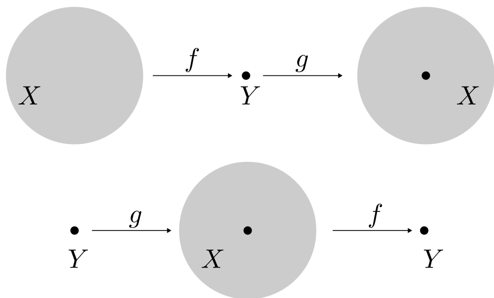


Figure 3.2.1: Homotopy equivalence between the unit disk and its center.

Example 3.2.4.
Similar, we consider X to be whole space $\mathbb{R}^2$ and Y be $\{O\}$ the origin. Consider the following two maps

$$\begin{aligned} f : X &\to Y \\ p &\mapsto O \end{aligned} \quad \text{and} \quad \begin{aligned} g : Y &\to X \\ O &\mapsto O \end{aligned}$$

Then same as in the previous example, we have

$$f \circ g = \mathrm{id}_Y,$$

and

$$g \circ f = f$$

which is homotopic to id_X by the following homotopy:

$$\begin{aligned} H : X \times [0,1] &\to X \\ (p,t) &\mapsto tp \end{aligned}.$$

If we examine the topological properties of these spaces, we have the following table:

	connected	path connected	compact	contractible
$\{O\}$	Yes	Yes	Yes	Yes
D^2	Yes	Yes	Yes	Yes
$\mathbb{R}^2$	Yes	Yes	No	Yes

Example 3.2.5.
Let X be the annulus in $\mathbb{C}$ defined by

$$X := \{z \in \mathbb{C} \mid 1 \le |z| \le 2\},$$

and Y be the unit circle.

First we would like show that X and Y are not homeomorphic to each other. This follows from the following facts: a restriction of a bijection is bijective to its image, and a restriction of a continuous map is continuous. Hence a restriction of a homeomorphism is a homeomorphism to its image.

We can remove a pair of antipodal points on Y and the resulting space is no longer connected. On the other hand, removing two points in X will not disconnect the space. Since being connected is preserved by continuous maps and in particular by homeomorphisms, we conclude that there is no homeomorphism between X and Y.

Secondly, we show that X and Y are homotopy equivalent to each other. Let f be the following map

$$\begin{aligned} f : X &\to Y \\ p &\mapsto \frac{p}{|p|}, \end{aligned}$$

where $|p|$ is the Euclidean distance between p and the origin, and g be the inclusion map

$$\begin{aligned} g : Y &\to X \\ p &\mapsto p \end{aligned}$$

See Figure 3.2.2.

We have on one hand

$$f \circ g = \mathrm{id}_Y.$$

At the same time, we can define the homotopy

$$\begin{aligned} H : X \times [0,1] &\to X, \\ (re^{i\theta}, t) &\mapsto (t + (1-t)r)e^{i\theta}, \end{aligned}$$

between $g \circ f$ and id_X (see Figure 3.2.3). Hence X and Y are homotopy equivalent.

Similar to homeomorphisms, just as its name suggests, in any set of topological spaces, homotopy equivalence induces an equivalence relation.

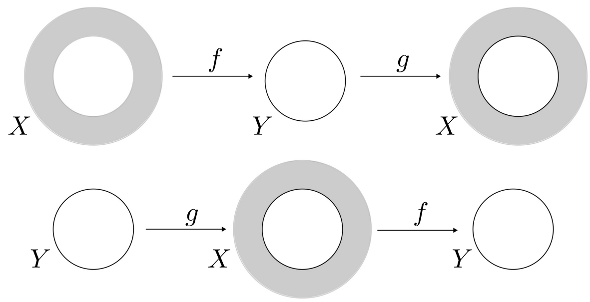


Figure 3.2.2: Homotopy equivalence between an annulus and the unit circle.

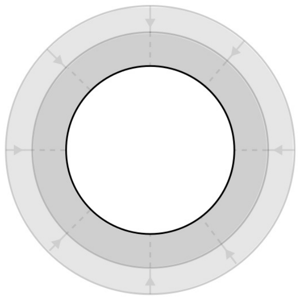


Figure 3.2.3: Homotopy between $g \circ f$ and id_X.

Proposition 3.2.6

Given any non empty set $\mathcal{X}$ of topological spaces, the following relation is an equivalent relation: for any $X, Y \in \mathcal{X}$

$$X \sim Y \Leftrightarrow X \text{ and } Y \text{ are homotopy equivalent.}$$

Proof. For any $X \in \mathcal{X}$, we consider the identity map on X. Using the constant homotopy, we have $X \sim X$.

By the symmetry of the definition of the homotopy equivalence, for any $X, Y \in \mathcal{X}$, if $X \sim Y$, then $Y \sim X$.

Let X, Y and Z be three topological spaces in $\mathcal{X}$. Assume that X and Y are homotopy equivalent, at the same time Y and Z are homotopy equivalent. By definition, there are continuous maps

$$f_1 : X \to Y, \quad g_1 : Y \to X,$$
$$f_2 : Y \to Z, \quad g_2 : Z \to Y,$$

such that

$$f_1 \circ g_1 \sim \mathrm{id}_Y, \quad g_1 \circ f_1 \sim \mathrm{id}_X,$$
$$f_2 \circ g_2 \sim \mathrm{id}_Z, \quad g_2 \circ f_2 \sim \mathrm{id}_Y.$$

Now we consider the following continuous maps

$$f_2 \circ f_1 : X \to Z,$$
$$g_1 \circ g_2 : Z \to X.$$

Here we need the following technical lemma.

Lemma 3.2.7

Let W, W' and W'' be three topological spaces. If h and $\widetilde{h}$ are two continuous map from W to W' homotopic to each other, and h' and $\widetilde{h}'$ be two continuous maps from W' to W'' homotopic to each other, then we have

$$h' \circ h \sim \widetilde{h}' \circ \widetilde{h}.$$

Proof of Lemma 3.2.7. Since a composition of continuous maps is still continuous, we have

$$h' \circ h, \ \widetilde{h}' \circ \widetilde{h} \in \mathcal{C}(W, W'')$$

Let H be a homotopy between h and h', and $\widetilde{H}$ be a homotopy between $\widetilde{h}$ and $\widetilde{h}'$. We now construct the following map

$$F : W \times [0,1] \to W''$$

by defining

$$F(a,t) = \begin{cases} (h' \circ H)(a, 2t) & t \in \left[0, \frac{1}{2}\right] \\ \widetilde{H}(\widetilde{h}(a), 2t-1) & t \in \left[\frac{1}{2}, 1\right] \end{cases}$$

It is a continuous map, such that for any $a \in W$, we have

$$F(a,0) = (h' \circ h)(a) \text{ and } F(a,1) = (\widetilde{h}' \circ \widetilde{h})(a).$$

See Figure 3.2.4 for an illustration. Hence the lemma.

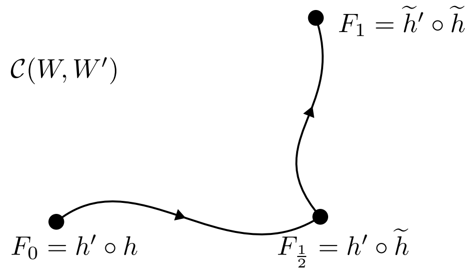


Figure 3.2.4: Homotopy in $\mathcal{C}(W, W')$ given by F.

□

By this lemma, in $\mathcal{C}(X, X)$ we have

$$\begin{aligned}(g_1 \circ g_2) \circ (f_2 \circ f_1) =& g_1 \circ (g_2 \circ f_2) \circ f_1 \\ \sim& g_1 \circ \mathrm{id}_Y \circ f_1 \\ =& g_1 \circ f_1 \\ \sim& \mathrm{id}_X.\end{aligned}$$

Similarly, in $\mathcal{C}(Z, Z)$ we have

$$\begin{aligned}(f_2 \circ f_1) \circ (g_1 \circ g_2) =& f_2 \circ (f_1 \circ g_1) \circ g_2 \\ \sim& f_2 \circ \mathrm{id}_Y \circ g_2 \\ =& f_2 \circ g_2 \\ \sim& \mathrm{id}_Z.\end{aligned}$$

Hence X and Y are homotopy equivalent. □

Definition 3.2.8

If a space Y is homotopy equivalent to a space X, we say that Y is of the ***homotopy type*** X.

To have an idea of properties preserved under homotopy equivalence, we give several non-examples to see what is preserved under the homotopy equivalence.

Example 3.2.9.
Consider the following subspaces of $\mathbb{C}$

$$X = \{1, -1\} \text{ and } Y = \{0\}.$$

We would like to show that X and Y are not homotopy equivalent.

Since Y has a single point, the only map f from X to Y is defined by sending both 1 and -1 to 0. On the other hand, a map from Y to X is determined by the image of 0. Without loss of generality, we may consider the map

$$g : Y \to X,$$

with $g(0) = 1$.

Notice that $f \circ g = \mathrm{id}_Y$. Now we turn to study $g \circ f$ and we will show that the composition $g \circ f$ is not homotopic to id_X. To see this, we assume that there is a homotopy

$$H : X \times [0, 1] \to X,$$

between $g \circ f$ and id_X, and see what would go wrong.

Consider the restriction H' of H on $\{-1\} \times [0, 1]$. Notice that this is a connected subspace of the product space $X \times [0, 1]$. Moreover, by the definition of H, we have

$$H(-1, 0) = 1, \quad H(-1, 1) = -1.$$

Hence the map

$$H' : \{-1\} \times [0, 1] \to X,$$

is continuous and surjective. However X is a not connected, which is a contradiction. (See Figure 3.2.5 for an illustration.)

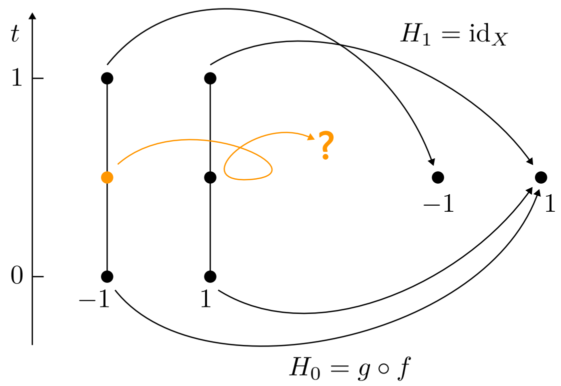


Figure 3.2.5: Where should the orange point go, if H is continuous?

In fact here we can use either $\{-1\} \times [0,1]$ is connected or path connected to get the contradiction. Now if we consider the point is a connected component or a path connected component, we can generalize this arguments to show that the connectedness and the path connectedness are both preserved by the homotopy equivalence.

Proposition 3.2.10

If two spaces X and Y are homotopy equivalent and X is connected, then Y is connected.

It can be considered as a corollary of the following proposition.

Proposition 3.2.11

If two spaces X and Y are homotopy equivalent, then they have the same number of connected components.

Proof. By definition of homotopy equivalence, there are continuous maps

$$f : X \to Y, \quad g : Y \to X,$$

such that $f \circ g \sim \mathrm{id}_Y$ and $g \circ f \sim \mathrm{id}_X$.

The decomposition of connected components of X and Y are denoted respectively by

$$X = \bigsqcup_{\alpha \in \mathcal{A}} X_\alpha, \quad Y = \bigsqcup_{\beta \in \mathcal{B}} Y_\beta.$$

Since f is continuous, for any $\alpha \in \mathcal{A}$, the image $f(X_\alpha)$ is connected hence is contained in some connected component Y_β of Y. This gives a map

$$\phi_f : \mathcal{A} \to \mathcal{B}.$$

We would like to show that this map is injective by contradiction. Assume that two distinct induces α and α' in $\mathcal{A}$ are mapped to β.

Since g is continuous, the image

$$(g \circ f)(X_\alpha \cup X_{\alpha'}) \subset g(Y_\beta)$$

is contained in the connected component $X_{\alpha''}$. Notice that we have either $\alpha \neq \alpha''$ or $\alpha' \neq \alpha''$. Without loss of generality, we may assume that $\alpha \neq \alpha''$.

Let H denote the homotopy between $g \circ f$ and id_X. Since

$$H(X_\alpha \times \{0\}) \subset X''_\alpha, \quad H(X_\alpha \times \{1\}) = X_\alpha,$$

we have

$$H(X_\alpha \times [0,1]) \cap X_\alpha \neq \emptyset, \quad H(X_\alpha \times [0,1]) \cap X''_\alpha \neq \emptyset.$$

Hence $H(X_\alpha \times [0,1])$ is not connected. On the other hand, the product space $X_\alpha \times [0,1]$ is connected. This contradict to the fact that H is continuous. Hence ϕ_f is injective.

Now we would like to show that ϕ_f is surjective. Assume that β is not in the image $\phi_f(\mathcal{A})$. Let $\alpha \in \mathcal{A}$ and $\beta' \in \mathcal{B}$, such that

$$g(Y_\beta) \subset X_\alpha \text{ and } f(X_\alpha) = Y_{\beta'}.$$

Notice that $\beta' \neq \beta$ by the hypothesis. Now we consider the homotopy F between $f \circ g$ and id_Y. Then since $Y_\beta \times [0,1]$ is connected, we have $F(Y_\beta \times [0,1])$ connected. Notice that

$$F(Y_\beta \times \{0\}) \subset Y_{\beta'} \text{ and } F(Y_\beta \times \{1\}) = Y_\beta,$$

which is a contradiction. Hence $\beta' = \beta$.

As a conclusion, the map ϕ_f is a bijective between $\mathcal{A}$ and $\mathcal{B}$. In particular, we have

$$|\mathcal{A}| = |\mathcal{B}|.$$

□

Remark 3.2.12.

In fact, we can consider

$$\phi_g : \mathcal{B} \to \mathcal{A}$$

constructed in a similar way as for ϕ_f. By a similar discussion, we can show that ϕ_g is a injective. Hence we have

$$|\mathcal{A}| = |\mathcal{B}|.$$

In the above proof, we show a stronger result that the map ϕ_f is bijective. In fact, we can work more to show that ϕ_f and ϕ_g are inverse to each other.

In particular, when X is connected, we have Y connected.

We also have a similar result for path connectedness and its generalization. They can be proved in an exact same way, by considering path connectedness instead of connectedness in the proof.

Proposition 3.2.13

If two spaces X and Y are homotopy equivalent and X is path connected, then Y is path connected.

Proposition 3.2.14

If two spaces X and Y are homotopy equivalent, they have the same number of path connected components.

Remark 3.2.15.

We have a remark for the path connectedness similar to Remark 3.2.12 for the connectedness.

Remark 3.2.16.

The above discussion shows that there are certain topological properties which are not only preserved by homeomorphisms, but also preserved by homotopy equivalence. Later we will see that the main object introduced in this chapter so called the fundamental group is also preserved under homotopy equivalence, up to isomorphism.

Remark 3.2.17.

We should mention that not all topological properties are preserved by homotopy equivalence. For example, the compactness is not always preserved by a homotopy equivalence. The real line and a single point are homotopic equivalent, yet the real line is not compact, while any space of a single point is.

Example 3.2.18.

Recall the topologist's sine curve. Let f defined on $I = (0, 1]$ by

$$f(x) = \sin\left(\frac{\pi}{x}\right).$$

We consider the following subspaces of $\mathbb{R}^2$

$$X = \text{Graph}(f), \quad Y = \overline{X}.$$

Notice that X is path connected, while Y is not. By Proposition 3.2.13, they are not homotopy equivalent.

3.3 Relative homotopy

Sometimes, we do not require a deformation globally on a space but only locally. Here comes the notion of relative homotopy.

Definition 3.3.1

Let X and Y be two topological spaces and A be a subspace of X. We say that maps

$$f : X \to Y \text{ and } g : X \to Y,$$

are homotopic relative to A, if there is a continuous map

$$H : X \times [0, 1] \to Y,$$

such that for any $x \in X$, we have

$$H(x, 0) = f(x) \text{ and } H(x, 1) = g(x),$$

moreover for any $x \in A$, for any $t \in [0, 1]$, we have $H(x, t) = f(x)$.

Remark 3.3.2.
Informally speaking, the map g is obtained from f by changing f-image of points in $X \setminus A$ in a continuous way.

One application of such homotopy is to simplify the space that we would like to discuss. In particular, we have the following several definition.

Definition 3.3.3

Let A ba a subspace of a topological space X, and

$$\iota : A \to X$$

be the inclusion map. We say that A is a ***retraction*** of X, if there is a continuous map

$$r : X \to A,$$

such that $r \circ \iota = \text{id}_A$. With the above notation,

1) if $\iota \circ r \sim \text{id}_X$, then we say that A is a ***deformation retraction*** of X;
2) if $\iota \circ r \sim \text{id}_X$ relative to A, then we say that A is a ***strong deformation retraction*** of X;

Example 3.3.4.
Let X be a two point subset of $\mathbb{R}^2$

$$X = \{p, q\},$$

and $A = \{p\}$. Then we consider

$$\begin{aligned} r : X &\to A \\ p &\mapsto p \\ q &\mapsto p \end{aligned}$$

and A is a retraction of X.

Example 3.3.5.
For any $n \in \mathbb{N}^*$, we consider the Euclidean space $\mathbb{R}^{n+1}$. Let X be the subspace $\mathbb{R}^{n+1} \setminus \{O\}$ and its subspace the n-sphere S^n. Consider the map

$$\begin{aligned} r : X &\to A, \\ x &\mapsto \frac{x}{|x|}, \end{aligned}$$

where $|x|$ is the Euclidean norm in $\mathbb{R}^{n+1}$.

Notice that $\iota \circ r = \mathrm{id}_A$, and we can verify that the following map:

$$\begin{aligned} H : X \times [0, 1] &\to X \\ (x, t) &\mapsto (1 - t)x + t\frac{x}{|x|}, \end{aligned}$$

is a homotopy relative to A between $\iota \circ r$ and id_X. Hence A is a strong deformation retraction of X.

The difference between retraction and deformation retraction is easy to tell. In particular, if A is a deformation retraction of X, then X is of the homotopy type of A, which is not always then case when A is only a retraction of X, as we can see in the above examples. Informally speaking, a retraction only care about the result, while a deformation retraction also care about how space retracts. In particular, when we focus on any point $x \in X$, its trace under this homotopy will be a path in X.

The difference between a deformation retraction and a strong deformation retraction is more difficult to tell.

Example 3.3.6.
We consider the topological space X in Example 2.4.32:

$$X = [0, 1] \times \{0\} \bigcup \{0\} \times [0, 1] \bigcup_{n \in \mathbb{N}^*} \left\{\frac{1}{n}\right\} \times [0, 1] \subset \mathbb{R}^2.$$

There is a strong deformation retraction of X to $\{(0, 0)\}$. This can be realized by first taking a strong deformation retraction from X to

$$[0, 1] \times \{0\},$$

then applying a strong deformation retraction of this horizontal segment to $(0, 0)$ (see Figure 3.3.1 for an illustration). In fact this also shows that there is a strong deformation retraction from X to any point on

$$[0, 1] \times \{0\}.$$

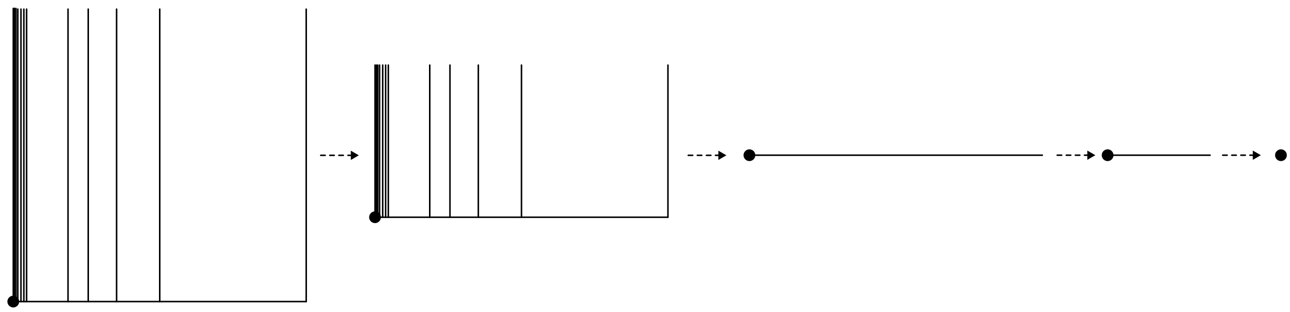


Figure 3.3.1: A strong deformation retraction to $(0,0)$.

Using this strong deformation retraction, we can moreover show that the space X has a deformation retraction to any of its point.

Let $A = \{(0,1)\}$. Now we would like to show that A is not a strong deformation retraction of X. For any $n \in \mathbb{N}^*$, we denote by p_n the point $(1/n, 1)$ and denote by p_∞ the point $(0,1)$.

Notice that the sequence $(p_n)_{n\in\mathbb{N}^*}$ converges to p_∞ as n goes to infinity in X.

Assume that the deformation retraction of X to A is given by some map r, and H is the homotopy between $\iota \circ r$ and id_X. If this is a homotopy relative to A, then we should have

$$H(p_\infty, t) = p_\infty$$

for all $t \in [0,1]$. We will show that this is impossible.

For each $n \in \mathbb{N}^*$, consider the subset

$$\{p_n\} \times [0,1],$$

whose image under H is path connected. Notice that

$$H(p_n, 0) = p_n,$$

and

$$H(p_n, 1) = p_\infty.$$

Hence there is a time $t_n \in [0,1]$, such that

$$H(p_n, t_n) = (0,0),$$

since every path connecting p_n and p_∞ must pass $(0,0)$ (see Figure 3.3.2). Now we consider the sequence

$$(p_n, t_n)_{n\in\mathbb{N}^*}.$$

Since X is a compact subspace of $\mathbb{R}^2$ (hence is sequential compact), there is a convergent subsequence whose limit point is

$$(p_\infty, s) \in X \times [0,1],$$

such that $H(p_\infty, s) = (0,0)$ by the continuity of H. Hence H is not a homotopy relative to $A = \{p_\infty\}$, and A is not a strong deformation retraction of X.

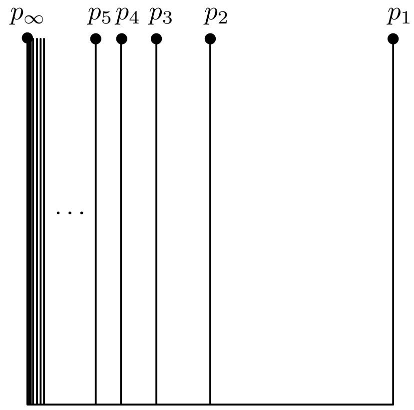


Figure 3.3.2: A deformation retraction to $\{p_\infty\}$ is not strong.

3.4 Path homotopy

Given one topological space, it is easy to tell that there is a difference between a (path) connected space and a space which is not (path) connected. However, only knowing a space is (path) connected is still not enough to characterize (even roughly) a space.

From now on, we will focus on the path connectedness. Let D^2 be the unit disk in $\mathbb{R}^2$. Notice that it is path connected. Even when we remove the center, the rest part denoted by X is still path connected. However, there is some difference between D^2 and X. For example, we consider p and q two points in D^2 different from the center. Notice that given any path in a space, it always has a strong deformation retraction to one of its end point. Hence either in D^2 or X, two paths from p to q are always homotopic to each other.

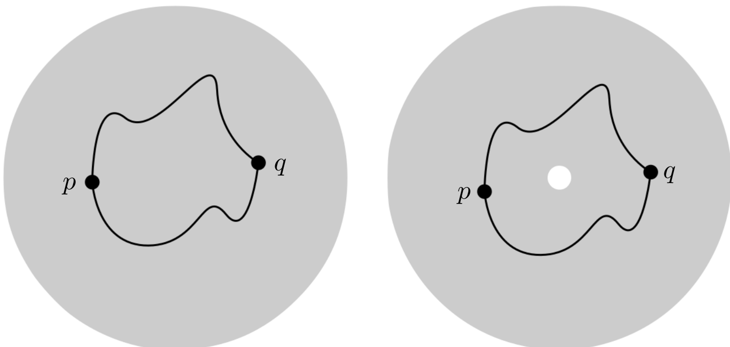


Figure 3.4.1: Different "path connected spaces".

However, if we fix endpoints when we perform a homotopy on a path, things are different. In D^2, we can still deform one path form p to q to another one in a continuous way. This is no long the case when we consider paths in X. There are certain pair of path from p to q, such that when we try to deform one to the other continuously, we must pass the center which has been removed.

This observation gives us a way to have a next level classification in the category of path connected spaces. Notice that when we deform a path fixing its endpoints, we actually perform a relative homotopy. To be more precise, recall that given any topological space X, a path in X is

a continuous map
$$\alpha : [0,1] \to X.$$

Definition 3.4.1

Consider two paths α and β in X with
$$\alpha(0) = \beta(0) = p,$$
$$\alpha(1) = \beta(1) = q.$$

A ***path homotopy*** between α and β is a continuous map
$$H : [0,1] \times [0,1] \to X,$$
such that for any $s, t \in [0,1]$, we have
$$H(s,0) = \alpha(s), \quad H(s,1) = \beta(s),$$
$$H(0,t) = p, \quad H(1,t) = q.$$

Remark 3.4.2.
In other words, a path homotopy between α and β is a homotopy relative to $\{0,1\}$. In Example 3.1.6, the two paths γ_0 and γ_1 are path homotopic.

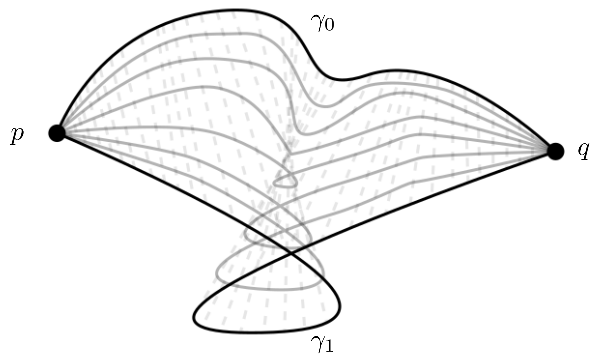


Figure 3.4.2: The paths γ_0 and γ_1 are path homotopic.

From now on, unless specified, when we say that two paths with the same endpoints are homotopic, we mean that they are path homotopic, and we denote this by $\alpha \sim \beta$.

Since the path homotopy is a special kind of homotopy, most of the discussions made before for homotopy still work here.

First notice that for each time parameter $t \in [0,1]$, the map H_t is also a path from p to q. Secondly, given two paths in X homotopic to each other, the homotopy is not unique for the same reasons as before.

Reparametrizations give new homotopies between two given paths.

Definition 3.4.3

A ***reparametrization*** of α is given by $\alpha \circ \varphi$, where

$$\varphi : [0,1] \to [0,1]$$

is a increasing continuous map with $\varphi(0) = 0$ and $\varphi(1) = 1$.

Figure 3.4.3 is an illustration of a reparametrization.

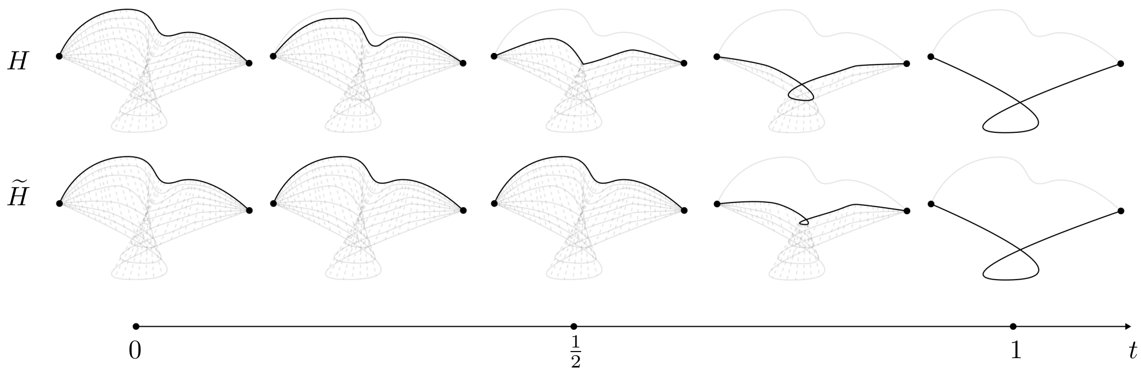


Figure 3.4.3: A reparametrization of a homotopy between two paths.

Two homotopies between two given paths can also be completely different, meaning that the collection of path $\{H_t\}_{t\in[0,1]}$ are different. See Figure 3.4.4 for an illustration.

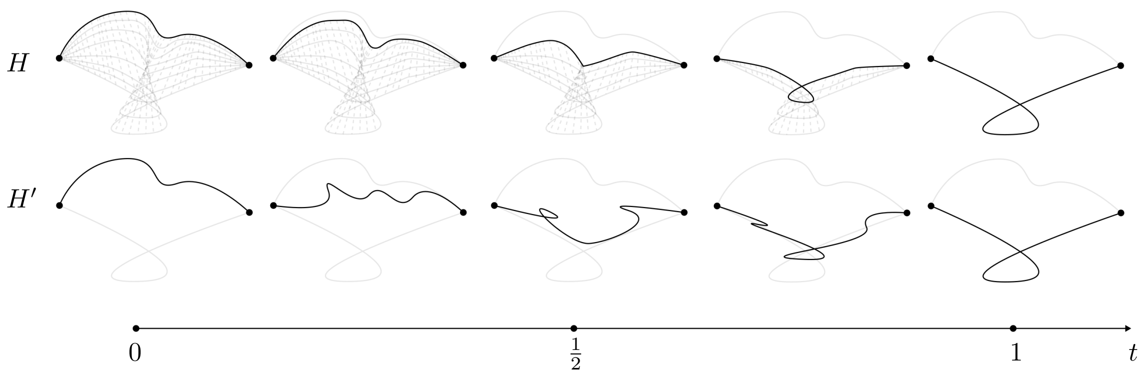


Figure 3.4.4: A "completely different" homotopy between two paths.

Inverse of a path homotopy

We can also define the inverse of a path homotopy.

Definition 3.4.4

The ***inverse*** of the path α is defined to be the following path

$$\overline{\alpha} : [0,1] \to X, \quad t \mapsto \alpha(1-t).$$

Figure 3.4.5 is an illustration of the inverse of a path homotopy.

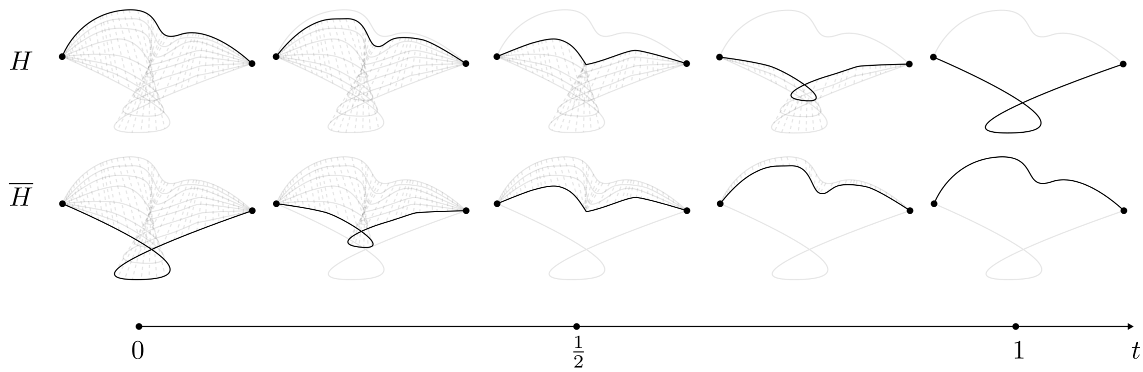


Figure 3.4.5: The inverse of a homotopy between two paths with end points fixed.

Induced equivalence relation on the set of path with same endpoints

Given any pair of points p and q in a path connected space X, we denote

$$\mathcal{P}(X,p,q) := \{\text{paths in } X \text{ going } p \text{ to } q\}.$$

As a special case of Proposition 3.1.9, we have

Corollary 3.4.5

The following relation on $\mathcal{P}(X,p,q)$ is an equivalence relation: for any paths α and β

$$\alpha \sim \beta \Leftrightarrow \alpha \text{ and } \beta \text{ are homotopic.}$$

For any $\alpha \in \mathcal{P}(X,p,q)$, we denote by $[\alpha]$ the equivalence class of α, and call it the *homotopy class* of α.

Composition between (the homotopy classes of) paths

Let X be a path connected space and p, p' and p'' be three points in it. Let α be a path going from p to p' and α' be a path going from p' to p'', then we can construct the following path in X going from p to p''.

$$\alpha * \alpha' : [0,1] \to X,$$

defined by

$$(\alpha * \alpha')(t) = \begin{cases} \alpha(2t) & t \in \left[0, \frac{1}{2}\right] \\ \alpha'(2t-1) & t \in \left[\frac{1}{2}, 1\right] \end{cases}$$

Roughly speaking, to get the path $\alpha * \alpha'$, we first go along α, then go along α'.

Definition 3.4.6

We call the path $\alpha * \alpha'$ the ***composition*** of α and α'.

Using the composition of paths, we can define the following map

$$\begin{aligned} * : \mathcal{P}(X,p,p') \times \mathcal{P}(X,p',p'') &\to \mathcal{P}(X,p,p'') \\ (\alpha, \alpha') &\mapsto \alpha * \alpha'. \end{aligned}$$

Moreover, if we have paths $\alpha, \beta \in \mathcal{P}(X,p,p')$ and $\alpha', \beta' \in \mathcal{P}(X,p',p'')$, such that $\alpha \sim \beta$ and $\alpha' \sim \beta'$, we denote by H and H' the two homotopy respectively.

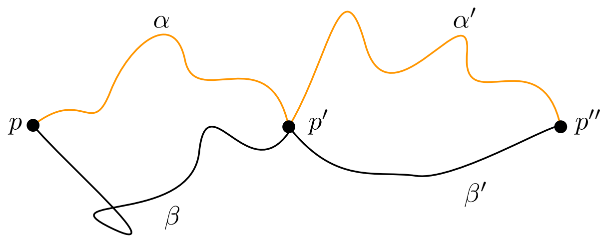


Figure 3.4.6: Composition of paths.

Then we have

$$\alpha * \alpha' \sim \beta * \beta',$$

for which a homotopy $\widetilde{H}$ satisfies

$$\widetilde{H}(s,t) = \widetilde{H}_t(s) = \begin{cases} (H_t * \alpha')(s) & t \in \left[0, \dfrac{1}{2}\right] \\ (\beta * H'_t)(s) & t \in \left[\dfrac{1}{2}, 1\right] \end{cases}$$

where s is the path parameter, and t is the time parameter for the homotopy.

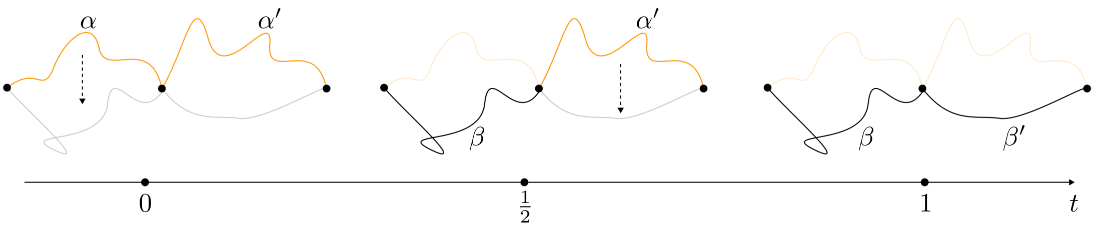


Figure 3.4.7: The homotopy $\widetilde{H}$.

We call this homotopy the composition of H and H'

$$\widetilde{H} = H * H'.$$

Remark 3.4.7.
The homotopy between $\alpha * \alpha'$ and $\beta * \beta'$ here can also be chosen to be the one such that for any $t \in [0,1]$, we have:

$$\widetilde{F}_t = H_t * H'_t.$$

In other words, we apply the homotopies H and H' at the same time, instead of first H, then H'.

Using the composition between two homotopy classes of two paths, we have the following map well-define:

$$\begin{aligned} * : \mathcal{P}(X,p,p')/\sim \times \mathcal{P}(X,p',p'')/\sim \quad &\to \mathcal{P}(X,p,p'')/\sim \\ ([\alpha],[\beta]) \quad &\mapsto [\alpha * \beta]. \end{aligned}$$

Definition 3.4.8

We call the class $[\alpha * \beta]$ the ***composition*** of $[\alpha]$ and $[\beta]$.

For any point $p \in X$, we call a constant map

$$\begin{aligned} c_x : [0,1] &\to X \\ t &\mapsto p \end{aligned}$$

a *constant path.*

Using reparametrizations of paths, we have the following facts.

Proposition 3.4.9

Let α, α' and α'' be three paths in X.

1) If $\alpha(1) = \alpha'(0)$ and $\alpha'(1) = \alpha''(0)$, then we have

$$(\alpha * \alpha') * \alpha'' \sim \alpha * (\alpha' * \alpha''),$$

2) If $p = \alpha(0)$ and $q = \alpha(1)$, we have

   (a) $c_p * \alpha \sim \alpha \sim \alpha * c_q$,

   (b) $\alpha * \overline{\alpha} \sim c_p$,

   (c) $\overline{\alpha} * \alpha \sim c_q$.

Proof. 1) The path $(\alpha * \alpha') * \alpha''$ comes from taking first the composition $\alpha * \alpha'$, then the composition $(\alpha * \alpha') * \alpha''$, while the path $\alpha * (\alpha' * \alpha'')$ comes from taking first the composition $\alpha' * \alpha''$, then the composition $\alpha * (\alpha' * \alpha'')$. Therefore for any $s \in [0,1]$, we have

$$((\alpha * \alpha') * \alpha'')(s) = \begin{cases} \alpha(4s), & s \in \left[0, \dfrac{1}{4}\right] \\ \alpha'(4s-1), & s \in \left[\dfrac{1}{4}, \dfrac{1}{2}\right] \\ \alpha''(2s-1), & s \in \left[\dfrac{1}{2}, 1\right] \end{cases}$$

and

$$(\alpha * (\alpha' * \alpha''))(s) = \begin{cases} \alpha(2s), & s \in \left[0, \frac{1}{2}\right] \\ \alpha'(4s-2), & s \in \left[\frac{1}{2}, \frac{3}{4}\right] \\ \alpha''(4s-3), & s \in \left[\frac{3}{4}, 1\right] \end{cases}$$

(See Figure 3.4.8 for an illustration)

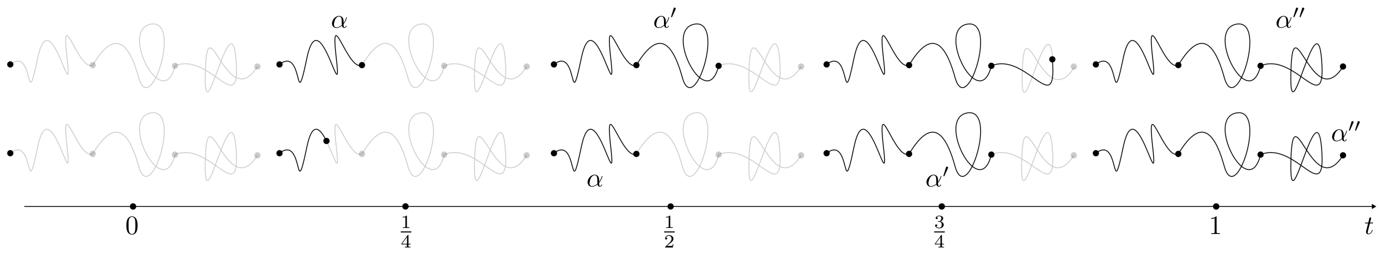


Figure 3.4.8: The paths $(\alpha * \alpha') * \alpha''$ (top) and $\alpha * (\alpha' * \alpha'')$ (bottom).

We consider the following continuous map H (illustrated in Figure 3.4.9) defined by

$$H(s,t) = \begin{cases} \alpha\left(\frac{4s}{1+t}\right), & 0 \leq s \leq \frac{1+t}{4} \\ \alpha'(4s-1-t), & \frac{1+t}{4} \leq s \leq \frac{2+t}{4} \\ \alpha''\left(\frac{4}{2-t}\left(s - \frac{2+t}{4}\right)\right), & \frac{2+t}{4} \leq s \leq 1 \end{cases}$$

for any $(s,t) \in [0,1] \times [0,1]$, which is a homotopy between $(\alpha * \alpha') * \alpha''$ and $\alpha * (\alpha' * \alpha'')$

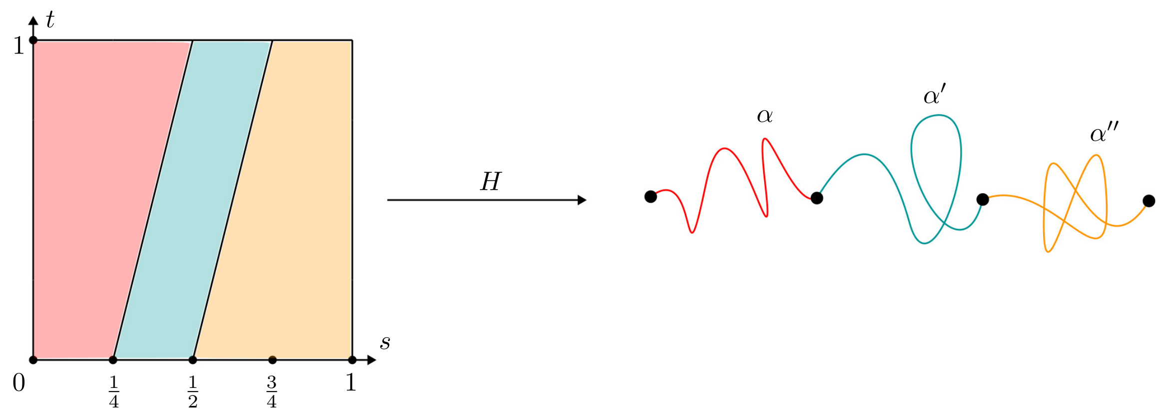


Figure 3.4.9: A homotopy between $(\alpha * \alpha') * \alpha''$ and $\alpha * (\alpha' * \alpha'')$.

2) We consider the path α. By the definition of the composition, we first give the maps involved

in these relations: for any $s \in [0,1]$, we have

$$(c_x * \alpha)(s) = \begin{cases} x, & s \in \left[0, \frac{1}{2}\right] \\ \alpha(2s-1), & s \in \left[\frac{1}{2}, 1\right] \end{cases}$$

$$(\alpha * c_y)(s) = \begin{cases} \alpha(2s), & s \in \left[0, \frac{1}{2}\right] \\ y, & s \in \left[\frac{1}{2}, 1\right] \end{cases}$$

$$(\alpha * \overline{\alpha})(s) = \begin{cases} \alpha(2s), & s \in \left[0, \frac{1}{2}\right] \\ \overline{\alpha}(2s-1), & s \in \left[\frac{1}{2}, 1\right] \end{cases}$$

$$(\overline{\alpha} * \alpha)(s) = \begin{cases} \overline{\alpha}(2s-1), & s \in \left[0, \frac{1}{2}\right] \\ \alpha(2s), & s \in \left[\frac{1}{2}, 1\right] \end{cases}$$

We consider the following homotopies.

(a) a homotopy between α and $c_p * \alpha$ (see Figure 3.4.10, spending more and more time in the beginning staying at p):

$$H(s,t) = \begin{cases} p, & 0 \le s \le \frac{t}{2} \\ \alpha\left(\frac{2}{2-t}(s - \frac{t}{2})\right), & \frac{t}{2} \le s \le 1 \end{cases};$$

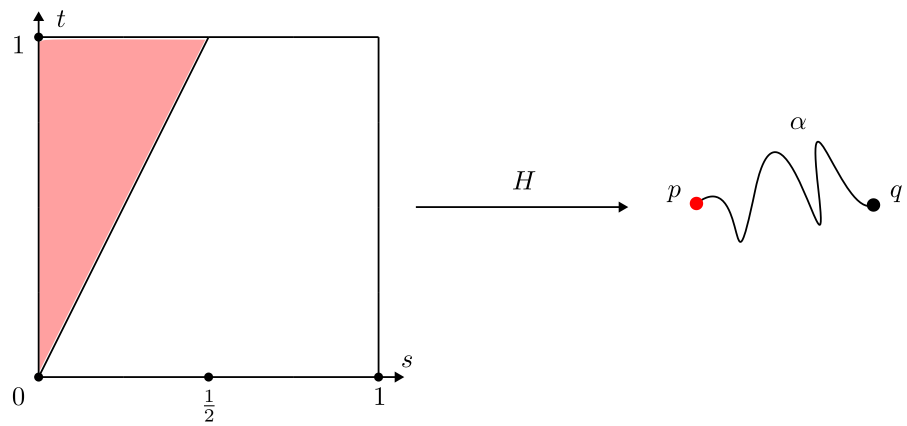


Figure 3.4.10: A homotopy between α and $c_p * \alpha$.

(b) a homotopy between α and $\alpha * c_q$ (see Figure 3.4.11, spending more and more time in the end staying at q):

$$H(s,t) = \begin{cases} \alpha\left(\frac{2s}{2-t}\right), & 0 \le s \le 1 - \frac{t}{2} \\ q, & 1 - \frac{t}{2} \le s \le 1 \end{cases};$$

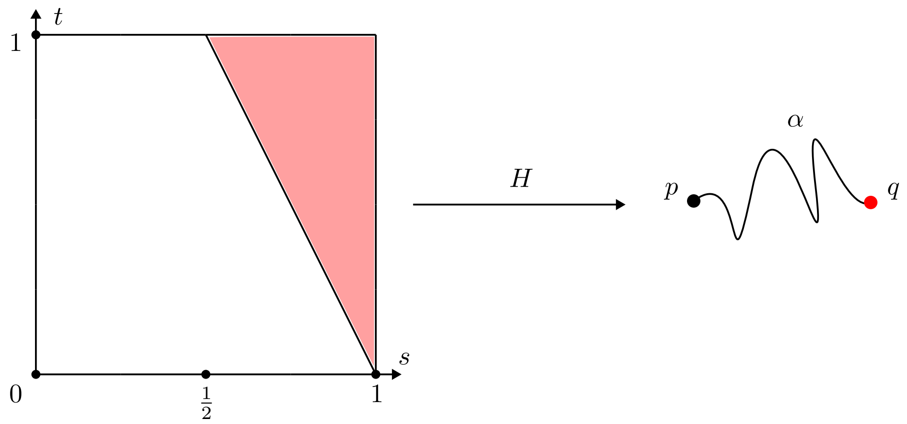


Figure 3.4.11: A homotopy between α and $\alpha * c_q$.

(c) a homotopy between $\alpha * \overline{\alpha}$ and c_p (see Figure 3.4.13, turning back earlier and earlier):

$$H(s,t) = \begin{cases} \alpha(2s(1-t)), & 0 \le s \le \dfrac{1}{2} \\ \overline{\alpha}(2s(1-t)), & \dfrac{1}{2} \le s \le 1 \end{cases};$$

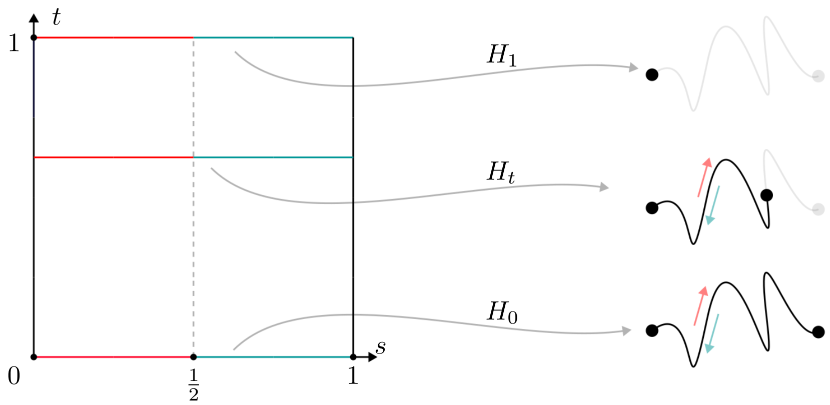


Figure 3.4.12: A homotopy between $\alpha * \overline{\alpha}$ and c_p.

(d) a homotopy between $\overline{\alpha} * \alpha$ and c_q (see Figure 3.4.13, turning back earlier and earlier):

$$H(s,t) = \begin{cases} \overline{\alpha}(2s(1-t)), & 0 \le s \le \dfrac{1}{2} \\ \alpha(2s(1-t)), & \dfrac{1}{2} \le s \le 1 \end{cases}.$$

□

This proposition can also be written as follows by considering homotopy classes of paths and their compositions.

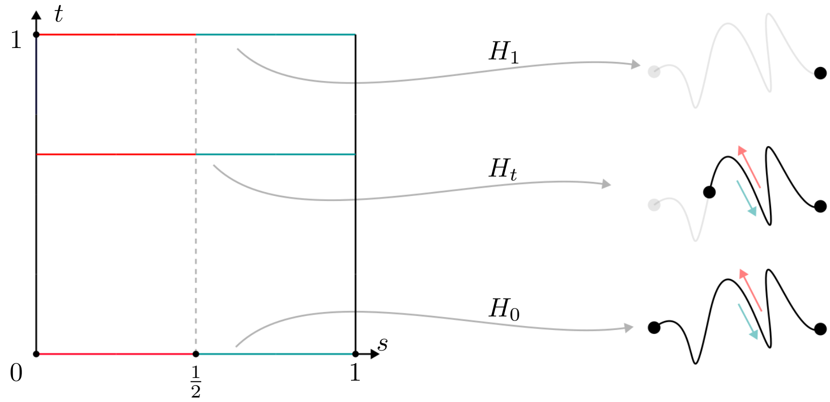


Figure 3.4.13: A homotopy between $\overline{\alpha} * \alpha$ and c_q.

Corollary 3.4.10

Let α, β and γ be three paths in X.

1) If $\alpha(1) = \beta(0)$ and $\beta(1) = \gamma(0)$, then we have

$$([\alpha] * [\beta]) * [\gamma] = [\alpha] * ([\beta] * [\gamma]),$$

2) If $x = \alpha(0)$ and $y = \alpha(1)$, we have

   (a) $[c_x] * [\alpha] = [\alpha] = [\alpha] * [c_y]$,

   (b) $[\alpha] * [\overline{\alpha}] = [c_x]$,

   (c) $[\overline{\alpha}] * [\alpha] = [c_y]$.

Remark 3.4.11.
With first point in the above proposition, when we taking a composition of finitely many paths, the order for which composition is done first is no longer important. Hence let $n > 1$ be a natural number, and $\alpha_1, ..., \alpha_n$ be n paths in X satisfying that for any $1 \leq i \leq n-1$, we have

$$\alpha_i(1) = \alpha_{i+1}(0).$$

Then we denote their composition by

$$[\alpha_1] * \cdots * [\alpha_n]$$

omitting the parentheses.

3.5 Fundamental Group

Now we are ready to introduce the fundamental group of a path connected space, which is constructed by studying loops based at a same point up to path homotopy.

Let X be a path connected space.

Definition 3.5.1

If a path α in X satisfies

$$\alpha(0) = \alpha(1) = p,$$

then we say that α is a ***loop based at*** $p \in X$.

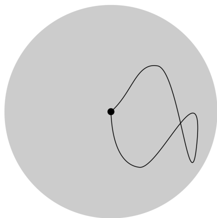


Figure 3.5.1: A loop in a disk based at its center.

We denote the space of loops in X based at $p \in X$ by

$$\mathcal{L}(X, p) := \{\text{loops in } X \text{ based at } p\}.$$

As special paths, the path homotopy induces an equivalence relation in $\mathcal{L}(X, p)$: for any loops α and α' in $\mathcal{L}(X, p)$, we define

$$\alpha \sim \alpha' \Leftrightarrow \alpha \text{ and } \alpha' \text{ are homotopic.}$$

Given any loop $\alpha \in \mathcal{L}(X, p)$, the equivalence class containing α is given by

$$[\alpha] := \{\alpha' \in \mathcal{L}(X, p) \mid \alpha' \sim \alpha\}$$

and is called the *homotopy class* of α, and any loop in $[\alpha]$ is a representative of $[\alpha]$. In particular, the loop α is a representative of $[\alpha]$.

Remark 3.5.2.
By its definition, if we have a path homotopy H of loops in a topological space X, all maps H_t's are loops in X based at a same point. In other words, when we deform a loop with a path homotopy, not only we have a loop at each time t, we also never move the base point.

We denote the space of homotopy classes of loops by

$$\pi_1(X, p) := \mathcal{L}(X, p)/\sim.$$

The composition between paths introduced previously then induces a binary operator on the $\mathcal{L}(X, p)$, then a binary operator on $\pi_1(X, p)$. We then have an immediate corollary of Corollary 3.4.10.

Corollary 3.5.3

The set $\pi_1(X, p)$ with the composition operator is a group.

Proof. By 1) of Corollary 3.4.10, the composition operator satisfies the associativity.

By 2.a) of Corollary 3.4.10, the homotopy class $[c_p]$ is an identity element.

By 2.b) and 2.c) of Corollary 3.4.10, for any loop α based at p, the homotopy class $[\overline{\alpha}]$ is the inverse of $[\alpha]$.

□

Definition 3.5.4

The group $\pi_1(X,p)$ is called the ***fundamental group of*** X ***based at*** p.

Remark 3.5.5.
This group is also called the *Poincaré group* or the *first homotopy group* of X based at p.

One important application of the fundamental group is to classifies topological spaces. In general, as we will see later that if two path connected spaces have non isomorphic fundamental groups, they are not homeomorphic, neither homotopy equivalent. The other direction is not true, two homotopy non-equivalent spaces may have isomorphic fundamental groups. As an elementary example, we can compare the 2-disk D^2 and the 2-sphere S^2. Both spaces have trivial fundamental groups based at any point. More information needs to be considered in order to distinguish them.

In fact the "first" stands for the dimension 1. We can also think a loop in X as the image of a continuous map from S^1 to X. The construction of $\pi_1(X,p)$ can also be generalized by considering continuous maps from n-sphere S^n to X and their maps. The resulting group is called the n-th homotopy group of X and denoted by $\pi_n(X,p)$.

Consider the above example where we compare D^2 amd S^2. Notice that their second homotopy groups are different. Any continuous map from S^2 to D^2 is homotopic to a constant map, which is not true for any continuous map from S^2 to S^2 (for example, the identity map). Even when we consider $\pi_n(X,p)$ for all $n \in \mathbb{N}^*$, we still cannot tell if two path connected spaces are homotopy equivalent. There are still more information needed.

Example 3.5.6 (Interval).
Consider $I=[0,1]$ the closed interval in $\mathbb{R}$. We consider the fundamental group of I based at $p=0$.

$$\pi_1(I,p) := \{\text{Loops in } I \text{ based at } p\}/\sim .$$

In the previous sections, we have seen that I is a contractible space. In particular, there is a strong deformation retraction of I to $\{p\}$. We denote a homotopy between id_I and c_p by H.

Given any loop $\alpha \in \mathcal{L}(I,p)$, we consider the following map

$$\begin{aligned} \widetilde{H} : [0,1]\times[0,1] &\to I \\ (s,t) &\mapsto H(\alpha(s),t) \end{aligned}$$

which is continuous, such that $\widetilde{H}_0 = \alpha$ and $\widetilde{H}_1 = c_p$ and for any $t \in [0,1]$

$$\widetilde{H}(0,t) = \widetilde{H}(1,t) = p,$$

thus a path homotopy between α and c_p. Therefore we have

$$\pi_1(I,p) := \{[c_p]\}$$

The fundamental group of I based at p is trivial.

Informally speaking, when we apply the strong deformation retraction of I to p, we retract the whole space to p, including the image of α. By slightly modifying the discussion, we can show that this also holds for any contractible space.

Proposition 3.5.7

If a topological space X is contractible, then for any $p \in X$, the fundamental group $\pi_1(X, p)$ is trivial.

Proof. The space X is contractible, by definition there is a homotopy H

$$H : X \times [0,1] \to X$$

between the identity map id_X and the constant map Const_{p_0} for some $p_0 \in X$. By considering the restriction of H to $\{q\} \times [0,1]$ for any $q \in X$, we have a path in X connecting q to p_0. Hence X is path connected.

Since X is path connected, for any $p, q \in X$, the constant maps Const_p and Const_q are homotopic to each other. By taking composition of homotopies, for any $p \in X$, the identity map id_X is homotopic to Const_p.

Therefore to show the proposition, it is enough to show that $\pi_1(X, p_0)$ is trivial. We first consider the path

$$\begin{aligned} \beta : [0,1] &\to X \\ t &\mapsto H(p_0, t) \end{aligned}.$$

For each $t \in [0,1]$, we denote

$$\begin{aligned} \beta_t : [0,1] &\to X \\ s &\mapsto \beta(st) \end{aligned}.$$

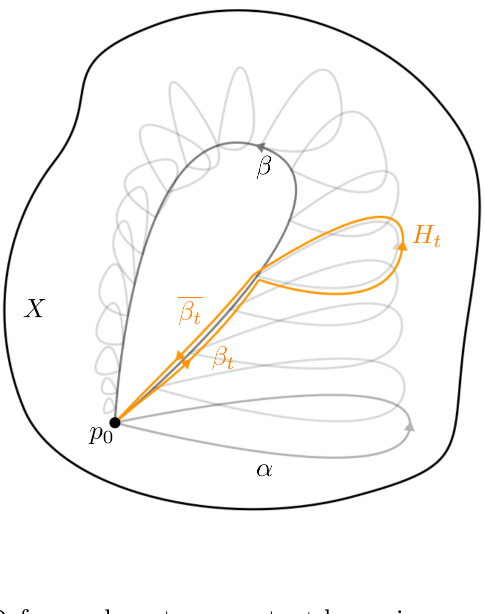


Figure 3.5.2: Deform a loop to a constant loop via a general homotopy.

Let α be any loop in X based at p_0. Consider the following map

$$\begin{aligned} F : [0,1] \times [0,1] &\to X \\ (s,t) &\mapsto (\beta_t * H_t * \overline{\beta_t})(s) \end{aligned}$$

We can verify that this is a path homotopy between $c_{p_0} * \alpha * \overline{c_{p_0}}$ and $\beta * c_{p_0} * \overline{\beta}$, where c_{p_0} is the constant path with image $\{p_0\}$. Hence we have

$$\alpha \sim c_{p_0}.$$

This shows that $\pi_1(X, p_0)$ is trivial. □

Remark 3.5.8.
This shows that although there could be many loops based at p_0 in X which look quite different from each other, their homotopy classes could be quite few.

Fundamental group of S^1

Now we consider the space S^1 as the next example, which is at the same time elementary and important in the whole story of fundamental groups.

For our convenience, we consider

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

Let $p = (1, 0)$. We would like to study loops in S^1 based at p.

The fundamental group $\pi_1(S^1, p)$ is cyclic.

Notice that a loop in S^1 could be quite arbitrary. For example, if as the time parameter t moves 0 to 1, the moving direction of the point $\alpha(t)$ can switch between counterclockwise and clockwise infinitely many times (See Figure 3.5.3 for an illustration).

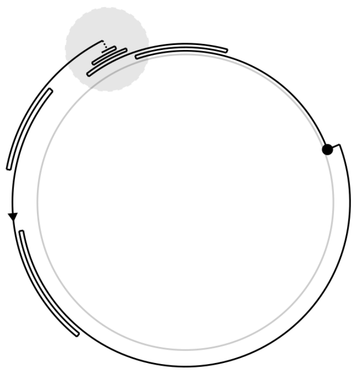


Figure 3.5.3: A loop with infinite backtracks in the shadowed area.

The first step is to show that all loops are homotopic to some loops standard in certain way. The compactness of $[0, 1]$ plays an essential role here.

Let α be such a loop. For any $t \in [0, 1]$, we consider $\epsilon > 0$ such that the neighborhood

$$U_t = (t - \epsilon, t + \epsilon) \cap [0, 1]$$

of t is contained in a half circle in S^1. Since $[0, 1]$ is compact, there are finitely many of these open sets in $[0, 1]$ forming an open cover of $[0, 1]$. We denote them by

$$\{U_1, ..., U_n\}.$$

Consider their end points

$$0 = t_0 < t_1 < ... < t_m = 1.$$

Hence the restriction of α to $[t_j, t_{j+1}]$ is contained in a half circle for any integer $0 \le j \le m-1$, since each (t_j, t_{j+1}) is contained in some U_k.

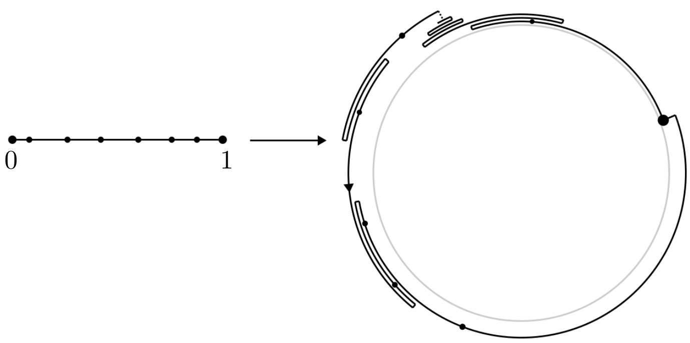


Figure 3.5.4: A partition of $[0, 1]$ with the desired property.

For each integer $0 \le j \le m-1$, we denote the restriction of α by

$$\alpha_j = \alpha|_{[t_j, t_{j+1}]}.$$

Denote

$$\alpha_j(t_j) = e^{2\pi i s_j}.$$

Then for any $t \in [t_j, t_{j+1}]$, there is a $s \in [s_j - 1/2, s_j + 1/2]$, such that

$$\alpha_j(t) = e^{2\pi i s},$$

moreover, since the image of α_j is contained in a half circle, the map

$$\begin{aligned} \varphi_j : \alpha_j([t_j, t_{j+1}]) &\to [s_j - 1/2, s_j + 1/2] \\ e^{2\pi i s} &\mapsto s \end{aligned}$$

is continuous. We consider the composition

$$\psi_j = \varphi_j \circ \alpha_j : [t_j, t_{j+1}] \to \mathbb{R},$$

and perform a homotopy relative to $\{t_j, t_{j+1}\}$ using linear maps in $\mathbb{R}$ to get the following one

$$\begin{aligned} \rho_j : [t_j, t_{j+1}] &\to [\theta_j, \theta_j + 1] \\ t &\mapsto \psi_j(t_j) + \frac{t - t_j}{t_{j+1} - t_j}(\psi_j(t_{j+1}) - \psi_j(t_j)) \end{aligned}$$

Consider the map

$$\begin{aligned} \Phi : \mathbb{R} &\to S^1 \\ s &\mapsto e^{2s\pi i} \end{aligned}.$$

Now we consider the following map α' such that for any j and any $t \in [t_j, t_j]$,

$$\alpha'(t) = \Phi \circ \rho_j(t).$$

Notice that for any $0 \leq j \leq m-1$, we have

$$\Phi \circ \rho_j(t_{j+1}) = \Phi \circ \rho_{j+1}(t_{j+1}),$$

hence α' is well defined and is a loop based at p. Now we consider

$$\alpha'^{-1}(\{\alpha'(t_0), ..., \{\alpha'(t_m)\}),$$

and denote it by

$$0 = s_0 < s_1 < ... < s_N = 1.$$

By our construction, for each $1 \leq j \leq N-1$, the images

$$\alpha'(s_{j-1}, s_j) \text{ and } \alpha'(s_j, s_{j+1})$$

are either same or disjoint.

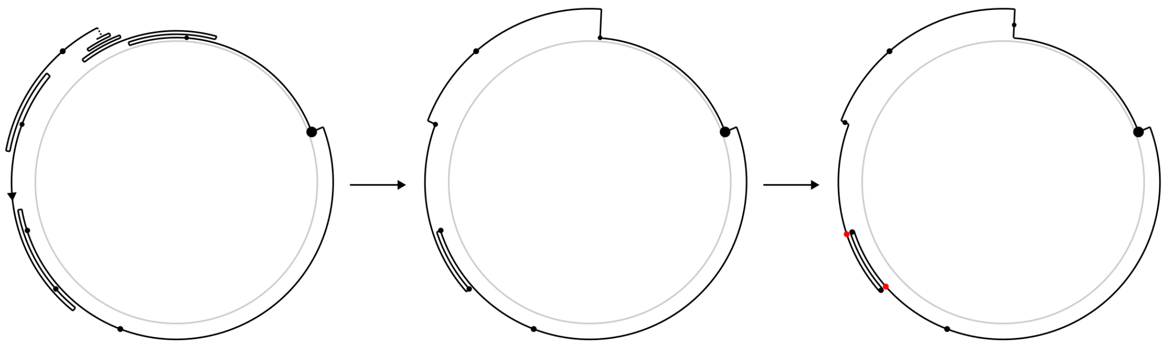


Figure 3.5.5: Pull the loop tight locally.

Now we start by comparing $\alpha'(s_0, s_1)$ and $\alpha'(s_1, s_2)$. If

$$\alpha'(s_0, s_1) = \alpha'(s_1, s_2)$$

then we can apply a homotopy to get a new loop

$$\alpha'_1(t) = \begin{cases} p, & t \in [0, s_2] \\ \alpha'(t), & t \in [s_2, 1] \end{cases}$$

Otherwise, let $\alpha'_1 = \alpha'$ and we consider s_1 and compare $\alpha'(s_1, s_2)$ and $\alpha'(s_2, s_3)$. If they are the same, we can apply a homotopy to get a new loop

$$\alpha'_2(t) = \begin{cases} \alpha'_1(t), & t \in [0, s_1] \\ \alpha'_1(s_1), & t \in [s_1, s_3] \\ \alpha'_1(t), & t \in [s_3, 1] \end{cases}$$

We repeat this process for all $1 \leq j \leq N-1$, then we have a path α'' such that in each interval (s_j, s_{j+1}), it either stays at $\alpha''(s_j)$ or moving to a fixed direction as parameter t increases. Then after a reparametrization, the loop α'' is homotopic to the following standard one

$$\gamma_k : [0,1] \to S^1$$
$$t \mapsto e^{2kt\pi i}$$

for some integer $k \in \mathbb{Z}$, such that $|k|$ is the times that α'' passes p and the sign of k is determined by how α'' passes p (clockwise or counterclockwise). This is the end of the first step.

From the first step, we can see that

$$\pi_1(S^1, p) = \{[\gamma_k] \mid k \in \mathbb{Z}\}.$$

Notice that for any integers k_1 and k_2, we can define an explicit path homotopy to show that

$$\gamma_{k_1} * \gamma_{k_2} \sim \gamma_{k_1+k_2}.$$

Hence this shows that

$$\pi_1(S^1, p) = \langle \gamma_1 \rangle.$$

The fundamental group $\pi_1(S^1, p)$ is isomorphic to $\mathbb{Z}$.

After the previous discussion, there is one last problem whose answer is unclear. Is it possible that there is a $k \in \mathbb{N}^*$ such that $\gamma_k \sim c_p$, or equivalently is the generator $[\gamma_1]$ of finite order?

Notice that the above construction depends on the choice of representative α in a homotopy class of loops in $\mathcal{L}(S^1, p)$ and the choice of finite covers of $[0, 1]$ associated to α. Hence we cannot get the answer directly.

To study this problem, we consider the continuous map

$$\begin{aligned} \Phi : \mathbb{R} &\to S^1 \\ \theta &\mapsto e^{2\theta\pi i} \end{aligned}.$$

Notice that for any $q \in S^1$, there is a unique point $\widetilde{q} \in [0, 1)$ such that $\Phi(\widetilde{q}) = q$, and

$$\Phi^{-1}(q) = \{\widetilde{q}_k = \widetilde{q} + k \mid k \in \mathbb{Z}\}.$$

Moreover the map Φ has the following property:

Observation 3.5.9
Given any point $q \in S^1$, it has a neighborhood V such that for any $\widetilde{q}_k \in \Phi^{-1}(q)$ it has a neighborhood U_k, such that

(i) for any $k \in \mathbb{Z}$, the restriction

$$\Phi|_{U_k} : U_k \to V$$

is a homeomorphism;

(ii) for different integers k_1 and k_2, we have $U_{k_1} \cap U_{k_2} = \emptyset$.

Remark 3.5.10.
Later we will see that with this property, the space $\mathbb{R}$ is called a *covering space* of S^1, and the map Φ is called a *covering map*.

For any point $q \in S^1$, we call a neighborhood of q satisfying the property in the definition a *covering neighborhood* of q (See Figure 3.5.6 for an illustration).

The following two facts are applications of two general results for covering maps to our case.

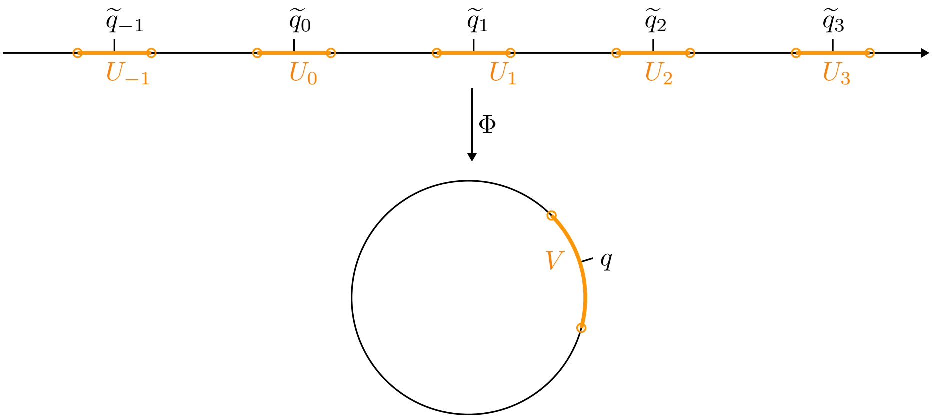


Figure 3.5.6: A covering neighborhood V of q.

Observation 3.5.11
Let α be any loop in S^1 based at p. Let $\widetilde{p} \in \Phi^{-1}(p)$. Then there is a unique path

$$\widetilde{\alpha} : [0,1] \to \mathbb{R},$$

such that $\widetilde{\alpha}(0) = \widetilde{p}$ and the following diagram commutes:

$$\begin{array}{ccc} [0,1] & \xrightarrow{\widetilde{\alpha}} & \mathbb{R} \\ & \searrow^{\alpha} & \downarrow \Phi \\ & & S^1 \end{array}$$

Observation 3.5.12
Let H be any path homotopy between loops in S^1 based at p. Let $\widetilde{p} \in \Phi^{-1}(p)$. Then there is a unique path homotopy

$$\widetilde{H} : [0,1] \times [0,1] \to \mathbb{R},$$

such that $\widetilde{H}(0,0) = \widetilde{p}$ and the following diagram commutes:

$$\begin{array}{ccc} [0,1] \times [0,1] & \xrightarrow{\widetilde{H}} & \mathbb{R} \\ & \searrow^{H} & \downarrow \Phi \\ & & S^1 \end{array}$$

Remark 3.5.13.
The objects $\widetilde{q}$, $\widetilde{p}$, $\widetilde{\alpha}$ and $\widetilde{H}$ are called *lifts* of p, q, α and H respectively.

The proof of these facts are constructive, and we skip them for the moment and only provide some rough idea of the proof for Observation 3.5.11. The complete proof for general cases will be given in the later sections.

The rough idea is to use Φ^{-1} to construct a map $\widetilde{\alpha}$ from $[0,1]$ to $\mathbb{R}$ which is a lift of a path α in S^1. However, since Φ is not injective, taking Φ-preimage is not a map. Hence we cannot obtain $\widetilde{\alpha}$ by just taking Φ-preimage.

To get over this problem, we use the property satisfied by the covering map Φ listed in Observation 3.5.9. Instead of lifting α directly as a whole, we may consider lifting restriction of α

on subintervals of $[0,1]$, and then show that all these lifts can be "glued" together and gives a lift $\widetilde{\alpha}$ of α.

More precisely, for any $t \in [0,1]$, let $q = \alpha(t)$, from Observation 3.5.9, there is V a covering neighborhood of q. Then by the continuity of α, the parameter t has a path connected open neighborhood I_t such that

$$\alpha(I_t) \subset V.$$

The collection $\{I_t \mid t \in I\}$ of such open subsets in $[0,1]$ is an open cover of $[0,1]$. By the compactness of $[0,1]$, there is a finite subcover

$$J_1, ..., J_m.$$

We denote by $V_1, ..., V_m$ the associated open covering neighborhood.

To simplify the argument, up to taking subinterval of J_j's and relabeling the indices, we can moreover assume that for any $1 \le j \le m-1$, we have

$$J_j \cap J_{j+1} \ne \emptyset$$

and for any $1 \le j, k \le m$ with $|j-k| \ge 2$, we have

$$J_j \cap J_k = \emptyset.$$

At same time, we may assume that $0 \in J_1$ and $1 \in J_m$.

To construct $\widetilde{\alpha}$, we start from $t_1 = 0$. Let $\widetilde{p}$ is a chosen lift of $\alpha(0)$. Then we have a unique lift U_1 of V_1. Consider the map

$$\Phi|_{U_1} : U_1 \to V_1$$

which is a homeomorphism. Then we consider the following composition

$$\widetilde{\alpha}_1 = (\Phi|_{U_1})^{-1} \circ \alpha|_{J_1}.$$

Then consider a point $t_2 \in J_1 \cap J_2$. There is a unique lift $\widetilde{\alpha}_1(t_2)$ of $\alpha(t_2)$ in U_1. Now we consider the lifts of V_2, there is a unique one containing $\widetilde{\alpha}_1(t_2)$ and we denote it by U_2. We have a homeomorphism

$$\Phi|_{U_2} : U_2 \to V_2.$$

We define

$$\widetilde{\alpha}_2 = (\Phi|_{U_2})^{-1} \circ \alpha|_{J_2}.$$

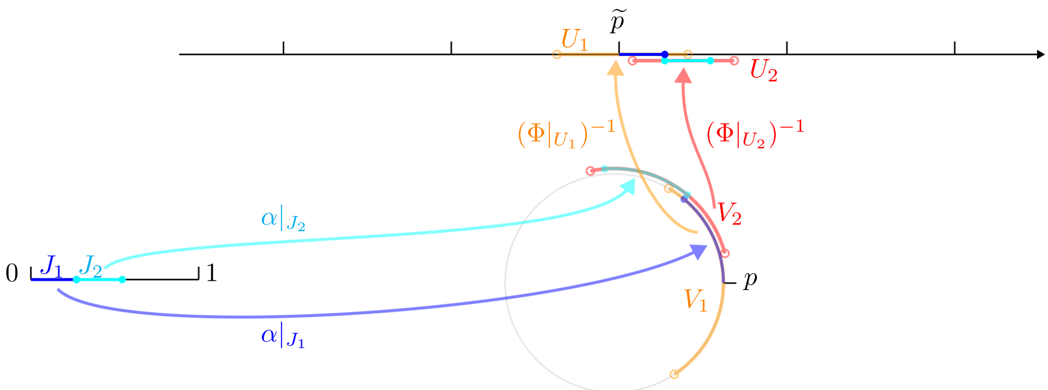


Figure 3.5.7: Construct $\widetilde{\alpha}$ piece by piece.

Repeating this process, we have a collection of maps $\widetilde{\alpha}_j$ for $1 \leq j \leq m$. Then we define

$$\widetilde{\alpha}(t) = \begin{cases} \widetilde{\alpha}_1(t), & t \in J_1 \\ \widetilde{\alpha}_2(t), & t \in J_2 \\ \quad \vdots & \\ \widetilde{\alpha}_m(t), & t \in J_m \end{cases}$$

By the construction of $\widetilde{\alpha}_j$'s, the map $\widetilde{\alpha}$ is well-defined and continuous, and satisfies the desired commutative diagram. The uniqueness is obtained by the uniqueness of each $\widetilde{\alpha}_j$.

The proof for the existence and uniqueness of homotopy lifting is similar. The difference is that instead of considering decomposing $[0, 1]$ into subinterval, we have to consider the square $[0, 1]^2$ and decompose it into subsquares. Using the compactness of $[0, 1]^2$, by a similar construction as above, we obtain a unique lift of a homotopy for a given lift of the base point.

By the definition of Φ and $p = 1 \in \mathbb{C}$, we have $\Phi^{-1}(p) = \mathbb{Z}$. We choose $\widetilde{p} = 0$. By Observation 3.5.11, we have a well defined map

$$\begin{aligned} f : \mathcal{L}(S^1, p) &\to \mathcal{P}(\mathbb{R}, \widetilde{p}) \\ \alpha &\mapsto \widetilde{\alpha} \end{aligned}.$$

By Observation 3.5.12, we have

$$\begin{aligned} f : \mathcal{L}(S^1, p)/\sim &\to \mathcal{P}(\mathbb{R}, \widetilde{p})/\sim \\ [\alpha] &\mapsto [\widetilde{\alpha}] \end{aligned}.$$

Here we consider the path homotopy (relative to $\{0, 1\}$). Hence the starting and the ending points of $\widetilde{\alpha}$ will be fixed during the homotopy. We consider the endpoint $\widetilde{\alpha}(1)$, and have a map

$$\begin{aligned} g : \mathcal{L}(S^1, p)/\sim &\to \mathbb{Z} \\ \alpha &\mapsto \widetilde{\alpha}(1) \end{aligned}.$$

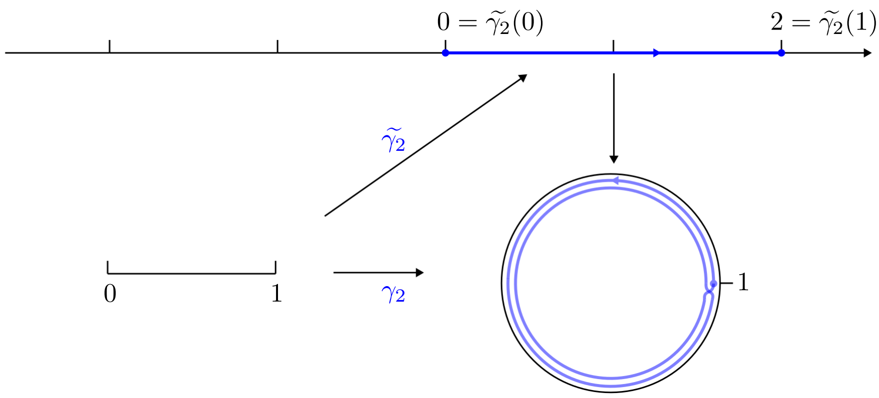


Figure 3.5.8: Lifting γ_2 to $\mathbb{R}$.

Now we consider the lift $\widetilde{\gamma_k}$ of γ_k with $\widetilde{\gamma_k}(0) = 0$. Then we have

$$\widetilde{\gamma_k}(1) = k \in \mathbb{Z}.$$

Therefore for different integers $k_1 \neq k_2$, we have

$$\gamma_{k_1} \not\simeq \gamma_{k_2},$$

for otherwise we should have $k_1 = \widetilde{\gamma_{k_1}}(1) = \widetilde{\gamma_{k_2}}(1) = k_2$, which is a contradiction.

Hence the generator $[\gamma_1]$ of $\pi_1(S^1, p)$ is of infinite order, and

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

Change the base point

From its construction, the group $\pi_1(X, p)$ depends on the choice of p. One may continue to ask what happens when we change the base point p to a different point, say $q \in X$, and if there is any relation between $\pi_1(X, x)$ and $\pi_1(X, y)$. To answer these questions, we consider the following construction to relate loops based at p and those based at q.

Let α be a path in X with $\alpha(0) = q$ and $\alpha(1) = p$, which is chosen once and for all. For any loop γ in $\mathcal{L}(X, p)$, we consider the composition $\alpha * \gamma * \overline{\alpha}$ which is a loop based at q (See Figure 3.5.9 for an illustration).

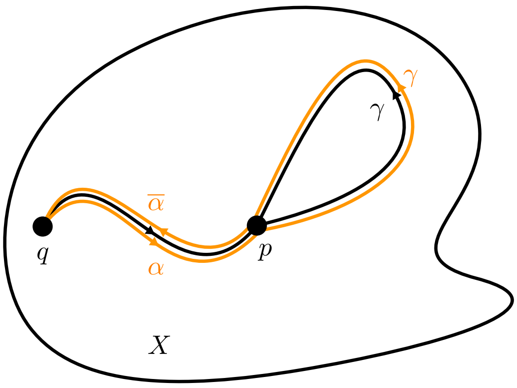


Figure 3.5.9: Change the base point from p to q.

This induces a map

$$\begin{aligned} \varphi_\alpha : \mathcal{L}(X, p) &\to \mathcal{L}(X, q), \\ \gamma &\mapsto \alpha * \gamma * \overline{\alpha}, \end{aligned}$$

which in turn induces a map

$$\begin{aligned} \Phi_\alpha : \pi_1(X, p) &\to \pi_1(X, q), \\ [\gamma] &\mapsto [\alpha * \gamma * \overline{\alpha}]. \end{aligned}$$

Proposition 3.5.14

The map Φ_α is an isomorphism.

Proof. It is enough to show that Φ_α is bijective and it is a group homomorphism.

We first show that the map Φ_α is bijective. We consider another map constructed in a similar way

$$\begin{aligned} \Phi_{\overline{\alpha}} : \pi_1(X, q) &\to \pi_1(X, p) \\ [\eta] &\mapsto [\overline{\alpha} * \eta * \alpha] \end{aligned}.$$

For any $[\gamma] \in \pi_1(X,p)$, we have

$$\begin{aligned}(\Phi_{\overline{\alpha}} \circ \Phi_\alpha)([\gamma]) &= \Phi_{\overline{\alpha}}([\alpha * \gamma * \overline{\alpha}]) \\ &= [\overline{\alpha} * (\alpha * \gamma * \overline{\alpha}) * \alpha] \\ &= [c_p * \gamma * c_p] = [\gamma].\end{aligned}$$

Hence

$$\Phi_{\overline{\alpha}} \circ \Phi_\alpha = \mathrm{id}_{\pi_1(X,p)}.$$

Similarly, we have

$$\Phi_\alpha \circ \Phi_{\overline{\alpha}} = \mathrm{id}_{\pi_1(X,q)}.$$

Therefore Φ_α is bijective.

To show that it is a group homomorphism, we consider any pair $[\gamma]$ and $[\gamma']$ in $\pi_1(X,p)$ and have

$$\begin{aligned}\Phi_\alpha([\gamma] * [\gamma']) &= \Phi_\alpha([\gamma * \gamma']) \\ &= [\alpha * \gamma * \gamma' * \overline{\alpha}] \\ &= [\alpha * \gamma * c_p * \gamma' * \overline{\alpha}] \\ &= [\alpha * \gamma * \overline{\alpha} * \alpha * \gamma' * \overline{\alpha}] \\ &= [\alpha * \gamma * \overline{\alpha}] * [\alpha * \gamma' * \overline{\alpha}] \\ &= \Phi_\alpha([\gamma]) * \Phi_\alpha([\gamma']).\end{aligned}$$

Hence Φ_α is a group homomorphism. □

Remark 3.5.15.
By this proposition, up to isomorphism, the fundamental group of X based at a point is independent of choice of the base point. Hence we may omit the base point and call it the *fundamental group* of X, denote it by $\pi_1(X)$.

However, when we try to do computation in details with loops, we have to choose a base point $p \in X$ and consider the corresponding fundamental group $\pi_1(X,p)$ based at p.

Notice the above construction of Φ_α seems to depend on the choice of α or at least the homotopy class of $[\alpha]$. If we choose another path β in X with $\beta(0) = q$ and $\beta(1) = p$, by the same construction, we have the map φ_β and the isomorphism Φ_β.

Proposition 3.5.16

For any $[\gamma]$, we have

$$\Phi_\beta([\gamma]) = [\beta * \overline{\alpha}] * \Phi_\alpha([\gamma]) * [\beta * \overline{\alpha}]^{-1}.$$

In another word, the two group homomorphisms Φ_α and Φ_β are different by a conjugation in $\pi_1(X,q)$ given by $[\beta * \overline{\alpha}]$.

Proof. The proof is a direct computation.

Notice that

$$[\beta * \overline{\alpha}] \in \pi_1(X,q).$$

Given any $\gamma \in \mathcal{L}(X,p)$, we have

$$\begin{aligned}\Phi_\beta([\gamma]) &= [\beta * \gamma * \overline{\beta}] \\ &= [\beta] * [c_x * \gamma * c_x] * [\overline{\beta}] \\ &= [\beta] * [\overline{\alpha} * \alpha] * [\gamma] * [\overline{\alpha} * \alpha] * [\overline{\beta}] \\ &= [\beta * \overline{\alpha}] * [\alpha * \gamma * \overline{\alpha}] * [\alpha * \overline{\beta}] \\ &= [\beta * \overline{\alpha}] * \Phi_\alpha([\gamma]) * [\beta * \overline{\alpha}]^{-1}\end{aligned}$$

Hence the two isomorphisms are different by a global conjugacy by $[\beta * \overline{\alpha}] \in \pi_1(X, q)$.

□

As a corollary, we have

Corollary 3.5.17

If $\pi_1(X, p)$ is abelian, the isomorphism between $\pi_1(X, p)$ and $\pi_1(X, q)$ given by changing the base point is canonical, i.e. for any paths α and β in X going from q to p, we have

$$\Phi_\alpha = \Phi_\beta.$$

Example 3.5.18.
Both circle S^1 and torus T^2 have abelian fundamental groups. On the other hand, the fundamental group of a disk with more than 1 holes is not abelian. For example, let X be a disk with 2 holes. Let p and q be two points in X. Consider two paths α and β going from q to p as in the picture, such that $\beta * \overline{\alpha}$ is a loop based at q going around another hole counterclockwise once. We consider the change of base point induced by α and that induced by β (See Figure 3.5.10 for an illustration).

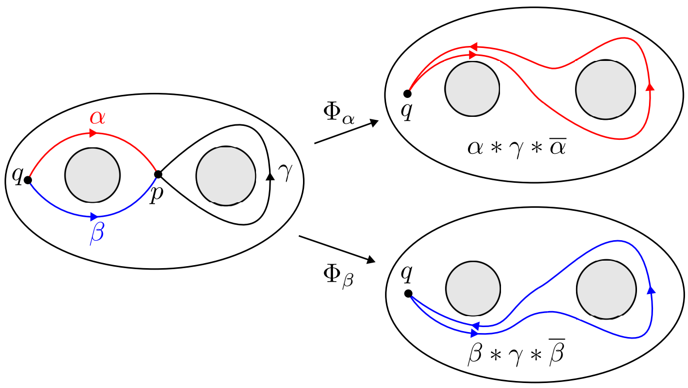


Figure 3.5.10: Different ways of change the base point from p to q.

Informally speaking to have a path homotopy between $\alpha * \gamma * \overline{\alpha}$ and $\beta * \gamma * \overline{\beta}$, we should be able to go over the hole on the left which is impossible. After we discuss the Seifert-Van-Kampen Theorem, we will be able to compute the fundamental group of X in a simple way. Then we will see the two elements $[\alpha * \gamma * \overline{\alpha}]$ and $[\alpha * \overline{\beta}]$ do not commute with each other, hence

$$[\beta * \gamma * \overline{\beta}] = [\alpha * \overline{\beta}]^{-1} * [\alpha * \gamma * \overline{\alpha}] * [\alpha * \overline{\beta}] \neq [\alpha * \gamma * \overline{\alpha}].$$

3.6 Fundamental groups and continuous maps

From the topological point of view, different spaces can be related through continuous maps. In this part, we would like to discuss this kind of relation in the fundamental group level.

Let X and Y be two path connected topological spaces. Denote by f a continuous map from X to Y. Since the composition of any pair of continuous maps (if possible) is still a continuous map, we have a natural way to relate paths in X to paths in Y which can be described by the following map (See Figure 3.6.1 for an illustration)

$$f_* : \mathcal{P}(X, p, q) \to \mathcal{P}(Y, f(p), f(q)),$$
$$\alpha \mapsto f \circ \alpha.$$

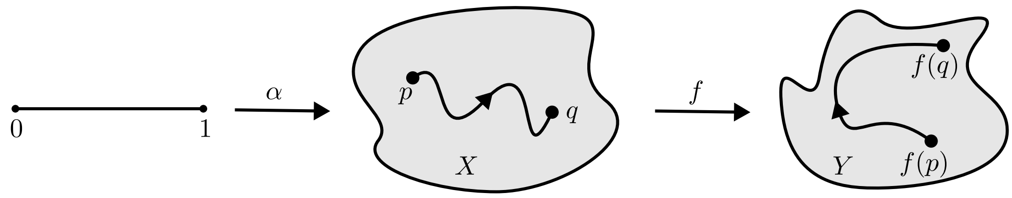


Figure 3.6.1: The continuous map f "sends" a path in X to a path in Y.

Moreover, if α and β are homotopic in X through the homotopy H, then $f \circ \alpha$ and $f \circ \beta$ are homotopic in Y through the homotopy $f \circ H$. Therefore, the map f_* can descend to a map between the two spaces of the homotopy classes of paths which we will still denote by f_*:

$$f_* : \mathcal{P}(X, p, q)/\sim \to \mathcal{P}(Y, f(p), f(q))/\sim,$$
$$[\alpha] \mapsto [f \circ \alpha].$$

In particular, if f is a homeomorphism, then f_* should be a bijective map between $\mathcal{P}(X, p, q)/\sim$ and $\mathcal{P}(Y, f(p), f(q))/\sim$.

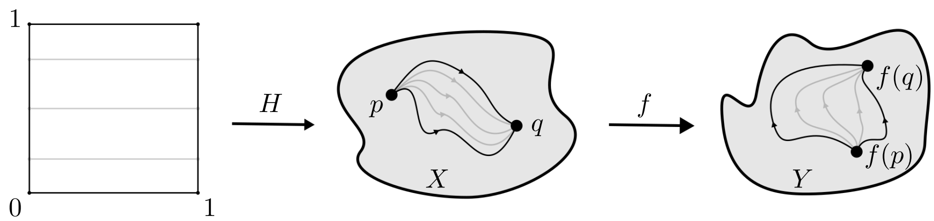


Figure 3.6.2: The continuous map f "sends" a path homotopy in X to a path homotopy in Y.

In particular, we consider loops in X and in Y, and have the following map between fundamental groups:

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

Moreover we can verify the following fact.

Proposition 3.6.1

The map

$$f_* : \pi_1(X, x) \to \pi_1(Y, f(x))$$

is a group homomorphism.

Proof. Given any loops α and β in $\mathcal{L}(X,p)$, we have

$$f \circ (\alpha * \beta) = (f \circ \alpha) * (f \circ \beta).$$

Now let $\alpha_1, \alpha_2, \beta_1, \beta_2$ be loops in $\mathcal{L}(X,p)$, such that $\alpha_1 \sim \alpha_2$ and $\beta_1 \sim \beta_2$. We denote by H_1 the homotopy between α_1 and α_2, and by H_2 the homotopy between β_1 and β_2. Then since we have

$$f \circ (H_1 * H_2) = (f \circ H_1) * (f \circ H_2),$$

we have

$$f_*([\alpha_1] * [\beta_1]) = f_*([\alpha_1]) * f_*([\beta_1]).$$

Hence f_* is a group homomorphism. □

Let X, Y and Z be three path connected spaces, and

$$f : X \to Y, \quad g : Y \to Z,$$

be two continuous maps. Let x be a point in X, then by the previous proposition, we have three group homomorphisms

$$f_* : \pi_1(X,x) \to \pi_1(Y, f(x)),$$
$$g_* : \pi_1(Y, f(x)) \to \pi_1(Z, g(f(x)),$$
$$(g \circ f)_* : \pi_1(X,x) \to \pi_1(Z, g(f(x))).$$

We can verify the following fact.

Proposition 3.6.2

The three homomorphisms satisfy the following relation:

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

Proof. Let α_1 and α_2 be two homotopic loops in $\mathcal{L}(X,x)$, and H be the path homotopy between them. Then consider the composition of maps, we have

$$(g \circ f) \circ H = g \circ (f \circ H),$$

which is a path homotopy between $g \circ f \circ \alpha_1$ and $g \circ f \circ \alpha_2$.

Hence, we have

$$(g_* \circ f_*)([\alpha_1]) = g_*(f_*([\alpha_1])).$$

□

Now we consider a simple case where $X = Y$ and f is the identity map. As a first guess, the corresponding homomorphism f_* should be an isomorphism, since nothing is changed under an identity map.

Lemma 3.6.3

For any path connected topological space X with a base point p, the identity map id_X induces the identity isomorphism

$$\mathrm{id}_{\pi_1(X,p)} = (\mathrm{id}_X)_* : \pi_1(X,p) \to \pi_1(X,p).$$

Proof of Lemma 3.6.3. Let α be a loop in $\mathcal{L}(X,p)$, we have

$$\mathrm{id}_X \circ \alpha = \alpha.$$

Hence for any $[\alpha] \in \pi_1(X,p)$, we have

$$(\mathrm{id}_X)_*([\alpha]) = ([\alpha]).$$

□

In fact, we have the following more general statement.

Proposition 3.6.4

If the map

$$f : X \to Y,$$

is a homeomorphism, then

$$f_* : \pi_1(X,p) \to \pi_1(Y, f(p)),$$

is an isomorphism.

Proof. Since f is a homeomorphism, it admits an inverse

$$f^{-1} : Y \to X,$$

which is also a homeomorphism. In the following, we will show that f_* and $(f^{-1})_*$ are inverse to each other.

Since we have

$$f \circ f^{-1} = \mathrm{id}_Y, \quad f^{-1} \circ f = \mathrm{id}_X.$$

Combining Proposition 3.6.2 and Lemma 3.6.3, we have

$$f_* \circ (f^{-1})_* = (f \circ f^{-1})_* = (\mathrm{id}_Y)_* = \mathrm{id}_{\pi_1(Y,f(x))}.$$
$$(f^{-1})_* \circ f_* = (f^{-1} \circ f)_* = (\mathrm{id}_X)_* = \mathrm{id}_{\pi_1(X,x)}.$$

this implies that the homomorphism f_* is both injective and surjective, hence an isomorphism. □

Invariance of fundamental group under homotopy

The fact that the fundamental group is invariant by an homeomorphism is not surprising, since two homeomorphic spaces are considered the same in the topological point of view and the fundamental group is an topological invariant.

In fact, the fundamental group is invariant under a homotopy equivalence which is strictly weaker than a homeomorphism. (Here by being invariant, we mean that up to isomorphism it is the same.)

Let X and Y be two path connected topological spaces. Assume that there are two continuous maps

$$f : X \to Y \text{ and } g : X \to Y.$$

Choose $p \in X$ as a base point. We consider the homomorphisms induced by them

$$f_* : \pi_1(X,p) \to \pi_1(Y, f(p)),$$
$$g_* : \pi_1(X,p) \to \pi_1(Y, g(p)).$$

We assume that f and g are homotopic to each other, and denote by H the homotopy between them. Then we have the following path in Y by considering the trace of p:

$$\begin{aligned} \beta : [0,1] &\to Y \\ t &\mapsto H(p,t) \end{aligned}.$$

With these notation, we have the following proposition.

Proposition 3.6.5

If f and g are homotopic, then $f_* = [\beta] * g_* * [\overline{\beta}]$.

Proof. For any $t \in [0,1]$, we consider the path

$$\begin{aligned} \beta_t : [0,1] &\to Y \\ s &\mapsto \beta(st) \end{aligned}.$$

Let α be a loop in $\mathcal{L}(X,p)$. The homotopy H between f and g induces a general homotopy between $f \circ \alpha$ and $g \circ \alpha$, denoted by H'. For each $t \in [0,1]$, we have

$$H'_t = H_t \circ \alpha.$$

Then we define the following map

$$\begin{aligned} F : [0,1] \times [0,1] &\to Y \\ (s,t) &\mapsto (\beta_t * H'_t * \overline{\beta_t})(s) \end{aligned}$$

This is a path homotopy between $f \circ \alpha$ and $\beta * (g \circ \alpha) * \overline{\beta}$ (See Figure 3.6.3 for an illustration).

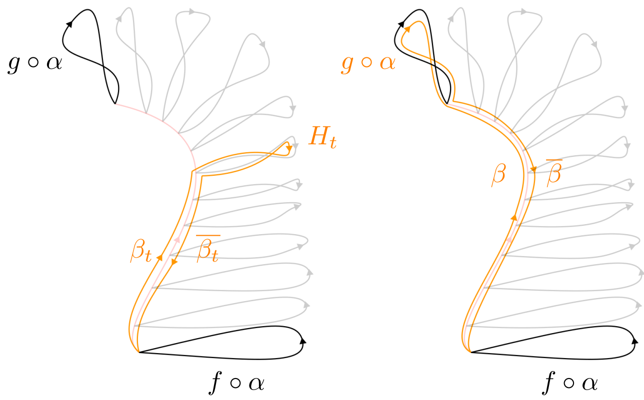


Figure 3.6.3: The homotopy between $f \circ \alpha$ and $\beta * (g \circ \alpha) * \overline{\beta}$.

If $\alpha' \in \mathcal{L}(X,p)$ be a loop homotopic to α, and β' be a path from $f(p)$ to $g(p)$ homotopic to β, then we have the following sequences of homotopic loops in Y:

$$f \circ \alpha' \sim f \circ \alpha \sim \beta * (g \circ \alpha) * \overline{\beta} \sim \beta' * (g \circ \alpha') * \overline{\beta'}.$$

Therefore, we have the desired identity

$$f_*([\alpha]) = [\beta] * g_*([\alpha]) * [\overline{\beta}],$$

which holds for any $[\alpha] \in \pi_1(X,p)$. Hence

$$f_* = [\beta] * g_* * [\overline{\beta}].$$

□

The above proposition tells us that the two maps f_* and g_* are different by a change of base point given by β. Using the same notation as in previous section, we denote by Φ_β the change of base point isomorphism between $\pi_1(Y, f(p))$ and $\pi_1(Y, g(p))$ induced by β. The above proposition is equivalence to the existence of the following commutative diagram.

Proposition 3.6.6

With the notation introduced above, we have the following commutative diagram

$$\begin{array}{ccc} \pi_1(X,p) & \xrightarrow{f_*} & \pi_1(Y,f(p)) \\ & \underset{g_*}{\searrow} & \uparrow \Phi_\beta \\ & & \pi_1(Y,g(p)) \end{array}$$

Applying the above proposition to maps in $C(X)$ and maps in $C(Y)$, we have the following invariance of fundamental groups under homotopy.

Theorem 3.6.7

If X and Y are two homotopy equivalent path connected topological spaces, we have

$$\pi_1(X,x) \cong \pi_1(Y,y),$$

for any $x \in X$ and $y \in Y$.

Proof. By the definition of homotopy equivalence, we have maps

$$f : X \to Y \text{ and } g : Y \to X.$$

satisfying

$$f \circ g \sim \mathrm{id}_Y \text{ and } g \circ f \sim \mathrm{id}_X.$$

By considering the isomorphism induced by a change of base point, it is enough to prove the statement for a special choice of x and y. For our convenience, we choose x to be in the image of g, and we denote by $y \in Y$ with $g(y) = x$. We would like to show that

$$f_* : \pi_1(X,x) \to \pi_1(Y,f(x))$$

is an isomorphism.

Let H denote the hotomopy between $g \circ f$ and id_X. Let β denote the path

$$\begin{aligned} \beta : [0,1] &\to X \\ t &\mapsto H(x,t) \end{aligned}$$

which is the trace of x under the homotopy H. Then by Proposition 3.6.6, we have

$$\Phi_\beta \circ (g \circ f)_* = (\mathrm{id}_X)_* = \mathrm{id}_{\pi_1(X,x)},$$

and the following commutative diagram

$$\begin{array}{ccccc} \pi_1(X,x) & \xrightarrow{f_*} & \pi_1(Y,f(x)) & \xrightarrow{g_*} & \pi_1(X,g(f(x))) \\ & & \underset{(\mathrm{id}_X)_*}{\searrow} & & \downarrow \Phi_\beta \\ & & & & \pi_1(X,x) \end{array}$$

Hence we have

$$(\Phi_\beta \circ g_*) \circ f_* = \mathrm{id}_{\pi_1(X,x)}.$$

which implies that f_* is injective.

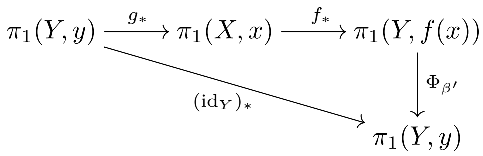
Let H' denote the homotopy between $f \circ g$ and id_Y. Let β' denote the path

$$\begin{aligned} \beta' : [0,1] &\to Y \\ t &\mapsto H'(y,t) \end{aligned}$$

which is the trace of $y = f(x)$ under the homotopy H'. Then by Proposition 3.6.6, we have

$$\Phi_{\beta'} \circ (f \circ g)_* = (\mathrm{id}_Y)_* = \mathrm{id}_{\pi_1(Y,y)},$$

and the following commutative diagram



Therefore, we have

$$(\Phi_{\beta'} \circ f_*) \circ g_* = (\mathrm{id}_Y)_* = \mathrm{id}_{\pi_1(Y,y)}.$$

This implies that

$$\Phi_{\beta'} \circ f_* : \pi_1(X,x) \to \pi_1(Y,y),$$

is surjective. Notice that $\Phi_{\beta'}$ is an isomorphism, hence if f_* is not surjective, then neither is $\Phi_{\beta'} \circ f_*$ which is a contradiction. Hence we have f_* is surjective. Together with the injectivity of f_*, we may conclude that f_* is an isomorphism, hence the theorem. □

As discussed before, applying deformation retraction is a special way to get two spaces which are homotopy equivalent to each other. We first consider certain examples of this kind.

Example 3.6.8.
Let $\mathbb{D}$ denote the closed unit disk in $\mathbb{C}$, and O denote its center. Then $\{O\}$ is a deformation retraction of $\mathbb{D}$. Hence for any point $z \in \mathbb{D}$, we have

$$\pi_1(\mathbb{D}, z) \cong \pi_1(\{O\}, O) \cong \{e\}.$$

Hence the fundamental group of $\mathbb{D}$ is trivial.

This example is in fact a special case for a more general fact.

Proposition 3.6.9

Let X be a path connected topological space. If X is contractible, then for any $x \in X$, we have $\pi_1(X,x)$ is trivial.

Proof. By Definition 3.1.11, the space X is contractible if and only if there is a homotopy between the identity map id_X and the constant map c_x for some point x. Consider the inclusion map

$$\begin{aligned} \iota_x : \{x\} &\to X, \\ x &\mapsto x. \end{aligned}$$

On one hand, we have

$$c_x \circ \iota_x = c_x,$$

at the same time, by definition, we have

$$\iota_x \circ c_x \sim \mathrm{id}_X.$$

Hence X and $\{x\}$ are homotopy equivalent. Therefore, their fundamental groups are isomorphic to each other. Hence for any $x \in X$, we have

$$\pi_1(X, x) = \{e\}.$$

□

Remark 3.6.10.
The converse is not true. For example, the 2-sphere has trivial fundamental group, but is not contractible. In fact two homotopy equivalent space have isomorphic nth homotopy group for any $n \in \mathbb{N}$. We can see that $\pi_2(S^2)$ is not trivial and a single point has trivial nth homotopy group for any $n \in \mathbb{N}$. One may continue to ask if having isomorphic nth homotopy group for any $n \in \mathbb{N}$ can tell the homotopy equivalence. The answer is not true. One counter-example is given by $S^3 \times \mathbb{RP}^2$ and $S^2 \times \mathbb{RP}^3$. Notice that one is orientable while the other is not. All these will be explained in details in the future.

Definition 3.6.11

A path connected topological space X is said to be ***simply connected*** if its fundamental group is trivial.

By the Proposition 3.6.9, we have the following fact.

Corollary 3.6.12

If a path connected topological space X is contractible, then it is simply connected.

Previously, we introduce the construction of a cone based on a space. Notice that topologically the disk can be consider as a cone based on S^1. With this observation, we have the following criteria to detect which loop in $\mathcal{L}(X, x)$ has trivial homotopy class.

Proposition 3.6.13

A loop

$$\alpha : S^1 \to X,$$

is homotopic to a point in X if and only if α can be extended to a continuous map

$$\widetilde{\alpha} : \mathbb{D} \to X.$$

Proof. A loop α is homotopic to the constant loop

$$\begin{aligned} c_x : S^1 &\to X, \\ t &\mapsto x, \end{aligned}$$

if and only if there is a homotopy

$$H : S^1 \times [0, 1] \to X,$$

between α and c_x for some $x \in X$, or equivalently there is a homotopy

$$H : S^1 \times [0, 1] \to X,$$

such that

$$H(S^1 \times \{1\}) = x.$$

This moreover is equivalent to the existence of the following commutative diagram

$$\begin{array}{ccc} S^1 \times [0,1] & \xrightarrow{H} & X \\ \big\downarrow \pi & \nearrow \overline{H} & \\ \mathrm{Cone}(S^1) & & \end{array}$$

Hence we have the proposition. □

Example 3.6.14.
Consider the region in $\mathbb{C}$ defined as follows:

$$A := \{z \in \mathbb{C} \mid 1 \leq |z| \leq 2\}.$$

Notice that

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

is a strong deformation retraction of A. With the retraction map

$$\begin{aligned} r : A &\to S^1 \\ z &\mapsto \frac{z}{|z|} \end{aligned}$$

The homotopy between $\iota \circ r$ and id_A can be given as follows:

$$\begin{aligned} H : A \times [0,1] &\to A \\ (z,t) &\mapsto \frac{z}{1 - t + t|z|}. \end{aligned}$$

Hence A and S^1 are homotopy equivalent. We have

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

3.7 First glance on Seifert-Van-Kampen's Theorem

As we can see from the previous discussion, the fundamental group is a both elementary and important invariant of a topological space. Once it is defined, an immediate question is how we can compute it.

Previously, we have computed fundamental groups of some spaces such as single point sets, disks, which are topologically simple. We also introduce the homotopy invariance of fundamental groups. With this property, we can deform th space without changing the isomorphic type of the fundamental group.

But we still have the problem: how to compute the fundamental group of a topologically complicated space. In the following, we will introduce an important tool for studying fundamental groups of complicated spaces which is called the *Seifert van Kampen Theorem*.

To study the fundamental group, we have to consider loops in a space up to homotopy, which, as one could imagine, is difficult in general. For example, in the previous part, we have spent quite some time to compute even the fundamental group of S^1.

When studying a complicated space, one natural ideal is to decompose it into simple pieces. Then we study the fundamental group of each piece. Finally, we study how to glue the fundamental groups of all pieces back to that for the entire space. The Seifert van Kampen Theorem (the

SVK Theorem) serves as a tool for the last step, i.e telling us how to glue fundamental groups of pieces together.

In this part, we will have a first glance to get a rough idea about the SVK theorem. The statement for general cases will be given after introducing some necessary background in the group theory. We start by discussing a simple case.

Proposition 3.7.1

Let X be a path connected topological space. Let U and V be two path connected open subset such that

- $U \cup V = X$,
- $U \cap V$ is non-empty and path connected,
- all U, V are simply connected.

Then X is simply connected.

Proof. The rough idea is to rewrite any path in X (after a homotopy) into a composition of finitely many paths, each one of which is in either U or V. Then we use the simple connectivity in either U, V or $U \cap V$ to show the path in X is homotopic to a point. A key property used here is the compactness of $[0, 1]$.

Since X is path connected, we will choose $p \in U \cap V$ to be a base point and consider the associated fundamental group with base point $\pi_1(X, p)$. With this choice, all fundamental groups $\pi_1(U, p)$, $\pi_1(V, p)$, $\pi_1(U \cap V, p)$ and $\pi_1(X, p)$ are well defined.

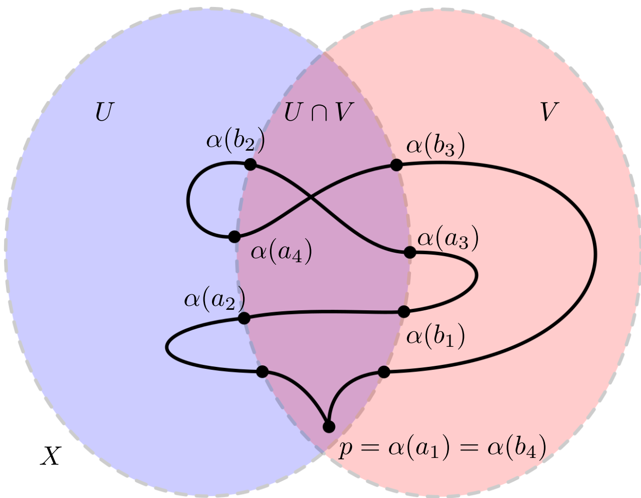


Figure 3.7.1: An open covering of the loop.

Let α be any loop in $\mathcal{L}(X, p)$. We consider $\alpha^{-1}(U)$ and $\alpha^{-1}(V)$ which are unions of intervals in $[0, 1]$, each of which is open in $[0, 1]$. These intervals form an open cover of $[0, 1]$. By the compactness of $[0, 1]$, there is a finite cover of $[0, 1]$ denote by

$$\{I_1, ..., I_k\},$$

with $k \in \mathbb{N}^*$, such that for any $1 \le i \le k$, we have $\alpha(I_i) \subset U$ or $\alpha(I_i) \subset V$. Denote by a_i and b_i the end points of I_i with $a_i < b_i$ (See Figure 3.7.1 for an illustration).

We will find finitely many points to cut $[0,1]$ into finitely many pieces. The cutting points are constructed inductively as follows. Assume that $0 \in I_{i_1}$. Let $t_0 = a_{i_0} = 0$. Let i_1 be the index such that $b_{i_0} \in I_{i_1}$. Then the intersection $I_{i_0} \cap I_{i_1}$ is non empty. Choose

$$t_1 \in I_{i_0} \cap I_{i_1}.$$

with $t_1 > t_0$.

If we determine $t_j \in I_{i_j}$, then consider the $I_{i_{j+1}}$ such that $b_{i_j} \in I_{i_{j+1}}$. Then if $b_{i_{j+1}} = 1$, let $t_{j+1} = 1$, otherwise, choose

$$t_{j+1} \in I_{i_j} \cap I_{i_{j+1}},$$

with $t_{j+1} > t_j$.

In this inductive way, we find a sequence of parameters

$$0 = t_0 < t_1 < \cdots < t_s = 1,$$

with $0 < s \le k+1$. Moreover, we have

$$\alpha([t_j, t_{j+1}]) \subset U \text{ or } \alpha([t_j, t_{j+1}]) \subset V.$$

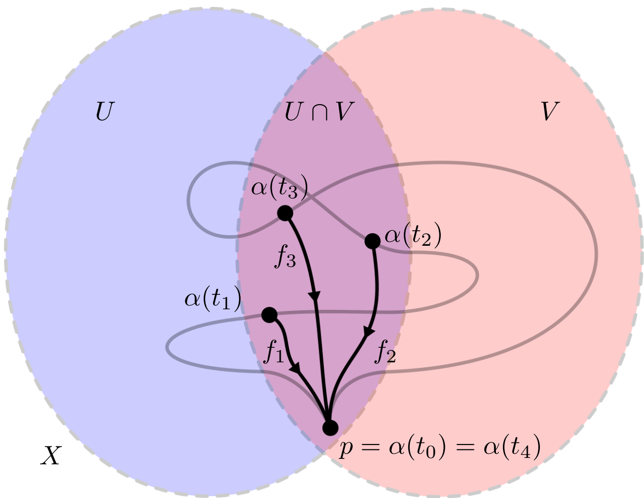


Figure 3.7.2: A partition of the loop.

For any $0 \le j \le s-1$, we denote by J_j the closed interval $[t_j, t_{j+1}]$, and by α_j the restriction of α on J_j. Next we would like to connect $\alpha(t_{j+1})$ to p by a path f_{j+1} of X. We choose the path in the following way.

- if α_j and α_{j+1} are both in U, then f_j is a path in U;
- if α_j and α_{j+1} are both in V, then f_j is a path in V;
- if α_j is in U and α_{j+1} is in V, or if α_j is in V or α_{j+1} is in U, then α_{j+1} is in $U \cap V$, and we choose f_j to be in $U \cap V$.

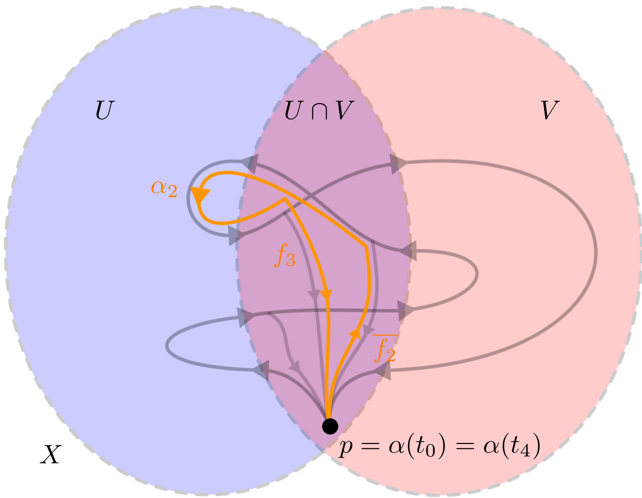


Figure 3.7.3: Decompose the loop.

Let f_0 and f_s be the constant path c_x. Then for any $0 \leq j \leq s-1$, we consider the composition

$$\beta_j = f_{j+1} * \alpha_j * \overline{f_j}.$$

(See Figure 3.7.3 for an illustration.)

From the construction, we have β_j either in $\mathcal{L}(U, p)$ or in $\mathcal{L}(V, p)$. Since both U and V are simply connected, we have β_j is homotopic to p.

On the other hand, in X, we know that α is homotopic to the path

$$\begin{aligned} &\beta_{s-1} * \cdots * \beta_0 \\ =&(f_s * \alpha_{s-1} * \overline{f_{s-1}}) * (f_{s-1} * \alpha_{s-2} * \overline{f_{s-2}}) * \cdots * (f_1 * \alpha_0 * \overline{f_0}) \sim \alpha \end{aligned}$$

Hence α is homotopic to c_p. We then can conclude that X is simply connected. □

The construction in the above proof can be used to show something more. Let X be a path connected topological space. For any path connected subset $A \subset X$, we have the inclusion map

$$\begin{aligned} \iota_A : A &\to X, \\ y &\mapsto y, \end{aligned}$$

which induces a group homomorphism:

$$\begin{aligned} (\iota_A)_* : \pi_1(A, p) &\to \pi_1(X, p), \\ [\alpha]_A &\mapsto [\alpha]_X, \end{aligned}$$

where $p \in A$, $\alpha \in \mathcal{L}(A, p) \subset \mathcal{L}(X, p)$, $[\alpha]_A$ the homotopy class of α in A, and $[\alpha]_X$ the homotopy class of α in X. Since A is a subset of X, and any path homotopy in A is also a path homotopy in X, the above homomorphism is well defined. (Also, one may notice that the inclusion map is a continuous map.)

With these notation, we have the following fact.

Proposition 3.7.2

Let U and V be two path connected open subsets of X such that

- $U \cup V = X$,
- $U \cap V$ is non-empty and path connected,

Let $p \in U \cap V$. Then the fundamental group $\pi_1(X, p)$ is generated by

$$(\iota_U)_*(\pi_1(U, p)) \cup (\iota_V)_*(\pi_1(V, p)).$$

Proof. Given any loop α in $\mathcal{L}(X, p)$, as in the proof of the previous proposition, it is homotopic to a composition

$$\beta_{s-1} * \cdots * \beta_0,$$

where fore each $0 \leq j \leq s-1$, we have β_j either in $\mathcal{L}(U, p)$ or in $\mathcal{L}(V, p)$, hence $[\beta_{s-1}]$ is contained in

$$(\iota_U)_*(\pi_1(U, p)) \cup (\iota_V)_*(\pi_1(V, p)).$$

Therefore, we have the proposition. □

A presentation of a group is in general a way of describing a group using generators. In order to know how to compute group operations, a presentation of a group also involves the information satisfied by the generators, which will be called relations.

With this being said, the SVK theorem is all about giving a presentation of $\pi_1(X, p)$ using $\pi_1(U, p)$ and $\pi_1(V, p)$. The above proposition gives a generating set of $\pi_1(X, p)$, hence only half of the actual SVK Theorem (for an open cover of two subsets). Another half is about the relation satisfied by the elements in $(\iota_U)_*(\pi_1(U, p)) \cup (\iota_V)_*(\pi_1(V, p))$.

In order to be able to talk about this, we have to introduce first some necessary background in group theory on amalgamations and HNN extensions of groups. Before that, let us first see how to use these baby versions of SVK theorem to study some topological spaces.

Corollary 3.7.3

For any $n \geq 2$, the n-sphere S^n is simply connected.

Proof. We would like to find a suitable pair of open subsets U and V for S^n. Recall that

$$S^n = \{(x_1, ..., x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1^2 + \cdots + x_{n+1}^2 = 1\}.$$

The topology on S^n is the subspace topology induced by the Euclidean metric topology on $\mathbb{R}^{n+1}$.

Let $p = (1, 0, ..., 0)$ and $q = (0, ..., 0, 1)$, and let

$$U = S^n \setminus \{p\}, \quad V = S^n \setminus \{q\}.$$

Both U and V are open subsets and form an open cover of S^n.

We consider the stereographic projection Pr_p from U to

$$P_p := \{(x_1, x_2, ..., x_n) \in \mathbb{R}^{n+1} \mid x_1 = 0\}.$$

Notice that P_p is homeomorphic to $\mathbb{R}^n$ which is simply connected.

Since Pr_p is a homeomorphism from U to P_p, we have U is simply connected. By considering the stereographic projection from q, we prove in the same way that V is simply connected.

The intersection $U \cap V$ is path connected. To see this, we consider its image under Pr_p which is

$$P_p \setminus \{(0, ..., 0)\} = \{(x_1, x_2, ..., x_n) \in \mathbb{R}^{n+1} \mid x_1 = 0\} \setminus \{(0, ..., 0)\}.$$

By Proposition 3.7.1, we have S^n is simply connected. □

Now we consider the wedge sum between spheres.

Corollary 3.7.4

For any $k \geq 2$, for any natural numbers $n_1, ..., n_k$ all greater or equal to 2, the space

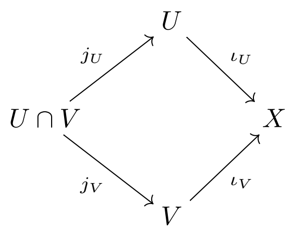
$$S^{n_1} \vee \cdots \vee S^{n_k}$$

is simply connected.

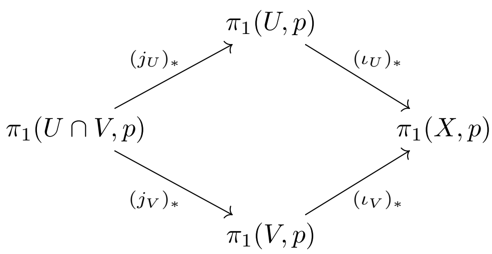
Proof. Exercise. □

3.8 SVK Theorem

As mentioned in the previous section, the whole SVK theorem is about how to glue fundamental groups of pieces of a space together to get the fundamental group of the entire space. More precisely, Let X be a path connected space. Let U and V be two path connected open subset of X which form a cover of X. Moreover, assume that $U \cap V$ is path connected. Let p be a point in $U \cap V$. The inclusion maps among these sets form a commutative diagram:



Since all inclusions are continuous map which induce homomorphisms between fundamental groups, we have another commutative diagram in the fundamental group level:



Previously, we have shown that the fundamental group $\pi_1(X, p)$ is generated by

$$(\iota_U)_*(\pi_1(U, p)) \cup (\iota_V)_*(\pi_1(V, p))$$

In this section, we will discuss the following general statement.

Theorem 3.8.1

We have the following group isomorphism

$$\pi_1(X, p) \cong \pi_1(U, p) \underset{\pi_1(U \cap V, p)}{*} \pi_1(V, p),$$

where the amalgamation through $(j_U)_*$ and $(j_V)_*$.

Remark 3.8.2.
Some necessary background in group theory can be found in Appendix A and Appendix B.

Proof. By the definition of an amalgamated free product between two groups, we have the following relation

$$\pi_1(U,p) \underset{\pi_1(U\cap V,p)}{*} \pi_1(V,p) = \pi_1(U,p) * \pi_1(V,p)/N$$

where

$$N = \langle\langle (j_U)_*(\gamma)((j_V)_*(\gamma))^{-1} \mid \gamma \in \pi_1(U\cap V,p)\rangle\rangle.$$

Consider the above commutative diagram about fundamental groups. First, there is a natural group homomorphism

$$\Phi : \pi_1(U,p) * \pi_1(V,p) \to \pi_1(X,p).$$

From the above discussion, to prove the theorem, it is enough to show that the kernel of Φ is N. We denote by

$$[\alpha_1]_{\epsilon_1} * \cdots * [\alpha_n]_{\epsilon_n}$$

an element in $\pi_1(U,p) * \pi_1(V,p)$, where for any $1 \leq j \leq n$,

$$\epsilon_j \in \{U, V\}$$

is a symbol to show where this element belongs to.

Notice that if a loop $\alpha \in \mathcal{L}(U,p)$ has image in $U \cap V$, then it can also be considered as a loop in $\mathcal{L}(V,p)$. We denote by $[\alpha]_U$ its associated element in $\pi_1(U,p)$ and by $[\alpha]_V$ its associated element in $\pi_1(V,p)$. Since we have

$$[\alpha]_U * [\alpha]_V^{-1}, [\alpha]_V * [\alpha]_U^{-1} \in N,$$

then

$$[\alpha]_U N = [\alpha]_V N.$$

From the commutative diagram above the theorem, we have $N \subset \ker\Phi$. In order to show the equality, it is enough to show that any element $[\alpha_1]_{\epsilon_1} * \cdots * [\alpha_n]_{\epsilon_n} \in \ker\Phi$ will be trivial in

$$\pi_1(U,p) * \pi_1(V,p)/N.$$

In the other words, up to changing the "identity" of a component (belongs to $\pi_1(U,p)$ or to $\pi_1(V,p)$), and computation in $\pi_1(U,p)$ and that in $\pi_1(V,p)$, it can be transformed to an element in N.

It would be easier to discuss on the side of $\pi_1(X,p)$. Notice that the Φ-image of $[\alpha_1]_{\epsilon_1} * \cdots * [\alpha_n]_{\epsilon_n}$ is

$$[\alpha_1 * \cdots * \alpha_n] = [c_p] \in \pi_1(X,p).$$

Hence there is a homotopy between $\alpha_1 * \cdots * \alpha_n$ and c_p in X. We would like to show that this homotopy can be realized by a sequence of homotopy in U, homotopy in V and change of "identity" (from a U-loop to a V-loop, or the other way around).

Assume that

$$H : [0,1] \times [0,1] \to X$$

be a homotopy between $\alpha_1 * \cdots * \alpha_n$ and c_p. (Reminder: all homotopies are path homotopies). Assume that the bottom side of this square corresponds to c_p, while the top side of this square corresponds to $\alpha_1 * \cdots * \alpha_n$. Since H is a path homotopy, the left and right sides of the square are sent to p by H.

Some preparation

Similar to the previous discussion on loop, since $[0,1] \times [0,1]$ is compact, we can decompose it into small rectangles such that the H-image of each closed rectangle is either entirely in U or entirely in V.

To be more precise, assume that the rectangle partition of the square is given by the partition of the first interval

$$0 = s_0 < s_1 < \cdots < s_k = 1$$

and the following partition of the second interval

$$0 = t_0 < t_1 < \cdots < t_l = 1.$$

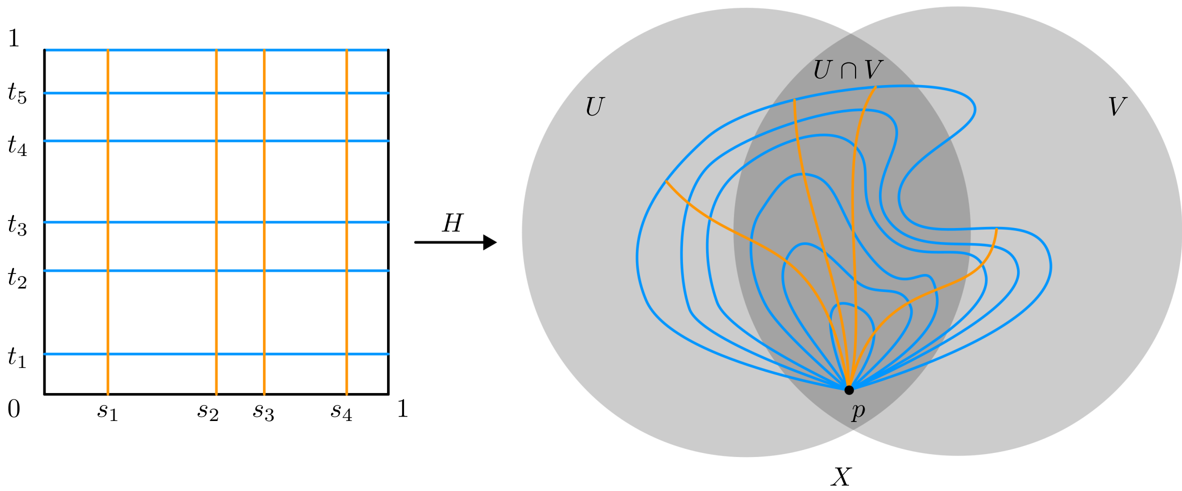


Figure 3.8.1: Partition of a square.

For each $0 \le i \le k$ and $0 \le j \le l$, we denote

$$v_i^j = (s_i, t_j).$$

Each closed rectangle is in the form of

$$R_i^j := [s_i, s_{i+1}] \times [t_j, t_{j+1}]$$

for some $0 \le i < k$ and $0 \le j < l$, with vertices v_i^j, v_i^{j+1},v_{i+1}^j and v_{i+1}^{j+1}. Since the image of each R_i^j is in U or V, we can label them with U or V depends on where they belong. If the image of a rectangle belongs to both U and V, we may choose one to label this rectangle.

We also label each v_i^j in the following way:

- if all rectangles adjacent to it are labeled by U, we label it by U;
- if all rectangles adjacent to it are labeled by V, we label it by V;
- otherwise, we label it by $U \cap V$.

For each v_i^j ($i \neq 0, 1$ or $j \neq 0$), we choose a path going from p to $p_i^j = H(v_i^j)$ and denote by β_i^j, such that, if v_i^j is labeled by U, β_i^j is a path in U; if v_i^j is labeled by V, β_i^j is a path in V; if v_i^j is labeled by $U \cap V$, β_i^j is a path in $U \cap V$.

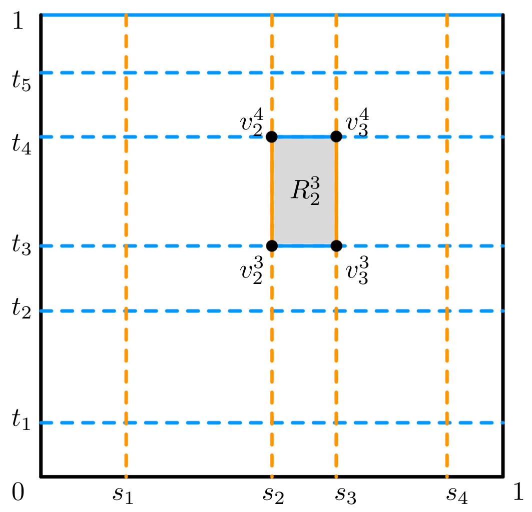


Figure 3.8.2: A small rectangle and its vertices.

Let J_L and J_R be the vertical sides of the square on the left and on the right respectively. Any path η in the square going from J_L to J_R along sides of rectangles corresponds to a loop

$$\gamma = H \circ \eta$$

in X based at p.

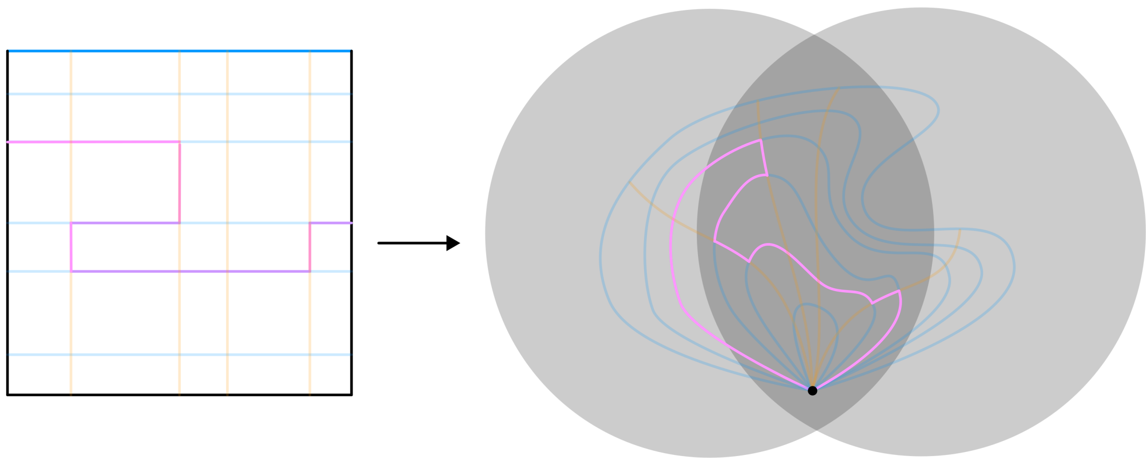


Figure 3.8.3: A path in the square from J_L to J_R and the loop in X associated to it.

Each such path η is a concatenation of a sequence of sides of rectangle R_i^j's. This decomposition of η induces a decomposition of γ into paths with endpoints in

$$\{H(v_i^j) \mid 0 \leq i \leq k, 0 \leq j \leq l\}.$$

By inserting $\overline{\beta_i^j} * \beta_i^j$, te path γ is homotopic to a composition of a sequence of loops each of which is in U or in V.

Another observation is that for any rectangle R_i^j and any two of its vertices, there are two ways to go from one vertex to another along the sides of R_i^j. Their H-image are two path in X homotopic to each other where the homotopy can be given by considering the restriction of H. Moreover this homotopy is either in U or in V depending on the label of R_i^j. Without loss of generality, we may assume that R_i^j is labeled by U. As mentioned above, by taking composition with β_i^j's and their inverse, these paths in X can be completed into loops in U, and the homotopy between paths can induce a homotopy between loops in U.

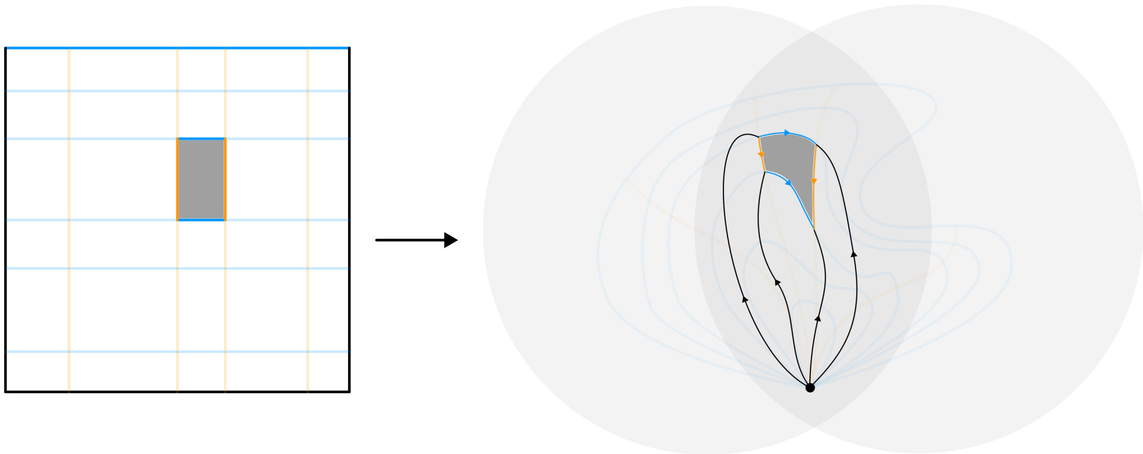


Figure 3.8.4: A homotopy associated to a small rectangle.

Main discussion

Let η_0 be the path with image the top side of the square corresponding to the parameter set for $\alpha_1 * \cdots * \alpha_n$:

$$[0,1] \times \{1\}.$$

Now we would like to modify it step by step. At step m, we obtain a path η_m going from J_L to J_R along the sides of rectangles in the partition. By inserting $\overline{\beta_i^j} * \beta_i^j$ for each v_i^j, the path $\gamma_m = H \circ \eta_m$ is homotopic to a composition of loops in U or in V:

$$\gamma_m \sim \alpha_1^{(m)} * \cdots * \alpha_{r_m}^{(m)}.$$

This gives us an element in $\pi_1(U,p) * \pi_1(V,p)$:

$$[\gamma_m] = [\alpha_1^{(m)}]_{\epsilon_1^m} * \cdots * [\alpha_{r_m}^{(m)}]_{\epsilon_{r_m}^m}.$$

We first describe how η_m changes in the parameter square. For a pair of adjacent vertices v and v' (endpoints of a same side of some R_i^j), we will denote by $\overrightarrow{vv'}$ to denote the path from v to v' along the side between them. Hence η_0 can be expressed as

$$\overrightarrow{v_0^l v_1^l} * \cdots * \overrightarrow{v_{k-2}^l v_{k-1}^l} * \color{red}{\overrightarrow{v_{k-1}^l v_k^l}}.$$

We start from the upper right corner R^{l-1}_{k-1}, modify η_0 to a new path

$$\eta_1 = \overrightarrow{v^l_0 v^l_1} * \cdots * \overrightarrow{v^l_{k-2} v^l_{k-1}} * \overrightarrow{v^l_{k-1} v^{l-1}_{k-1}} * \overrightarrow{v^{l-1}_{k-1} v^{l-1}_k}.$$

Then we consider R^{l-1}_{k-2}, and modify η_1 to

$$\eta_2 = \overrightarrow{v^l_0 v^l_1} * \cdots * \overrightarrow{v^l_{k-3} v^l_{k-2}} * \overrightarrow{v^l_{k-2} v^{l-1}_{k-2}} * \overrightarrow{v^{l-1}_{k-2} v^{l-1}_{k-1}} * \overrightarrow{v^{l-1}_{k-1} v^{l-1}_k}.$$

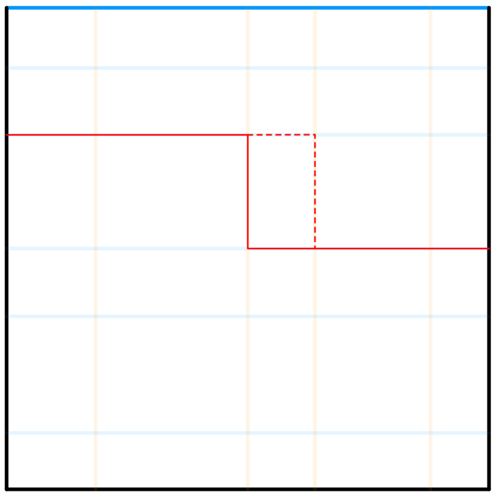


Figure 3.8.5: Modify a path step by step.

By applying this for all rectangle on the top first row, we move the path with image

$$[0, 1] \times \{1\}$$

to a path with image

$$[0, 1] \times \{t_{k-1}\}.$$

We repeat this process, until we meet the bottom line

$$[0, 1] \times \{0\}.$$

Notice that when we go from step m to step $m+1$, we modify the elements associated to one rectangle R^j_i. By our assumption, we have $H(R^{j_0}_{i_0}) \subset U$ or $H(R^j_i) \subset V$. Without loss of generality, we may assume that it is the former.

Then by our choice of β^j_i's, those paths associated to vertices of $R^{j_0}_{i_0}$ are either in $U \cap V$ or in U, hence all of them are path in U. Hence all loops associated to the sides of $R^{j_0}_{i_0}$ are all in U. Moreover when we change from a pair of these loops to another pair through a homotopy, the homotopy can be chosen in U. To simplify the discussion, let us denote by α'_1, α'_2, α'_3 and α'_4 the loops associated to the upper, right, left and lower sides of $R^{j_0}_{i_0}$ respectively.

Now we consider change the element in $\pi_1(U, p) * \pi_1(V, p)$ associated to γ_m and that associated to γ_{m+1}, the last thing we have to be careful is to which fundamental group each letter in those elements belongs to. When we try to write

$$[\alpha'_1]_{\epsilon_1} * [\alpha'_2]_{\epsilon_2} = [\alpha'_3]_{\epsilon_3} * [\alpha'_4]_{\epsilon_4}$$

we should have all these elements are in a same group, hence all markings ϵ_1, ϵ_2, ϵ_3 and ϵ_4 should be the same.

However $[\alpha'_1]_{\epsilon_1}$ and $[\alpha'_2]_{\epsilon_2}$ are obtained from the previous steps. Notice that each side of $R_{i_0}^{j_0}$ has another adjacent rectangle which could possible labeled with V. Without loss of generality, we may assume that this happens for $[\alpha'_1]_{\epsilon_1}$, hence we have $\epsilon_1 = V$. In this case, we have $v_{i_0}^{j_0+1}$ and $v_{i_0+1}^{j_0+1}$ both in $U \cap V$, hence the loop α'_1 is in $\mathcal{L}(U \cap V, p)$. So we can change $[\alpha'_1]_V$ to $[\alpha'_1]_U$ in the quotient group $\pi_1(U,p) * \pi_1(V,p)/N$.

Hence we can make the whole homotopy happen by only using homotopies in U, homotopies in V and change the "identity" (between U-loops and V-loops). This tells us the whole kernel is in N. Hence the theorem. □

3.9 Application of the SVK Theorem

In this part, we use the SVK theorem to compute the fundamental groups of some spaces.

Wedge sum among circles

We first consider a wedge sum between two circles.

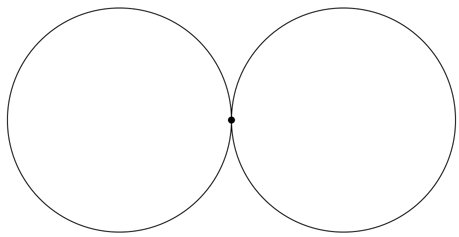


Figure 3.9.1: $S^1 \vee S^1$.

Let R_2 denote the whole space. We denote by A and B the two circles in X, which share a same point $p \in X$. Let I be an open circular arc containing p in A and J be an open circular arc containing p in B. We consider

$$U = A \cup J \quad \text{and} \quad V = B \cup I.$$

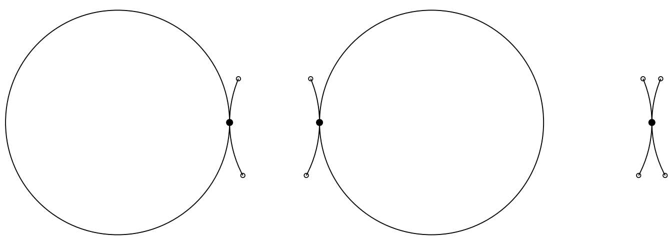


Figure 3.9.2: The open subsets U and V, and their intersection for R_2.

Now we check if U can V satisfy the hypothesis of the SVK theorem. Notice that both U and V are open and path connected. Their intersection

$$U \cap V = I \cup J.$$

is also path connected. In fact it is simply connected, since $\{p\}$ is its deformation retraction.

Now we consider the fundamental group of each piece. Notice that A (resp. B) is a deformation retraction of U (resp. V), hence we have

$$\pi_1(U,p) \cong \pi_1(V,p) \cong \pi_1(A,p) \cong \mathbb{Z}.$$

Since $U \cap V$ is simply connected, it has trivial fundamental group, hence

$$\pi_1(R_2,p) \cong \pi_1(U,p) * \pi_1(V,p) \cong \mathbb{Z} * \mathbb{Z}$$

which is a rank 2 free group.

Now we consider taking the wedge sum between R_2 and one more circle C at p. We denote by R_3 the whole space.

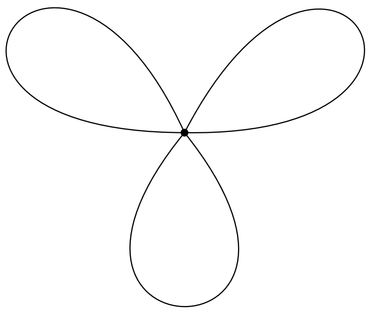


Figure 3.9.3: The 3-rose.

Let K be an open circular arc containing p in C. We consider

$$U' = R_2 \cup K \quad \text{and} \quad V' = C \cup I \cup J.$$

Now we check if U' and V' satisfy the hypothesis of the SVK theorem. Notice that both U' and V' are open and path connected. Their intersection

$$U' \cap V' = I \cup J \cup K.$$

is also path connected. In fact it is simply connected, since $\{p\}$ is its deformation retraction.

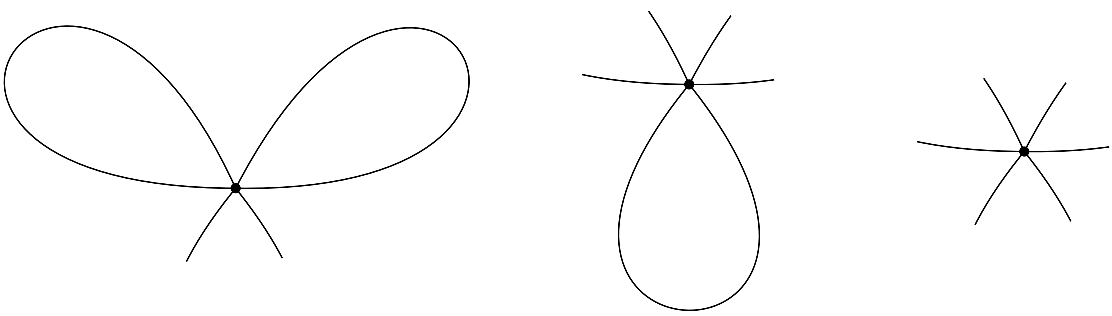


Figure 3.9.4: The open subsets U and V, and their intersection for R_3.

Now we consider the fundamental group of each piece. Notice that R_2 (resp. C) is a deformation retraction of U (resp. V), hence we have

$$\pi_1(U', p) \cong \pi_1(R_2, p) \cong \mathbb{Z} * \mathbb{Z},$$

and

$$\pi_1(V', p) \cong \pi_1(C, p) \cong \mathbb{Z}.$$

Since $U' \cap V'$ is simply connected, it has trivial fundamental group, hence

$$\pi_1(R_3, p) \cong \pi_1(R_2, p) * \pi_1(C, p) \cong \mathbb{Z} * \mathbb{Z} * \mathbb{Z}.$$

which is a rank 3 free group.

We call the wedge sum of n circle the *n-rose*. Its fundamental group is rank n free group.

Graphs

One way of consider graph is to a result of identification of endpoints of several closed compact interval. We can consider it as a quotient space, and in this way we can easily understand its topology. The image of endpoints are called *vertices* and the image of each interval is called an *edge*. A graph is *finite* if it has finitely many vertices and edges. In this part, we assume that all graphs are finite and connected.

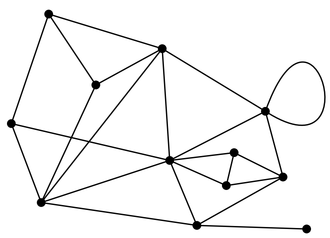


Figure 3.9.5: A graph.

Let G be a graph. We denote $G = (V, E)$ where V is the set of vertices and E is the set of edges. A *subgroup* of G is a graph whose vertices and edges are also vertices and edges of G. A graph is a *tree* if it is simply connected, i.e. there is no loop in it which is homotopically non trivial. A subgroup of G is called a *maximal subtree* if it is a tree and it is maximal with respect to the partial order induced by inclusion of subsets. Notice that maximal subtrees of a graph are not unique.

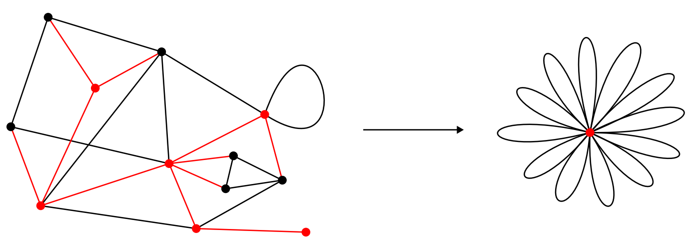


Figure 3.9.6: Collapse a maximal tree of a graph to get a rose.

Let $T \subset G$ be a maximal subtree. Let n denote the number edges in $G \setminus T$. We can collapse T to a point, then each edge which is in $G \setminus T$ becomes a loop, hence we obtain an n-rose R_n.

From this, we can show that G and R_n are homotopy equivalent and has the same fundamental group isomorphic to F_n.

Torus

We consider torus T obtained from the unit square by identify its opposite sides. Let R denote the unit square determined by $(0,0)$, $(1,0)$, $(0,1)$ and $(1,1)$. Let p denote the middle point of R. Let D denote a radius $1/2$ open disk neighborhood of p.

Let $\pi : R \to T$ be the quotient map. Then we set

$$U = \pi(R \setminus \{p\}) \quad \text{and} \quad V = \pi(D).$$

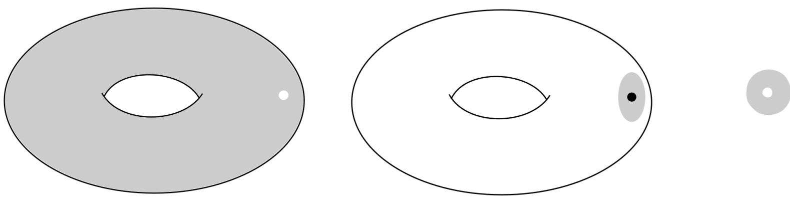


Figure 3.9.7: The open sets U and V, and their intersection for a torus.

Notice that ∂R is a strong deformation retraction of $R \setminus \{p\}$, moreover, this retraction can be realized on T. Hence

$$\pi_1(R \setminus \{p\}) \cong \pi_1(\partial R).$$

The image $\pi(\partial R)$ is a 2-rose (figure eight), hence has fundamental group $\mathbb{Z} * \mathbb{Z}$. Moreover, We denote by q_0 the π-image of vertices of R. Then each pair of opposite sides of R induces a generator of $\pi_1(\partial R)$ by choosing one orientation. We denote them by a and b.

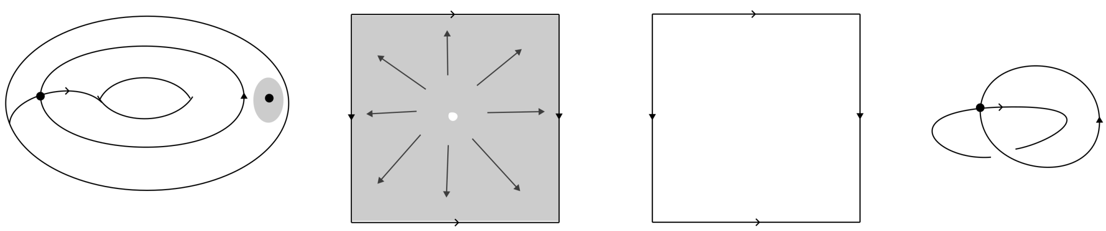


Figure 3.9.8: The fundamental group of U.

Since V is a topologically a disk, hence has trivial fundamental group. Hence by the SVK theorem

$$\pi_1(T, q) \cong \pi_1(U, q)/\langle\langle (j_U)_*([\gamma]) \mid [\gamma] \in \pi_1(U \cap V, q)\rangle\rangle.$$

Hence all we have to study is the $(j_U)_*(\pi_1(U \cap V, q))$. By changing the base point to q_0, the generator of the fundamental group of the punctured disk is then $aba^{-1}b^{-1}$.

Hence a presentation of the fundamental group of $\pi_1(T, q_0)$ is then

$$\langle a, b \mid aba^{-1}b^{-1}\rangle.$$

One may prove moreover that this group is isomorphic to $\mathbb{Z}^2$ the free abelian group of 2 generators.

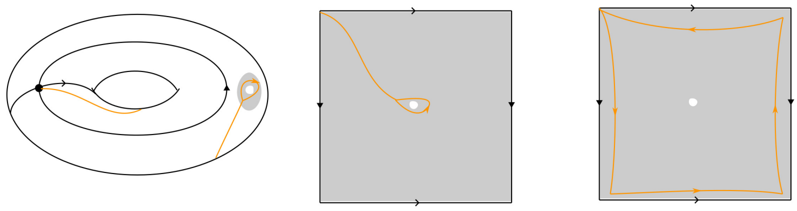


Figure 3.9.9: A generator of $(j_U)_*(\pi_1(U \cap V, p))$.

Connected sums among torus

Now we consider connected sums among torus. We only discuss a connected sum of two torus as an example. Let T and T' be two torus. By make a connected sum, we remove an open disk in T and an open disk in T', then glue the boundary with certain choice of orientation.

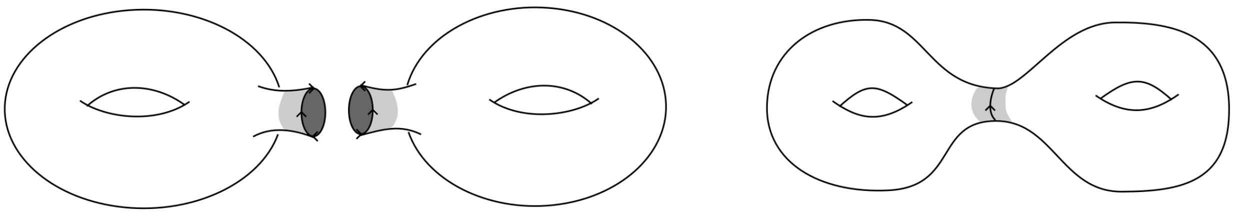


Figure 3.9.10: The connected sum Σ_2 between two torus.

We denote by T_1 and T_1' be the resulting surfaces after removing disks from T and T' respectively. Let A (resp. A') be an open cylinder neighborhood of ∂T_1 (resp. $\partial T_1'$) in T (resp. T').

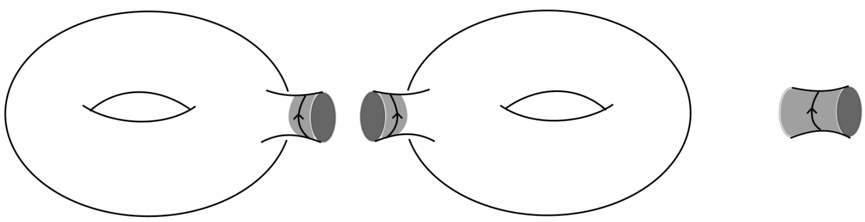


Figure 3.9.11: The connected sum Σ_2 between two torus.

Let $\Sigma_2 = T \# T'$. We choose

$$U = T_1 \cup A' \quad \text{and} \quad V = T_1' \cup A.$$

Then

$$U \cap V = A \cup A'$$

is homeomorphic to a cylinder. Let $p \in U \cap V$. By well choosing presentations of $\pi_1(U, p)$ and $\pi_1(V, p)$, The generator of $\pi_1(U \cap V, p)$ has a presentation $a^{-1}bab^{-1}$ in $\pi_1(U, p)$ and has the presentation $c^{-1}dcd^{-1}$ in $\pi_1(V, p)$. Hence the presentation of $\pi_1(\Sigma_2, p)$ can be given as follows:

$$\langle a, b, c, d \mid a^{-1}bab^{-1}cd^{-1}c^{-1}d\rangle.$$

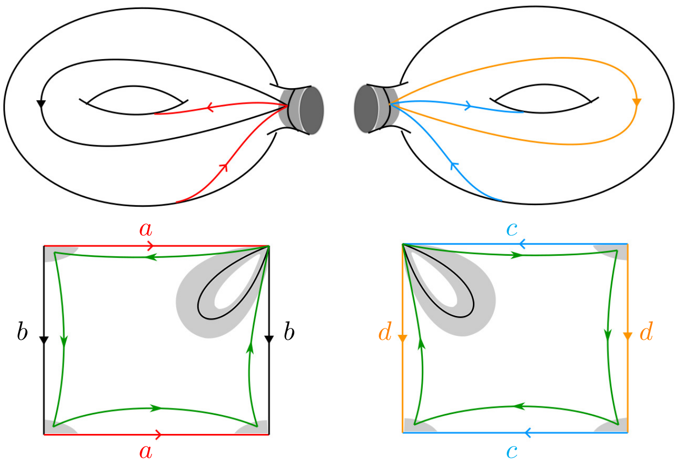


Figure 3.9.12: Identification of boundary elements.

Chapter 4

Covering spaces

Consider the map

$$f : \mathbb{C} \to \mathbb{C},$$
$$z \mapsto z^2.$$

If we look at its restriction on the unit circle S^1, the map f winds the circle around itself twice. Topologically, a loop going around S^1 once is not homotopic to a loop going around S^1 twice. But they still have certain things in common. For example, locally we cannot distinguish them, the homomorphism between fundamental groups induces by f is injective, etc. Such phenomenon relates to a notion called "covering". In fact we see this a lot in daily life. For example, we may consider what happens when we try to roll a paper into a straw. We may consider a straw as a cylinder. When we cut either a disk or a disk with center removed out of the straw, we get several pieces of paper homeomorphic to the part of the straw being cut out. In this chapter, we would like to study covering maps and covering spaces in details.

4.1 Covering maps

Intuitively, if a space X covers another one Y, locally they looks like each other.

Definition 4.1.1

Let X and Y be two topological spaces. A continuous map

$$f : X \to Y,$$

is a ***local homeomorphism***, if for any $x \in X$, it has an open neighborhood $U \subset X$ such that

(i) $f(U) \subset Y$ is open;

(ii) the restriction $f|_U$ is a homeomorphism to its image.

As we will see later, this definition is weaker than what we need for discussing "covering". Let us first give the formal definition of a covering map.

Definition 4.1.2

Let X and Y be two topological spaces. A continuous map

$$f : X \to Y$$

is a ***covering map***, if it is surjective and for any $p \in Y$, it has a neighborhood $V \subset Y$ such that

1. $f^{-1}(V)$ can be written as a disjoint union of sets
$$f^{-1}(V) = \bigsqcup_{\alpha \in \Omega} U_\alpha,$$
2. for each $\alpha \in \Omega$, we have
$$f|_{U_\alpha} : U_\alpha \to V,$$
is a homeomorphism.

In this case, we call X is a ***covering space*** of Y (or simply a ***cover*** of Y) with covering map f.

Remark 4.1.3.
For our convenience, for any $p \in Y$, we will call any neighborhood V satisfying the condition in the definition a *covering neighborhood* of p. We remark that a point p may have many covering neighborhoods. For any example, if V is a covering neighborhood of p, then any neighborhood of p contained in V is again a covering neighborhood of p. Another remark is that if V is a covering neighborhood of $p \in Y$, then V is also a covering neighborhood of any $p' \in \mathring{V}$.

Consider the covering map $f : X \to Y$, given any point $p \in Y$, if V is a covering neighborhood of p, then its intersection with any neighborhood U of p is still a covering neighborhood of p. Hence, we have the following observation.

Proposition 4.1.4

Let $f : X \to Y$ be a covering map, then the open covering neighborhoods of points in Y form a basis of the topology of Y. The connected components of preimages of open covering neighborhoods of points in Y form a basis of the topology of X.

We first see several example to see what this definition requires.

Example 4.1.5.
We consider the example given in the beginning of this chapter. Consider S^1 the unit circle in $\mathbb{C}$ and the map from S^1 to S^1 given by

$$\begin{aligned} f : \mathbb{C} &\to \mathbb{C} \\ z &\mapsto z^2. \end{aligned}$$

For any $\theta \in \mathbb{R}$, let

$$V := \left\{ e^{it} \in S^1 \middle| \, t \in \left(\theta - \frac{\pi}{2}, \theta + \frac{\pi}{2}\right) \right\}.$$

The its preimage under f can be written as $U_1 \sqcup U_2$, where

$$U_1 := \left\{ e^{it} \in S^1 \middle| t \in \left(\frac{\theta}{2} - \frac{\pi}{4}, \frac{\theta}{2} + \frac{\pi}{4} \right) \right\},$$
$$U_2 := \left\{ e^{it} \in S^1 \middle| t \in \left(\frac{\theta}{2} + \frac{3\pi}{4}, \frac{\theta}{2} + \frac{5\pi}{4} \right) \right\}.$$

Notice that

$$U_1 \cap U_2 = \emptyset.$$

Moreover

$$f|_{U_1} : U_1 \to V \quad \text{and} \quad f|_{U_2} : U_2 \to V,$$

are both homeomorphisms. Therefore V is a covering neighborhood of $e^{i\theta}$ (See Figure 4.1.1 for an illustration).

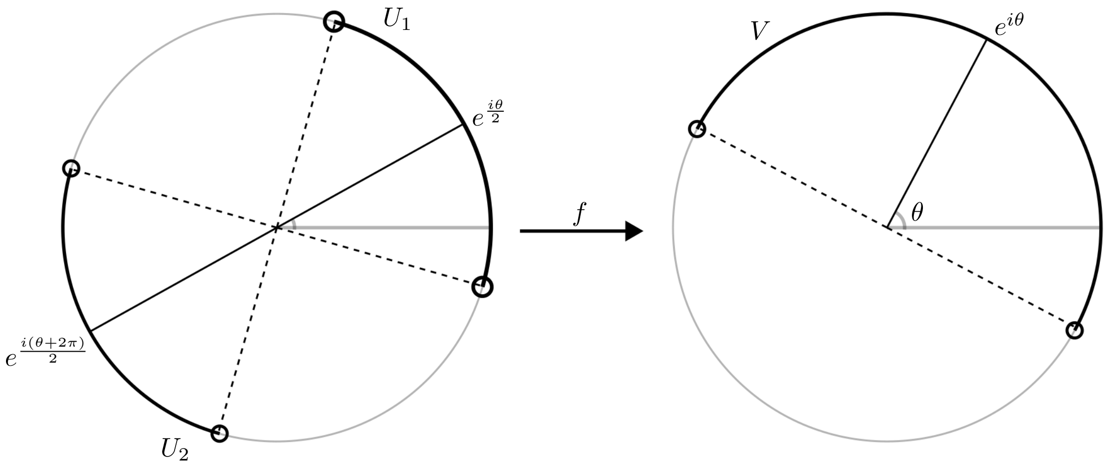


Figure 4.1.1: A covering neighborhood V of $e^{i\theta}$.

Therefore f is a covering map from S^1 to S^1.

The discussion in this example can be used to discuss any map

$$\begin{aligned} f_n : \mathbb{C} &\to \mathbb{C} \\ z &\mapsto z^n \end{aligned}$$

with $n \in \mathbb{N}^*$. The differences between the case $n = 2$ and other cases are the choice of the covering neighborhood V for each θ and the number of U_α's. See Figure 4.1.2 for an illustration for f_n's with $n = 2, 4, 7$.

Example 4.1.6.
This is an example which is local homeomorphism, but not a covering map. We consider

$$\begin{aligned} f : (0, 4\pi) &\to S^1 \\ t &\mapsto \exp(it) \end{aligned}.$$

Notice that given $1 \in S^1$, we consider an open neighborhood V of 1 in S^1 which is an open circular arc. It preimage will contain a connected component which is an interval $U = (4\pi - \epsilon, 4\pi)$ for some $\epsilon > 0$. Notice that the restriction of f on U is not a homeomorphism to V (See Figure 4.1.3 for un illustration).

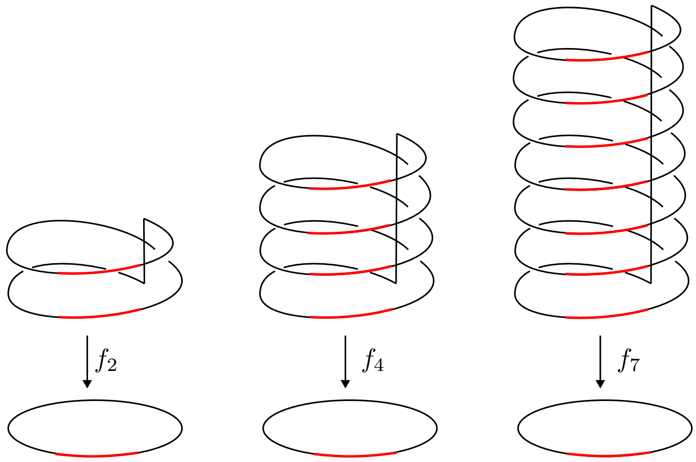


Figure 4.1.2: Covering maps for $n = 2, 4, 7$.

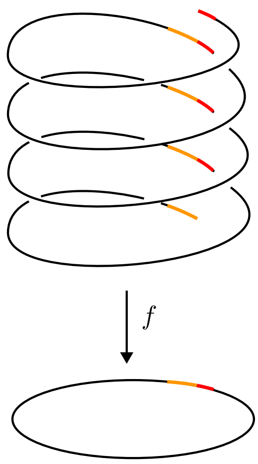


Figure 4.1.3: An illustration of f.

One thing that we may notice is that the number of f_n^{-1}-preimage of each point in S^1 is the same. In fact this is not a surprise.

Proposition 4.1.7

Let X be a cover of a topological space Y with covering map f. Assume that Y is connected, and there is natural number $n \in \mathbb{N}^*$, such that $|f^{-1}(p_0)| = n$ for some $p_0 \in Y$. Then for any $p \in Y$, we have $|f^{-1}(p)| = n$.

Proof. We consider the map

$$\begin{aligned} \varphi : Y &\to \mathbb{N} \cup \{\infty\} \\ p &\mapsto |f^{-1}(p)|. \end{aligned}$$

As a convention, if $f^{-1}(p)$ is not finite, we define its value under φ to be ∞. We consider the discrete topology on $\mathbb{N} \cup \{\infty\}$.

For any $k \in \mathbb{N} \cup \{\infty\}$, we consider its preimage

$$\varphi^{-1}(k) = \{p \in Y \mid |f^{-1}(p)| = k\}.$$

For any $p \in \varphi^{-1}(k)$, by the definition of a covering map, there is a neighborhood V of p such that

1. $f^{-1}(V)$ can be written as a disjoint union of open sets
$$f^{-1}(V) = \bigsqcup_{\alpha \in \Omega} U_\alpha,$$
2. for each $\alpha \in \Omega$, we have
$$f|_{U_\alpha} : U_\alpha \to V,$$
is a homeomorphism.

For any $q \in V$, we have

$$f^{-1}(q) \subset f^{-1}(V),$$

and for any $\alpha \in \Omega$, we have

$$|f^{-1}(q) \cap U_\alpha| = 1.$$

Hence we have

$$\varphi(q) = \varphi(p).$$

Hence we have

$$V \subset \varphi^{-1}(k).$$

This implies that $\varphi^{-1}(k)$ is open. Hence φ is a continuous map.

Since Y is connected, we have $\varphi(Y) \subset \mathbb{N} \cup \{\infty\}$ is connected. Hence φ is constant.

Since we have a point p_0 with $\varphi(p_0) = n$, for any $p \in Y$, we have

$$|f^{-1}(p)| = \varphi(p) = n.$$

□

Remark 4.1.8.
Since path connectivity implies connectivity, the above proposition holds in particular for path connected space Y.

Definition 4.1.9

Let X be a cover of a connected topological space Y with covering map f. If there is $p \in Y$, such that

$$|f^{-1}(p)| = n \in \mathbb{N}^*,$$

then we call X a ***finite cover*** of Y, and n is called the ***cover index*** or ***cover degree***.

Remark 4.1.10.
A cover of a topological space is not always finite. Here is one example. Let S^1 be the unit circle in $\mathbb{C}$. We consider the map

$$f : \mathbb{R} \to S^1,$$
$$\theta \mapsto e^{i\theta}.$$

This is a covering map. Notice that

$$f^{-1}(1) = 2\pi\mathbb{Z}.$$

Hence with the covering map f, the space $\mathbb{R}$ is not a finite cover of S^1.

Remark 4.1.11.
A cover of a space Y need not to be different from Y. For example, for any $n \in \mathbb{N}^*$, with the covering map f_n, the unit circle S^1 is a index n cover of itself.

Remark 4.1.12.
Recall that the definition of a covering space of a space Y includes two parts of information: the space X and the covering map f.

Remark 4.1.13.
The above example reminds us that when we discuss the notion of quotient space we use exactly the same example with $\mathbb{R}$ and S^1. In fact, in some cases, instead of seeing Y and its covering space X as two spaces with special relation, we can also consider Y as a quotient space of X and the covering map can be consider as a quotient map. We will discuss this later in details.

Proposition 4.1.14

A degree 1 covering map is a homeomorphism.

Proof. If a covering map f from X to Y has degree 1, then it is bijective by the definition of degree. The covering neighborhood of points in Y form a basis of the topology in Y. Hence f is continuous. On the other hand, the connected components of preimage of a open covering neighborhood of points in Y form a basis of X. Hence f^{-1} is also continuous. Therefore f is a homeomorphism. □

4.2 Lifting

Let X and Y be two path connected spaces. Assume that X is a cover of Y with covering map f. For any $p \in Y$, we call a point $\widetilde{p} \in f^{-1}(p)$ a ***lift*** of p. By the definition of a covering map, for each p, we can have a covering neighborhood V of p, such that

$$f^{-1}(V) = \bigsqcup_{\alpha \in \Omega} U_\alpha,$$

such that for each $\alpha \in \Omega$, U_α and V are homeomorphic through f. We then call U_α a ***lift*** of V.

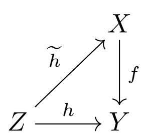
We can also talk about lifts of continuous maps. Let

$$h : Z \to Y$$

be a continuous map. If we have a map

$$\widetilde{h} : Z \to X,$$

such that we have the following commutative diagram



we then call $\widetilde{h}$ a ***lift*** of h.

Remark 4.2.1.
One possible advantage to lift a map is that it is possible that the topology of X is simpler than that of Y. Then it would be easier to study maps to X than those to Y. For example, when we compute the fundamental group of S^1, we try to lift maps to $\mathbb{R}$ which is simply connected and the linear structure in $\mathbb{R}$ can be use to construct homotopies which would be quite helpful.

Remark 4.2.2.
Later we will see that given a space Y and its cover X, not every continuous map to Y can be lift to a map to X. We will also give a criterion to tell in which situation a continuous map to Y is liftable.

Lifts of paths

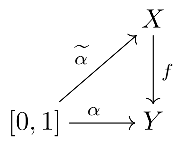
We start by discussing lifts of paths. Let X and Y be two path connected spaces, and

$$f : X \to Y,$$

be a covering map.

Definition 4.2.3

Let α be a path in Y, then a ***lift*** of α in X is a path $\widetilde{\alpha}$ in X satisfying the following commutative diagram



We first show that lifts of any path do exists.

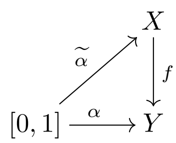
Proposition 4.2.4

Let α be a path in Y with $\alpha(0) = p$. Then for each $\widetilde{p} \in f^{-1}(p)$, there is a unique lift $\widetilde{\alpha}$ of α, such that $\widetilde{\alpha}(0) = \widetilde{p}$.

Proof. The goal is to find a path

$$\widetilde{\alpha} : [0, 1] \to X,$$

such that the following diagram commutes



For any $t \in [0, 1]$, by the definition of the covering map, the point $\alpha(t)$ has a open covering neighborhood V_t in Y, such that each $\widetilde{\alpha(t)} \in f^{-1}(t)$ has a neighborhood U_t satisfying that

$$f|_{U_t} : U_t \to V_t$$

is a homeomorphism. For each t, we have a interval open neighborhood I_t of t such that

$$I_t \subset \alpha^{-1}(V_t).$$

These interval I_t's form an open cover of $[0,1]$. Since $[0,1]$ is compact, we can choose finitely many of them to cover $[0,1]$. Let n be the number of these intervals. The endpoints of these intervals form a sequence

$$0 = s_0 < s_1 < \cdots < s_n < s_{n+1} = 1.$$

Choose

$$0 = s_0 = t_0 < s_1 < t_1 < s_2 < t_2 < \cdots < t_{n-1} < s_n < t_n < s_{n+1} = t_{n+1} = 1$$

Then we have a finite partition of $[0,1]$:

$$0 = t_0 < t_1 < \cdots < t_n < t_{n+1} = 1,$$

such that for each $0 \le i \le n$, $\alpha([t_i, t_{i+1}])$ is contained in one of the covering neighborhood V_t's chosen in the beginning of this construction, and we denote it by V_i. In the following, for any $0 \le i \le n$, we will denote

$$\alpha_i = \alpha|_{[t_i, t_{i+1}]}.$$

We will construct $\widetilde{\alpha}$ piece by piece. Let $\widetilde{p} \in f^{-1}(p)$. Then we have a neighborhood of $\widetilde{p}$ denoted by U_0 homeomorphic to V_0 via f. Denote by

$$f_0 = f|_{U_0} : U_0 \to V_0.$$

Then we define

$$\widetilde{\beta}_0 : [0, t_1] \to X,$$

by $\widetilde{\beta}_0 = f_0^{-1} \circ \alpha_0$.

Assume that we have $\widetilde{\beta}_0, ..., \widetilde{\beta}_i$ with $t_{i+1} < 1$. We consider $\alpha(t_{i+1})$ and the covering neighborhood V_{i+1} containing $\alpha([t_{i+1}, t_{i+2}])$.

Let U_{i+1} be the neighborhood of $\widetilde{\beta}_i(t_{i+1})$ which is homeomorphic to V_{i+1} via f. We denote by

$$f_{i+1} = f|_{U_{i+1}} : U_{i+1} \to V_{i+1}.$$

Then we define

$$\widetilde{\beta}_{i+1} : [t_{i+1}, t_{i+2}] \to X,$$

by $\widetilde{\beta}_{i+1} = f_{i+1}^{-1} \circ \alpha_{i+1}$.

In this way, we obtain a sequence of maps $\widetilde{\beta}_i$ for $0 \le i \le n$. Since for any $0 \le i \le n-1$, we have

$$\widetilde{\beta}_i(t_{i+1}) = \widetilde{\beta}_{i+1}(t_{i+1}),$$

we can "glue" all these maps together and define the following map

$$\widetilde{\alpha} : [0,1] \to X,$$

such that for any $t \in [0,1]$, we have

$$\widetilde{\alpha}(t) = \begin{cases} \widetilde{\beta}_0(t), & t \in [0, t_1], \\ \widetilde{\beta}_1(t), & t \in [t_1, t_2], \\ & \vdots \\ \widetilde{\beta}_n(t), & t \in [t_n, 1]. \end{cases}$$

(See Figure 4.2.1 for an illustration.)

From the continuity of each $\widetilde{\beta}_i$, we may get the continuity of $\widetilde{\alpha}$. We can check the $\widetilde{\alpha}$-preimage of any closed subset of X is closed in $[0,1]$, since it is a union of some closed subsets in each $[t_i, t_{i+1}]$.

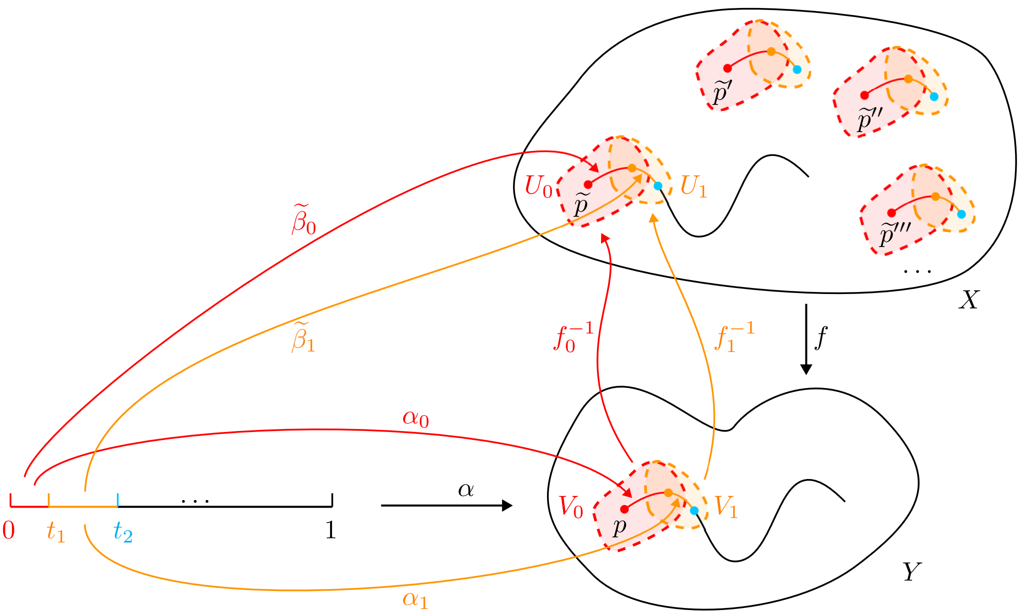


Figure 4.2.1: Construct $\widetilde{\alpha}$ piece by piece.

Next we would like to show the uniqueness of $\widetilde{\alpha}$. Assume that $\widetilde{\alpha}'$ is another lift α, such that

$$\widetilde{\alpha}'(0) = \widetilde{p}.$$

Hence we have

$$f \circ \widetilde{\alpha} = \alpha = f \circ \widetilde{\alpha}'.$$

We consider first the interval $[0, t_1]$. Since

$$\widetilde{\alpha}(0) = \widetilde{p} = \widetilde{\alpha}'(0),$$

we have

$$\widetilde{\alpha}([0, t_1]) \subset U_0 \quad \text{and} \quad \widetilde{\alpha}'([0, t_1]) \subset U_0.$$

Hence

$$\widetilde{\alpha}|_{[0,t_1]} = f_0^{-1} \circ \alpha_0 = \widetilde{\alpha}'|_{[0,t_1]},$$

the two maps coincide on $[0, t_1]$.

Now assume that they coincide on $[0, t_i]$ with $t_i < 1$. Then for $\widetilde{\alpha}(t_i) = \widetilde{\alpha}'(t_i)$. By a similar argument as for $[0, t_1]$, we can show that $\widetilde{\alpha}$ and $\widetilde{\alpha}'$ coincide on $[t_i, t_{i+1}]$. Therefore by induction, we have

$$\widetilde{\alpha} = \widetilde{\alpha}'$$

on $[0, 1]$. This shows the uniqueness of the lift for a chosen lift of $\alpha(0)$.

□

Example 4.2.5.
A lift of a loop may not be a loop. Consider S^1 as the unit circle of $\mathbb{C}$, and the covering map from S^1 to itself given by

$$f : z \mapsto z^2.$$

Consider $p = 1$, then it has two lifts 1 and -1. Consider the loop

$$\begin{aligned} \alpha : [0,1] &\to S^1 \\ t &\mapsto e^{2\pi i t} \end{aligned}$$

It has a lift

$$\begin{aligned} \widetilde{\alpha} : [0,1] &\to S^1 \\ t &\mapsto e^{\pi i t} \end{aligned}$$

with $\widetilde{\alpha}(0) = 1$ and $\widetilde{\alpha}(1) = -1$ which is not a loop.

On the other hand, a covering map always sends loops to loops.

Now we consider the path homotopy.

Proposition 4.2.6

Let α and β be two homotopic paths in Y and H be the homotopy between them. Denote

$$p = \alpha(0) = \beta(0).$$

For any $\widetilde{p} \in f^{-1}(y)$, there is a unique lift

$$\widetilde{H} : [0,1] \times [0,1] \to X,$$

of H with $\widetilde{H}(0,0) = \widetilde{p}$.

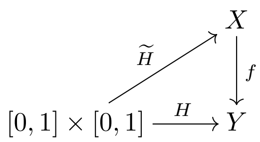
Moreover, the map $\widetilde{H}$ is a path homotopy between $\widetilde{\alpha}$ and $\widetilde{\beta}$ lifts of α and β respectively with

$$\widetilde{\alpha}(0) = \widetilde{\beta}(0) = \widetilde{p}.$$

Proof. The goal is to find a continuous map

$$\widetilde{H} : [0,1] \times [0,1] \to X,$$

such that the following diagram commutes



By the definition of a covering map, each point $H(s,t) \in Y$ has a covering neighborhood $V_{s,t}$ in Y.

Notice that the topology of $[0,1] \times [0,1]$ has a basis consisting of only open rectangles. Hence for any (s,t), there is a rectangle open neighborhood $R_{s,t}$ of (s,t) with

$$R_{s,t} \subset H^{-1}(V_{s,t}).$$

These rectangles $R_{s,t}$'s form an open cover of $[0,1] \times [0,1]$. Since $[0,1] \times [0,1]$ is compact, we can choose finitely many of them to cover $[0,1] \times [0,1]$.

We consider the vertices of these rectangles. Their horizontal coordinates and their vertical coordinates give partitions of $[0,1] \times \{0\}$ and $\{0\} \times [0,1]$ respectively, denoted by

$$0 = s_0 < s_1 < \cdots < s_m < s_{m+1} = 1,$$
$$0 = t_0 < t_1 < \cdots < t_n < t_{n+1} = 1.$$

We consider x_i's and y_i's in $[0,1]$, such that

$$0 = s_0 = x_0 < s_1 < x_1 < \cdots < s_m < x_m < s_{m+1} = x_{m+1} = 1,$$
$$0 = t_0 = y_0 < t_1 < y_1 < \cdots < t_n < y_n < t_{n+1} = y_{n+1} = 1.$$

Then for each $0 \le i \le m$ and $0 \le j \le n$, the image

$$H([x_i, x_{i+1}] \times [y_j, y_{j+1}]) \subset V_{i,j}.$$

where $V_{i,j}$ is the H-image of a rectangle in the finite rectangle cover of $[0,1] \times [0,1]$ obtained above.

We now construct the lift of H piece by piece in an inductive way. Let $\widetilde{p} \in f^{-1}(p)$. We have a neighborhood of $U_{0,0}$ of $\widetilde{p}$ homeomorphic to $V_{0,0}$. Denote by

$$f_{0,0} := f|_{U_{0,0}} : U_{0,0} \to V_{0,0}.$$

Then we define

$$\widetilde{F}_{0,0} : [0, x_1] \times [0, y_1] \to X$$

by $\widetilde{F}_{0,0} = f_{0,0}^{-1} \circ H|_{[0,x_1]\times[0,y_1]}$.

Assume that we have defined $\widetilde{F}_{0,i}$ with $y_{i+1} < 1$, then we consider $H(0, y_{i+1})$ and its covering neighborhood $V_{0,i+1}$ containing $H([0, x_1] \times [y_{i+1}, y_{i+2}])$. Let $U_{0,i+1}$ be the neighborhood of $H(0, y_{i+1})$ homeomorphic to $V_{0,i+1}$. Denote by

$$f_{0,i+1} := f|_{U_{0,i+1}} : U_{0,i+1} \to V_{0,i+1}.$$

Then we define $\widetilde{F}_{0,i+1} = f_{0,i+1}^{-1} \circ H|_{[0,x_1]\times[y_{i+1},y_{i+2}]}$.

For any $0 \le i \le m$, if we have defined $\widetilde{F}_{j,k}$ for any (j,k) with $0 \le j \le i$ and $0 \le k < 1$. Assume that $x_{i+1} < 1$. Then we consider $H(x_{i+1}, 0)$ and its covering neighborhood $V_{i+1,0}$ containing $H([x_{i+1}, x_{i+2}] \times [0, y_1])$. Let $U_{i+1,0}$ be the neighborhood of $H(x_{i+1}, 0)$ homeomorphic to $V_{i+1,0}$. Denote by

$$f_{i+1,0} := f|_{U_{i+1,0}} : U_{i+1,0} \to V_{i+1,0}.$$

Then we define $\widetilde{F}_{i+1,0} = f_{i+1,0}^{-1} \circ H|_{[x_i,x_{i+1}]\times[0,y_1]}$.

For each $(s,t) \in [x_i, x_{i+1}] \times [y_j, y_{j+1}]$, we define

$$\widetilde{H}(s,t) = \widetilde{F}_{i,j}(s,t).$$

By the construction of $\widetilde{F}_{i,j}$, for any two adjacent rectangles, the images of $\widetilde{F}_{i,j}$'s on their common part are the same. Hence $\widetilde{H}$ is well defined. Moreover, the union of closed subsets in each small rectangles is a closed subset in $[0,1] \times [0,1]$, by the continuity of $\widetilde{F}_{i,j}$'s (which are composition between continuous maps), we have the continuity of $\widetilde{H}$. Hence we have a lift of H.

The uniqueness of $\widetilde{H}$ with $\widetilde{H}(0,0) = \widetilde{p}$ can also be proved in an inductive way. Assume that $\widetilde{H}'$ is a lift of H with

$$\widetilde{H}'(0,0) = \widetilde{p} = \widetilde{H}(0,0).$$

Therefore on $[0, x_1] \times [0, y_1]$, we have

$$\widetilde{H}([0, x_1] \times [0, y_1]) \subset U_{0,0} \quad \text{and} \quad \widetilde{H}'([0, x_1] \times [0, y_1]) \subset U_{0,0}.$$

Hence

$$\widetilde{H}'|_{[0,x_1]\times[0,y_1]} = f_{0,0}^{-1} \circ H|_{[0,x_1]\times[0,y_1]} = \widetilde{H}|_{[0,x_1]\times[0,y_1]}.$$

Now assume that $\widetilde{H}'$ and $\widetilde{H}$ coincide on $[x_i, x_{i+1}] \times [y_j, y_{j+1}]$ with $y_{j+1} < 1$, then since

$$\widetilde{H}'(x_i, y_{j+1}) = \widetilde{H}(x_i, y_{j+1}),$$

the two lifts $\widetilde{H}$ and $\widetilde{H}'$ coincide over $[x_i, x_{i+1}] \times [y_{j+1}, y_{j+2}]$.

If $\widetilde{H}'$ and $\widetilde{H}$ coincide on $[x_i, x_{i+1}] \times [y_m, 1]$ with $x_{i+1} < 1$, then since

$$\widetilde{H}'(x_{i+1}, 0) = \widetilde{H}(x_{i+1}, 0),$$

the two lifts $\widetilde{H}$ and $\widetilde{H}'$ coincide over $[x_{i+1}, x_{i+2}] \times [0, y_1]$. By induction, we conclude that we have

$$\widetilde{H} = \widetilde{H}'$$

on $[0, 1] \times [0, 1]$.

We denote

$$\widetilde{\alpha} = \widetilde{H}_0$$

and

$$\widetilde{\beta} = \widetilde{H}_1.$$

Then we have

$$f \circ \widetilde{\alpha} = f \circ \widetilde{H}_0 = H_0 = \alpha,$$

and

$$f \circ \widetilde{\beta} = f \circ \widetilde{H}_1 = H_1 = \beta.$$

Hence $\widetilde{\alpha}$ and $\widetilde{\beta}$ are lifts of α and β respectively with

$$\widetilde{\alpha}(0) = \widetilde{\beta}(0) = \widetilde{H}(0, 0) = \widetilde{p}.$$

Moreover for any $q \in Y$, the preimage set $f^{-1}(q)$ has discrete topology as subspace of X, and each point in $f^{-1}(q)$ is a connected component of $f^{-1}(q)$. Since $\widetilde{H}$ is a continuous map and

$$f \circ \widetilde{H} = H,$$

we have

$$(f \circ \widetilde{H})(\{0\} \times [0, 1]) = H(\{0\} \times [0, 1]) = \{p\},$$

and

$$(f \circ \widetilde{H})(\{1\} \times [0, 1]) = H(\{1\} \times [0, 1]) = \{\alpha(1)\},$$

Hence

$$\widetilde{H}(\{0\} \times [0, 1]) = \widetilde{p},$$

and

$$\widetilde{H}(\{1\} \times [0, 1]) = \widetilde{\alpha}(1).$$

As a conclusion, the map $\widetilde{H}$ is a path homotopy between the paths $\widetilde{\alpha}$ and $\widetilde{\beta}$ in X. □

Remark 4.2.7.

Same as preciously for lifts of paths, the lift $\widetilde{H}$ is determined by $\widetilde{H}(0, 0)$.

Remark 4.2.8.

When we have a continuous map φ from some topological space Z to Y, and X is a cover of Y, we can also ask if we can lift φ to a map from Z to X. The story would be a little bit more complicated, and we will come back to this question later (See Proposition 4.3.16).

4.3 Covers and fundamental groups

In this section, we would like to study the relation between the fundamental group of a path connected space and that of its cover.

Let X and Y be two path connected spaces. Assume that X is a cover of Y with covering map f.

Embedding the fundamental group of X into that of Y

Let $p \in Y$ be a point and $\widetilde{p}$ be one of its lifts in X. A covering map f is in particular continuous, hence there is a homomorphism

$$f_* : \pi_1(X, \widetilde{p}) \to \pi_1(Y, p).$$

Proposition 4.3.1

The homomorphism f_* is injective.

Proof. Since we can always lift a path homotopy map H in Y to one in X, if $\widetilde{\alpha}$ and $\widetilde{\alpha'}$ are two loops in $\mathcal{L}(X, \widetilde{p})$ such that

$$f \circ \widetilde{\alpha} = \alpha \sim \alpha' = f \circ \widetilde{\alpha}'$$

in Y through the homotopy H, then by lifting H to

$$\widetilde{H} : [0,1] \times [0,1] \to X,$$

with $\widetilde{H}(0,0) = \widetilde{\alpha}(0)$. Since $\widetilde{\alpha}'(0) = \widetilde{\alpha}(0) = \widetilde{p}$, by Proposition 4.2.4, we have $\widetilde{H}_0 = \widetilde{\alpha}$ and $\widetilde{H}_1 = \widetilde{\alpha}'$. The map $\widetilde{H}$ is a homotopy between $\widetilde{\alpha}$ and $\widetilde{\alpha}'$.

□

Example 4.3.2.
Consider the circle

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

and the covering map

$$\begin{aligned} f : S^1 &\to S^1 \\ z &\mapsto z^2 \end{aligned}$$

Here we have $X = Y = S^1$. We consider $z = 1$. It has two lifts: 1 and -1. In the fundamental group level, we have

$$\begin{aligned} f_* : \pi_1(S^1, 1) &\to \pi_1(S^1, 1) \\ [\alpha] &\mapsto [\alpha]^2 \end{aligned}$$

where $[\alpha]$ is a generator of $\pi_1(S^1, 1)$.

Example 4.3.3.
Consider $S^1 \vee S^1$. We may identify one of the S^1 with the unit circle in $\mathbb{C}$ and the common point with $1 \in \mathbb{C}$. Then the covering map f_2, f_5 from S^1 to S^1 or more generally the covering map from $\mathbb{R}$ to S^1 given by $t \mapsto e^{2\pi i t}$ can be extended to the following covering maps

Action of the fundamental group of Y on $f^{-1}(p)$

From the algebraic point of view, we can identify $\pi_1(X, \widetilde{p})$ with a subgroup of $\pi_1(Y, p)$. To make this more precise, we consider the other direction and discuss lifts of a loop in Y.

First we consider the following observation. Assume that the cover from X to Y is not of index 1, i.e. not a homeomorphism. Let $\widetilde{p}'$ be a lift of p different from $\widetilde{p}$. Since X is path connected, there is a path $\widetilde{\eta}$ with $\widetilde{\eta}(0) = \widetilde{p}$ and $\widetilde{\eta}(1) = \widetilde{p}'$. When we project this path into Y, we notice that

$$f(\widetilde{\eta}(0)) = f(\widetilde{p}) = p = f(\widetilde{p}') = f(\widetilde{\eta}(1)).$$

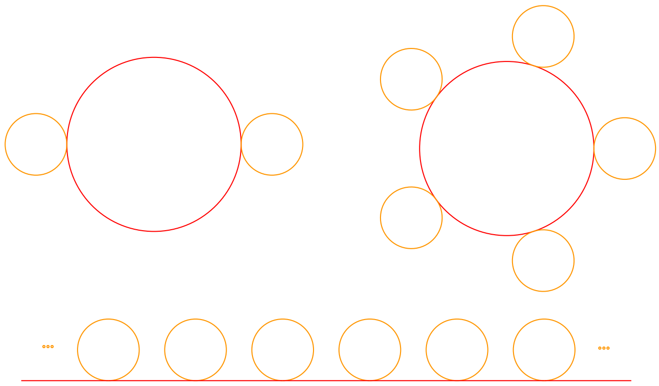


Figure 4.3.1: 3 covering spaces of $S^1 \vee S^1$ (of degree 2, 5 and ∞ respectively).

Hence $\eta = f \circ \widetilde{\eta}$ is a loop in Y. This suggest that the homomorphism f_* may not be surjective, and the lifts of base point p may play an important role in this study. In fact, the above discussion can be use to construct a group action of $\pi_1(Y, p)$ on the set $f^{-1}(p)$.

Let us first recall what a group action is. Let G be a group and A be a set. A *left group action* G on A is map

$$\begin{aligned} \Phi : G \times A &\to A, \\ (g, a) &\mapsto g.a \end{aligned}$$

satisfying the following properties:

1. for any $a \in A$, for any $g, g' \in G$, we have $g.(g'.a) = (gg').a$;
2. for any $a \in A$, we have $e.a = a$ where $e \in G$ is the identity.

if the first condition is replaced by the following one

3. for any $a \in A$, for any $g, g' \in G$, we have $g.(g'.a) = (g'g).a$.

we call it a *right group action* G on A.

We may consider walking along a path

$$\eta : [0, 1] \to X,$$

as pushing the point $\eta(0)$ along $\operatorname{Im} \eta$ until $\eta(1)$. Then given any lift $\widetilde{p} \in f^{-1}(p)$, the lifts of loops in $\mathcal{L}(Y, p)$ to $\mathcal{L}(X, \widetilde{p})$ can be considered as different ways to move $\widetilde{p}$ to some lift $\widetilde{p}' \in f^{-1}(q)$ in X along paths.

More precisely, we choose a lift $\widetilde{p} \in f^{-1}(p)$ of p. Let α be a loop in Y based at p. Let $\widetilde{\alpha}$ be the lift of α with $\widetilde{\alpha}(0) = \widetilde{p}$. We denote

$$\widetilde{\alpha}(1) = \widetilde{p}',$$

which is also a lift of p.

Let α' be another loop in Y based at p and homotopic to α. Let $\widetilde{\alpha'}$ be its lift in X with $\widetilde{\alpha'}(0) = \widetilde{p}$, then since any homotopy in Y can be lifted to a homotopy in X, we have

$$\widetilde{\alpha'}(1) = \widetilde{\alpha}(1) = \widetilde{p}'.$$

Hence the following map is well defined

$$\begin{aligned} \Phi : \pi_1(Y,p) \times f^{-1}(p) &\to f^{-1}(p) \\ ([\alpha], \widetilde{p}) &\mapsto \widetilde{\alpha}(1) \end{aligned}$$

where $\widetilde{\alpha}$ is a lift of a representative α in $[\alpha]$ with $\widetilde{\alpha}(0) = \widetilde{p}$.

Proposition 4.3.4

The map Φ induces a right group action of $\pi_1(Y,p)$ on the set $f^{-1}(p)$.

Proof. Let α and β be two loops in Y based at p. Their composition $\alpha * \beta$ is also a loop based at p. For any $\widetilde{p} \in f^{-1}(p)$, consider a lift $\widetilde{\alpha * \beta}$ of $\alpha * \beta$ with $\widetilde{\alpha * \beta}(0) = \widetilde{p}$. We have

$$f \circ (\widetilde{\alpha * \beta}) = \alpha * \beta.$$

Hence for any $t \in [0,1]$, we define paths in X:

$$\widetilde{\alpha}(t) = \widetilde{\alpha * \beta}\left(\frac{t}{2}\right)$$

and

$$\widetilde{\beta}(t) = \widetilde{\alpha * \beta}\left(\frac{t+1}{2}\right)$$

Then for any $t \in [0,1]$, we have

$$(f \circ \widetilde{\alpha})(t) = \left(f \circ (\widetilde{\alpha * \beta})\right)\left(\frac{t}{2}\right) = (\alpha * \beta)\left(\frac{t}{2}\right) = \alpha(t),$$

and

$$\left(f \circ \widetilde{\beta}\right)(t) = \left(f \circ (\widetilde{\alpha * \beta})\right)\left(\frac{t+1}{2}\right) = (\alpha * \beta)\left(\frac{t+1}{2}\right) = \beta(t).$$

From this we can conclude that $\widetilde{\alpha}$ is a lift of α with $\widetilde{\alpha}(0) = \widetilde{p}$ and $\widetilde{\beta}$ is a lift of β with

$$\widetilde{\beta}(0) = \widetilde{\alpha}(1).$$

and moreover

$$\widetilde{\alpha * \beta} \sim \widetilde{\alpha} * \widetilde{\beta},$$

Hence by the definition of Φ, we have

$$[\alpha * \beta].\widetilde{p} = [\beta].([\alpha].\widetilde{p})$$

(See Figure 4.3.2 for an illustration.) □

As we have seen previously, since X is path connected by assumption, for any $\widetilde{p}$ and $\widetilde{p}'$ two lifts of p, there is a path $\widetilde{\eta}$ connecting them whose composition with f as a path in Y is a loop η in Y based at p. Hence we have

$$[\eta].\widetilde{p} = \widetilde{p}'.$$

Hence we have the following observation.

Proposition 4.3.5

The action of $\pi_1(Y,p)$ on $f^{-1}(p)$ is transitive.

Now we consider the stabilizer of a lift $\widetilde{p}$ in $f^{-1}(p)$, and have the following proposition.

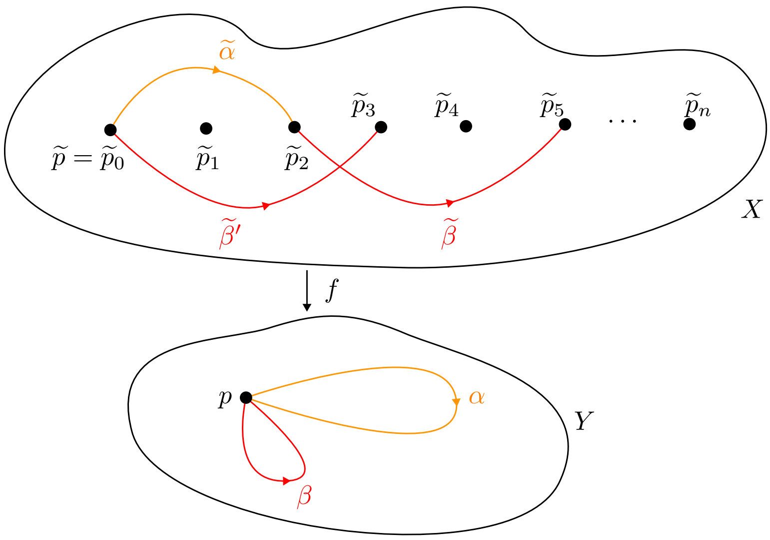


Figure 4.3.2: $[\alpha * \beta].\widetilde{p} = [\beta].([\alpha].\widetilde{p})$.

Proposition 4.3.6

For any $\widetilde{p} \in f^{-1}(p)$, the stabilizer of $\widetilde{p}$ satisfies

$$\mathrm{Stab}(\widetilde{p}) = f_*(\pi_1(X, \widetilde{p})).$$

Proof. For any $[\alpha] \in f_*(\pi_1(X, \widetilde{p}))$, let α be a representative of $[\alpha]$. Let $\widetilde{\alpha}$ be the unique lift of α with $\widetilde{\alpha}(0) = \widetilde{p}$.

Since $[\alpha] \in f_*(\pi_1(X, \widetilde{p}))$, there is an element $[\widetilde{\beta}] \in \pi_1(X, \widetilde{p})$, such that

$$[f \circ \widetilde{\beta}] = f_*([\widetilde{\beta}]) = [\alpha].$$

Hence $\beta = f \circ \widetilde{\beta}$ is homotopic to α and β is a representative of $[\alpha]$. Therefore, we have

$$[\alpha].\widetilde{p} = \widetilde{\beta}(1) = \widetilde{p}.$$

This shows the inclusion of one direction

$$\mathrm{Stab}(\widetilde{p}) \supset f_*(\pi_1(X, \widetilde{p})).$$

Let α be a loop in Y based at p such that

$$[\alpha] \in \mathrm{Stab}(\widetilde{p}).$$

Let $\widetilde{\alpha}$ be the lift of α in X with $\widetilde{\alpha}(0) = \widetilde{p}$. Since $[\alpha]$ is in the stabilizer of $\widetilde{p}$, we have

$$\widetilde{\alpha}(1) = \widetilde{p}.$$

Therefore $\widetilde{\alpha}$ is a loop in X based at $\widetilde{p}$, and we have

$$f_*([\widetilde{\alpha}]) = [f \circ \widetilde{\alpha}] = [\alpha].$$

Hence we have

$$[\alpha] \in f_*(\pi_1(X, \widetilde{p})).$$

□

An immediate corollary is as follows, which is the relation between the orbit and the stabilizer when studying a group action.

Corollary 4.3.7

Let $\widetilde{p} \in f^{-1}(p)$ be a lift of $p \in Y$. There is a bijection

$$\psi : f_*(\pi_1(X, \widetilde{p}))\backslash\pi_1(Y, p) \to f^{-1}(p)$$
$$\overline{[\alpha]} \mapsto [\alpha].\widetilde{p}$$

where

$$\overline{[\alpha]} := f_*(\pi_1(X, \widetilde{p}))[\alpha]$$

is a right coset of $f_*(\pi_1(X, \widetilde{p}))$ in $\pi_1(Y, p)$.

Notice that $|f^{-1}(y)|$ is also the index of the cover.

Corollary 4.3.8

Let $\widetilde{p} \in f^{-1}(p)$ be a lift of $p \in Y$. We have the following identity.

$$[\pi_1(Y, p) : f_*(\pi_1(X, \widetilde{p}))] = \deg f.$$

The map f is homeomorphism if and only if $|f^{-1}(y)| = 1$. Hence using Proposition 4.1.14, we have the following corollary.

Corollary 4.3.9

The covering map

$$f : X \to Y,$$

is an homeomorphism if and only if

$$f_* : \pi_1(X, x) \to \pi_1(Y, y),$$

is an isomorphism.

Remark 4.3.10.
If a covering map is not homeomorphism, we cannot say that $\pi_1(X, x)$ and $\pi_1(Y, y)$ are not isomorphic. A group can be isomorphic to its proper subgroup. For example, when we consider the 2-cover of S^1 to itself given by the map $f : z \mapsto z^2$ on the complex plane, it is not isomorphism, however, since both X and Y are S^1, hence they have the same fundamental group, which of course are isomorphic. Here what we have is the following isomorphism: $\mathbb{Z} \cong 2\mathbb{Z}$. Therefore what the second condition really needs is the surjectivity of f_* by considering Proposition 4.3.1

Now we consider a special case when Y is simply connected. In this case, we have $\pi_1(Y, p)$ is trivial. Let X be any cover of Y with covering map f, and let $\widetilde{p} \in f^{-1}(p)$ be a lift of p. The map

$$f_* : \pi_1(X, \widetilde{p}) \to \pi_1(Y, p),$$

is an isomorphism. With the above corollary, we have the following conclusion.

Corollary 4.3.11

Any covering map over a simply connected space is a homeomorphism.

Remark 4.3.12.
This means that there is no non-trivial cover of a simply connected space.

Covers of a space

Given a space Y, there are several questions that one could ask regarding the covers of Y.

1) How many different covers of Y are there?

2) Is it possible to have a degree n cover of Y for any $n \in \mathbb{N}^*$?

3) Is there any relation between different covers of Y?

4) Is there a biggest cover of Y?

We know that a covering map from a space X to Y induces an embedding of fundamental group of X to that of Y.

1) Can we read the information of covering map from the information of the fundamental groups?

2) Is it possible to have a cover for any subgroup of the fundamental group of Y?

In order to be able to compare different covers of a same space, we first introduce the notation of morphism between covers.

Definition 4.3.13

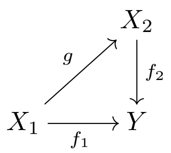
Let X_1 and X_2 be two covering of Y with covering maps:

$$f_1 : X_1 \to Y \quad \text{and} \quad f_2 : X_2 \to Y.$$

A continuous map

$$g : X_1 \to X_2$$

is said to be a ***morphism*** between the two covers X_1 and X_2, if we have the following commutative diagram



If moreover the morphism g is a homeomorphism, we call it an ***isomorphism*** between covers X_1 and X_2.

From the commutative diagram, it seems that the cover X_1 is bigger than X_2 in some sense. Let us make this clear in the following way.

Proposition 4.3.14

If Y is locally connected, then a surjective morphism g between two covers of Y, if exists, is a covering map.

Proof. We use the same notation. Since Y is locally connected, for any $p \in Y$, it has an open connected covering neighborhood V_1 for the covering map f_1 and am open connected covering neighborhood of V_2 for the covering map f_2.

Let $V \subset V_1 \cap V_2$ be an open connected neighborhood of p. Then V is a connected covering neighborhood of p for both f_1 and f_2. From the definition of a covering neighborhood, we have

$$f_1^{-1}(V) = \bigsqcup_{i \in \Omega} U_i,$$

such that the restriction of f_1 to each U_i is a homeomorphism to V, and

$$f_2^{-1}(V) = \bigsqcup_{j \in \Theta} W_j,$$

such that the restriction of f_2 to each W_j is a homeomorphism to V. Notice that all U_i's and W_j's are connected. Moreover for each $i \in \Omega$, U_i is a connected component of $f_1^{-1}(V)$ and for each $j \in \Theta$, W_j is a connected component of $f_2^{-1}(V)$.

Let $\widetilde{p}$ be a lift of p in X_1, then there is a unique neighborhood U_i such that

$$\widetilde{p} \in U_i.$$

We consider $g(\widetilde{p}) \in X_2$, since $f_1 = f_2 \circ g$, we have

$$g(\widetilde{p}) \in f_2^{-1}(p),$$

hence a lift of p in X_2. There is a unique W_j, such that

$$g(\widetilde{p}) \in W_j.$$

Now we would like to show that the restriction of g to U_i is a homeomorphism to W_j. Notice that

$$f_1|_{U_i} = f_2 \circ g|_{U_i}.$$

Since $f_1(U_i) = V$, we have

$$g(U_i) \subset f_2^{-1}(V).$$

Since U_i is connected, so is $g(U_i)$, hence we have

$$g(U_i) \subset W_j.$$

Therefore, we have

$$f_1|_{U_i} = f_2|_{W_j} \circ g|_{U_i}$$

which gives

$$g|_{U_i} = (f_2|_{W_j})^{-1} \circ f_1|_{U_i},$$

hence a homeomorphism. This shows that W_j is a covering neighborhood of $g(\widetilde{p})$.

Since g is surjective, it is a covering map. □

Remark 4.3.15.

When we consider manifolds, the locally path connectedness and the locally connectedness are satisfied naturally, since locally a manifold is the same as the Euclidean space.

Another observation is the following one. If we consider the definition of a lift of a map, then a morphism g between

$$f_1 : X_1 \to Y \quad \text{and} \quad f_2 : X_2 \to Y,$$

is a lift of f_1 with respect to the cover X_2 of Y. Therefore, to see which pair of covers have a morphism between them is equivalent to ask when a covering map can be lifted with respect to the other covering map. Here is a general statement for topological spaces which are path connected an locally path connected. This can be considered as a generalization of Proposition 4.2.4 and Proposition 4.2.6

Proposition 4.3.16

Let X and Y be two topological space, and

$$f : X \to Y$$

be a covering map. Let Z be a topological space which is path connected and locally path connected. Let

$$g : Z \to Y$$

be a continuous map, then g has a lift $\widetilde{g}$ if and only if in $\pi_1(X, p)$, we have

$$g_*(\pi_1(Z, u)) \subset f_*(\pi_1(X, \widetilde{p}))$$

where $p \in Y$, $\widetilde{p} \in X$ and $u \in Z$ such that $g(u) = f(\widetilde{p}) = p$.

Proof. One direction is clear. If there exists a lift

$$\widetilde{g} : Z \to X$$

of g, then we have $g = f \circ \widetilde{g}$.

Let $p \in Y$, $\widetilde{p} \in X$ and $u \in Z$ with $\widetilde{g}(u) = \widetilde{p}$, and $f(\widetilde{p}) = p$. We have the following commutative diagram

$$\begin{array}{ccc} & & \pi_1(X, \widetilde{p}) \\ & \overset{\widetilde{g}_*}{\nearrow} & \downarrow f_* \\ \pi_1(Z, u) & \xrightarrow{g_*} & \pi_1(Y, p) \end{array}$$

The relation $g_* = f_* \circ \widetilde{g}_*$ yields

$$g_*(\pi_1(Z, u)) \subset f_*(\pi_1(X, \widetilde{p})).$$

Conversely, let $p \in Y$, $\widetilde{p} \in X$ and $u \in Z$ with $g(u) = f(\widetilde{p}) = p$ and assume that

$$g_*(\pi_1(Z, u)) \subset f_*(\pi_1(X, \widetilde{p})).$$

Now we would like to construct a lift $\widetilde{g}$ of g. Using covering neighborhoods of points in Y, we can locally lift g. The problem left is whether all these lifts can be chosen so that they can be glued to one continuous map from Z to X.

The precise construction is as follows. Since $g(u) = f(\widetilde{p})$, we set

$$\widetilde{g}(u) = \widetilde{p}.$$

Now for any point $v \in Z$, choose a path η in Z with $\eta(0) = u$ and $\eta(1) = v$. Then $g \circ \eta$ is a path in Y. We lift it to a path

$$\widetilde{g \circ \eta} : [0, 1] \to X,$$

with $\widetilde{g \circ \eta}(0) = \widetilde{p}$. Then we define
$$\widetilde{g}(v) = \widetilde{g \circ \eta}(1).$$

Notice that this indeed gives a map
$$\begin{aligned} \widetilde{g} : Z &\to X \\ v &\mapsto \widetilde{g \circ \eta}(1) \end{aligned}.$$
satisfies the condition $g = f \circ \widetilde{g}$.

We first show that for each $v \in Z$, the image $\widetilde{g}(v)$ is independent of choice of η.

Let v be a point in X. Let η and η' be two path with
$$\eta(0) = \eta'(0) = u \quad \text{and} \quad \eta(1) = \eta'(1) = v.$$
We consider the lifts $\widetilde{g \circ \eta}$ and $\widetilde{g \circ \eta'}$ of $g \circ \eta$ and $g \circ \eta'$ respectively with
$$\widetilde{g \circ \eta}(0) = \widetilde{g \circ \eta'}(0) = \widetilde{p}.$$
We would like to show that
$$\widetilde{g \circ \eta}(1) = \widetilde{g \circ \eta'}(1).$$
Notice that $\eta * \overline{\eta'}$ is a loop in Z based at u, hence
$$(g \circ \eta) * (g \circ \overline{\eta'}) = g \circ (\eta * \overline{\eta'})$$
is also a loop in Y based at p.

We consider the lift of $g \circ (\eta * \overline{\eta'})$ which is a path in X. Since
$$[g \circ (\eta * \overline{\eta'})] \in g_*(\pi_1(Z, u)),$$
by the hypothesis, we have
$$[g \circ (\eta * \overline{\eta'})] \in f_*(\pi_1(X, \widetilde{p})).$$
Hence there is a loop $\widetilde{\alpha}$ in X based at $\widetilde{p}$ with
$$f_*([\widetilde{\alpha}]) = [g \circ (\eta * \overline{\eta'})].$$
Denote $\alpha = f \circ \widetilde{\alpha}$, we have
$$\alpha \sim g \circ (\eta * \overline{\eta'}).$$
let H be a path homotopy between α and $g \circ (\eta * \overline{\eta'})$. Then we can lift it to a path homotopy $\widetilde{H}$ between $\widetilde{\alpha}$ and $\widetilde{g \circ (\eta * \overline{\eta'})}$. Hence we have
$$\widetilde{\alpha} \sim \widetilde{g \circ \eta} * \widetilde{g \circ \overline{\eta'}},$$
where
$$\widetilde{g \circ \eta}(0) = \widetilde{p}$$
and
$$\widetilde{g \circ \overline{\eta'}}(0) = \widetilde{g \circ \eta}(1).$$
Since
$$\widetilde{\alpha}(1) = \widetilde{\alpha}(0),$$
we have
$$\widetilde{g \circ \overline{\eta'}}(1) = \widetilde{g \circ \eta}(0).$$
Hence the lift $\widetilde{g \circ \eta'}$ of $g \circ \eta'$ with
$$\widetilde{g \circ \eta'}(0) = \widetilde{p}$$

satisfies the condition

$$\widetilde{g\circ\eta'}(1)=\widetilde{g\circ\eta}(1).$$

With these discussion, we would like to show next that the map $\widetilde{g}$ is continuous. Let $\widetilde{V}$ be an open subset of X. We consider its preimage $\widetilde{g}^{-1}(\widetilde{V})$. For any $v\in\widetilde{g}^{-1}(\widetilde{V})$, we denote

$$q=g(v).$$

Consider a covering neighborhood V_1 of q such that there is a neighborhood $\widetilde{V_1}$ of $\widetilde{g}(v)=\widetilde{q}$ contained in $\widetilde{V}$ homeomorphic to V_1 through f.

Since g is continuous, we can take W a path connected neighborhood of v in Z, such that

$$g(W)\subset V_1.$$

Let η be a path in Z, such that $\eta(0)=u$ and $\eta(1)=v$. For any $w\in W$, there is a path ζ with $\zeta(0)=v$ and $\zeta(1)=w$. Hence we have $\eta*\zeta$ is a path in Z from u to w.

The map g sends all above paths to paths in Y. In particular, we have $g\circ\eta$ a path from $p=g(u)$ to $q=g(v)$. Then the path

$$g\circ(\eta*\zeta)=(g\circ\eta)*(g\circ\zeta)$$

is from q to $r=g(w)$.

We consider the lift of the above paths. Then $\widetilde{g\circ\eta}$ is a path from $\widetilde{p}$ to $\widetilde{g}(v)=\widetilde{q}$ by the definition of $\widetilde{g}$. Let $\widetilde{g*\zeta}$ be a lift of $g\circ\zeta$ with

$$\widetilde{g*\zeta}(0)=\widetilde{q}=\widetilde{g\circ\eta}.$$

Hence

$$\widetilde{g\circ(\eta*\zeta)}=\widetilde{g\circ\eta}*\widetilde{g\circ\zeta}.$$

Notice that

$$(g\circ\eta)([0,1])\subset V_1,$$

and

$$f_1:=f|_{V_1}:V_1\to\widetilde{V_1}$$

is a homeomorphism, we have

$$\widetilde{g\circ\zeta}=f_1^{-1}\circ(g\circ\zeta).$$

Therefore, we have

$$\widetilde{g}(w)=\widetilde{g\circ(\eta*\zeta)}(1)=\widetilde{g\circ\zeta}(1)=(f_1^{-1}\circ(g\circ\zeta))(1)\in\widetilde{V_1}\subset V.$$

Hence

$$W\subset\widetilde{g}^{-1}(V),$$

and $\widetilde{g}^{-1}(V)$ is a neighborhood of v. Since v can be chosen arbitrarily, we have $\widetilde{g}^{-1}(V)$ open. Therefore $\widetilde{g}$ is continuous. □

Regarding the uniqueness of a lift of a continuous map, we have the following statement.

Proposition 4.3.17

If X and Y are two topological space with

$$f:X\to Y,$$

a covering map. Assume that Z is a topological space which is path connected and locally path connected and

$$g:Z\to Y$$

is a continuous map.

If $\widetilde{g}_1$ and $\widetilde{g}_2$ are two lifts of g with $\widetilde{g}_1(u)=\widetilde{g}_2(u)$ for some $u\in Z$, then $\widetilde{g}_1=\widetilde{g}_2$.

Proof. Notice that the construction of lifts of g based on lifting paths in Y to paths in X. Hence the uniqueness result of lifts of g is a consequence of the uniqueness result of lifts of paths in Y. □

Remark 4.3.18.
The proofs of the above two proposition give a different way to describe the image of Z using points and paths in Y. This of course works for $Z = X$, when we assume that X is path connected and locally path connected (so is Y as an immediate consequence of the fact that the covering map is in particular continuous).

In the next part, this will be the key idea in the construction of covering spaces, in particular the universal cover, which shows that they do exist.

Universal cover

We consider topological spaces which are path connected and locally path connected in this section. We will show that for any such space, there is a largest cover which is unique up to morphism between covers.

For technical reason, other than path connected and locally path connected, we also assume that the space will be studied is semilocally simply connected.

Definition 4.3.19

Let X be a topological space which is path connected and locally path connected. We say that X is ***semilocally simply connected*** if for any $p \in X$, there is a path connected neighborhood U of p, such that the homomorphism between the fundamental groups

$$\pi_1(U, p) \to \pi_1(X, p)$$

induced by the inclusion of U in X is trivial (all elements are sent to the identity of $\pi_1(X, p)$).

Example 4.3.20.
The unit disk D in $\mathbb{C}$ is a semilocally simply connected. Notice that D is contractible, and any point in D is a strong deformation retraction of D. Hence given any neighborhood U of any $p \in D$, any loop in U based at p is homotopic to the constant loop based at p through a homotopy in D.

For example, let U be the annulus defined by

$$U = \{z \in \mathbb{C} \mid 0.2 < |z| < 0.5\}.$$

Although $\pi_1(U) \cong \mathbb{Z}$, and a loop in U may not be homotopically trivial in U, it is homotopically trivial in D.

In the rest part of this section, let X denote a topological space which is path connected, locally path connected and semilocally simply connected.

Let $p \in X$ be a point. We consider the following abstract set

$$\widetilde{X} := \{[\gamma] \mid \gamma \text{ is a path in } X \text{ with } \gamma(0) = p\}.$$

Remark 4.3.21.
Here the homotopy is path homotopy. Inspired by the proof of Proposition 4.3.16, by using $[\gamma]$, we get not only a point $\gamma(1)$ in X, but also the information about how we go to $\gamma(1)$ from p.

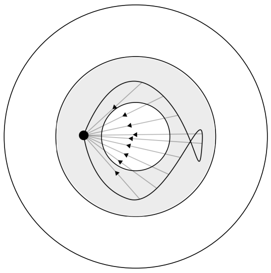


Figure 4.3.3: A loop in U can be deformed to a constant path in D through the linear homotopy in D.

We can check the following example to get a more clear idea what this construction is about. We consider the covering map

$$f : \mathbb{R} \times [0,1] \to S^1 \times [0,1],$$
$$(s,t) \mapsto (e^{2\pi i s}, t).$$

As shown in Figure 4.3.4, let p be a base point in $S^1 \times [0,1]$ and q be another point of it. Let γ

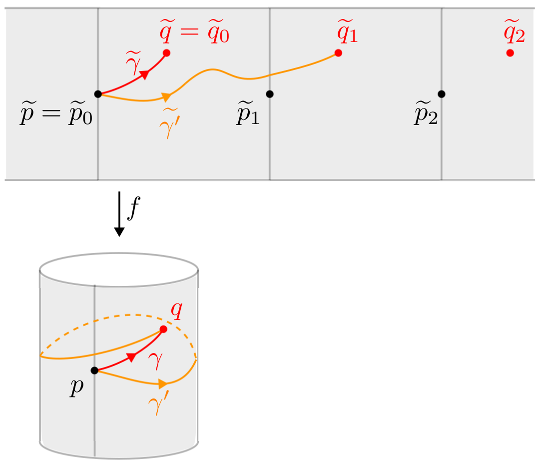


Figure 4.3.4: A pair of non-homotopic paths connecting p to q give different lifts of q.

and γ' be two paths going from p to q, such that $[\gamma' * \overline{\gamma}]$ is a generator of $\pi_1(S^1 \times [0,1], p)$. Then by considering their lifts in $\mathbb{R} \times [0,1]$ starting at a same lift $\widetilde{p}$ of p, they end at different lifts of q. Or alternatively, we may consider $\widetilde{q}_0$ and $\widetilde{q}_1$ associated to $[\gamma]$ and $[\gamma']$ respectively.

Now we would like to equip it with a topology. We first consider the topology $\mathcal{T}$ on X.

Proposition 4.3.22

The topology $\mathcal{T}$ has a basis $\mathcal{B}$ where each $U \in \mathcal{B}$ is path connected and the homomorphism of fundamental groups based at $p \in U$

$$\pi_1(U, p) \to \pi_1(X, p)$$

induced by the inclusion of U in X is trivial.

Proof. Since X is locally path connected, each point $p \in X$ has a open neighborhood basis $\mathcal{C}_p$ where every open set in $\mathcal{C}_p$ is path connected.

On the other hand, since X is semilocally simply connected, there is a path connected neighborhood U_p such that the homomorphism

$$\pi_1(U_p, p) \to \pi_1(X, p)$$

induced by the inclusion of U_p in X is trivial.

For any open neighborhood W of p, we consider $W \cap U$ which is again a neighborhood of p. There is an open set $V \in \mathcal{C}_p$ such that

$$V \subset U_p \cap W.$$

Consider the inclusions

$$V \to U_p \to X,$$

we have

$$\pi_1(V, p) \to \pi_1(U_p, p) \to \pi_1(X, p)$$

a trivial homomorphism.

Hence p has a neighborhood basis $\mathcal{B}_p$ where each $V \in \mathcal{B}_p$ is open and path connected, and the homomorphism

$$\pi_1(V, p) \to \pi_1(X, p)$$

induced by the inclusion of V to X is trivial.

Hence $\mathcal{T}$ has a basis

$$\mathcal{B} := \bigcup_{p \in X} \mathcal{B}_p$$

with the desired property. □

In the following, we will use the basis $\mathcal{B}$ of $\mathcal{T}$ constructed in the proof of Proposition 4.3.22 to construct a topology on $\widetilde{X}$.

For any $p \in X$, let $\mathcal{B}_q$ be a neighborhood basis constructed in the proof of Proposition 4.3.22. For any $V \in \mathcal{B}_q$, and for any $[\gamma] \in \widetilde{X}$ with $\gamma(1) = q$, we define

$$U([\gamma], V) := \{[\gamma * \eta] \mid \eta \text{ is a path in } V \text{ with } \eta(0) = \gamma(1)\}.$$

Notice that $[\gamma * \eta]$ is only depends on $\eta(1)$. If η' is another path in V such that $\eta'(0) = q$ and $\eta(1) = \eta'(1)$, then $\eta * \eta'$ is a loop in V based at q. Since

$$\pi_1(V, q) \to \pi_1(X, q)$$

is trivial, we have

$$\eta * \overline{\eta'} \sim c_q,$$

in X. Hence we have

$$\eta \sim c_q * \eta \sim \eta' * (\overline{\eta} * \eta) \sim \eta' * c_{\eta(1)} \sim \eta'.$$

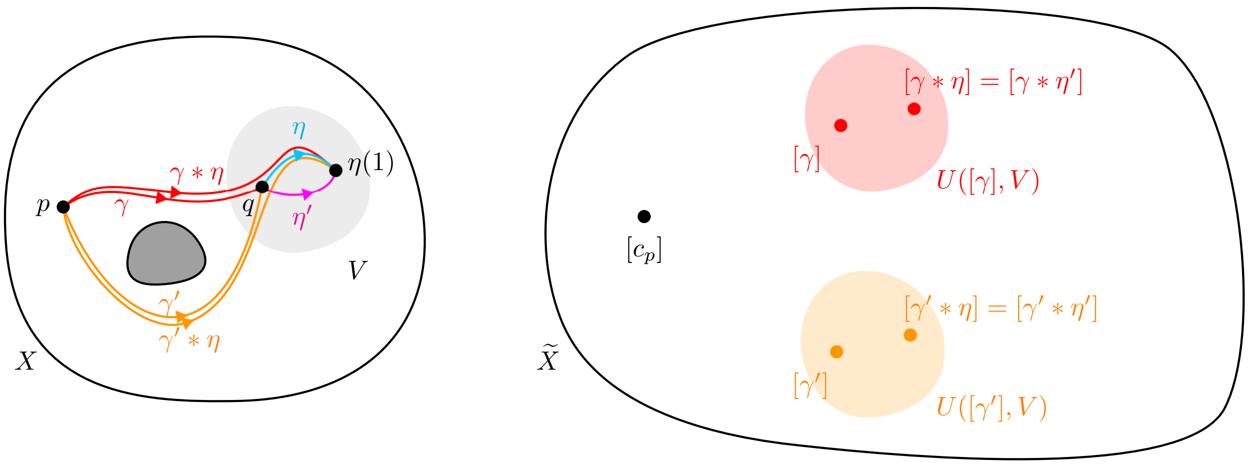


Figure 4.3.5: A pair of non-homotopic paths connecting p to q give different lifts of q.

from which we obtain

$$[\gamma * \eta] = [\gamma * \eta'].$$

Let $\widetilde{\mathcal{T}}$ be the topology generated by

$$\widetilde{\mathcal{B}} := \{U([\gamma], V) \mid [\gamma] \in \widetilde{X},\ V \in \mathcal{B}_{\gamma(1)}\}.$$

Proposition 4.3.23

The subbasis $\widetilde{\mathcal{B}}$ is a basis of $\widetilde{\mathcal{T}}$.

Proof. Let $U([\gamma], V)$ and $U([\gamma'], V')$ be two elements of $\widetilde{\mathcal{B}}$, such that

$$U([\gamma], V) \cap U([\gamma'], V') \neq \emptyset.$$

Let $[\alpha] \in U([\gamma], V) \cap U([\gamma'], V')$. By definition, there are η and η' paths in V and V' respectively with $\eta(0) = \gamma(1)$ and $\eta'(0) = \gamma'(1)$, such that

$$\alpha \sim \gamma * \eta \sim \gamma' * \eta'.$$

Since V and V' are open in X, the intersection $V \cap V'$ is also open in X with

$$\alpha(1) \in V \cap V'.$$

Hence there is an open path connected set W in $\mathcal{B}_{\alpha(1)}$, such that

$$W \subset V \cap V'.$$

Now we would like to show that

$$U([\alpha], W) \subset U([\gamma], V) \cap U([\gamma'], V').$$

For each $[\beta] \in U([\alpha], W)$, there is a path ζ in W with $\zeta(0) = \alpha(1)$, such that

$$\beta \sim \alpha * \zeta.$$

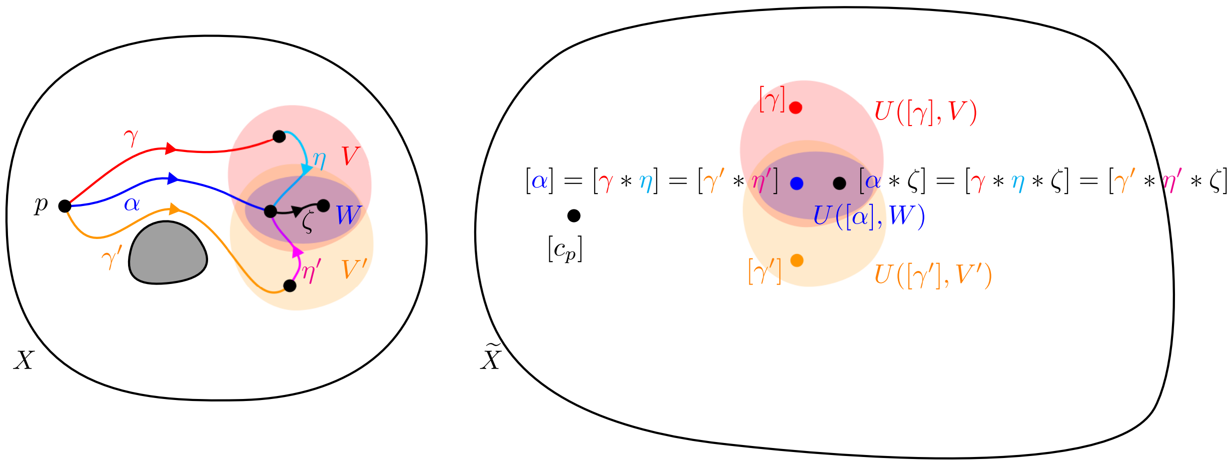


Figure 4.3.6: A pair of non-homotopic paths connecting p to q give different lifts of q.

Hence we have
$$\beta \sim \alpha * \zeta \sim \gamma * (\eta * \zeta) \sim \gamma' * (\eta' * \zeta).$$
Notice that $\eta * \zeta$ is a path in V with
$$(\eta * \zeta)(0) = \eta(0) = \gamma(1),$$
hence we have
$$[\beta] = [\gamma * (\eta * \zeta)] \in U([\gamma], V).$$
Similarly, we have
$$[\beta] = [\gamma' * (\eta' * \zeta)] \in U([\gamma'], V').$$
Hence
$$[\beta] \in U([\gamma], V) \cap U([\gamma'], V').$$
This shows that
$$U([\alpha], W) \subset U([\gamma], V) \cap U([\gamma'], V').$$
We conclude that $\widetilde{\mathcal{B}}$ is a basis of $\widetilde{\mathcal{T}}$. □

Theorem 4.3.24

The topological space $(\widetilde{X}, \widetilde{\mathcal{T}})$ is a cover of X with trivial fundamental group.

Proof. Notice that there is a natural way to define a map from $\widetilde{X}$ to X:
$$\begin{aligned} f : \widetilde{X} &\to X \\ [\gamma] &\mapsto \gamma(1) \end{aligned}.$$
The surjectivity is given by the fact that X is path connected.

We now try to show that it is a covering map. Consider the restriction f_1 of f to each $U([\gamma], V)$. We would like to show that the following two facts

(i) The map f_1 is a homeomorphism.

(ii) For any γ' with $\gamma(1) = \gamma'(1)$, if $U([\gamma], V) \cap U([\gamma'], V) \neq \emptyset$, then we have $[\gamma] = [\gamma']$.

With these two facts, for any q, we consider the V in $\mathcal{B}_q$, then we have

$$f^{-1}(V) = \bigsqcup_{[\gamma] \text{ with } \gamma(1)=q} U([\gamma], V).$$

Hence f is a covering map.

We first show that f_1 to $U([\gamma], V)$ is a homeomorphism. By the definition of f_1, it is surjective on $U([\gamma], V)$. Now let $[\alpha]$ and $[\beta]$ be two points in $U([\gamma], V)$. Hence there are η and ζ paths in V such that

$$\alpha \sim \gamma * \eta \quad \text{and} \quad \beta \sim \gamma * \zeta.$$

Assume that

$$\alpha(1) = f_1([\alpha]) = f_1([\beta]) = \beta(1).$$

Then we have that $\eta(1) = \zeta(1)$, hence $\eta * \overline{\zeta}$ is a loop based at $\gamma(1)$. Since the homomorphism

$$\pi_1(V, \gamma(1)) \to \pi_1(X, \gamma(1))$$

is trivial, we have

$$\eta \sim \zeta,$$

hence

$$[\alpha] = [\gamma * \eta] = [\gamma * \zeta] = [\beta].$$

This shows that f_1 on $U([\gamma], V)$ is bijective.

Now we show that f_1 is continuous. For any $W \subset V$ open set, for any $[\alpha] \in f_1^{-1}(W)$, hence there is a path η in V with $\eta(0) = \gamma(1)$, such that

$$\alpha \sim \gamma * \eta.$$

Consider a neighborhood $W' \in \mathcal{B}_{\alpha(1)}$ with

$$W' \subset W.$$

Given any path ζ in W' with $\zeta(0) = \alpha(1)$, we have

$$\alpha * \zeta \sim \gamma * (\eta * \zeta).$$

Hence we have

$$U([\alpha], W') \subset U([\gamma], V).$$

Since f_1 is bijective on $U([\gamma], V)$, we have

$$U([\alpha], W') = f^{-1}(W') \subset f^{-1}(W).$$

Hence $f^{-1}(W)$ is a neighborhood of $[\alpha]$ for any $[\alpha] \in f^{-1}(W)$ and $f^{-1}(W)$ is open. Hence f is continuous.

On the other hand, $U([\gamma], V)$ has a basis

$$\widetilde{\mathcal{B}}_V = \{U([\alpha], W) \mid [\alpha] \in U([\gamma], V), W \subset V, W \in \mathcal{B}_{\alpha(1)}\}.$$

By the definition of f_1, for any $U([\alpha], W) \in \widetilde{\mathcal{B}}_V$, we have

$$f(U([\alpha], W)) = W$$

which is open in V. Hence f^{-1} is continuous. Therefore f_1 is a homeomorphism.

Now we show the second fact. Let γ' be a path in X with $\gamma'(0) = p$ and $\gamma'(1) = \gamma(1)$. Assume that

$$U([\gamma], V) \cap U([\gamma'], V) \neq \emptyset,$$

then there is

$$[\alpha] \in U([\gamma], V) \cap U([\gamma'], V).$$

Hence there are path η and η' in V such that

$$\alpha \sim \gamma * \eta \sim \gamma' * \eta'.$$

Notice that $\eta(1) = \eta'(1)$, hence $\eta * \overline{\eta'}$ is a loop in V based at $\gamma(1)$. By hypothesis on V, we have the homomorphism

$$\pi_1(V, \gamma(1)) \to \pi_1(X, \gamma(1))$$

is trivial. Hence $\eta * \overline{\eta'}$ is homotopic to $c_{\gamma(1)}$ in X. Therefore, we have

$$\gamma \sim \gamma * c_{\gamma(1)} \sim \gamma * (\eta * \overline{\eta'}) \sim \gamma' * (\eta' * \overline{\eta'}) \sim \gamma' * c_{\gamma(1)} \sim \gamma'.$$

Hence

$$U([\gamma], V) = U([\gamma'], V).$$

We now show that the space $\widetilde{X}$ is path connected. Let γ be a path in X with $\gamma(0) = p$. For any $t \in [0, 1]$, we define

$$\begin{aligned} \gamma_t : [0,1] &\to X \\ s &\mapsto \gamma(ts) \end{aligned}$$

Then we can define the map

$$\begin{aligned} \widetilde{\gamma} : [0,1] &\to \widetilde{X} \\ t &\mapsto [\gamma_t] \end{aligned}.$$

A direct verification shows that the map $\widetilde{\gamma}$ is the lift of γ with $\widetilde{\gamma}(0) = [c_p]$, hence a path connecting $[c_p]$ with $[\gamma]$.

Therefore, the space $\widetilde{X}$ is path connected.

Now let $\widetilde{\gamma}$ be a loop in $\widetilde{X}$ based at c_p:

$$\begin{aligned} \widetilde{\gamma} : [0,1] &\to \widetilde{X} \\ t &\mapsto [\gamma_t] \end{aligned},$$

where $\gamma = f \circ \widetilde{\gamma}$ is a path in X.

Notice that $\widetilde{\gamma}$ is a loop, hence we have

$$[c_p] = [\gamma_0] = [\gamma_1] = [\gamma],$$

or equivalently

$$\gamma \sim c_p.$$

Let H be a homotopy in X with $H_0 = \gamma$ and $H_1 = c_p$, then its lift $\widetilde{H}$ is a homotopy in $\widetilde{X}$ between $\widetilde{\gamma}$ and $\widetilde{c_p} = c_{[c_p]}$, hence

$$[\widetilde{\gamma}] = [c_{[c_p]}].$$

Therefore, we have

$$\pi_1(\widetilde{X}, [c_p]) = [c_{[c_p]}],$$

and $\widetilde{X}$ is simply connected. □

Definition 4.3.25

A path connected cover Z of X with covering map g is called a **universal cover** if it satisfies the following universal property: for any path connected cover X_1 of X with covering map f, the map f has a lift

$$\widetilde{g} : Z \to X_1,$$

such that $g = f \circ \widetilde{g}$.

Remark 4.3.26.
We have the following commutative diagram:

$$\begin{array}{ccc} & & X_1 \\ & \overset{\widetilde{g}}{\nearrow} & \downarrow f \\ Z & \xrightarrow[g]{} & X \end{array}$$

By the proposition 4.3.16, the cover $\widetilde{X}$ of X constructed above is a universal cover, since its fundamental group is trivial.

Example 4.3.27 (S^1)**.**
The universal cover of S^1 can be identified with $\mathbb{R}$. For each path α in S^1 starting at 1, there is a natural number $l \in \mathbb{R}$, such that we can find a standard path

$$\begin{aligned} \gamma_l : [0,1] &\to S^1, \\ t &\mapsto e^{2l\pi it}, \end{aligned}$$

homotopic to α. Here $2l\pi$ can be considered as the total angle passed when we go from 1 to $e^{2l\pi i}$ along α. Then the identification between $\widetilde{S^1}$ with $\mathbb{R}$ can be given as follows (See Figure 4.3.7 for an illustration):

$$\begin{aligned} h : \widetilde{S^1} &\to \mathbb{R}, \\ \gamma_l &\mapsto l. \end{aligned}$$

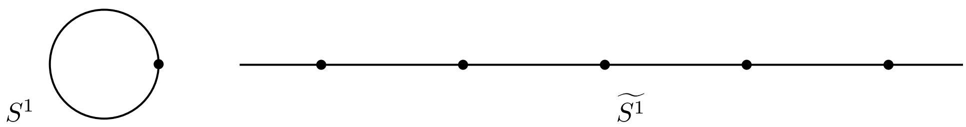


Figure 4.3.7: The universal cover of S^1.

Example 4.3.28 (Figure-8).
The universal cover of the figure-8 graph can be identified with the 4-valence regular infinite tree (See Figure 4.3.8 for an illustration).

Example 4.3.29 (Torus).
The universal cover of the torus T can be identified with the plane $\mathbb{R}^2$ (See Figure 4.3.9 for an illustration).

Remark 4.3.30.
The above construction of universal cover does not work for Hawaii earring (See Example 2.3.4). Notice that it is not semilocally simply connected.

Classification of covers

Let X be a topological space which is path connected, locally path connected and semilocally simply connected. In this part, we would like to give a classification of all path connected covers

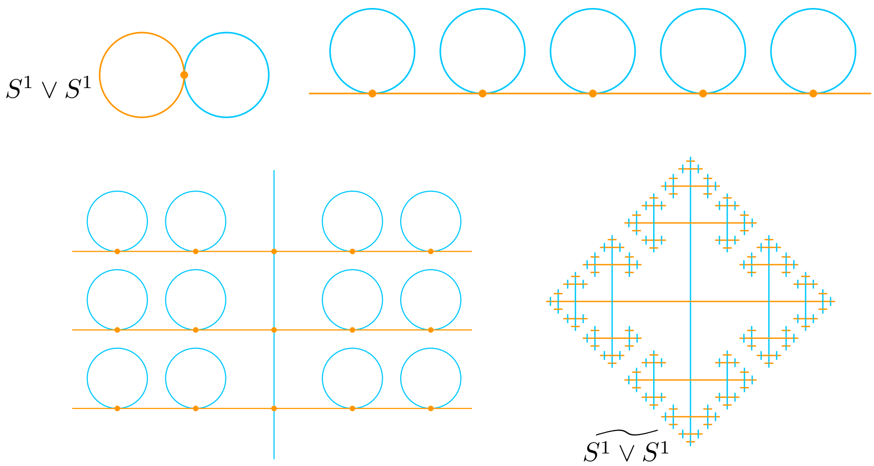


Figure 4.3.8: The universal cover of $S^1 \vee S^1$.

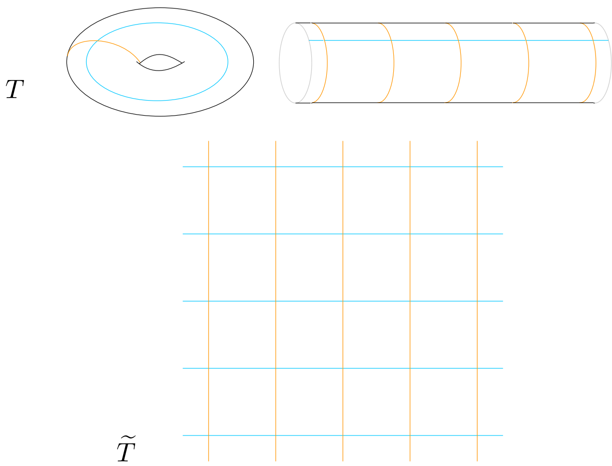


Figure 4.3.9: The universal cover of T.

of X.

Let $p \in X$ be a base point. We first show the existence result for every subgroup of $\pi_1(X, p)$.

Proposition 4.3.31

For any subgroup H of $\pi_1(X, p)$, there is a path connected cover X_H of X with covering map

$$f_H : X_H \to X,$$

such that $(f_H)_*(\pi_1(X_H, p_H)) = H$ where p_H is a lift of p in X_H.

Proof. The proof is constructive. We consider the universal cover constructed previously:

$$\widetilde{X} = \{[\gamma] \mid \gamma \text{ is a path in } X \text{ with } \gamma(0) = p\}.$$

We then define the following equivalent relation on $\widetilde{X}$: for any $[\gamma]$ and $[\gamma']$ in $\widetilde{X}$

$$[\gamma]\mathcal{R}[\gamma'] \Leftrightarrow \gamma(1) = \gamma'(1) \text{ and } [\gamma * \overline{\gamma'}] \in H.$$

The reflexivity is trivial, since

$$\gamma(1) = \gamma(1)$$

and

$$[\gamma * \overline{\gamma}] = [c_p] \in H.$$

If $[\gamma]$ and $[\gamma']$ two points in $\widetilde{X}$ satisfy $[\gamma]\mathcal{R}[\gamma']$, then

$$\gamma'(1) = \gamma(1)$$

and

$$[\gamma' * \overline{\gamma}] = [\gamma * \overline{\gamma'}]^{-1} \in H.$$

Hence we have $\gamma'\mathcal{R}\gamma$.

If $[\gamma], [\gamma']$ and $[\gamma'']$ three points in $\widetilde{X}$ satisfy

$$[\gamma]\mathcal{R}[\gamma'] \quad \text{and} \quad [\gamma']\mathcal{R}[\gamma''],$$

then we have

$$\gamma(1) = \gamma'(1) = \gamma''(1)$$

and

$$[\gamma * \overline{\gamma''}] = [\gamma * \overline{\gamma'}] * [\gamma' * \overline{\gamma''}] \in H.$$

Hence $\mathcal{R}$ is an equivalence relation on $\overline{X}$.

We then define

$$X_H := \widetilde{X}/\mathcal{R}.$$

The path connectedness of X_H comes from the fact that $\widetilde{X}$ is path connected.

Now we would like to show that X_H is a cover of X. We consider the map

$$\begin{aligned} f_H : X_H &\to X \\ \widehat{[\gamma]} &\mapsto \gamma(1) \end{aligned}.$$

It is well defined due to the fact that for any $[\gamma]\mathcal{R}[\gamma']$, we have

$$\gamma(1) = \gamma'(1).$$

Hence the image of $\widehat{[\gamma]}$ is independent of choice of γ in this $\mathcal{R}$-equivalence class $\widehat{[\gamma]}$.

Consider any pair $[\gamma], [\gamma'] \in \overline{X}$ with $[\gamma]\mathcal{R}[\gamma']$. Let η be a path in $V \in \mathcal{B}_{\gamma(1)}$ with $\eta(0) = \gamma(1)$. Then we have

$$(\gamma * \eta)(1) = (\gamma' * \eta)(1)$$

and

$$[(\gamma * \eta) * \overline{(\gamma' * \eta)}] = [\gamma * \overline{\gamma'}] \in H.$$

Hence

$$[\gamma * \eta]\mathcal{R}[\gamma' * \eta].$$

We denote

$$\widehat{U}([\gamma], V) = \mathrm{pr}(\widehat{U}([\gamma], V)).$$

The above discussion implies that

$$\widehat{U}([\gamma], V) = \widehat{U}([\gamma'], V).$$

Denote by pr the projection from $\overline{X}$ to X_H, we have

$$\mathrm{pr}^{-1}(\widehat{U}([\gamma],V)) = \bigcup_{[\gamma'] \text{ with } [\gamma']\mathcal{R}[\gamma]} U([\gamma'],V).$$

Notice that the above discussion shows one inclusion

$$\bigcup_{[\gamma'] \text{ with } [\gamma']\mathcal{R}[\gamma]} U([\gamma'],V) \subset \mathrm{pr}^{-1}(\widehat{U}([\gamma],V)).$$

To see the other inclusion, let $[\alpha] \in \widetilde{X}$ be in $\mathrm{pr}^{-1}(\widehat{U}([\gamma],V))$, there is $[\beta] \in U([\gamma],V)$, such that

$$[\beta]\mathcal{R}[\alpha],$$

we have $[\beta * \overline{\alpha}] \in H$. Notice that there is η path in V with $\eta(0) = \gamma(1)$ such that

$$[\beta] = [\gamma * \eta].$$

This implies that

$$[\beta * \overline{\eta}]\mathcal{R}[\gamma].$$

Let $\gamma' = \alpha * \overline{\eta}$, we have $[\gamma']\mathcal{R}[\gamma]$ and

$$[\alpha] \in U([\gamma'],V)$$

Hence we have the other inclusion

$$\mathrm{pr}^{-1}(\widehat{U}([\gamma],V)) \subset \bigcup_{[\gamma'] \text{ with } [\gamma']\mathcal{R}[\gamma]} U([\gamma'],V).$$

Hence $\widehat{U}([\gamma],V)$ is open.

Next denote pr_1 a restriction on $U([\gamma],V)$, and we would like to show that pr_1 is a homeomorphism.

The map pr_1 is surjective by its definition. On the other hand, for each $[\alpha]$ and $[\beta]$ in $U([\gamma],V)$, there is η and ζ paths in V with $\eta(0) = \zeta(0) = \gamma(1)$ such that

$$\alpha \sim \gamma * \eta, \quad \beta \sim \gamma * \zeta,$$

If $[\alpha]\mathcal{R}[\beta]$, then we have

$$\eta(1) = \alpha(1) = \beta(1) = \zeta(1),$$

hence $\eta * \overline{\zeta}$ is a loop in V based at $\gamma(1)$. Since $V \in \mathcal{B}_q$, we have

$$[\eta * \overline{\zeta}] = [c_q],$$

hence $\eta \sim \zeta$, and moreover

$$[\alpha] = [\gamma * \eta] = [\gamma * \zeta] = [\beta].$$

Hence pr_1 is injective. Notice that for any $q' \in V$, and any $W \in \mathcal{B}_{q'}$ such that $W \subset V$, we have

$$\mathrm{pr}_1(U([\gamma * \eta],W)) = \widehat{U}([\gamma * \eta],W), \quad \mathrm{pr}_1^{-1}(\widehat{U}([\gamma * \eta],W)) = U([\gamma * \eta],W).$$

Hence pr_1 is a homeomorphism between $U([\gamma * \eta],W)$ and $\widehat{U}([\gamma * \eta],W)$.

Now we would like to show that the map f_H is a covering map. The continuity comes from the fact that for any $q \in X$ and for any $V \in \mathcal{B}_q$, we have

$$f_H^{-1}(V) = \bigcup_{\overline{[\gamma]} \text{ with } \gamma(1)=q} \widehat{U}([\gamma],V).$$

Now we would like to show that for any γ and γ' path in X with $\gamma(0) = \gamma'(0) = p$ and $\gamma(1) = \gamma'(1) = q$, let $V \in \mathcal{B}_q$, if

$$\widehat{U}([\gamma], V) \cap \widehat{U}([\gamma], V) \neq \emptyset,$$

then

$$\widehat{U}([\gamma], V) = \widehat{U}([\gamma], V).$$

Let $\widehat{[\alpha]} \in \widehat{U}([\gamma], V) \cap \widehat{U}([\gamma], V)$, then there are paths η and η' in V with $\eta(0) = \zeta(0) = \gamma(1)$ and

$$[\alpha]\mathcal{R}[\gamma * \eta], \quad [\alpha]\mathcal{R}[\gamma' * \eta'].$$

Hence

$$\eta(1) = \eta'(1).$$

Since $V \in \mathcal{B}_q$, we have $\eta * \overline{\eta'}$ a loop based at q and

$$[\eta * \overline{\eta'}] = [c_q].$$

Consider the following relation

$$[\gamma * \overline{\gamma'}] = [\gamma * c_q * \overline{\gamma'}] = [\gamma * (\eta * \overline{\eta'})\overline{\gamma'}] = [\gamma * \eta] * [\overline{\eta'} * \overline{\gamma'}] = ([\gamma * \eta] * [\overline{\alpha}]) * ([\alpha] * [\overline{\eta'} * \overline{\gamma'}]) \in H.$$

Hence we have

$$[\gamma]\mathcal{R}[\gamma'],$$

and

$$\widehat{U}([\gamma], V) = \widehat{U}([\gamma], V).$$

As a conclusion, we show that the map f_H is a covering map.

The last thing to show is that

$$(f_H)_*(\pi_1(X_H), \overline{[c_p]}) = H.$$

Let γ be a loop based at p. Then we consider its lift γ_H in X_H with $\gamma_H(0) = \widehat{[c_p]}$. If γ_H is a loop, we have

$$\gamma_H(1) = \widehat{[\gamma]} = \widehat{[c_p]}.$$

Therefore we have

$$[\gamma]\mathcal{R}[c_p],$$

and this is equivalent to

$$[\gamma] \in H$$

by considering the definition of $\mathcal{R}$. □

Remark 4.3.32.
In a path connected space, identifying points may creating loops which are homotopically non trivial (See).

Hence for any subgroup H of $\pi_1(X, p)$, there is a path connected cover of X associated to H. Next we would like to show that such a cover is unique up to isomorphism between covers. More precisely, we show

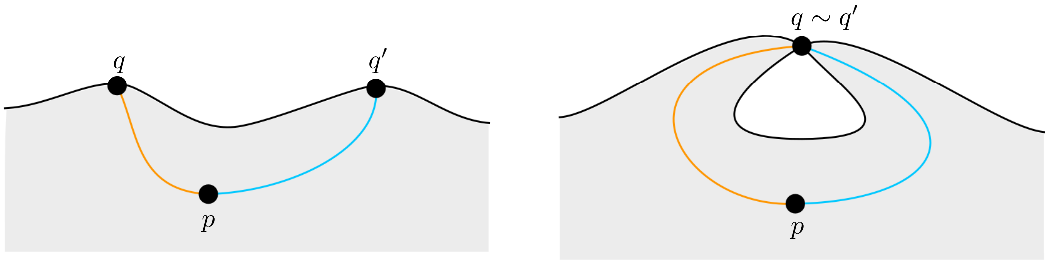


Figure 4.3.10: Identifying points creates loops.

Proposition 4.3.33

Let X_1 and X_2 be two path connected covers of X with covering maps f_1 and f_2 respectively. Let u_1 and u_2 be lifts of $p \in X$ in X_1 and X_2 respectively. Then X_1 and X_2 are isomorphic as covers of X if and only if

$$(f_1)_*(\pi_1(X_1, u_1)) = (f_2)_*(\pi_1(X_2, u_2)).$$

Proof. If two covers X_1 and X_2 are isomorphic to each other, then we have a homeomorphism

$$g : X_1 \to X_2$$

such that the following diagram commutes

$$\begin{array}{ccc} & & X_2 \\ & \overset{g}{\nearrow} & \downarrow f_2 \\ X_1 & \xrightarrow[f_1]{} & X \end{array}$$

Let u_1 and u_2 be lifts of p in X_1 and X_2 respectively, such that $g(u_1) = u_2$, then since g is a homeomorphism, we have

$$g_*(\pi_1(X_1, u_1)) = \pi_1(X_2, u_2).$$

Hence we have

$$(f_1)_*(\pi_1(X_1, u_1)) = ((f_2)_* \circ g_*)(\pi_1(X_1, u_1)) = (f_2)_*(\pi_1(X_2, u_2)).$$

Conversely, assume that

$$(f_1)_*(\pi_1(X_1, u_1)) = (f_2)_*(\pi_1(X_2, u_2)).$$

Using Proposition 4.3.16, there are maps

$$g_1 : X_1 \to X_2 \quad \text{and} \quad g_2 : X_2 \to X_1$$

which are lift of f_1 and lift of f_2 respectively with $g_1(u_1) = u_2$ and $g_2(u_2) = u_1$. We have the following commutative diagram:

$$\begin{array}{ccccc} & & X_2 & & \\ & \overset{g_1}{\nearrow} & \downarrow f_2 & \overset{g_2}{\searrow} & \\ X_1 & \xrightarrow[f_1]{} & X & \xleftarrow[f_1]{} & X_1 \end{array}$$

Therefore $g_2 \circ g_1$ is a lift of f_1 with respect to f_1:

$$\begin{array}{ccc} & & X_1 \\ & \overset{g_2 \circ g_1}{\nearrow} & \big\downarrow f_1 \\ X_1 & \xrightarrow[f_1]{} & X \end{array}$$

On the other hand, we have

$$\begin{array}{ccc} & & X_1 \\ & \overset{\mathrm{id}_{X_1}}{\nearrow} & \big\downarrow f_1 \\ X_1 & \xrightarrow[f_1]{} & X \end{array}$$

By Proposition 4.3.17, we have

$$g_2 \circ g_1 = \mathrm{id}_{X_1}.$$

Similarly, we have

$$g_1 \circ g_2 = \mathrm{id}_{X_2}.$$

Hence g_1 and g_2 are homeomorphisms, and the two covers X_1 and X_2 are isomorphic. □

As a corollary, we have the following relation between covers of X and subgroups of $\pi_1(X, p)$.

Corollary 4.3.34

There are a bijection

$$\{\text{subgroups of } \pi_1(X,p)\} \leftrightarrow \{\text{covers of } (X,p), \text{ up to base point preserving isomorphisms}\}.$$

and a bijection

$$\{\text{conjugacy class of subgroups of } \pi_1(X,p)\} \leftrightarrow \{\text{covers of } (X,p), \text{ up to isomorphisms}\}.$$

Now we give some examples to illustrate these results.

Example 4.3.35 (S^1).
Since S^1 has fundamental group isomorphic to $\mathbb{Z}$, let $p \in S^1$ be a base points, and

$$[\gamma] \in \pi_1(S^1, p)$$

be a generator. Then all subgroups of $\pi_1(S^1, p)$ will have the form

$$\langle [\gamma^k] \rangle,$$

for some $k \in \mathbb{N}$. In Figure 4.3.11, we show the covering spaces of S^1 for $k = 0, 1, 2, 3, 4$.

Example 4.3.36 (Figure eight).
We have seen the universal cover of $S^1 \vee S^1$. Since the fundamental group of $S^1 \vee S^1$ is a free group of 2 letters, its subgroups are all free, and the rank could be any finite natural number or infinite.

Let p denote the vertex and be the base point. Let α and β denote the loops based at p associated to the two copies of S^1 respectively. Hence

$$a := [\alpha] \quad \text{and} \quad b := [\beta]$$

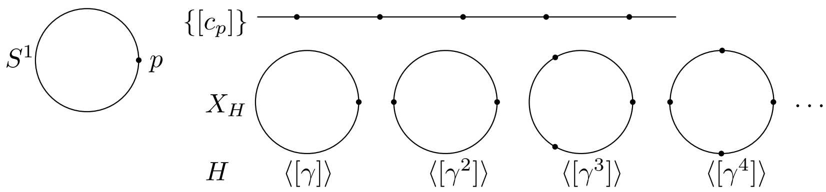


Figure 4.3.11: Covering spaces of S^1.

form a free basis of $\pi_1(S^1 \vee S^1, p)$. Denote by e the identity element.

In the following figures, we show some covering spaces of $S^1 \vee S^1$ and their corresponding subgroups of $\pi_1(S^1 \vee S^1, p)$.

1) $H = \langle a \rangle$.

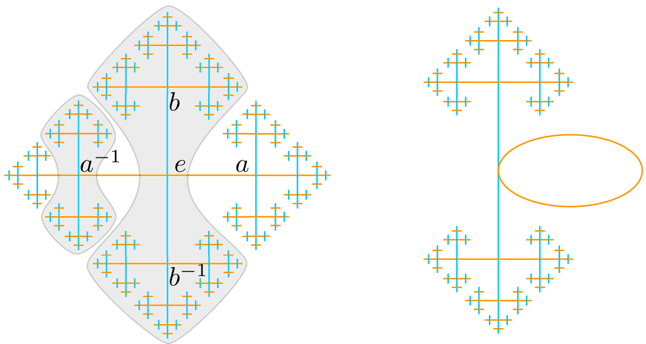


Figure 4.3.12: Covering space of $S^1 \vee S^1$ for $\langle a \rangle$.

2) $H = *_{k\in\mathbb{Z}}\langle b^k a b^{-k} \rangle$.

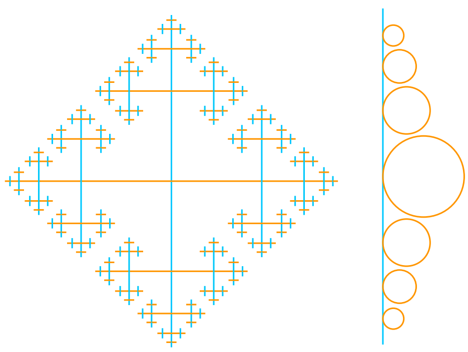


Figure 4.3.13: Covering space of $S^1 \vee S^1$ for $*_{k\in\mathbb{Z}}\langle b^k a b^{-k} \rangle$.

3) $H = \langle a^2, b^2 \rangle$.

4) $H = \langle a, b^2, bab \rangle$.

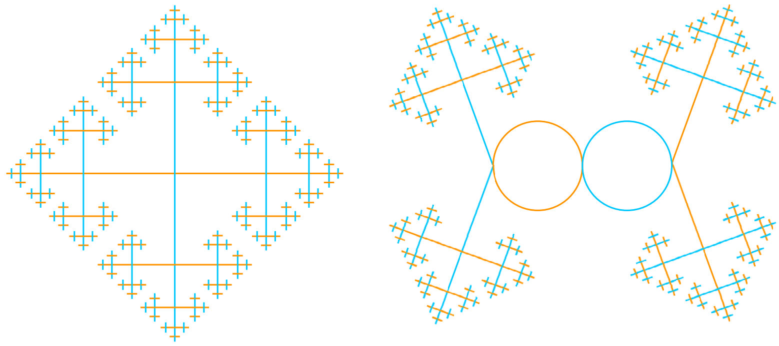


Figure 4.3.14: Covering space of $S^1 \vee S^1$ for $\langle a^2, b^2 \rangle$.

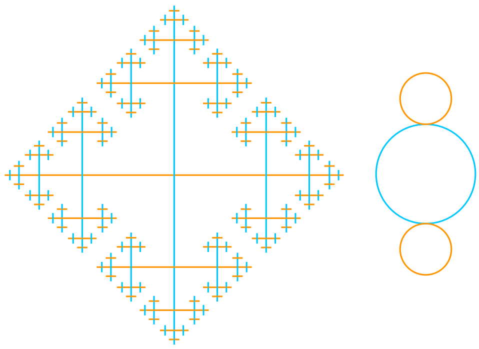


Figure 4.3.15: Covering space of $S^1 \vee S^1$ for $\langle a, b^2, bab \rangle$.

Notice that the first three subgroups are infinite index subgroups, while the last one is a subgroup of index 2.

Deck transformations

In this part we consider a space X which is path connected locally path connected and semilocally simply connected. We consider its universal cover

$$\widetilde{X} := \{[\gamma] \mid \gamma \text{ path in } X \text{ with } \gamma(0) = p\},$$

and denote the covering map by

$$\widetilde{f} : (\widetilde{X}, \widetilde{p}) \to (X, p).$$

There is a natural action of $\pi_1(X, p)$ on $\widetilde{X}$ induced by the following map

$$\begin{aligned} \Phi : \pi_1(X, p) \times \widetilde{X} &\to \widetilde{X}, \\ ([\alpha], [\gamma]) &\mapsto [\alpha * \gamma]. \end{aligned}$$

We may directly check the following two facts to see that it is indeed a left action

1) $[c_p].[\gamma] = [c_p * \gamma] = [\gamma]$;

2) $[\alpha].([\beta].[\gamma]) = [\alpha].[\beta * \gamma] = [\alpha * \beta * \gamma] = [\alpha * \beta].[\gamma]$.

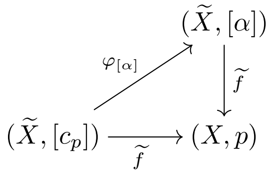
For any $[\alpha] \in \pi_1(X, p)$, the map

$$\varphi_{[\alpha]} : \widetilde{X} \to \widetilde{X}$$
$$[\gamma] \mapsto [\alpha * \gamma]$$

is a homeomorphism. Notice that for any $U([\gamma], V)$ where $V \in \mathcal{B}_{\gamma(1)}$, we have

$$\varphi_{[\alpha]}^{-1}(U([\gamma], V)) = U([\overline{\alpha} * \gamma], V),$$

which is still open by definition. It is an isomorphism between covers:



We generalize this discussion for any cover of X and give the following definition.

Definition 4.3.37

Let X_1 be a cover of X. We denote the covering map by

$$f : (X_1, p_1) \to (X, p).$$

A ***deck transformation*** on X_1 is an isomorphism

$$g : X_1 \to X_1$$

between covers.

Notice that all deck transformations on X_1 form a group under the composition operation. We denote by $\mathrm{Deck}(X_1)$ the ***deck transformation group*** of X_1.

From its definition, a deck transformation is in particular a lift of f with respect to f. By the uniqueness of the lift of a continuous map, such a deck transformation is determined by $g(p_1) \in f^{-1}(p)$. There is then an immediate question: given any lift p_1' of p in X_1, do we have a deck transformation such that $g(p_1) = p_1'$?

Notice that if there is a deck transformation

$$g : X_1 \to X_1,$$

with $g(p_1) = p_1'$. By Proposition 4.3.16, we have

$$f_*(\pi_1(X_1, p_1)) = f_*(\pi_1(X_1, p_1')).$$

Considering the change of base point in X_1, we have a path α_1 such that

$$\alpha(0) = p_1 \quad \text{and} \quad \alpha_1(1) = p_1'.$$

Hence we have

$$\pi_1(X_1, p_1') = [\overline{\alpha_1}] * \pi_1(X_1, p_1) * [\alpha_1].$$

We then have

$$f_*(\pi_1(X_1, p_1)) = f_*(\pi_1(X_1, p_1')) = f_*(\overline{\alpha_1} * \pi_1(X_1, p_1) * \alpha_1) = [\alpha]^{-1} * f_*(\pi_1(X_1, p_1)) * [\alpha],$$

where $\alpha = f \circ \alpha_1$ which is a loop based at p in X, since α_1 starts and ends at lifts of p. If we denote

$$H = f_*(\pi_1(X_1, p_1)) < \pi_1(X, p),$$

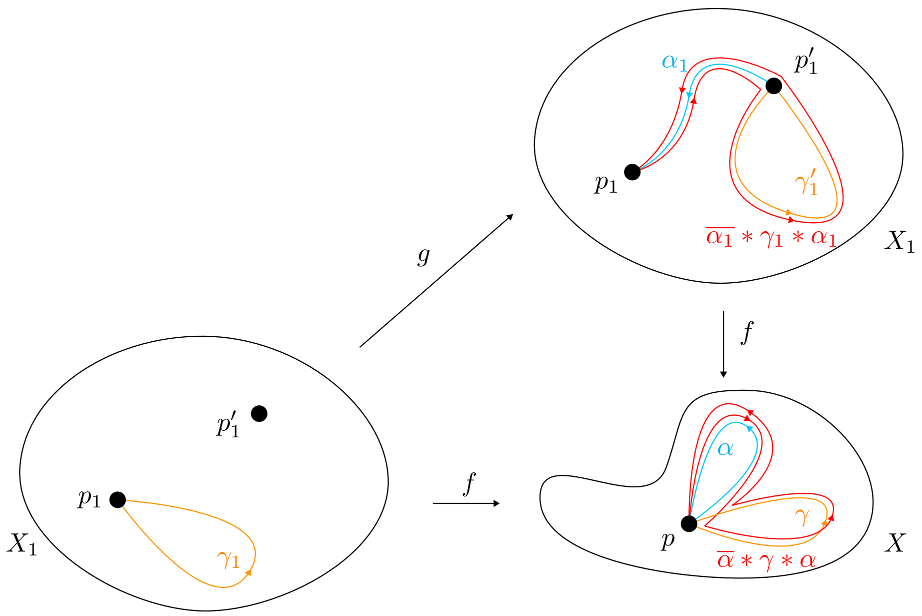


Figure 4.3.16: Deck transformations change an element in the f_*-image by a conjugacy.

then

$$[\alpha] \in N(H),$$

the normalizer of H.

We can have another changes of base point in X_1 denoted by β_1, and we have

$$[\alpha_1] = [\alpha_1 * \overline{\beta_1}] * [\beta_1],$$

where $[\alpha_1 * \overline{\beta_1}] \in \pi_1(X_1, p_1)$. If we consider $\beta = f \circ \beta_1$, then since

$$[\alpha * \overline{\beta}] \in H,$$

we have

$$[\alpha] * H = [\beta] * H.$$

Therefore, $[\alpha]$ and $[\beta]$ are representative of a same element in $N(H)/H$. From this observation, we have the following proposition.

Proposition 4.3.38

There is an isomorphism from $\text{Deck}(X_1)$ to $N(H)/H$.

Proof. Denote the quotient map from $N(H)$ to $N(H)/H$ by

$$\pi : N(H) \to N(H)/H.$$

Then we may construct the following map

$$\begin{aligned} \Psi : \text{Deck}(X_1) &\to N(H)/H \\ g &\mapsto \pi([\alpha]) \end{aligned}$$

where α has a lift α_1 in X_1 with

$$\alpha(0) = p_1 \quad \text{and} \quad \alpha(1) = g(p_1).$$

By the previous discussion, g is determined by $g(p_1)$, and all homotopy classes of paths in X_1 from p_1 to $g(p_1)$ are mapped to $N(H)$ by f_* which are representatives of a same element in $N(H)/H$. Hence this is a well-defined map.

Let g and h be two deck transformations on X_1. Let p_1' and p_1'' be lifts of p in X_1 with

$$g(p_1) = p_1' \quad \text{and} \quad h(p_1') = p_1''.$$

Let $[\alpha]$ and $[\beta]$ be elements in $N(H)$ corresponding to g and h respectively. Then we have the lift $h \circ \alpha_1$ of α with

$$\alpha_1(0) = p_1' \quad \text{and} \quad \alpha_1(1) = p_1'',$$

and the lift β_1 of β with

$$\beta_1(0) = p_1 \quad \text{and} \quad \beta_1(1) = p_1'.$$

Hence $\beta_1 * \alpha_1$ is a path in X_1 going from p_1 to

$$p_1'' = h(g(p_1)).$$

which is a lift of $[\beta * \alpha]$. Hence we have

$$\Psi(h \circ g) = \pi([\alpha * \beta]) = \pi([\alpha]) * \pi([\beta]) = \Psi(h) * \Psi(g).$$

The surjectivity comes from Proposition 4.3.16. For any $[\alpha] \in N(H)$, there is a lift α_1 of α in X_1 with $\alpha_1(0) = p_1$. We denote by $p_1' = \alpha_1(1)$. Then since $[\alpha] \in N(H)$, we have

$$[\alpha]^{-1} * H * [\alpha] = H.$$

This is equivalent to

$$f_*(\pi_1(X_1, p_1)) = f_*(\pi_1(X_1, p_1')).$$

Hence by Proposition 4.3.16, we have a deck transformation sending p_1 to p_1'. Hence we have

$$\pi([\alpha]) = \Psi(g).$$

The injectivity comes from the uniqueness of the lift. If we have $[\alpha] \in N(H)$, such that $\pi([\alpha])$ is trivial in $N(H)/H$, then we have

$$[\alpha] \in H.$$

Hence the lift α_1 of α in X_1 with $\alpha_1(0) = p_1$ will have

$$\alpha_1(1) = p_1.$$

Hence any deck transformation corresponding to $\pi([\alpha])$ will satisfies $g(p_1) = p_1$. On the other hand, the identity map id_{X_1} is a lift of f. By the uniqueness of the lift, we have $g = \mathrm{id}_{X_1}$.

As a conclusion, the map Ψ is an group isomorphism. □

Given any lift p_1' of p_1 in X_1 and any path α_1 from p_1 to p_1' in X_1, its projection

$$\alpha = f \circ \alpha'$$

is a loop based at p in X. On the other hand, given any loop α in X based at p, it can always be lifted to α_1 a path in X_1 with $\alpha_1(0) = p_1$. These two observations shows the following facts.

Corollary 4.3.39

The group $\text{Deck}(X_1)$ acts on $f^{-1}(p)$ transitively, i.e.

$$\text{Deck}(X_1).p_1 = f^{-1}(p),$$

if and only if $G = N(H)$, i.e. $H \lhd G$.

Definition 4.3.40

A cover X_1 of X is called a ***normal cover*** if

$$\text{Deck}(X_1).p_1 = f^{-1}(p).$$

Example 4.3.41 (Figure eight degree 2).
By Corollary 4.3.8, the index of the subgroup associated to any degree 2 cover is 2. Since in any group, any index 2 subgroup is normal, we may conclude that any index 2 cover of $S^1 \vee S^1$ is a normal cover. All index 2 covers are as follows (Figure 4.3.17). Notice that the deck transformation

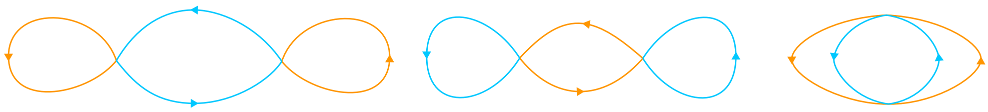


Figure 4.3.17: All degree 2 covers of $S^1 \vee S^1$.

can be constructed by considering first how lifts of the vertex are mapped. Notice that this also tells us all index 2 subgroup of F_2 up to conjugacy.

Example 4.3.42 (Figure eight degree 3).
We list all degree 3 covers of $S^1 \vee S^1$ below in Figure 4.3.18. Notice that the one on the left are normal covers, while the ones on the right are not.

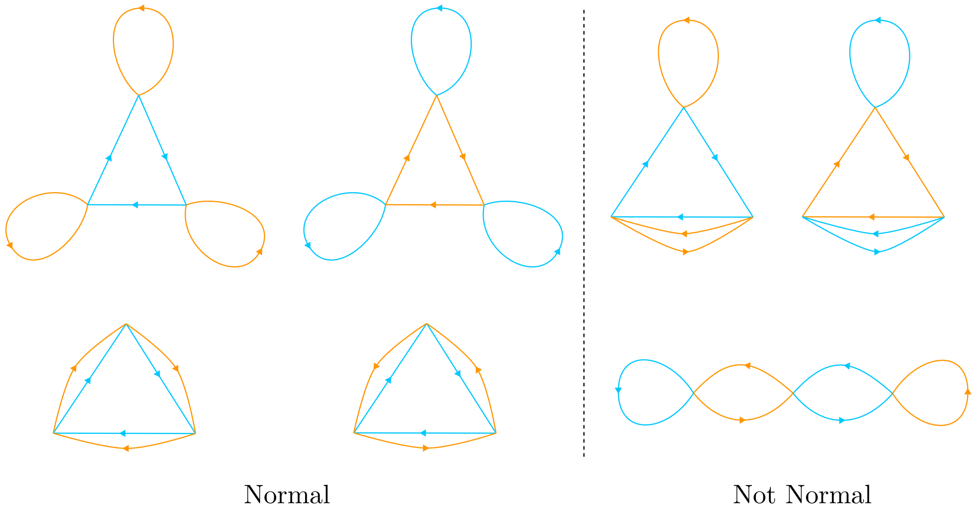


Figure 4.3.18: All degree 3 covers of $S^1 \vee S^1$.

Chapter 5

Surface

The topological spaces familiar to everyone the most would be the Euclidean space $\mathbb{R}^n$. A topological manifold can be consider as a generalization $\mathbb{R}^n$. Roughly speaking, a *manifold of dimension n*, or simply an *n-manifold*, is a topological space which locally looks like $\mathbb{R}^n$. In another words, one may consider a n-manifold is constructed by gluing open sets of $\mathbb{R}^n$ together. For example, the circle can be considered as a result of gluing two intervals together. Hence it is a 1-dimensional manifold.

In this chapter, we will focus on the dimension 2 case, i.e. the surface case. We would like to consider this case as an example to review the content which was introduced previously. In particular, we will give the classification of closed compact surfaces. Moreover, we will discuss the triangulation of surfaces to give an idea of what are simplicial structures for a topological space, and how to use them to obtain topological invariants such that Euler characteristics and orientations.

5.1 Surfaces in various contents

Surfaces are elements objects studied in many area. In the following, we take the torus as an example to illustrate this fact.

Example 5.1.1 (Torus in Different Geometry).
In the Euclidean space $\mathbb{R}^3$, the following formula define a torus

$$\begin{aligned} &f : \mathbb{R}^2 \to \mathbb{R}^3, \\ &(\theta, \eta) \mapsto (\cos\theta(\cos\eta + 2), \cos\theta(\sin\eta + 2), \sin\theta). \end{aligned}$$

This formula gives a local charts on torus. We could use it to compute quantities such as area, curvature etc.

Example 5.1.2 (Torus as a Riemann Surface).
We consider the algebraic equation in $\widehat{\mathbb{C}}^2$

$$w^2 = z(z-1)(z-\lambda),$$

where $\lambda \in \widehat{\mathbb{C}} \setminus \{0, 1, \infty\}$. The solution set in $\widehat{\mathbb{C}}^2$ is topological a torus. Notice that given any $z \in \widehat{\mathbb{C}}$, there are two distinct roots for w:

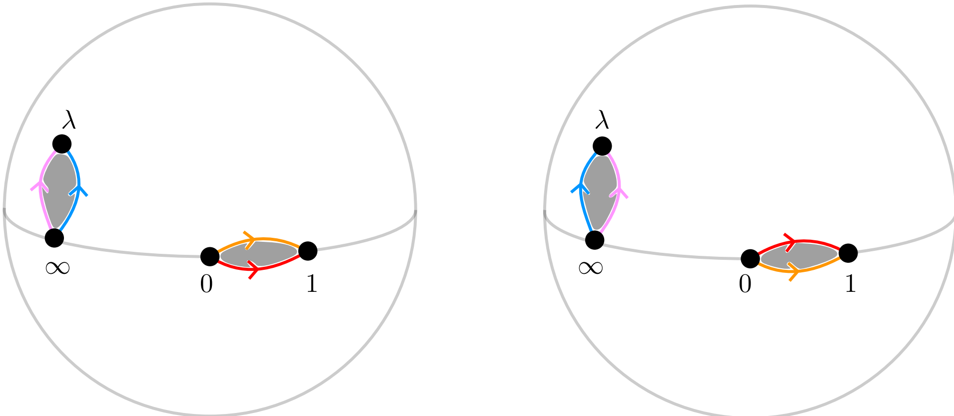
$$\begin{aligned} w &= \sqrt{z(z-1)(z-\lambda)} \\ w &= -\sqrt{z(z-1)(z-\lambda)} \end{aligned}$$

except when $z = 0, 1, \lambda, \infty$.

Consider the projection of $\widehat{\mathbb{C}}^2$ to the copy of $\widehat{\mathbb{C}}$ for z, and the above discussion shows that the restriction of this map to the solution set is a 2-cover branched over $\{0, 1, \lambda, \infty\}$.

If we take a circle path separate $z = 0$ from other three branched points, when we go around once along this path, we change the root $w = \sqrt{z(z-1)(z-\lambda)}$ to $w = -\sqrt{z(z-1)(z-\lambda)}$. On the contrary, if we take a circle path separate $z = 0, 1$ from $z = \lambda, \infty$, then after going around once along this path, we still get the same root for w.

One may roughly understand this phenomenon in the following way. Since we have a 2-cover, there are two copies of $\widehat{\mathbb{C}}$ for w. If the z parameter walks along a path around 0 only, then the w parameter goes from one copy of $\widehat{\mathbb{C}}$ to another copy. If the z parameter walks along a path around 0 and 1, then the w parameter stays in the same copy of $\widehat{\mathbb{C}}$ for w.



The copy of z-plane for (z, w) The copy of z-plane for $(z, -w)$

Figure 5.1.1: Gluing two copies of $\widehat{\mathbb{C}}$ along two slices connecting 0 to 1 and ∞ to λ respectively.

Hence one may consider cut a slip in $\widehat{\mathbb{C}}$ along the interval $[0, 1]$ in $\mathbb{R}$ and glue the two copies of $\widehat{\mathbb{C}}$ along this slice. The same discussion works for ∞ and λ, we may take a path with no self-intersection going from λ to ∞ and disjoint from $[0, 1]$. We cut a slice along it, and glue the two copies of $\widehat{\mathbb{C}}$ along the slice. As a result, the solution set is a torus.

Notice that the above construction of torus depends on a choice of 4 point $0, 1, \lambda, \infty$. Using fractional linear map, we can send any triple of distinct points in $\widehat{\mathbb{C}}$ to $\{0, 1, \infty\}$. Hence up to holomorphism, the complex structure on a torus is determined by λ.

Example 5.1.3 (Torus from group actions).
Another way to describe a torus is by considering it as a quotient space of $\mathbb{R}^2$ under a group action. This has been described previously in Example 2.3.21.

Figure 5.1.2 lists some surfaces that one may meet in various occasions.

5.2 Construction of surfaces using polygons

Before we start, it should be remark that all discussions from now on are based on a result which we will admit and will not give a proof. It says that any surface can have a triangulation (which we will introduce later). For the proof of this result, one may read.

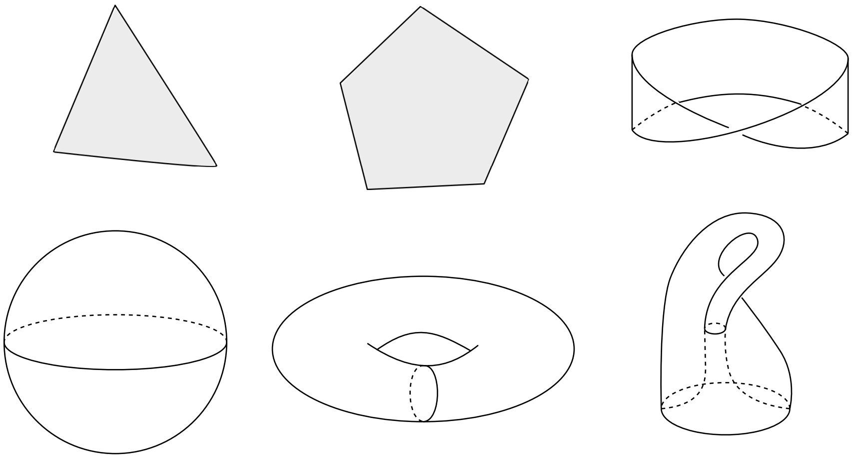


Figure 5.1.2: First row: triangle, pentagon, Möbius band; second row: sphere, torus, Klein bottle.

Polygon

Recall that in the context of Euclidean geometry, a ***convex Euclidean polygon*** is a compact region, geometrically a intersection of finitely many half plans (see Figure 5.2.1).

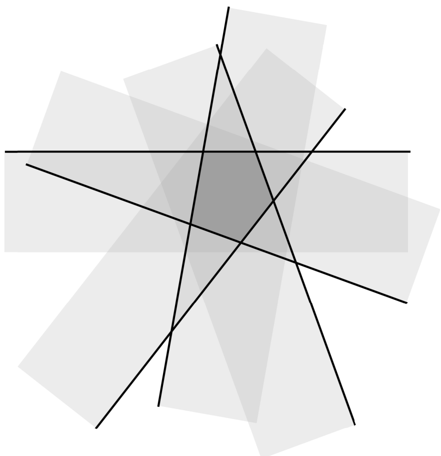


Figure 5.2.1: A polygon as the intersection of half planes in $\mathbb{R}^2$.

The boundary of a polygon is piecewise straight. Each straight piece is usually called an ***edge*** of the polygon, and each pair of edges meet at a ***vertex*** of the polygon. One may notice that a polygon is a topologically a disk, with some points on its boundary marked special.

A topological ***polygon*** is topologically a closed disk with finitely many marked points on its boundary called vertices. The vertices separate the boundary into connected components, each one of which is called an edge of the polygon.

Labels

The goal is to obtain surfaces by sewing edges of a polygon together. To do this, we have to clarify which edges are glued together in which way. For the first "which", we given each edge of a polygon a letter. Then two edges labeled by a same letter will be glued. For the second "which", we give each edge an orientation. Since each edge is an interval. Up to homotopy, there are two homeomorphism between two intervals, corresponding two ways of identifying them. We use 1 and -1 to represent the two orientation, then we consider the map from one interval to another preserving the chosen orientation.

More precisely, let P be a polygon with the set of vertices

$$V = \{v_1, ..., v_n\},$$

and the set of edges

$$E = \{e_1, ..., e_n\}.$$

We orient ∂P with the counterclockwise direction, and the vertices and the edges are ordered following this orientation.

Let S be a finite set of letters. A ***label*** on P with letters in S is a map

$$\begin{aligned} L : E &\to S \times \{1, -1\} \\ e &\mapsto (x, \epsilon) \end{aligned}$$

such that for each letter $x \in S$, the preimage $L^{-1}(\{(x, \pm 1)\})$ contains are most 2 elements (See the left figure in Figure 5.2.2 for an illustration).

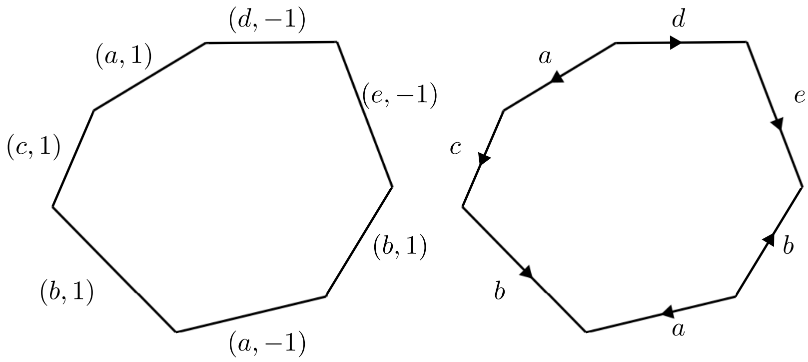


Figure 5.2.2: A labeled polygon.

Remark 5.2.1.
Since the edges labeled by a same letter will be glued together and our goal is to get a surface, the last requirement is natural. Otherwise, if we identify three or more half disks along their diameter, consider the point on the diameter and it has not neighborhood homeomorphic to a disk.

Another observation is that in order to avoid the existence of boundary, all edges of P should be paired, i.e. for any $e \in E$ labeled by (x, ϵ), there should be another edge labeled by either $(x, 1)$ or $(x, -1)$.

In the rest of this chapter, we will not discuss surface with boundaries, hence all polygons will have even number of edges. The orientation will be labeled using arrows (See the right polygon in Figure 5.2.2).

In the following, starting from the edge e_1, if the label is given following the counterclockwise direction by

$$(a_1, \epsilon_1), (a_2, \epsilon_2), ..., (a_{2n}, \epsilon_{2n}),$$

then the label for e_i is denoted by $a_i^{\epsilon_i}$, and the label is denoted by a word

$$a_1^{\epsilon_1} \cdots a_{2n}^{\epsilon_{2n}}.$$

Moreover, if $\epsilon_i = 1$, we omit it. As an example, in Figure 5.2.2, the edge labels are

$$a, c, b, a^{-1}, b, e^{-1}, d^{-1},$$

and the label is written as

$$acba^{-1}be^{-1}d^{-1}.$$

When a polygon P is given a label L, we call it a ***labeled polygon*** and denote it by (P, L). Two edges of P are said to be ***paired*** if there is a letter a, such that the two labels of the two edges are in

$$\{a, a^{-1}\}.$$

From polygon to surface

Consider a polygon P with $2n$ edges with $n \in \mathbb{N} \setminus \{0, 1\}$. We denote its vertices by $v_1, ..., v_{2n}$ following a cyclic order induces by the counterclockwise direction of ∂P. As a convention, we consider indices of vertices and edges up to mod $2n$, and assume that the vertices of e_i is v_i and v_{i+1}.

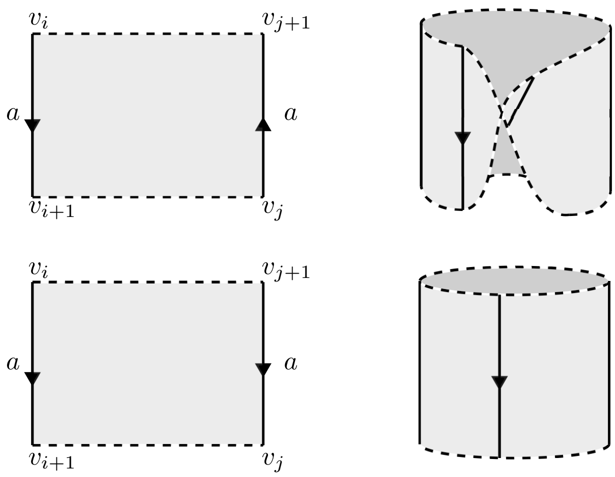


Figure 5.2.3: Different identifications of paired sides associated to different ways of labeling.

Let L be a label of P using n letters. Hence all edges of P are paired. Assume that e_i and e_j are labeled by a^ϵ and $a^{\epsilon'}$ for some letter $a \in S$.

If $\epsilon = \epsilon'$, we consider a homeomorphism

$$\varphi_a : e_i \to e_j,$$

such that $\varphi_a(v_i) = v_j$ and $\varphi_a(v_{i+1}) = v_{j+1}$.

If $\epsilon \neq \epsilon'$, we consider a homeomorphism

$$\varphi_a : e_i \to e_j,$$

such that $\varphi_a(v_i) = v_{j+1}$ and $\varphi_a(v_{i+1}) = v_j$.

Notice that such homeomorphism is unique up to homeomorphism. For any edge $v_i v_{i+1}$ labeled by x, we identify $p \in v_i v_{i+1}$ with $\varphi_x(p)$, then we denote the associated quotient space by Σ. In this case, we call the labeled polygon (P, L) a ***polygonal presentation*** of Σ. We denote by π the projection map (gluing map)

$$\pi : P \to \Sigma.$$

Quotient topology on Σ

The quotient topology on Σ gives a 2-manifold structure on it, i.e. each point admits a neighborhood which is homeomorphic to a disk in $\mathbb{R}^2$. More precisely, consider Σ as a quotient space of a labeled polygon (P, L), where P is a $2n$-gon for some $n \in \mathbb{N} \setminus \{0, 1\}$ and its edges are all paired through L. There are three types points in P, interior points, edge points and vertices. To describe the topology around each point in P, we may identity P with an Euclidean polygon for the moment and consider the subspace topology. We still denote by $v_1, ..., v_{2n}$ the vertices and $e_1, ..., e_{2n}$ the edges following the counterclockwise direction on ∂P, such that e_i is adjacent to v_i for any $i \in \{1, ..., 2n\}$.

There is not much to say about interior points, since the restriction of π to $\mathring{P}$ is a homeomorphism to its image. To be more precise, notice that the topology on P is Hausdorff, and for any $p \in \mathring{P}$, there is a neighborhood basis of p contained in $\mathring{P}$. Hence the restriction of π to each such neighborhood is an homeomorphic to image and the image of these neighborhoods form a neighborhoods of $\pi(p)$.

Now consider a point $p \in \mathring{e}_i$ for some edge e_i. It has a neighborhoods formed by half disks in $\mathbb{R}^2$. If e_i is labeled by a^ϵ and another edge e_j is labeled by $a^{\epsilon'}$, then there is a point $q \in \mathring{e}_j$ with $\pi(p) = \pi(q)$. We consider U a half disk neighborhood of p and V a half disk neighborhood of q, such that $\pi(U \cap e_i) = \pi(V \cap e_j)$. Then a neighborhood of $\pi(p)$ can be given by

$$\pi(U \cup V) \cong U \sqcup V / (x \sim y \Leftrightarrow \pi(x) = \pi(y))$$

which is homeomorphic to a disk in $\mathbb{R}^2$. Notice that such neighborhoods form a neighborhood basis of $\pi(p)$.

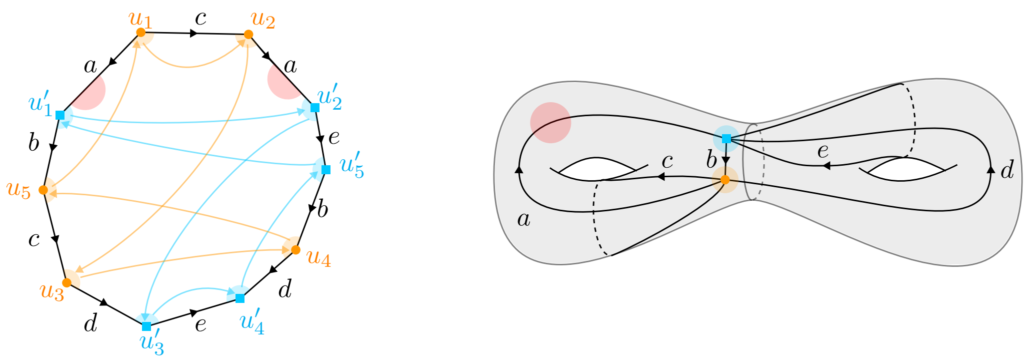


Figure 5.2.4: Glue neighborhoods of boundary points together. There are two vertex cycles: $u_1 u_2 u_3 u_4 u_5$ and $u'_1 u'_2 u'_3 u'_4 u'_5$.

To see how neighborhoods of vertices are glued together, we consider the following process. Let $u_1 = v_{i_1}$ be a vertex of P, it has two adjacent edges e_{i_1} and $e_{i_1 - 1}$. We denote $f_1 = e_{i_1}$. Let e_{i_2} be the edge paired with e_{i_1}, the we have

$$\pi_1(u_1) = \pi_1(v_{i_2}) \quad \text{or} \quad \pi_1(u_1) = \pi_1(v_{i_2 - 1}).$$

We denote $f_1' = e_{i_2}$ and by u_2 the vertex of e_{i_2} identified with u_1 through π. If we have obtained u_j, then we consider f_j' be the edge adjacent to u_j different from f_j with which we obtained u_j. Denote by f_{j+1} the edge paired with f_j', and by u_{j+1} the vertex of f_{j+1} identified with u_j through π. Notice that there are only $2n$ vertices, there will be a step $k \in \mathbb{N}^*$ such that for any $1 < j < k$, $u_j \neq u_1$, and

$$u_k = u_1.$$

We call

$$u_1 \cdots u_k$$

a ***vertex cycle*** for (P, L) (See Figure 5.2.4 for an example). Through π, the neighborhoods of u_j's in a vertex cycle are glued together to give a neighborhood of $\pi(u_1)$ in Σ. More precisely, recall that we have identified P with an Euclidean polygon. Consider a sector neighborhood S_j for each u_j with a same radius, then we identify their radius sides together to get a space

$$\pi\left(\bigcup_{j=1}^{k} S_j\right) \cong \bigsqcup_{j=1}^{k} S_j/(x \sim y \Leftrightarrow \pi(x) = \pi(y)).$$

Since sectors with different angles are homeomorphic and the homeomorphic can be given by rescaling the central angles. Hence we can identify all S_j's with sectors of central angle $2\pi/k$. Then the resulting space

$$\bigsqcup_{j=1}^{k} S_j/(x \sim y \Leftrightarrow \pi(x) = \pi(y)),$$

is homeomorphic to an Euclidean disk, which gives a neighborhood of $\pi(u_1)$.

As a conclusion, the quotient space Σ is a 2-manifold, i.e. a surface (See Figure 5.2.4 for an illustration).

Remark 5.2.2.
The whole story also works for a finite collection of polygons. Let $P_1, ..., P_k$ be a collection of polygons. Then we can define a label on it and obtained a surface from them by gluing according the the label. The only difference is that the resulting surface may not be connected.

Some examples

We gives some examples to illustrate the above discussion.

1) The quadrilateral labeled by $a^{-1}ab^{-1}b$ is glued into the 2-sphere S^2 (see Figure 5.2.5 for an illustration):

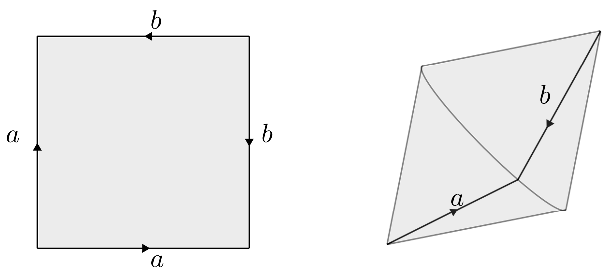


Figure 5.2.5: A polygon labeled by $a^{-1}ab^{-1}b$ is glued into the 2-phere.

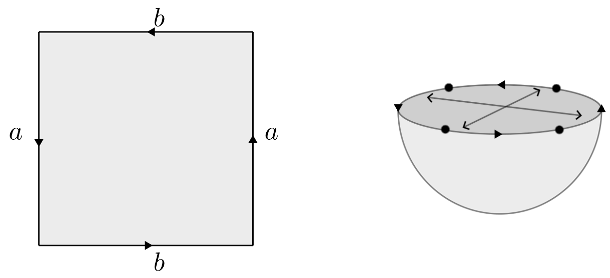


Figure 5.2.6: A polygon labeled by $abab$ is glued into the projective plane.

2) The quadrilateral labeled by $abab$ is glued into the projective plane $\mathbb{RP}^2$ (see Figure 5.2.6 for an illustration):

3) The quadrilateral labeled by $aba^{-1}b^{-1}$ is glued into the torus T (see Figure 5.2.7 for an illustration):

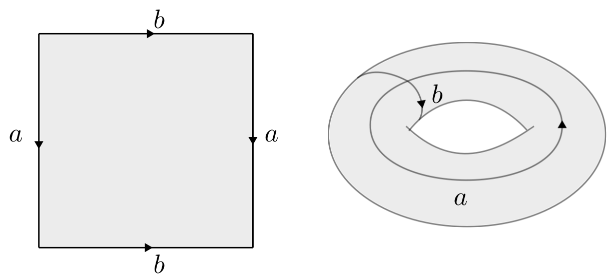


Figure 5.2.7: A polygon labeled by $aba^{-1}b^{-1}$ is glued into the torus.

4) The quadrilateral labeled by $abab^{-1}$ is glued into the Klein bottle K (see Figure 5.2.8 for an illustration):

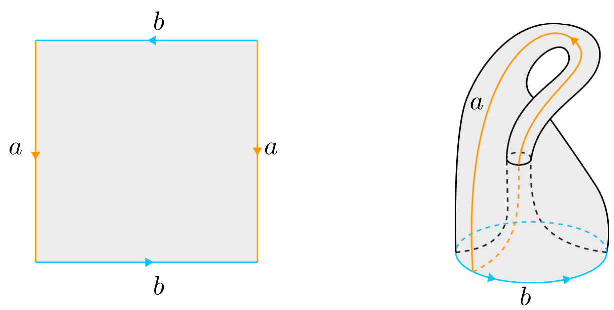


Figure 5.2.8: A polygon labeled by $abab^{-1}$ is glued into the Klein bottle.

Remark 5.2.3.
Klein bottles can only appear in dimension 4 and higher. The figure is a 2-dimension illustration of its projection to $\mathbb{R}^3$. This is why it looks like intersecting itself. To understand this, one may consider the projection of a non-trivial knot in $\mathbb{R}^3$ to a plane to understand this (see Figure 5.2.9 for an illustration of a projection of a trefoil knot in $\mathbb{R}^3$ to a plane).

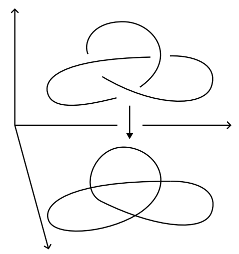


Figure 5.2.9: Projection of a trefoil knot in $\mathbb{R}^3$ to a plane.

5.3 Classification of closed surfaces

Equivalent labeled polygons

When trying to find polygonal presentation of a surface Σ, one may notice that such a presentation may not be unique. For example, the following two labeled polygons are both presentation for torus.

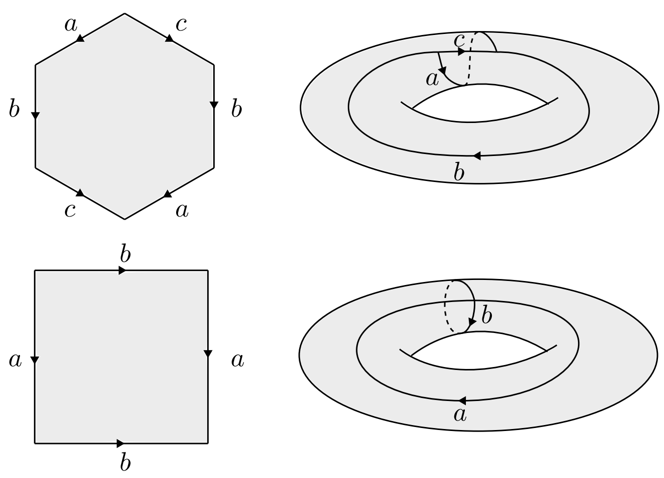


Figure 5.3.1: Two polygonal presentations of torus: $abca^{-1}b^{-1}c^{-1}$ and $aba^{-1}b^{-1}$.

Two labeled polygons are said to be ***equivalent*** if the surfaces induced by them are homeomorphic to each other. We may consider a labeled polygon as a result of "cutting" a surface along a graph in it whose edges are labeled. Hence if (P_1, L_1) and (P_2, L_2) are equivalent, then we can consider first glue P_1 with respect to L_1 and obtain a surface Σ, then cut along a labeled graph in Σ to get (P_2, L_2). Hence intuitively, there should be a way to relate equivalent labeled polygons through cutting and gluing, which will be precised in the following.

We now introduce geometrically the ***elementary operation*** which can be applied to a labeled polygon (P, L):

1) Cut and Glue This is done in 3 steps:

(i) Add a diagonal to P. Associate to it a letter x different from all letters used in L and a orientation.

(ii) Cut P along the diagonal x. Rigorously speaking, consider the closure of each connected component of the complement of this edge in P, and we obtain two polygons Q_1 and Q_2. The label on P and the letter with the orientation on the diagonal induces a label on Q_1 and a label on Q_2. Hence we have a disjoint union $(Q_1, L_1) \sqcup (Q_2, L_2)$.

(iii) Assume that there is a paired edges e_i and e_j of P which are not both in Q_1 or in Q_2. Let φ be a orientation preserving homeomorphism between them. Then we consider

$$P' := Q_1 \sqcup Q_2/(p \sim \varphi(p), \forall\, p \in e_i)$$

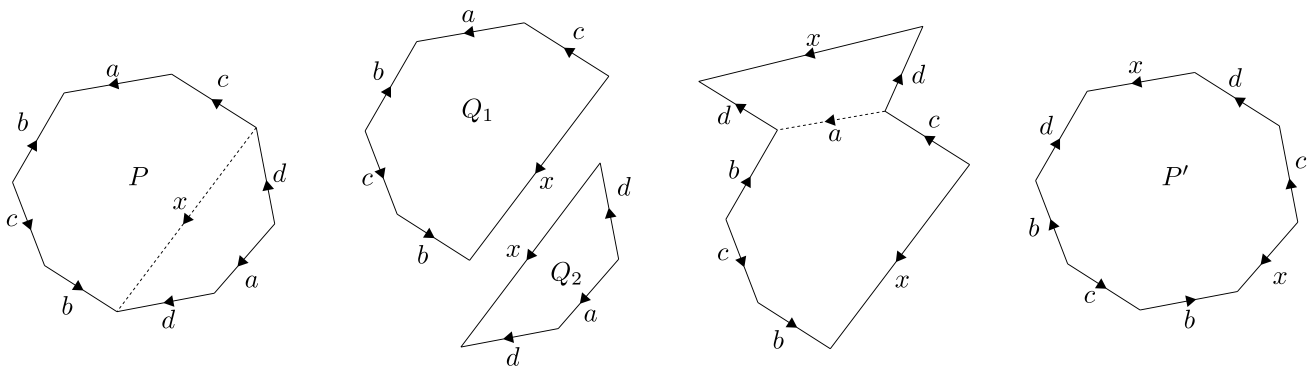


Figure 5.3.2: Cut and glue.

2) Relabel Replace all copies of a letter by a letter which does not appear anywhere in the label L.

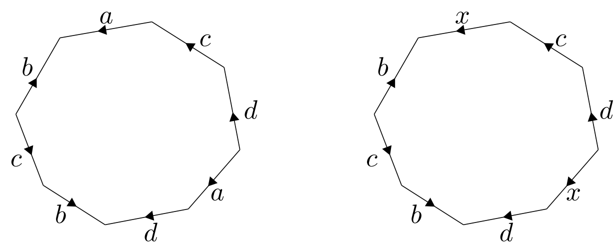


Figure 5.3.3: Relabel letter a by x.

3) Flip Reverse the orientation on all edges of P at the same time.

4) Cancel If there are two successive edge which labeled by a same letter with different orientation, we can glue them to get a polygon P' with two edges less. The label L on the other edges induces a label L' on P'.

5) Cut a slit Let α be an oriented segment in P with one end point at a vertex and the other one in $\mathring{P}$, then we cut P along α. One may consider the complement of α in $\mathring{P}$. It is homeomorphic to $\mathring{D}$ the open disk. Then this map can be extends to a map from the closed disk D to P. The preimages of vertices of P and the end points of α gives marked point on ∂D,

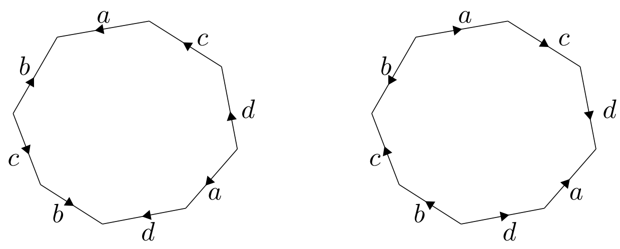


Figure 5.3.4: Flip.

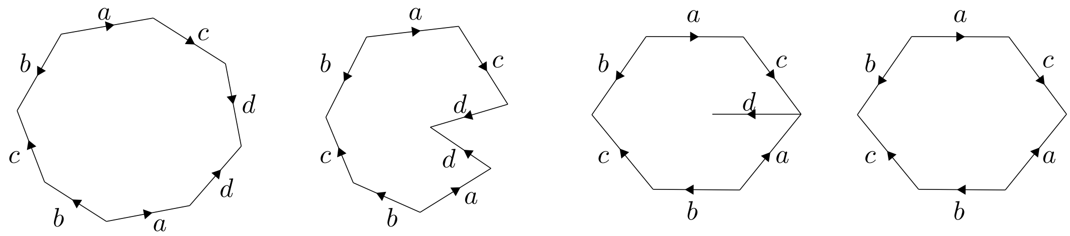


Figure 5.3.5: Cancel.

which make D a P' polygon with 2 more sides than P. We associate to α a letter x different from all letters appearing in L. Then the label L and the letter with the orientation associated to α induce a label on P'.

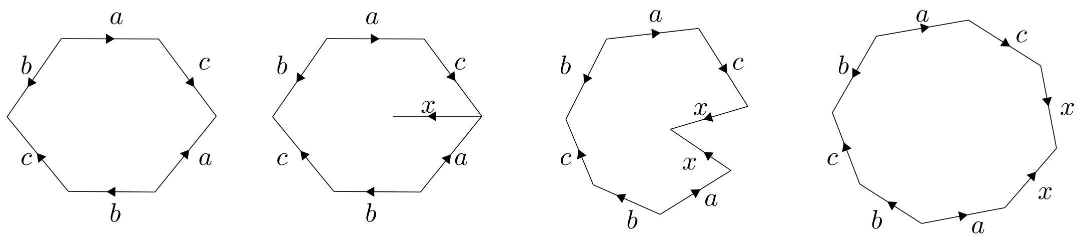


Figure 5.3.6: Cut a slit.

Now we consider the words associated to a labeled polygon. Notice that if we cyclically permute the word, we may consider this as choosing a different vertex to start writing the word. Let w be the word associated to (P, L). Given any word subword of w

$$[y] = a_1^{\epsilon_1} \cdots a_k^{\epsilon_k},$$

we will denote by y^{-1} the following word

$$[y^{-1}] = a_k^{-\epsilon_k} \cdots a_1^{-\epsilon_1}.$$

The elementary operation can be described as follows.

1) Cut and glue This is done in 3 steps:

(i) Add a diagonal to P. Let u and $\widetilde{u}$ be the two vertex of this diagonal. Since each vertex is an end point of an edge, we consider the counterclockwise direction on ∂P. Let e and e' be the two edges with u and u' as an end point respectively. Assume that the label of e is in the end of a subword $[y_1]$ of w and the label of e' is in the end of a subword $[y_2]$ of w, such that we have

$$w = [y_1][y_2][y_3].$$

We apply the permutation to get a new word w'

$$w' = [y_3][y_1][y_2]$$

(ii) Associated to this diagonal a letter x and an orientation, and cut P along the diagonal. Here x does not appear in L. Up to reverse the chosen orientation on the diagonal, we get two words

$$w_1 = [y_3][y_1]x, \quad w_2 = x^{-1}[y_2].$$

They correspond to two labeled polygon (Q_1, L_1) and (Q_2, L_2).

(iii) Assume that there is a letter b different from x, such that b^ϵ and $b^{\epsilon'}$ appearing w_1 and w_2 respectively. If $\epsilon = -\epsilon'$, up to a cyclic permutation on w_1 and w_2 respectively, we have

$$w_1' = [z_1]x[z_2]b^\epsilon, \quad w_2' = b^{-\epsilon'}[z_3]x^{-1}[z_4].$$

We take the concatenation and get

$$w_3 = [z_1]x[z_2][z_3]x^{-1}[z_4],$$

which is the word associated to the labeled polygon (P', L').

If $\epsilon = \epsilon'$, up to a cyclic permutation on w_1 and w_2 respectively, we have

$$w_1' = [z_1]x[z_2]b^\epsilon, \quad w_2' = [z_3]x^{-1}[z_4]b^\epsilon.$$

We apply the flip (see 3) for more details) on (Q_2, L_2) to get a new labeled polygon (Q_2', L_2') corresponding to the word

$$w_2'' = b^{-\epsilon}[z_4^{-1}]x[z_3^{-1}].$$

The take the concatenation and we get

$$w_3 = [z_1]x[z_2][z_4^{-1}]x[z_3^{-1}],$$

which is the word associated to the labeled polygon (P', L').

2) Relabel Replace all copies of a letter by a letter which does not appear anywhere in the label L. Assume that a is a letter appear in w:

$$w = [y_1]a^\epsilon[y_2]a^{\epsilon'}[y_3].$$

Then we take a letter x different from all letters appearing in w, and replace a by x

$$w' = [y_1]x^\epsilon[y_2]x^{\epsilon'}[y_3]$$

3) Flip Reverse the orientation on all edges of P at the same time. If

$$w = a_1^{\epsilon_1} \cdots a_{2n}^{\epsilon_{2n}},$$

then we flip it and get

$$w^{-1} = a_{2n}^{-\epsilon_{2n}} \cdots a_1^{-\epsilon_1}.$$

4) Cancel If there is a letter a such that

$$w = [y_1]a^{\epsilon}a^{-\epsilon}[y_2],$$

we cancel it and get

$$w' = [y_1][y_2].$$

5) Cut a slit Let x be a letter different from all letters appearing in w. Insert xx^{-1} in to w between two successive letters in it and change it from

$$w = [y_1][y_2].$$

to

$$w' = [y_1]xx^{-1}[y_2].$$

An immediate observation is the following one.

Proposition 5.3.1

Two labeled polygons different by a sequence of elementary operations are equivalent.

Let (P, L) be a labeled polygon. We apply an elementary operation on it and obtain (P', L'). Let Σ and Σ' be the surfaces associated to them respectively. From the geometrical description of each elementary operation, one can construct the homeomorphism between Σ and Σ' directly. Here we omit the details.

A less obvious fact is the following one.

Proposition 5.3.2

Any two equivalent labeled polygons can be transform from one to the other by a sequence of elementary operations.

In the following, we are going to prove s stronger statement and show that all labeled polygons can be transform to a standard one by a sequence of elementary operation. This can moreover be used to give a classification of compact closed surfaces. We use words to represent labeled polygons.

Theorem 5.3.3

All labeled polygons are equivalent to one of the following ones:

1) $aa^{-1}bb^{-1}$

2) $abab$;

3) $aba^{-1}b^{-1}$;

4) $a_1a_1a_2a_2 \cdots a_na_n, \quad n \geq 2$;

5) $a_1b_1a_1^{-1}b_1^{-1} \cdots a_nb_na_n^{-1}b_n^{-1}, \quad n \geq 2$.

Remark 5.3.4.
The word of type 1) is associated to the sphere. The word of type 2) is associated to the projective

plane. The gluing pattern induced by the label is in fact the same as the one given by a antipodal map. The word of type 3) is associated to a torus.

A word of type 4) or type 5) is a concatenation of several copies of type 2) word or type 3) word. We will call a word of type 4) a word of the ***projective type*** and a word of type 5) a word of the ***torus type***.

Let w be the word associated to a labeled polygon (P, L). This theorem can be proved by showing the following lemmas.

Lemma 5.3.5

If there is a letter a, such that

$$w = [y_1]a[y_2]a[y_3],$$

then it is equivalent to

$$w' = aa[y_1][y_2^{-1}][y_3].$$

Lemma 5.3.6

If there is a letter a, such that

$$w = [y_1]a[y_2]a[y_3],$$

then it is equivalent to

$$w' = aa[y_1][y_2^{-1}][y_3].$$

Proof. We consider the following cut and glue process (See Figure 5.3.7 for an illustration).

$$\begin{array}{ll} [y_1]a[y_2]a[y_3] & \\ \to[y_1]ab, \quad b^{-1}[y_2]a[y_3] & \text{cut} \\ \to b[y_1]a, \quad a^{-1}[y_2^{-1}]b[y_3^{-1}] & \text{cyclically permute and flip} \\ \to b[y_1][y_2^{-1}]b[y_3^{-1}] & \text{glue} \\ \to[y_1][y_2^{-1}]bc, \quad c^{-1}[y_3^{-1}]b & \text{cyclically permute and cut} \\ \to c[y_1][y_2^{-1}]b, \quad b^{-1}[y_3]c & \text{cyclically permute and flip} \\ \to cc[y_1][y_2^{-1}][y_3] & \text{glue and cyclically permute} \end{array}$$

□

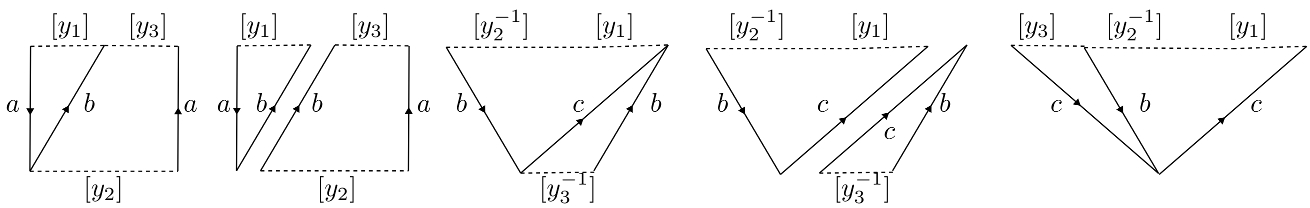


Figure 5.3.7

Corollary 5.3.7

The word w is equivalent to

$$a_1 a_1 \cdots a_k a_k w',$$

such that if a appears in w', so is a^{-1}.

Proof. By the previous lemma, if for some letter a_1, we have $a_1^{\epsilon_1}$ appears twice in w with $\epsilon_1 \in \{1, -1\}$, then w is equivalent to $a_1^{\epsilon_1} a_1^{\epsilon_1} w_1$. Then if for some letter a_2, we have $a_2^{\epsilon_2}$ appears twice in w_1 with $\epsilon_2 \in \{1, -1\}$, then w is equivalent to $a_1^{\epsilon_1} a_1^{\epsilon_1} a_2^{\epsilon_2} a_2^{\epsilon_2} w_2$. We repeat this process, and get $w_1, w_2, ...$ in sequence.

Notice that there are finitely many letters in w, and for each i, there are two letters less in w_{i+1} than in w_i. Hence the above process will stop in finitely many steps. Assume that it stops at step k, then w is equivalent to

$$a_1^{\epsilon_1} a_1^{\epsilon_1} \cdots a_k^{\epsilon_k} a_k^{\epsilon_k} w_{k+1}$$

Up to replace $a_i^{\epsilon_i}$ by $a_i^{-\epsilon_i}$, and denote $w' = w_{k+1}$, we have w equivalent to

$$a_1 a_1 \cdots a_k a_k w'.$$

If w' is not empty, then if a appears in w', so is a^{-1}. □

Lemma 5.3.8

If there are two distinct letters a and b, such that

$$w = [y_1]a[y_2]b[y_3]a^{-1}[y_4]b^{-1}[y_5],$$

then it is equivalent to

$$w' = aba^{-1}b^{-1}[z].$$

for some word z.

Proof. We consider the following cut and glue process (See Figure 5.3.8 for an illustration).

$[y_1]a[y_2]b[y_3]a^{-1}[y_4]b^{-1}[y_5]$	
$\to [y_1]a[y_2]c, \quad c^{-1}b[y_3]a^{-1}[y_4]b^{-1}[y_5]$	cut
$\to [y_2]c[y_1]a, \quad a^{-1}[y_4]b^{-1}[y_5]c^{-1}b[y_3]$	cyclically permute
$\to [y_2]c[y_1][y_4]b^{-1}[y_5]c^{-1}b[y_3]$	glue
$\to [y_1][y_4]b^{-1}[y_5]c^{-1}b[y_3][y_2]c$	cyclically permute
$\to [y_1][y_4]b^{-1}[y_5]d, \quad d^{-1}c^{-1}b[y_3][y_2]c$	cut
$\to [y_5]d[y_1][y_4]b^{-1}, \quad b[y_3][y_2]cd^{-1}c^{-1}$	cyclically permute
$\to [y_5]d[y_1][y_4][y_3][y_2]cd^{-1}c^{-1}$	glue
$\to [y_1][y_4][y_3][y_2]cd^{-1}c^{-1}[y_5]d$	cyclically permute
$\to [y_1][y_4][y_3][y_2]ce, \quad e^{-1}d^{-1}c^{-1}[y_5]d$	cut
$\to e[y_1][y_4][y_3][y_2]c, \quad c^{-1}[y_5]de^{-1}d^{-1}$	cyclically permute
$\to de^{-1}d^{-1}e[y_1][y_4][y_3][y_2][y_5]$	glue and cyclically permute
$\to aba^{-1}b^{-1}[y_1][y_4][y_3][y_2][y_5]$	relabel

□

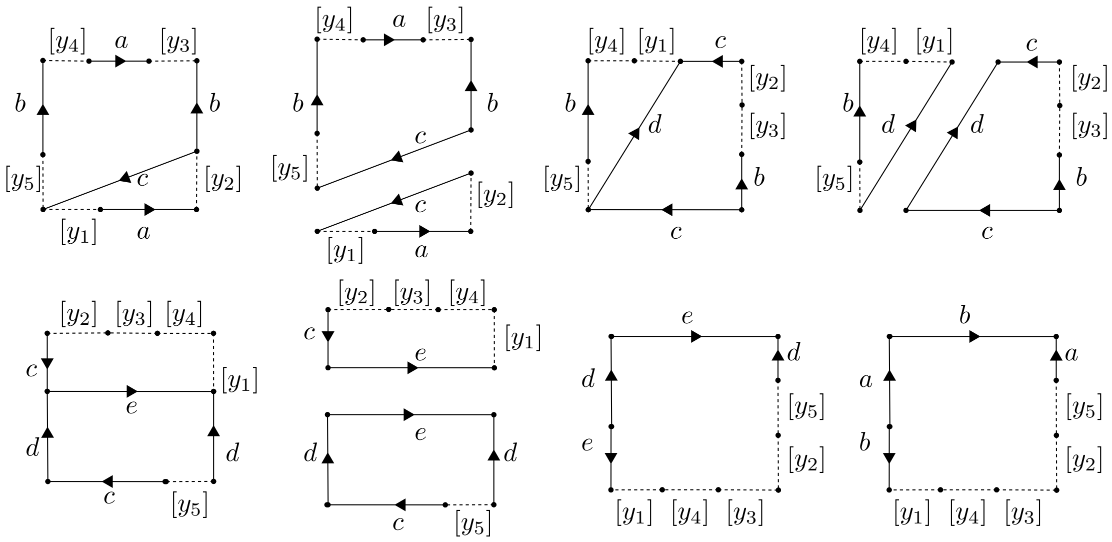


Figure 5.3.8

Corollary 5.3.9

The word w is equivalent to

$$a_1a_1\cdots a_ka_kc_1d_1c_1^{-1}d_1^{-1}\cdots c_nd_nc_n^{-1}d_n^{-1}.$$

Proof. From the above two lemmas, the word w is equivalent to the following one:

$$a_1a_1\cdots a_ka_kc_1d_1c_1^{-1}d_1^{-1}\cdots c_nd_nc_n^{-1}d_n^{-1}[y],$$

where $[y]$ is a subword such that for any letter h, both h and h^{-1} appear in $[y]$, and moreover given any two letters g and h, either there is no letter g between h and h^{-1} or no h between g and g^{-1}.

Without loss of generality, we may assume that

$$[y]=[z_1]h[z_2]h^{-1}[z_3].$$

Assume that $[z_2]$ starts with h_1^{ϵ}, then we have

$$[y]=[z_1]hh_1^{\epsilon}[z_4]h_1^{-\epsilon}[z_5]h^{-1}[z_3].$$

We repeat this process, since there are only finitely many letters in $[z_4]$, after finite steps, there will be a letter g, such that g and g^{-1} are adjacent to each other:

$$[y]=[z_1]hh_1^{\epsilon}\cdots g^{\delta}g^{-\delta}\cdots h_1^{-\epsilon}[z_5]h^{-1}[z_3].$$

We may cancel $g^{\delta}g^{-\delta}$. Then we start over again the discussion. By out assumption, all letters in $[y]$ will be canceled out. Hence

$$a_1a_1\cdots a_ka_kc_1d_1c_1^{-1}d_1^{-1}\cdots c_nd_nc_n^{-1}d_n^{-1}[y]\sim a_1a_1\cdots a_ka_kc_1d_1c_1^{-1}d_1^{-1}\cdots c_nd_nc_n^{-1}d_n^{-1}.$$

□

Lemma 5.3.10

Let a, b, c be three distinct letters. The labeled polygon associated to the word $aabbcc$ is equivalent to the one associated to $aabcb^{-1}c^{-1}$.

Proof. We consider the following cut and glue process (See Figure 5.3.9 for an illustration).

$aabbcc$	
$\to abbcca$	cyclically permute
$\to abd, \quad d^{-1}bcca$	cut
$\to a^{-1}d^{-1}b^{-1}, \quad bccad^{-1}$	flip and cyclically permute
$\to a^{-1}d^{-1}ccad^{-1}$	glue
$\to d^{-1}ccad^{-1}a^{-1}$	cyclically permute
$\to cad^{-1}a^{-1}e, \quad e^{-1}d^{-1}c$	cut
$\to a^{-1}ecad^{-1}, \quad dec^{-1}$	flip and cyclically permute
$\to a^{-1}ecaec^{-1}$	glue
$\to ecaec^{-1}a^{-1}$	cyclically permute
$\to ecaef, \quad f^{-1}c^{-1}a^{-1}$	cut
$\to efeca, \quad a^{-1}f^{-1}c^{-1}$	cyclically permute
$\to efecf^{-1}c^{-1}$	glue
$\to f^{-1}c^{-1}efec$	cyclically permute
$\to f^{-1}c^{-1}efg, \quad g^{-1}ec$	cut
$\to efgf^{-1}c^{-1}, \quad cg^{-1}e$	cyclically permute
$\to efgf^{-1}g^{-1}e$	glue
$\to aabcb^{-1}c^{-1}$	cyclically permute and relabel

□

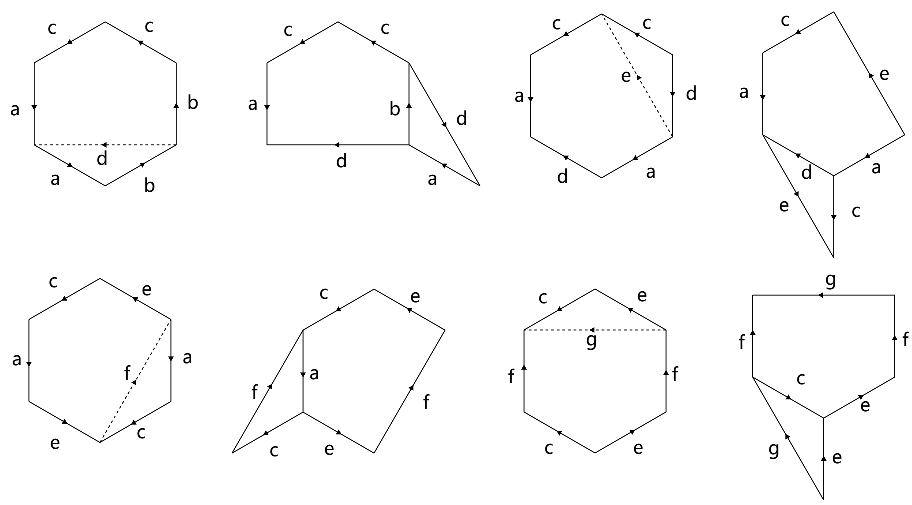


Figure 5.3.9

Hence given any word

$$a_1a_1\cdots a_ka_kc_1d_1c_1^{-1}d_1^{-1}\cdots c_nd_nc_n^{-1}d_n^{-1},$$

if $k=0$ or $n=0$, then we have the desired form

$$c_1d_1c_1^{-1}d_1^{-1}\cdots c_nd_nc_n^{-1}d_n^{-1}$$

or

$$a_1a_1\cdots a_ka_k.$$

If both k and n are not 0, then by induction the above lemma shows that it is equivalent to

$$a_1a_1\cdots a_{k+2n}a_{k+2n}.$$

Hence Theorem 5.3.3 is proved.

Connected sums and labeled polygons

Now we would like to see which surface is associated to each standard word, and this is enough to describe all closed compact surfaces by Theorem 5.3.3. We will discuss the case for a standard word w of the torus type, and the case for words of the projective type can be treated in a similar way.

Given a polygon P whose label is given by a standard word

$$w=a_1b_1a_1^{-1}b_1^{-1}\cdots a_nb_na_n^{-1}b_n^{-1},\quad n\geq 2,$$

we denote by Σ_n the resulting surface. The first observation is that all vertices of P are in one vertex cycle. We still use π to denote the projection map. Then the π-image of each edge or diagonal of P induces a loop in Σ.

We cut P along a labeled diagonal as done previously, such that we obtain

$$a_1b_1a_1^{-1}b_1^{-1}c,\quad c^{-1}a_2b_2a_2^{-1}b_2^{-1}\cdots a_nb_na_n^{-1}b_n^{-1}.$$

One may consider glue a triangle labeled by $c^{-1}dd^{-1}$ to $a_1b_1a_1^{-1}b_1^{-1}c$ and obtain

$$a_1b_1a_1^{-1}b_1^{-1}dd^{-1},$$

which is equivalent to

$$a_1b_1a_1^{-1}b_1^{-1}.$$

Notice that the surface associated to $c^{-1}dd$ is a disk. In the other words, the surface associated to $a_1b_1a_1^{-1}b_1^{-1}c$ can be considered as being obtained by removing a disk from the torus associated to $a_1b_1a_1^{-1}b_1^{-1}$.

Similarly, the surface associated to $c^{-1}a_2b_2a_2^{-1}b_2^{-1}\cdots a_nb_na_n^{-1}b_n^{-1}$ can be considered as being obtained by removing a disk from the surface assocaited to $c^{-1}a_2b_2a_2^{-1}b_2^{-1}\cdots a_nb_na_n^{-1}b_n^{-1}$.

Hence the surface Σ_n is a connected sum

$$T\#\Sigma_{n-1}.$$

where T denote the torus (See Figure 5.3.10). By induction, for any $n\in\mathbb{N}\setminus\{0,1\}$, we have the homeomorphism

$$\Sigma_n\cong\underbrace{T\#\cdots\#T}_{n}.$$

A similar discussion shows that the surface associated to the word of a projective type

$$a_1a_1a_2a_2\cdots a_na_n,\quad n\geq 2,$$

is homeomorphic to the connected sum

$$\underbrace{\mathbb{RP}^2\#\cdots\#\mathbb{RP}^2}_{n}.$$

Hence as a corollary of Theorem 5.3.3, we have the following statement.

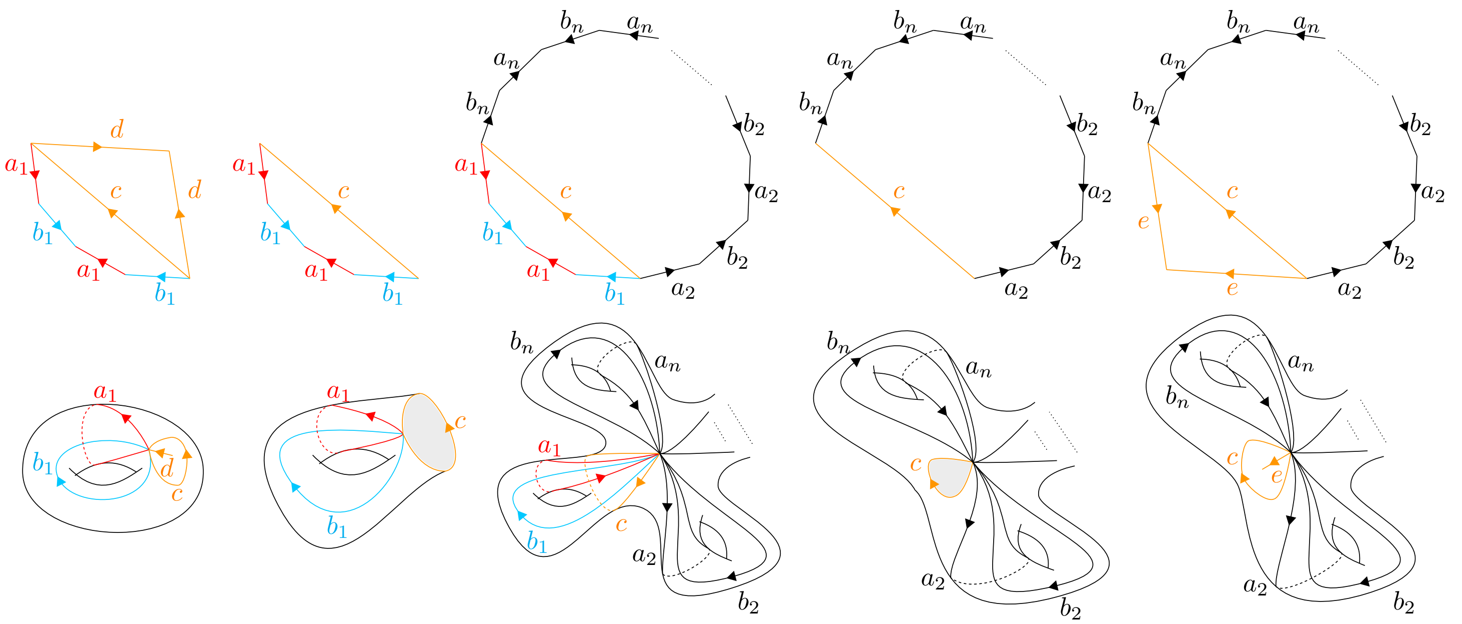


Figure 5.3.10: $\Sigma_n \cong T\#\Sigma_{n-1}$.

Theorem 5.3.11

Any closed compact surface Σ is homeomorphic to one of the following surfaces:

1) S^2;

2) $\mathbb{RP}^2$;

3) T;

4) $\underbrace{\mathbb{RP}^2\#\cdots\#\mathbb{RP}^2}_{n}$, $n \geq 2$;

5) $\underbrace{T\#\cdots\#T}_{n}$, $n \geq 2$.

Classification theorem

Theorem 5.3.11 shows that the list in the statement contains all closed compact surfaces up to homeomorphism. The question left is if two surfaces in the list could be homeomorphic to each other. To answer this question, we will use fundamental groups of surfaces and their abelianizations.

Fundamental groups of surfaces

Theorem 5.3.3 also provides the information of fundamental groups of the surfaces associated to the labeled polygons of each type. The key tool used in this discussion is of course the Seifert-Van-Kampen Theorem.

Let P be a polygon labeled by one of the words w in the list in Theorem 5.3.3. Notice that P

is topologically a disk. Denote D a closed disk contained in $\mathring{P}$. Let $p \in \mathring{D}$, then we denote

$$U = P \setminus \{p\} \quad \text{and} \quad V = \mathring{D}.$$

We still denote by Σ the surface associated to P and by π the projection from P to Σ. Notice that U is homotopy equivalent to ∂P, whose image under π is a graph. Moreover each edge is sent to a loop by π. Hence the image of ∂P under π is a rose whose fundamental group is isomorphic to a free group generated by letters used in the label of P (See Figure 5.3.11 for an illustration for computing $\pi_1(\pi(U))$ for a torus type labeled polygon).

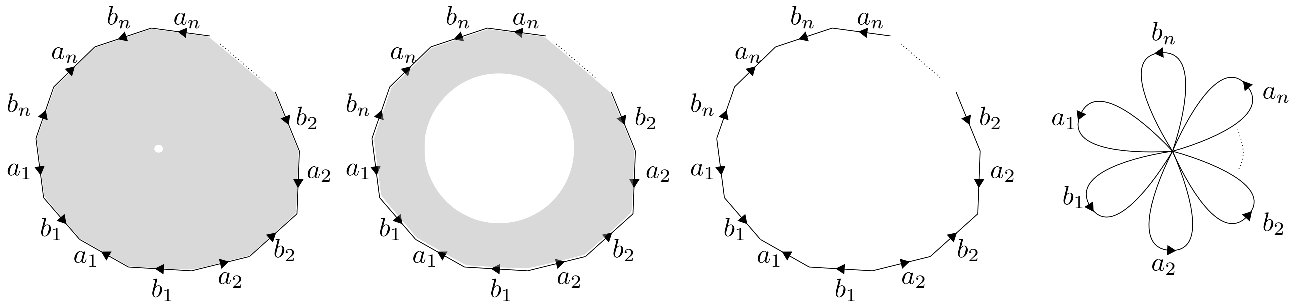


Figure 5.3.11

Using SVK theorem, the fundamental group of Σ is given by a free group quotient by the subgroup normally generated by the word associated to the boundary. As a result, we have the following presentations of fundamental groups for surfaces in the above list:

1) $w = aa^{-1}bb^{-1}$ and $\Sigma \cong S^2$:

$$\pi_1(\Sigma) \cong 0.$$

2) $w = abab$ and $\Sigma \cong \mathbb{RP}^2$: (let $c = ab$)

$$\pi_1(\Sigma) \cong \langle c \mid c^2 \rangle \cong \mathbb{Z}_2.$$

3) $w = aa^{-1}bb^{-1}$ and $\Sigma \cong T$;

$$\pi_1(\Sigma) \cong \langle a, b \mid aa^{-1}bb^{-1} \rangle$$

4) $w = a_1a_1a_2a_2 \cdots a_na_n$ and $\Sigma = \underbrace{\mathbb{RP}^2 \# \cdots \# \mathbb{RP}^2}_{n}$ $(n \geq 2)$:

$$\pi_1(\Sigma) \cong \langle a_1, ..., a_n \mid a_1^2 \cdots a_n^2 \rangle.$$

5) $w = a_1b_1a_1^{-1}b_1^{-1} \cdots a_nb_na_n^{-1}b_n^{-1}$ and $\Sigma = \underbrace{T \# \cdots \# T}_{n}$ $(n \geq 2)$:

$$\pi_1(\Sigma) \cong \langle a_1, b_1, ..., a_n, b_n \mid [a_1, b_1] \cdots [a_n, b_n] \rangle.$$

First homology groups

Since fundamental groups are topological invariants, if two surfaces have non-isomorphic fundamental groups, then the two surfaces are not homeomorphic. However presentations of a group are not unique. To overcome this difficulty, we consider their abelianizations and discuss in the category of abelian groups. Notice that all fundamental groups are finitely generated, hence their abelianizations are finitely generated abelian groups, for which we have the classification theorem.

Definition 5.3.12

Let G be a group. Its ***abelianization*** is defined to be

$$G^{\mathrm{ab}} := G/[G, G].$$

Let F_n be a free group of n letters. Let N_1 and N_2 be two normal subgroup of F_n. Then by the fundamental theorem of group homomorphism, we have

$$(F/N_1)/(N_1N_2/N_1) \cong F_n/N_1N_2.$$

Therefore, if a group G has the following presentation

$$\langle a_1, ..., a_k \mid w_1, ..., w_l\rangle,$$

then G^{ab} has the presentation

$$\langle a_1, ..., a_k \mid w_1, ..., w_l, [a_1, a_2], ..., [a_{k-1}, a_k]\rangle,$$

By this discussion, we have the abelianizations of fundamental groups of surfaces

1) $w = aa^{-1}bb^{-1}$ and $\Sigma \cong S^2$:
$$\pi_1(\Sigma)^{\mathrm{ab}} \cong 0.$$
2) $w = abab$ and $\Sigma \cong \mathbb{RP}^2$:
$$\pi_1(\Sigma)^{\mathrm{ab}} \cong \mathbb{Z}_2.$$
3) $w = aba^{-1}b^{-1}$ and $\Sigma \cong T$;
$$\pi_1(\Sigma)^{\mathrm{ab}} \cong \mathbb{Z}^2$$
4) $w = a_1a_1a_2a_2 \cdots a_na_n$ and $\Sigma = \underbrace{\mathbb{RP}^2\# \cdots \#\mathbb{RP}^2}_{n}$ $(n \geq 2)$:
$$\pi_1(\Sigma)^{\mathrm{ab}} \cong \mathbb{Z}^{n-1} \oplus \mathbb{Z}_2.$$
5) $w = a_1b_1a_1^{-1}b_1^{-1} \cdots a_nb_na_n^{-1}b_n^{-1}$ and $\Sigma = \underbrace{T\# \cdots \#T}_{n}$ $(n \geq 2)$:
$$\pi_1(\Sigma)^{\mathrm{ab}} \cong \mathbb{Z}^{2n}.$$

Now we consider the classification of finitely generated abelian groups and have the following classification of surfaces:

Theorem 5.3.13

Any closed compact surface Σ is homeomorphic to exact one of the following surfaces:

1) S^2;
2) $\mathbb{RP}^2$;
3) T;
4) $\underbrace{\mathbb{RP}^2\# \cdots \#\mathbb{RP}^2}_{n}$, $n \geq 2$;
5) $\underbrace{T\# \cdots \#T}_{n}$, $n \geq 2$.

Remark 5.3.14.
After introducing the homology groups for a topological space, we will come back to the abelianization of fundamental group of a surface and moreover for a path connected space. We will show that the abelianization of the fundamental group is isomorphic to the first homology group.

5.4 Euler characteristic

As introduced in the introduction, Euler characteristic is a topological invariant for surfaces. In this part, we will analysis it more closely still in a combinatorial way. A generalization of this topological invariant for any CW-complex will be discussed later.

Triangulations and cellulations of a surface

The Euler characteristic of a surface can be computed by counting number of vertices, edges and faces in a triangulation of the surface.

As suggested by its name, a triangulation of a surface is a decomposition of the surface into triangles. More precisely, let $T = (V, E)$ be a graph embedded in Σ, where V is the set of vertices and E is the set of edges. If the closure of each connected component of the complement of T is a triangle, then we call T a ***triangulation*** of the surface Σ. Notice that each connected component Δ of $\Sigma \setminus T$ is homeomorphic to an open disk $\mathring{D}$. We denote by

$$f : \mathring{D} \to \Delta.$$

Then the map f can be extends to a continuous map

$$\overline{f} : D \to \overline{\Delta}.$$

Hence $\overline{f}(\partial D) \subset T$. By a triangle, we mean that $\overline{f}^{-1}(\partial\Delta \cap V)$ has three points. Each connected component of the complement of T is called a ***face*** for T. We denote by F the collection of all faces for T.

Another way to describe a triangulation is first marking a finite collection V of points in Σ as vertices, then connect them by simple paths, i.e. paths such that the restriction to $(0, 1)$ is injective. Two paths are said to be ***disjoint*** if the images of their restriction to $(0, 1)$ are disjoint. Then a triangulation of Σ is a maximal collection of simple paths connecting points in V which are pairwise non homotopic and pairwise disjoint. For example, given a n-gon with $n > 3$, we can add diagonals to get a triangulation of the n-gon. This can be done in finite time.

Let T and T' be two triangulations of Σ. If up to homotopy, there is an embedding of T into T', then we call T' a refinement of T. In this case, all faces and edges of T' come from taking a subdivision of those of T.

If we require only the interior closure of each connected component of the complement of T is simply a polygon, then we obtain a ***cellulation***. A refinement of a cellulation can be defined in a similar fashion.

Euler characteristic

Definition 5.4.1

The Euler characteristic $\chi(\Sigma)$ of a surface Σ is defined to be the following quantity:

$$\chi(\Sigma) = \#V - \#E + \#F,$$

where V, E and F are sets of vertices, edges and faces of a triangulation of Σ.

Proposition 5.4.2

The Euler characteristic of a surface Σ is independent of choice of the triangulation.

There are two steps that should be completed to prove this proposition:

1) Euler characteristic is invariant under refinements;

2) Up to homotopy, any two triangulations of Σ have a common refinement.

To see the point 2), we use the fact that the surface is compact, and locally homeomorphic to $\mathbb{R}^2$, hence it has a finite cover consisting only closed subsets homeomorphic to closed disk in $\mathbb{R}^2$. Moreover, since all edges are compact, we may assume that each edge is cut into finitely many segments each of which is contained in one subset in the covering. By identifying each disk set with an Euclidean disk, we can pulling tight each segment, and get new triangulation T_0 and T_0', such that T_0 and T_0' intersect at finitely many points. Now we consider all these intersection points, vertex in T_0 and those in T_0' and add edges if necessary to get a common refinement for T_0 and T_0' at the same time.

For the point 1), notice that a refinement can be obtained by adding one vertex a time. If we add one vertex on an edge, then we have one more vertex, two more edges and one more face. If we add one vertex in the interior of a face, then we have one more edge, three more edges and two more face. As a conclusion, the Euler characteristic is invariant.

To compute the Euler characteristic, we use the formula

$$\chi(\Sigma\#\Sigma') = \chi(\Sigma) + \chi(\Sigma') - 2.$$

Then by the classification of closed compact surfaces, we have

$$\begin{aligned}
&\chi(S^2) = 2;\\
&\chi(\mathbb{RP}^2) = 1;\\
&\chi(K) = 0;\\
&\chi(T) = 0;\\
&\chi\left(\underbrace{\mathbb{RP}^2\#\cdots\#\mathbb{RP}^2}_{n}\right) = 2-n;\\
&\chi\left(\underbrace{T\#\cdots\#T}_{n}\right) = 2-2n.
\end{aligned}$$

5.5 Orientation

The orientation is another object that we consider when studying manifolds. In $\mathbb{R}^2$, there are precisely two orientations. We consider a frame in $\mathbb{R}^2$ which is formed by two vectors. Consider the matrix formed by these vectors. The sign of the determinant is called an ***orientation*** of $\mathbb{R}^2$. Given an triangle, we can talk about its orientation.

For a general surface, we will use triangulation to study its orientability. Consider a collection of triangles glued together to get a surface Σ. We start by choosing an orientation on one triangle, then we may try to extend it to the entire Σ. However, this is not always doable. For example, one may consider the Möbius band.

To make this more precise, we give the following definitions. Let T be a triangulation of Σ. To each face f, we associate to it an orientation O_f. Let $\mathcal{O}$ be the collection of all O_f's.

Definition 5.5.1

The collection $\mathcal{O}$ is ***coherent*** if the two orientation on each edge induced by its adjacent triangles are inverse to each other. In this case, we say that the triangulation is ***orientable***.

Definition 5.5.2

A surface is ***orientable*** if it admits an orientable triangulation.

Proposition 5.5.3

If a surface admits an orientable triangulation, then all its triangulations are orientable.

The idea of the proof is the same as the one above for showing that the Euler characteristic is independent of choice of the triangulation. We should show that if a triangulation is orientable, so are all its refinements. This can be shown by observing that any refinement of T is obtained by taking subdivision of T.

Given any two triangulations T and T' of Σ, up to homotopy, they admits a common refinement T''. Let $\mathcal{O}$ be collections of orientations chosen for faces of T. Denote by $\mathcal{O}_0$ be the collection of orientations on faces of T'' induced by $\mathcal{O}$. If $\mathcal{O}$ is coherent, so is $\mathcal{O}_0$. Since a face f' of T' is subdivided into faces of T''. The coherent collection $\mathcal{O}_0$ induces a collection $\mathcal{O}'$ of orientations of faces of T' which is coherent. Hence T' is also orientable.

Chapter 6

Simplicial and singular homology

In this chapter, we will give an elementary introduction to the homology theory for a general topological space.

6.1 Rough idea of homology after Poincaré

Recall that a space is path connected if any two points can be connected by a path. Or in the other words, any two points are boundary of some 1-submanifold, i.e. a subset of the space which is a 1-manifold. For a general topological space X, we can try to use this as an equivalence relation and the number of corresponding equivalence classes is exactly the number of path connected components of X.

Topologically there are many different topological spaces which are path connected. In order to have a finer classification, we have to considering a somewhat higher level "connectivity". In the world of manifolds, there is a natural way to do this which is called "bordism". This is a generalization of the above observation. More precisely, let M be a n-manifold. We consider 1-submanifolds of M. If α and β are two 1-submanifolds, which form a boundary of a 2-submanifold of M, we say that they are equivalent. Then we consider the spaces of equivalence classes of 1-submanifolds of M, which could give more information about the connectivity of M. We can continue to study k-submanifolds of M, and define that two k-submanifolds are equivalent if and only if they form the boundary of a $(k+1)$-submanifold of M.

One observation made by Poincaré is that "the boundary of boundary of a manifold is empty." (See Figure 6.1.1 for an illustration.)

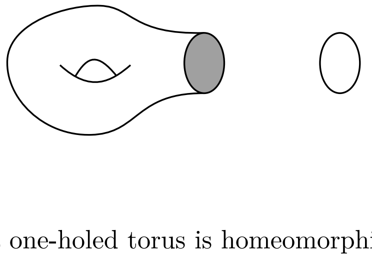


Figure 6.1.1: The boundary of a one-holed torus is homeomorphic to S^1 which has no boundary.

Instead of considering all k-submanifolds, we consider only those equivalence classes of closed ones. In this way, we obtain a set which is topologically invariant by considering its definition. This is the rough idea of the homology in the beginning.

We are going to introduce three homology for a space: simplicial homology, singular homology and cellular homology. For a space where all of them are well defined, they are algebraically the same.

6.2 Simplicial homology

As suggested by its name, we consider a space which admits a decomposition into a collection of "triangles" of different dimensions.

Euclidean simplices

For any $n \in \mathbb{N}$, a "triangle" of dimension n is usually called an *n-simplex.* We call the following subset of $\mathbb{R}^{n+1}$ the *standard n-simplex*

$$\Delta^n = \{(x_0, ..., x_n) \in \mathbb{R}^{n+1} \mid x_0 + \cdots + x_n = 1,\ x_0 \geq 0,\ ...,\ x_n \geq 0\}.$$

(See Figure 6.2.1 for an illustration.)

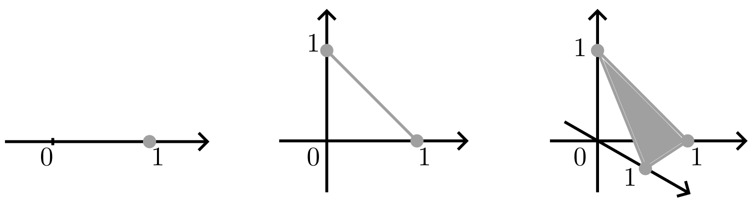


Figure 6.2.1: From left to right: Δ^0, Δ^1 and Δ^2.

For any $0 \leq i \leq n$, Let $\underline{e}_i$ denote the point of $\mathbb{R}^{n+1}$ with i-th coordinate 1 and all others 0. Then we have

$$\Delta^n = \{x_0\underline{e}_0 + \cdots + x_n\underline{e}_n \in \mathbb{R}^{n+1} \mid x_0 + \cdots + x_n = 1,\ x_0 \geq 0,\ ...,\ x_n \geq 0\}.$$

This construction can be done for any $n+1$ points of $\mathbb{R}^{n+1}$ in general position. When $n = 0$, any point in $\mathbb{R}$ is in general position. When $n > 0$, let $v_0, ..., v_n$ be $n+1$ points in $\mathbb{R}^{n+1}$. We say that they are in general position if the vectors

$$v_1 - v_0, ..., v_n - v_0,$$

are linearly independent. An *n-simplex* in $\mathbb{R}^{n+1}$ determined by $v_0, ..., v_n$ in general position is the following subset

$$\{x_0 v_0 + \cdots + x_n v_n \in \mathbb{R}^{n+1} \mid x_0 + \cdots + x_n = 1,\ x_0 \geq 0,\ ...,\ x_n \geq 0\}.$$

Points $v_0, ..., v_n$ are called *vertices* of this n-simplex. Notice that for any $m \geq n$, we can define n-simplices in $\mathbb{R}^m$. If a simplex Δ_1 determined by all but one vertex of another simplex Δ_2 with positive dimension, we say that Δ_1 is a *face* of Δ_2 (see Figure 6.2.2 for an illustration). For any simplex Δ determined by some vertices of another simplex Δ', we denote $\Delta \leq \Delta'$.

Let $n \geq 1$. Consider the simplex in $\mathbb{R}^m$ determined by vertices $v_0, ..., v_n$. It can be associated with an orientation by first giving an order among its vertices:

$$(v_0, ..., v_n),$$

then consider the sign of the determinant of the matrix

$$[v_1 - v_0 \ \cdots \ v_n - v_0].$$

If the sign is positive, we say that the orientation is *positive*; if the sign is negative, we say the orientation is *negative.* The simplex equipped with this orientation is denoted by $[v_0, ..., v_n]$. The

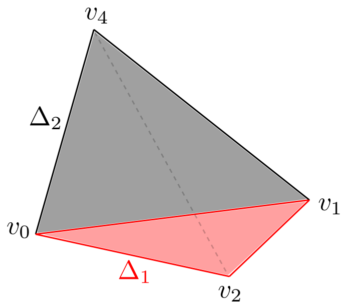


Figure 6.2.2: The simplex Δ_1 determined by v_0, v_1 and v_2 is a face of Δ_2 determined by v_0, v_1, v_2 and v_3.

order among $v_0, ..., v_n$ is not unique and there is a natural action of the permutation group of S_{n+1} on $\{v_0, ..., v_n\}$ by permuting the index $\{0, ..., n\}$. As a convention, we may use $-[v_0, ..., v_n]$ to denote the same simplex with the opposite orientation. For any τ a permutation on $\{0, ..., n\}$, we denote.

$$[v_0, ..., v_n] = \mathrm{sgn}(\tau)[v_{\tau(0)}, ..., v_{\tau(n)}].$$

A *simplicial complex* is a finite collection K of simplices in some Euclidean space $\mathbb{R}^m$, such that

(i) if a simplex $\Delta \in K$, so are all its faces;

(ii) if two simplices Δ and Δ' in K have non empty intersection, then $\Delta \cap \Delta' \leq \Delta$ and $\Delta \cap \Delta' \leq \Delta'$.

The union of simplices in K will be denoted by $\cup K$ and equipped with the subspace topology from $\mathbb{R}^m$.

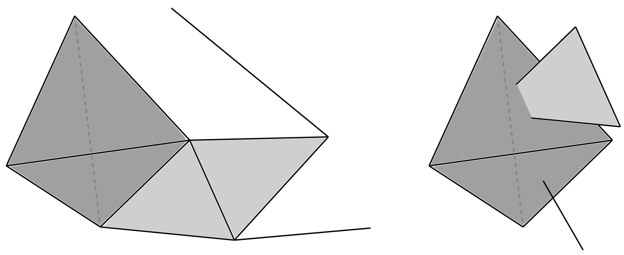


Figure 6.2.3: The left is a simplicial complex formed by one 3-simplex, six 2-simplices, twelve 1-simplices and eight 0-simplices, while the right is not a simplicial complex.

Simplicial chain complex and simplicial homology groups

Let X be a topological space. A *simplicial complex structure* on X is a homeomorphism

$$f : \cup K \to X,$$

where K is a simplicial complex in an Euclidean space $\mathbb{R}^m$ for some $m \in \mathbb{N}^*$ (see Figure 6.2.4 and 6.2.5 for illustrations).

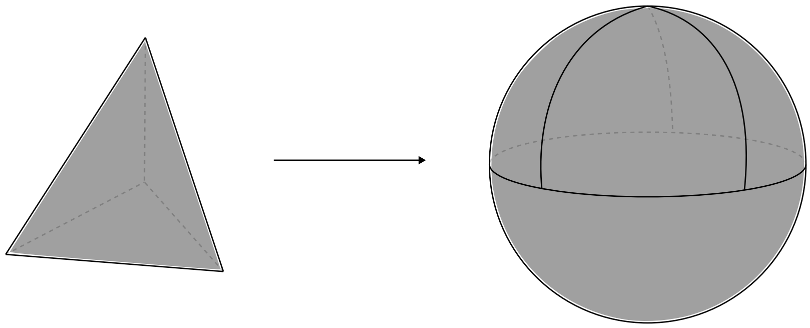


Figure 6.2.4: A simplicial complex structure on a 3 dimensional ball.

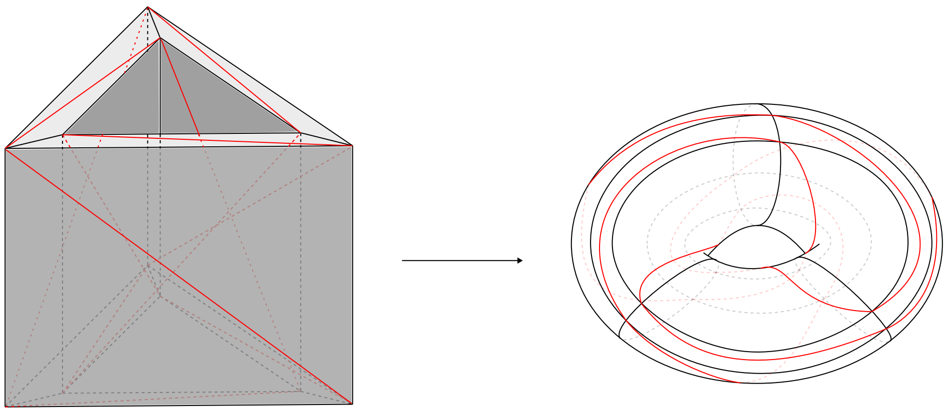


Figure 6.2.5: A simplicial complex structure on a torus.

A simplicial complex structure on X can also be defined to be a collection of compatible continuous maps from standard simplices to X. More precisely, consider the standard n-simplex Δ^n. There are orientation preserving affine maps to identity faces of Δ^n with standard simplices of same dimensions. We choose once for all such identification. A simplicial complex structure is then defined to be a collection of maps indexed by Ω:

$\Lambda = \{\sigma_\alpha : \Delta^{n_\alpha} \to X \mid \alpha \in \Omega,\ \Delta^{n_\alpha}$ is a standard simplex of dimension n_α, and σ_α is continuous$\}$,

such that

1) The restriction of each $\sigma_\alpha \in \Lambda$ to $\mathring{\Delta}^{n_\alpha}$ is injective.

2) Each $p \in X$ belongs to the image of an unique map $\sigma_\alpha|_{\mathring{\Delta}^{n_\alpha}}$.

3) The restriction of each σ_α to a face of Δ^{n_α} is one map

$$\sigma_\beta : \Delta^{n_\beta} \to X$$

in Λ.

4) For any $A \subset X$, the subset A is open if and only if for any $\sigma_\alpha \in \Lambda$, the preimage $\sigma_\alpha^{-1}(A)$ is open in Δ^{n_α}.

Remark 6.2.1.
Roughly speaking, the space X can be considered as a result of gluing of a collection of simplies. Moreover the topology on X is induced by the topology on each simplies.

We call the image of each σ_α an n_α-*simplex* in X, and denote it by e_α. By considering each such simplex e_α as a formal generator, we associated to it a copy of $\mathbb{Z}$ whose elements are denoted as ne_α. In this way, we have the following abelian group

$$C_n^\Delta(X) := \left\{ \sum_{i=0}^{k} m_i e_{\alpha_i} \,\middle|\, k \in \mathbb{N}, m_0, ..., m_k \in \mathbb{Z} \right\}.$$

Each element $\sigma \in C_n^\Delta$ is a finite $\mathbb{Z}$-coefficient formal sum

$$\sum_{i=0}^{k} m_i e_{\alpha_i},$$

which is called an n-*chain*. For $n_\alpha > 0$, the orientation on Δ^{n_α} induces an orientation on e_α by σ_α. The elements $-e_\alpha$ can be considered as the same simplex with opposite orientation. As a convention a 0-chain in X is a finite $\mathbb{Z}$-coefficient formal sum of 0-simplicies in X.

Notice that if $n \geq 2$, by restricting σ_α to the faces of Δ^{n_α}, there is a natural map from an n-chain to an $(n-1)$-chain. More precisely, assume $\Delta^{n_\alpha} = [v_0, ..., v_n]$, for any simplex $e_\alpha = \sigma_\alpha([v_0, ..., v_n])$ in X, its boundary $\partial_n(e_\alpha)$ is given as follows:

$$\partial_n(\sigma_\alpha([v_0, ..., v_n])) = \sum_{i=0}^{n} (-1)^i \sigma_\alpha([v_0, ..., \widehat{v_i}, ..., v_n]).$$

Here $[v_0, ..., \widehat{v_i}, ..., v_n]$ stands for oriented face determined by all vertices but v_i. With the coefficient $(-1)^i$, the orientation associated to each face is induced by that on $[v_0, ..., v_n]$. We can formally define the boundary of a 1-simplex in a same way. For any simplex $e_\alpha = \sigma_\alpha([v_0, v_1])$ in X, its boundary is defined to be

$$\partial_1(\sigma_\alpha([v_0, v_1])) = -\sigma_\alpha([v_0]) + \sigma_\alpha([v_1]).$$

For any $n > 0$, the boundary of an n-chain is then defined as follows:

$$\partial_n \left(\sum_{i=0}^{k} m_i e_{\alpha_i} \right) = \sum_{i=0}^{k} m_i \partial_n(e_{\alpha_i}).$$

A direct computation can show that this induces a group homomorphism

$$\partial_n : C_n^\Delta(X) \to C_{n-1}^\Delta(X)$$

for $n > 0$. When $n = 0$, we set

$$\partial_0 : C_0^\Delta(X) \to 0$$

where 0 stands for the trivial group.

Put all pieces together, we have the following chain of group homomorphisms:

$$\cdots \xrightarrow{\partial_{n+2}} C_{n+1}^\Delta(X) \xrightarrow{\partial_{n+1}} C_n^\Delta(X) \xrightarrow{\partial_n} C_{n-1}^\Delta(X) \xrightarrow{\partial_{n-1}} \cdots \xrightarrow{\partial_2} C_1^\Delta(X) \xrightarrow{\partial_1} C_0^\Delta(X) \xrightarrow{\partial_0} 0$$

Proposition 6.2.2

The composition of two successive boundary homomorphisms is trivial, i.e. for any $n \in \mathbb{N}$, we have

$$\partial_n \circ \partial_{n+1} = \mathbf{0},$$

where $\mathbf{0}$ is the trivial homomorphism.

Proof. For any $n \in \mathbb{N}$, it is enough to show that this holds for any $n+2$-simplex.

Let $\Delta^{n+1} = [v_0, ..., v_{n+1}]$ and $e = \sigma([v_0, ..., v_{n+1}]^o]$ be an open $n+1$-simplex in X. Hence we have

$$\partial_{n+1}(e) = \partial(\sigma([v_0, ..., v_{n+1}])) = \sum_{i=0}^{n+1} (-1)^i \sigma([v_0, ..., \widehat{v_i}, ..., v_{n+1}]).$$

We then compute its image under ∂_n:

$$\begin{aligned}
\partial_n(\partial_{n+1}(e)) &= \sum_{i=0}^{n+1} (-1)^i \partial_n(\sigma([v_0, ..., \widehat{v_i}, ..., v_{n+1}])). \\
&= \sum_{i=0}^{n+1} \left(\sum_{j<i} (-1)^i (-1)^j (\sigma([v_0, ..., \widehat{v_j}, ..., \widehat{v_i}, ..., v_{n+1}])) \right. \\
&\qquad \left. + \sum_{j>i} (-1)^i (-1)^{j-1} (\sigma([v_0, ..., \widehat{v_i}, ..., \widehat{v_j}, ..., v_{n+1}])) \right) \\
&= 0.
\end{aligned}$$

Hence we have the proposition. □

Equivalently, we have the following fact.

Corollary 6.2.3

For any $n \in \mathbb{N}$, we have

$$\mathrm{Im}(\partial_{n+1}) \subset \ker(\partial_n).$$

We have a name for such a diagram. Let $(C_n)_{n \in \mathbb{N}}$ be a sequence of abelian groups and for each $n \in \mathbb{N}^*$, we have a group homomorphism

$$\partial_n : C_n \to C_{n-1}.$$

We denote by ∂_0 the trivial homomorphism from C_0 to 0 the trivial group. Hence we have the following diagram:

$$\cdots \xrightarrow{\partial_{n+2}} C_{n+1} \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\partial_n} C_{n-1} \xrightarrow{\partial_{n-1}} \cdots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0$$

Definition 6.2.4

We say that $(C_n)_{n \in \mathbb{N}}$ with $(\partial_n)_{n \in \mathbb{N}}$ is a ***chain complex***, if for any $n \in \mathbb{N}$, we have

$$\partial_n \circ \partial_{n+1} = 0.$$

Given a chain complex

$$\cdots \xrightarrow{\partial_{n+2}} C_{n+1} \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\partial_n} C_{n-1} \xrightarrow{\partial_{n-1}} \cdots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0$$

for any $n \in \mathbb{N}$, we denote

$$Z_n := \ker \partial_n \quad \text{and} \quad B_n := \mathrm{Im}\partial_{n+1},$$

by the definition of chain complex, we have

$$B_n \subset Z_n.$$

Definition 6.2.5

For any $n \in \mathbb{N}$, an element of C_n is called an ***n*-chain**. Let $\sigma \in C_n$ be an n-chain. If we have $\sigma \in Z_n$, we call σ an ***n*-cycle**. If we have $\sigma \in B_n$, we call σ an ***n*-boundary**. The quotient group

$$H_n := Z_n / B_n$$

is called the ***n*-th homology group**.

For any $n \in \mathbb{N}$, an element in H_n is called a ***homology class***. If two cycles z and z' belong to a same homology class, we say that they are ***homologous***.

We now consider the abelian groups $(C_n^\Delta)_{n\in\mathbb{N}}$ and $(\partial_n)_{n\in\mathbb{N}}$ constructed previously for X with a simplicial complex structure, we have a chain complex.

Definition 6.2.6

For any $n \in \mathbb{N}$, the homology group $H_n^\Delta(X)$ for this chain complex is called the ***n*-th simplicial homology group**.

Example 6.2.7 (Standard simplices).
Let $n \in \mathbb{N}^*$ and let $X = \Delta^n = [v_0, ..., v_n]$ be the standard n-simplex. The numbers of simplices in Δ^n of different dimensions are as follows:

dim	n	$n-1$	$n-2$	$\cdots$	1
#	1	$\frac{n+1}{1!}$	$\frac{(n+1)n}{2!}$	$\cdots$	n

re are one n-simplex, $n+1$ $(n-1)$-simplices, $\frac{n(n+1)}{2}$ $(n-2)$-simplices,..., n 0-simplicies. The abelian groups $C_k^\Delta(\Delta^n)$'s are as follows:

$$C_k^\Delta(\Delta^n) = \begin{cases} 0 & k > n \\ \mathbb{Z}[v_0, ..., v_n] & k = n \\ \bigoplus_{0\le i\le n} \mathbb{Z}[v_0, ..., \widehat{v_i}, ...v_n] & k = n-1 \\ \bigoplus_{0\le i_1<i_2\le n} \mathbb{Z}[v_0, ..., \widehat{v_{i_1}}, ..., \widehat{v_{i_2}}...v_n] & k = n-2 \\ \cdots & \\ \bigoplus_{0\le i\le n} \mathbb{Z}[v_i] & k = 0 \end{cases}$$

Example 6.2.8 (Triangulated torus with boundary).
We consider the torus with a simplicial complex structure in Figure 6.2.5. Denote the torus by T. Let S be the interior of a triangle in T which is the image of $\sigma_\alpha|_{\Delta^2}$ for some α. Topologically we remove an open disc from the torus. The resulting surface is a torus with one boundary component, denoted by T'. By removing the map σ_α from the simplicial complex structure of T, we obtain a simplicial complex structure of T'.

Let $e_{\alpha_1}, ..., e_{\alpha_k}$ be all 2-simplicies in T'. An orientation on Δ^2 induces orientations on $e_{\alpha_1}, ..., e_{\alpha_k}$ by $\sigma_{\alpha_1}, ..., \sigma_{\alpha_k}$. If two 2-simplices e_{α_j} and e_{α_l} are adjacent, then by the definition of the simplicial complex structure σ_{α_j} and σ_{α_l} induces a same orientation on the common 1-simplex $e_{\alpha_j} \cap e_{\alpha_l}$. Hence the orientation on e_{α_j} and that on e_{α_l} are not coherent.

Consider the 2-chain

$$\sigma = \bigoplus_{i=1}^{k} \epsilon_i e_{\alpha_i},$$

such that $\epsilon_1, ..., \epsilon_k \in \{1, -1\}$, and for any two given neighbor simplicies e_{α_j} and e_{α_l}, we have $\epsilon_j = -\epsilon_l$. In this way, we use this chain to represent the torus T' with an orientation. The simplex removed from T to get T' is given by σ_α. Up to a sign, the boundary of chain σ is then

$$\partial\sigma = \sigma_\alpha([v_0, v_1]) + \sigma_\alpha([v_1, v_2]) - \sigma_\alpha([v_0, v_2]).$$

Consider its boundary, we then have

$$\partial(\partial\sigma) = \sigma_\alpha(v_1) - \sigma_\alpha(v_0) + \sigma_\alpha(v_2) - \sigma_\alpha(v_1) - \sigma_\alpha(v_2) + \sigma_\alpha(v_0) = 0.$$

Topologically the boundary of T' is a circle which has empty boundary.

6.3 Singular homology

Recall that our initial goal is to study the topological space X and a simplicial complex structure is an additional structure on X. Given a space X, its simplicial complex structures are not unique in general. For example, a compact surface X may have different triangulations which are different combinatorically. On the other hand the simplicial homology group depends on the existence and choice of a simplicial complex structiure on X. Hence we may face two immediate questions:

1) Does a space X has a simplicial complex structure?
2) Are the homology groups for different simplicial complex structures on X isomorphic to each other?

In fact the construction of simplicial homology groups can be generalized to avoid these problems. We now present the construction of singular homology groups for a space. Let X be a topological space. For any $n \in \mathbb{N}$, we still denote by Δ^n the standard n-simplex. Instead of focus on simplices as subsets in X, we consider the map from Δ^n to X.

Definition 6.3.1

For any $n \in \mathbb{N}$, a ***singular*** n***-simplex*** in X is a continuous map

$$\sigma : \Delta^n \to X.$$

As in the construction of simplicial chain complex, we also associated to Δ^n an orientation. Then define

$$C_n(X) := \left\{ \sum_{i=0}^{k} m_i\sigma_i \,\middle|\, k \in \mathbb{N}, m_0, ..., m_k \in \mathbb{Z} \right\}.$$

The construction of boundary homomorphisms still works here for $C_n(X)$. For any $n \in \mathbb{N}^*$, we may express Δ^n with an orientation as an ordered sequence of its vertices

$$[v_0, ..., v_n].$$

Then for any singular n-simplex σ, we define

$$\partial_n \sigma = \sum_{i=0}^{n} (-1)^n \sigma|_{[v_0,...,\widehat{v_i},...,v_n]}.$$

Here we identify a $(n-1)$-face $[v_0, ..., \widehat{v_i}, ..., v_n]$ with the standard $(n-1)$-simplex Δ^{n-1} by identifying the vertices in order and extending this identification to $[v_0, ..., \widehat{v_i}, ..., v_n]$ to Δ^{n-1} using linear maps. Then for any n-chain

$$\sigma' = \sum_{i=1}^{k} \sigma_i \in C_n(X),$$

we define

$$\partial_n(\sigma') = \sum_{i=1}^{k} \partial \sigma_i.$$

This gives a group homomorphism

$$\partial_n : C_n(X) \to C_{n-1}(X).$$

For $n = 0$, we define ∂_0 to be the trivial homomorphism from $C_0(X)$ to the trivial group 0. In this way, we have the following diagram

$$\cdots \xrightarrow{\partial_{n+2}} C_{n+1}(X) \xrightarrow{\partial_{n+1}} C_n(X) \xrightarrow{\partial_n} C_{n-1}(X) \xrightarrow{\partial_{n-1}} \cdots \xrightarrow{\partial_2} C_1(X) \xrightarrow{\partial_1} C_0(X) \xrightarrow{\partial_0} 0$$

A similar computation shows the following fact.

Proposition 6.3.2

For any $n \in \mathbb{N}$, we have

$$\partial_n \circ \partial_{n+1} = 0.$$

Hence $(C_n(X))_{n\in\mathbb{N}}$ with $(\partial_n)_{n\in\mathbb{N}}$ is a chain complex.

Definition 6.3.3

The homology group $H_n(X)$ for this chain complex is called the ***n-th singular homology group of X***.

Compare with simplicial homology groups, singular homology groups are independent of choice of simplicial complex structure and can be defined for the space even without any simplicial complex structures. Moreover by considering composition of continuous maps, we may immediately conclude that singular homology groups are topological invariants. More precisely, if two spaces are homeomorphic to each other, they have isomorphic n-th singular homology groups for any $n \in \mathbb{N}$.

Of course, there is a price to pay. We consider too many simplices. It is not easy to compute it directly by definition in general. We will study it more closely in the next several sections. But before that, let us check some simple cases first.

Proposition 6.3.4

For any path connected space X, we have $H_0(X) \cong \mathbb{Z}$.

Proof. Notice that ∂_0 is trivial, hence

$$Z_0(X) = C_0(X).$$

We will show next that all singular 0-simplices in X are homologous to each other. Notice that Δ^0 is a single point set. Hence an 0-simplex is determined by its image. For any p and q two points in X, we have

$$\sigma : \Delta^0 \to \{p\} \quad \text{and} \quad \tau : \Delta^0 \to \{q\}.$$

Since X is path connected, there is a path

$$\alpha : [0,1] \to X$$

with $\alpha(0) = p$ and $\alpha(1) = q$. We identify $[0,1]$ with

$$\Delta^1 = [v_0, v_1],$$

where 0 and 1 are identified with v_0 and v_1 respectively. Then α can be considered as a singular 1-simplex in X, and we have

$$\partial\alpha = \tau - \sigma.$$

Hence τ and σ are different by a 0-boundary, and we have in the homology group

$$[\tau] = [\sigma].$$

Hence $H_0(X)$ has one generator $[\tau]$ and

$$H_0(X) \cong \mathbb{Z}.$$

□

Another obvious fact comes from the fact that a continuous map sends a path connected space to a path connected space.

Proposition 6.3.5

Let X be a topological space with a path connected component decomposition

$$X = \bigsqcup_{\alpha\in\Omega} X_\alpha.$$

Then for any $n \in \mathbb{N}$, we have

$$H_n(X) \cong \bigoplus_{\alpha\in\Omega} H_n(X).$$

Proof. For any $n \in \mathbb{N}$, for any singular n-simplex σ in X, since Δ^n is path connected and σ is continuous, there is an index $\alpha \in \Omega$, such tha

$$\sigma(\Delta^n) \subset X_\alpha.$$

Hence we have

$$C_n(X) = \bigoplus_{\alpha\in\Omega} C_n(X_\alpha).$$

Moreover, for any $n \in \mathbb{N}^*$, any $\alpha \in \Omega$, and any $\sigma \in C_n(X_\alpha)$, we have

$$\partial_n \sigma \in C_{n-1}(X).$$

Hence we have

$$Z_n(X) = \bigoplus_{\alpha \in \Omega} Z_n(X_\alpha),$$

and

$$B_n(X) = \bigoplus_{\alpha \in \Omega} B_n(X_\alpha).$$

Hence by the fundamental theorem of group homomorphism, we have

$$H_n(X) = Z_n(X)/B_n(X) \cong \bigoplus_{\alpha \in \Omega} Z_n(X_\alpha)/B_n(X_\alpha) = \bigoplus_{\alpha \in \Omega} H_n(X_\alpha).$$

□

Corollary 6.3.6

Let X be a topological space with a path connected component decomposition

$$X = \bigsqcup_{\alpha \in \Omega} X_\alpha.$$

Then we have

$$H_0(X) \cong \bigoplus_{\alpha \in \Omega} \mathbb{Z}$$

The "simplest" topological space is the single point space. In this case, we can explicitly compute its singular homology groups using definition.

Proposition 6.3.7

Let X be a single point space. Then we have

$$H_n(X) \cong \begin{cases} \mathbb{Z}, & n = 0, \\ 0, & n > 0. \end{cases}$$

Proof. We denote by p the point in X. For any $n \in \mathbb{N}$, there is a unique singular n-simplex in X:

$$\begin{aligned} \sigma_n : \Delta^n &\to X \\ w &\mapsto p \end{aligned}$$

which is a constant map. We compute the boundary of σ_n. For any $n \in \mathbb{N}^*$, we have

$$\partial_n \sigma_n = \sum_{i=0}^{n} (-1)^i \sigma_{n-1}.$$

Hence we have

$$\partial_n \sigma_n = \begin{cases} \sigma_{n-1}, & n \text{ even}, \\ 0, & n \text{ odd}. \end{cases}$$

Therefore up to isomorphism, the chain complex

$$\cdots \xrightarrow{\partial_{n+2}} C_{n+1}(X) \xrightarrow{\partial_{n+1}} C_n(X) \xrightarrow{\partial_n} C_{n-1}(X) \xrightarrow{\partial_{n-1}} \cdots \xrightarrow{\partial_2} C_1(X) \xrightarrow{\partial_1} C_0(X) \xrightarrow{\partial_0} 0$$

can be identified with the following one

$$\cdots \xrightarrow{\text{id}} \mathbb{Z} \xrightarrow{0} \mathbb{Z} \xrightarrow{\text{id}} \mathbb{Z} \xrightarrow{0} \cdots \xrightarrow{\text{id}} \mathbb{Z} \xrightarrow{0} \mathbb{Z} \xrightarrow{0} 0$$

where n is even.

Hence for any $n \in \mathbb{N}^*$, we have

$$Z_n(X) = \begin{cases} 0, & n \text{ even}, \\ C_n(X), & n \text{ odd}, \end{cases}$$

and

$$B_n(X) = \begin{cases} 0, & n \text{ even}, \\ C_n(X), & n \text{ odd}. \end{cases}$$

Hence for any $n \in \mathbb{N}^*$, we have

$$Z_n(X) = B_n(X),$$

and

$$H_n(X) = Z_n(X)/B_n(X) \cong 0.$$

Since X is path connected, we have $H_0(X) = 0$. We can also get it from the above discussion which shows that

$$Z_0(X) = C_0(X) \quad \text{and} \quad B_0(X) = 0.$$

Therefore, we have

$$H_0(X) \cong C_0(X) \cong \mathbb{Z}.$$

□

Remark 6.3.8.
To simplify the notation, we will denote ∂ for all ∂_n. The meaning will be clear by considering the context

6.4 Homotopy invariance of singular homology groups

We have seen that the singular homology is invariant under homeomorphisms. In fact, it is also invariant under homotopy equivalence.

Consider two topological spaces X and Y. Let

$$f : X \to Y$$

be a continuous map. Then for any $n \in \mathbb{N}$ and any n-simplex in X:

$$\sigma : \Delta^n \to X,$$

we have

$$f \circ \sigma : \Delta^n \to Y$$

an n-simplex in Y (see Figure 6.4.1 for an illustration).

Hence for each $n \in \mathbb{N}$, the map f induces a group homomorphism

$$f_\# : C_n(X) \to C_n(Y)$$
$$\sum_{i=1}^{k} n_i\sigma_i \mapsto \sum_{i=1}^{k} n_i(f \circ \sigma_i).$$

Here to avoid too many subscription, we use $f_\#$ for all $n \in \mathbb{N}$. The meaning will be clear by considering the context.

This homomorphism satisfies the following property:

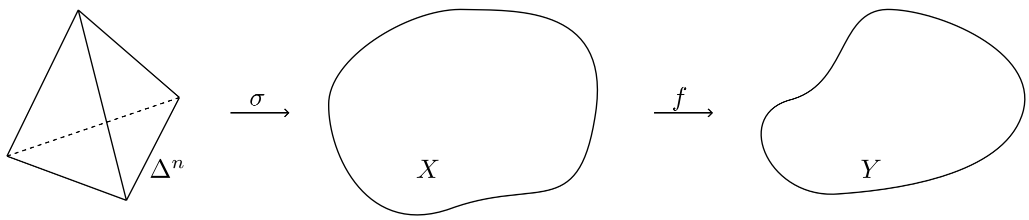


Figure 6.4.1

Proposition 6.4.1

We have

$$\partial \circ f_\# = f_\# \circ \partial.$$

Here the ∂ on the left hand side is the boundary homomorphism for $(C_n(Y))_{n\in\mathbb{N}}$ and the one on the right hand side is the boundary homomorphism for $(C_n(X))_{n\in\mathbb{N}}$

It comes from the observation that f sends an $(n-1)$-face of an n-simplex σ in X to an $(n-1)$-face of $f \circ \sigma$ in Y. The proof of the fact that $f_\#$ is a group homomorphism and the proof of the above proposition are left to readers.

Hence we have the following commutative diagram

$$\begin{array}{ccccccccccccc}
\cdots & \xrightarrow{\partial} & C_{n+1}(X) & \xrightarrow{\partial} & C_n(X) & \xrightarrow{\partial} & C_{n-1}(X) & \xrightarrow{\partial} \cdots \xrightarrow{\partial} & C_1(X) & \xrightarrow{\partial} & C_0(X) & \xrightarrow{\partial} & 0 \\
 & & \downarrow f_\# & & \downarrow f_\# & & \downarrow f_\# & & \downarrow f_\# & & \downarrow f_\# & & \\
\cdots & \xrightarrow{\partial} & C_{n+1}(Y) & \xrightarrow{\partial} & C_n(Y) & \xrightarrow{\partial} & C_{n-1}(Y) & \xrightarrow{\partial} \cdots \xrightarrow{\partial} & C_1(Y) & \xrightarrow{\partial} & C_0(Y) & \xrightarrow{\partial} & 0
\end{array}$$

Definition 6.4.2

The homomorphisms $f_\#$'s are called the ***chain maps induced by*** f.

One important property of the chain maps $f_\#$'s is that it builds the connection in the homology group level.

Proposition 6.4.3

For each $n \in \mathbb{N}$, the chain maps $f_\#$ induces a group homomorphism

$$f_* : H_n(X) \to H_n(Y).$$

Proof. It is enough to show that for any $n \in \mathbb{N}$, the map $f_\#$ sends n-cycles and n-boundaries in $C_n(X)$ to n-cycles and n-boundaries in $C_n(Y)$ respectively.

Let $n \in \mathbb{N}$. For any $z \in Z_n(X)$, we have

$$\partial(f_\#(z)) = f_\#(\partial(z)) = f_\#(0) = 0.$$

Hence $f_\#(Z) \in Z_n(Y)$. Therefore we have a homomorphism

$$\pi_Y \circ f_\# : Z_n(X) \to Z_n(Y) \to H_n(Y) := Z_n(Y)/B_n(Y).$$

For any $b \in B_n(X)$, there is a $(n+1)$-chain $\sigma \in C_{n+1}(X)$, such that

$$b = \partial\sigma.$$

We have

$$f_\#(b) = f_\#(\partial(\sigma)) = \partial(f_\#(\sigma)).$$

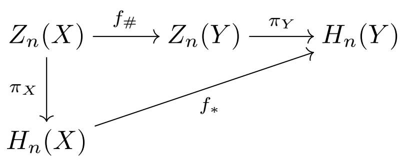
Therefore $f_\#(b) \in B_n(X)$, and

$$B_n(X) \subset \ker(\pi_Y \circ f_\#).$$

We have a group homomorphism

$$f_* : H_n(X) \to H_n(Y),$$

which satisfies the following commutative diagram



□

Now we consider two continuous maps

$$f : X \to Y \quad \text{and} \quad g : X \to Y,$$

homotopic to each other. We should like to show the following theorem.

Theorem 6.4.4

For any $n \in \mathbb{N}$, we have $f_* = g_*$.

Proof. Since f and g are homotopic, there is a homotopy

$$H : X \times [0,1] \to Y,$$

between them, such that $H_0 = f$ and $H_1 = g$.

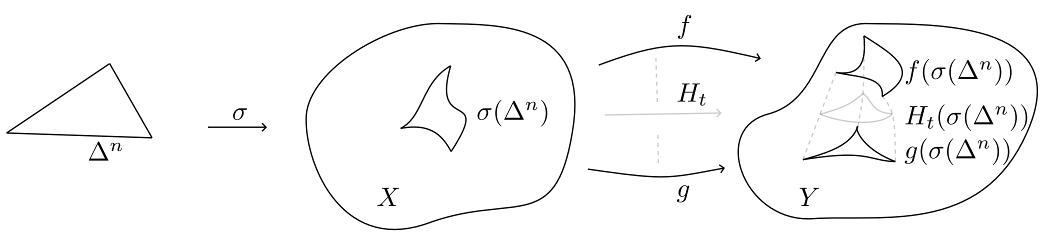


Figure 6.4.2: The relation between $f_\#$ and $g_\#$.

Let $n \in \mathbb{N}$. We consider the induced chain maps $f_\#$ and $g_\#$ (see Figure 6.4.2 for an illustration). It is enough to show that for any $z \in Z_n(X)$, we have

$$f_\#(z) - g_\#(z) \in B_n(Y).$$

For any $k \in \mathbb{N}$, Given any k-simplex τ in X, we can extend this to a continuous map

$$\widetilde{\tau} : \Delta^k \times [0,1] \to X \times [0,1],$$
$$(a,t) \mapsto (\tau(a), t).$$

Notice that $\Delta^k \times [0,1]$ is a prism. We denote the simplex Δ^k using its vertices

$$[u_0, ..., u_k].$$

Then we denote the simplex $\Delta^k \times \{0\}$ still by

$$[v_0, ..., v_k].$$

and the simplex $\Delta^k \times \{1\}$ by

$$[w_0, ..., w_k],$$

such that for any $0 \le i \le k$, we have $v_i = (u_i, 0)$ and $w_i = (u_i, 1)$.

We consider $\Delta^n \times [0,1]$ and the map $H \circ \widetilde{\sigma}$ (see Figure 6.4.3 for an illustration).

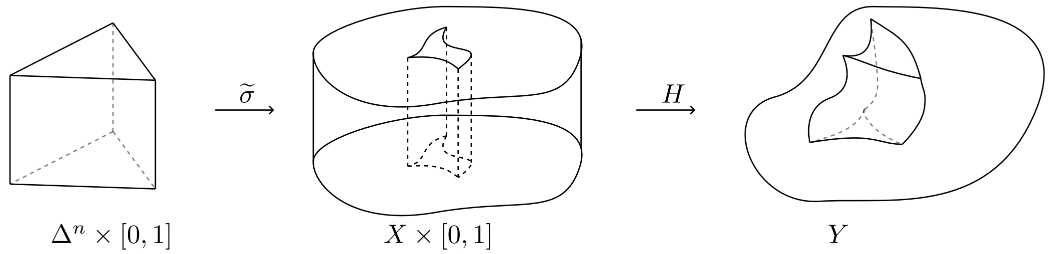


Figure 6.4.3: An illustration of the map $H \circ \widetilde{\sigma}$.

Then the prime can be decomposed in to the union of the following n-simplices whose interiors have empty intersections:

$$\{[v_0, ..., v_i, w_i, ..., w_n] \mid 0 \le i \le n\}.$$

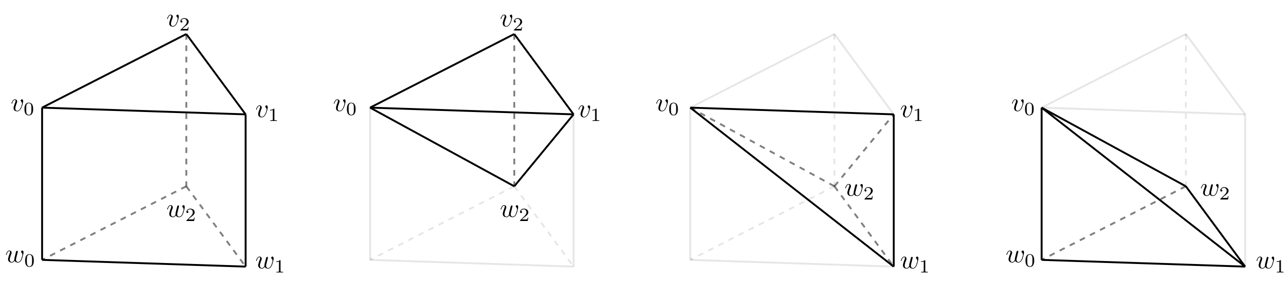


Figure 6.4.4: An illustration of the decomposition of the prism $\Delta^2 \times [0,1]$.

For any n-simplex σ in X, we then can define

$$P(\sigma) = \sum_{i=0}^{n} (-1)^i H \circ \widetilde{\sigma}|_{[v_0,...,v_i,w_i,...,w_n]}.$$

This can be extend to a group homomorphism

$$P : C_n(X) \to C_{n+1}(Y)$$

and P is called the prime operator.

One important property satisfied by P is the following identity

$$g_\# - f_\# = \partial \circ P + P \circ \partial.$$

To prove this identity, it is enough to check it for one n-simplex. We first compute $(\partial \circ P)(\sigma)$:

$$\begin{aligned}
(\partial \circ P)(\sigma) &= \partial \left(\sum_{i=0}^{n} (-1)^i H \circ \widetilde{\sigma}|_{[v_0,\dots,v_i,w_i,\dots,w_n]} \right) \\
&= \sum_{i=0}^{n} \left(\sum_{j \le i} (-1)^i (-1)^j H \circ \widetilde{\sigma}|_{[v_0,\dots,\widehat{v_j},\dots,v_i,w_i,\dots,w_n]} \right. \\
&\left. + \sum_{j \ge i} (-1)^i (-1)^{j+1} H \circ \widetilde{\sigma}|_{[v_0,\dots,v_i,w_i,\dots,\widehat{w_j},\dots,w_n]} \right).
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
(P \circ \partial)(\sigma) &= \sum_{j=0}^{n} (-1)^j P(\sigma|_{[v_0,\dots,\widehat{v_i},\dots,v_n]}) \\
&= \sum_{j=0}^{n} \left(\sum_{j<i} (-1)^j (-1)^{i-1} H \circ \widetilde{\sigma}|_{[v_0,\dots,\widehat{v_j},\dots,v_i,w_i,\dots,w_n]} \right. \\
&\left. + \sum_{j>i} (-1)^j (-1)^i H \circ \widetilde{\sigma}|_{[v_0,\dots,v_i,w_i,\dots,\widehat{w_j},\dots,w_n]} \right).
\end{aligned}$$

Hence we have

$$\begin{aligned}
(\partial \circ P + P \circ \partial)(\sigma) &= H \circ \widetilde{\sigma}|_{[w_0,\dots,w_n]} + \sum_{i=0}^{n-1} (-1)^{2i} H \circ \widetilde{\sigma}|_{[v_0,\dots,v_i,w_{i+1},\dots,w_n]} + \\
&+ \sum_{i=0}^{n-1} (-1)^{2i+1} H \circ \widetilde{\sigma}|_{[v_0,\dots,v_i,w_{i+1},\dots,w_n]} - H \circ \widetilde{\sigma}|_{[v_0,\dots,v_n]} \\
&= H \circ \widetilde{\sigma}|_{[w_0,\dots,w_n]} - H \circ \widetilde{\sigma}|_{[v_0,\dots,v_n]} \\
&= g_\#(\sigma) - f_\#(\sigma).
\end{aligned}$$

For any $z \in Z_n(X)$, we have

$$(g_\# - f_\#)(z) = (\partial \circ P + P \circ \partial)(z) = \partial(P(z)) \in B_n(Y).$$

Hence the theorem. □

Definition 6.4.5

Any chain map L mapping $C_n(X)$ to $C_{n+1}(Y)$ for any $n \in \mathbb{N}$ and satisfying

$$g_\# - f_\# = \partial \circ L + L \circ \partial,$$

is called a ***chain homotopy*** between chain maps $f_\#$ and $g_\#$.

With this theorem, we can show the homotopy invariance of homology groups.

Theorem 6.4.6

Let X and Y be two spaces homotopy equivalent to each other. Then for any $n \in \mathbb{N}$, we have
$$H_n(X) \cong H_n(Y).$$

Proof. Since X and Y are homotopy equivalent, there are continuous maps
$$f : X \to Y \quad \text{and} \quad g : Y \to X,$$
such that $f \circ g \sim \mathrm{id}_Y$ and $g \circ f \sim \mathrm{id}_X$.

By the previous theorem, for any $n \in \mathbb{N}$, we have
$$f_* \circ g_* = (f \circ g)_* = (\mathrm{id}_Y)_* = \mathrm{id}_{H_n(Y)},$$
and
$$g_* \circ f_* = (g \circ f)_* = (\mathrm{id}_X)_* = \mathrm{id}_{H_n(X)}.$$
This implies that both f_* and g_* are bijective. Therefore both of them are isomorphisms and we have
$$H_n(X) \cong H_n(Y).$$
□

As an application, we consider contractible spaces and have the following statement.

Corollary 6.4.7

If X be a contractible space, then for any $n \in \mathbb{N}$, we have
$$H_n(X) \cong \begin{cases} \mathbb{Z}, & n = 0 \\ 0, & n > 0 \end{cases}$$

Proof. Since X is contractible, it is homotopy equivalent to a single point space. By Proposition 6.3.7, we have the corollary. □

6.5 Singular homology and subspaces

When studying topological spaces, sometimes the whole space is difficult to study but it has some subspace which is easy to study and the quotient space is also easy to study. Sometimes it also happens that a space is difficult to study, but can be viewed as part of a space easy to be studied. Hence studying homology groups of a space relative to it subspace would be useful in these cases.

Let X be a topological space and A be a non-empty subspace in X. In order to make the homology machinery work, we assume that A has an open neighborhood U in X, such that A is a strong deformation contraction of U. We define the relative chain complex in the following way.

For any $n \in \mathbb{N}$, we define
$$C_n(X, A) := C_n(X)/C_n(A).$$
Hence we have the short exact sequence
$$0 \longrightarrow C_n(A) \xrightarrow{\ i\ } C_n(X) \xrightarrow{\ \mathrm{pr}\ } C_n(X, A) \longrightarrow 0$$
where i is the inclusion map and pr is the quotient map.

Notice that ∂ maps an n-chain in A to an $(n-1)$-chain in A. Hence for any $n \in \mathbb{N}$, the boundary homomorphism for $C_n(X)$ induces a homomorphism

$$\partial : C_{n+1}(X, A) \to C_n(X, A).$$

Then we have a chain complex

$$\cdots \xrightarrow{\partial} C_{n+1}(X, A) \xrightarrow{\partial} C_n(X, A) \xrightarrow{\partial} \cdots \xrightarrow{\partial} C_1(X, A) \xrightarrow{\partial} C_0(X, A) \xrightarrow{\partial} 0$$

The last ∂ from $C_0(X, A)$ is unique trivial homomorphism to the trivial group. Since the identity element in $C_n(X)$ is sent to the identity element in $C_n(X, A)$, hence we have

$$\partial \circ \partial = 0,$$

and $(C_n(X, A), \partial)_{n\in\mathbb{N}}$ is a chain complex.

Definition 6.5.1

For any $n \in \mathbb{N}$, the homology groups $H_n(X, A)$'s associated to this chain complex is called the ***n-th homology group of X relative to A***.

For any $n \in \mathbb{N}$, for any $\alpha \in C_n(X)$, we denote

$$\overline{\alpha} = \mathrm{pr}(\alpha).$$

Definition 6.5.2

For any $n \in \mathbb{N}$, for any $\alpha \in C_n(X)$, we call α an ***n-cycle relative to A***, if

$$\partial\alpha \in C_{n-1}(A),$$

and we call α an ***n-boundary relative to A***, if there is an $(n+1)$-chain $\beta \in C_{n+1}(X)$ and an n-chain $\gamma \in C_n(A)$, such that

$$\alpha = \partial\beta + \gamma.$$

For any $n \in \mathbb{N}$, by the definition, we have

$$Z_n(X, A) := \{\overline{\alpha} \in C_n(X, A) \mid \alpha \in C_n(X) \text{ is an } n\text{-cycle relative to } A\},$$
$$B_n(X, A) := \{\overline{\alpha} \in C_n(X, A) \mid \alpha \in C_n(X) \text{ is an } n\text{-boundary relative to } A\}.$$

and then

$$H_n(X, A) := Z_n(X, A)/B_n(X, A).$$

Another observation is that the map i and pr commute with ∂. In particular, we have the following commutative diagram which extends the above short exact sequence to the whole chain:

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
\cdots & \xrightarrow{\partial} & C_{n+1}(A) & \xrightarrow{\partial} & C_n(A) & \xrightarrow{\partial} & C_{n-1}(A) & \xrightarrow{\partial} & \cdots \\
 & & \downarrow i & & \downarrow i & & \downarrow i & & \\
\cdots & \xrightarrow{\partial} & C_{n+1}(X) & \xrightarrow{\partial} & C_n(X) & \xrightarrow{\partial} & C_{n-1}(X) & \xrightarrow{\partial} & \cdots \\
 & & \downarrow \mathrm{pr} & & \downarrow \mathrm{pr} & & \downarrow \mathrm{pr} & & \\
\cdots & \xrightarrow{\partial} & C_{n+1}(X,A) & \xrightarrow{\partial} & C_n(X,A) & \xrightarrow{\partial} & C_{n-1}(X,A) & \xrightarrow{\partial} & \cdots \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & & 0 & &
\end{array}$$

Definition 6.5.3

The above diagram is called a ***short exact sequence of chain complex***.

We would like to show the following long exact sequence which relates homology groups $H_n(X)$'s, $H_n(A)$'s and $H_n(X,A)$'s. To distinguish the homology in X and that in A, for any $\alpha \in Z_n(A)$, we will denote by $[\alpha]_A$ the homology class in $H_n(A)$, and by $[\alpha]_X$ the homology class in $H_n(X)$.

Theorem 6.5.4

There is a long exact sequence

$$\cdots \xrightarrow{\partial} H_{n+1}(A) \xrightarrow{i_*} H_{n+1}(X) \xrightarrow{\mathrm{pr}_*} H_{n+1}(X,A) \xrightarrow{\partial} H_n(A) \xrightarrow{i_*} \cdots$$

$$\cdots \xrightarrow{\mathrm{pr}_*} H_0(X,A) \xrightarrow{\partial} 0,$$

where the partial map is defined as follows: for any $n \in \mathbb{N}$,

$$\begin{aligned}\partial : H_n(X,A) &\to H_{n-1}(A),\\ [\overline{\alpha}] &\mapsto [\partial\alpha]_A.\end{aligned}$$

Proof. First we have check that the map

$$\partial : H_n(X,A) \to H_{n-1}(A),$$

is well defined.

Let $\alpha \in C_n(X)$ be a relative n-cycle, then we have

$$\partial\alpha \in C_{n-1}(A).$$

Since as an n-chain in X, we have

$$\partial(\partial(\alpha)) = 0 \in C_{n-2}(A),$$

hence $\partial\alpha \in Z_{n-1}(A)$. Hence the map is well defined.

Notice that we do not know if $\partial\alpha \in B_{n-1}(A)$ is true, since α is in $C_n(X)$, not necessary in $C_n(A)$.

Now we try to show the following three equality

$$\begin{aligned} \operatorname{Im}\partial &= \ker i_* \\ \operatorname{Im} i_* &= \ker \operatorname{pr}_* \\ \operatorname{Im}\operatorname{pr}_* &= \ker\partial \end{aligned}$$

We firs consider for any $n \in \mathbb{N}$

$$H_{n+1}(X,A) \xrightarrow{\partial} H_n(A) \xrightarrow{i_*} H_n(X)$$

$\operatorname{Im}\partial \subset \ker i_*$: Let $[\alpha] \in \operatorname{Im}\partial$, then there is a relative $(n+1)$-cycle

$$\beta \in C_{n+1}(X),$$

such that

$$[\alpha]_A = \partial[\overline{\beta}] = [\partial\beta]_A.$$

Now we consider the n-chain $\partial\beta$ in $C_n(X)$. Since

$$\partial\beta \in B_n(X),$$

we have

$$i_*([\alpha]_A) = [\alpha]_X = [\partial\beta]_X = [0]_X.$$

Hence

$$[\alpha]_A \in \ker i_*,$$

and we have

$$\operatorname{Im}\partial \subset \ker i_*.$$

$\operatorname{Im}\partial \supset \ker i_*$: Let $[\alpha]_A \in \ker i_*$, we have

$$[\alpha]_X = [0]_X.$$

Hence $\alpha \in B_n(X)$, or equivalently, there is an $(n+1)$-chain $\beta \in C_{n+1}(X)$, such that

$$\alpha = \partial\beta.$$

We consider $\overline{\beta} \in C_{n+1}(X,A)$. Since

$$\partial\beta = \alpha \in Z_n(A) \subset C_n(A),$$

the $(n+1)$-chain β is a relative $(n+1)$-cycle, hence

$$\overline{\beta} \in Z_{n+1}(X,A).$$

We then have

$$[\alpha]_A = [\partial\beta]_A = \partial([\overline{\beta}]) \in \operatorname{Im}\partial.$$

Hence

$$\operatorname{Im}\partial \supset \ker i_*.$$

Now for any $n \in \mathbb{N}$, we consider

$$H_n(A) \xrightarrow{i_*} H_n(X) \xrightarrow{\operatorname{pr}_*} H_n(X,A)$$

$\operatorname{Im} i_* \subset \ker \operatorname{pr}_*$: Let $[\alpha]_X \in \operatorname{Im} i_*$, then there is an n-cycle

$$\gamma \in Z_n(A),$$

such that

$$[\gamma]_X = [\alpha]_X,$$

or equivalently

$$\alpha - \gamma \in B_n(X).$$

Therefore, there is a $(n+1)$-chain $\beta \in C_{n+1}(X)$, such that

$$\alpha - \gamma = \partial\beta.$$

Hence

$$\alpha = \partial\beta + \gamma,$$

which shows that α is a relative n-boundary. Hence

$$\overline{\alpha} \in B_n(X, A),$$

and

$$\operatorname{pr}_*([\alpha]_X) = [\overline{\alpha}] = [\overline{0}].$$

Hence we have

$$[\alpha]_X \in \ker \operatorname{pr}_*,$$

and

$$\operatorname{Im} i_* \subset \ker \operatorname{pr}_*.$$

$\operatorname{Im} i_* \supset \ker \operatorname{pr}_*$: Let $[\alpha]_X \in \ker \operatorname{pr}_*$, we have

$$\overline{\alpha} \in B_n(X, A),$$

and α is a relative n-boundary. There are an $(n+1)$-chain $\beta \in C_{n+1}(X)$ and $\gamma \in C_n(A)$, such that

$$\alpha = \partial\beta + \gamma.$$

Hence we have

$$[\alpha]_X = [\gamma]_X = i_*([\gamma]_A) \in \operatorname{Im} i_*.$$

Therefore we have

$$\operatorname{Im} i_* \supset \ker \operatorname{pr}_*.$$

Finally for any $n \in \mathbb{N}$, we consider

$$H_{n+1}(X) \xrightarrow{\operatorname{pr}_*} H_{n+1}(X, A) \xrightarrow{\partial} H_n(A)$$

$\operatorname{Im} \operatorname{pr}_* \subset \ker \partial$: Let $[\overline{\alpha}] \in \operatorname{Im} \operatorname{pr}_*$, then there is an $(n+1)$-cycle

$$\gamma \in Z_{n+1}(X),$$

such that

$$\operatorname{pr}_*([\gamma]_X) = [\overline{\gamma}] = [\overline{\alpha}].$$

Hence

$$\overline{\alpha} - \overline{\gamma} \in B_n(X, A),$$

or equivalently $\alpha - \gamma$ is a relative $(n+1)$-boundary. Hence there is an $(n+2)$-chain $\beta \in C_{n+2}$, and an $(n+1)$-chain $\eta \in C_{n+1}(A)$, such that in $C_{n+1}(X)$, we have

$$\alpha - \gamma = \partial\beta + \eta.$$

Now we have

$$\partial\alpha - \partial\gamma = \partial\eta \in B_n(A).$$

Hence

$$[\partial\alpha]_A = [\partial\gamma]_A.$$

Notice that γ is an $(n+1)$-cycle, hence $\partial\gamma = 0$ in $C_n(A)$. Therefore we have

$$\partial[\overline{\alpha}] = [\partial\alpha]_A = [0]_A.$$

and

$$[\overline{\alpha}] \in \ker\partial.$$

Hence we have

$$\operatorname{Im}\operatorname{pr}_* \subset \ker\partial.$$

$\operatorname{Im}\operatorname{pr}_* \supset \ker\partial$: Let $[\overline{\alpha}] \in \ker\partial$, then

$$[\partial\alpha]_A = [0]_A.$$

Hence $\partial\alpha \in B_n(A)$ and there is an $(n+1)$-chain $\gamma \in C_{n+1}(A)$, such that

$$\partial\alpha = \partial\gamma.$$

Therefore

$$\alpha - \gamma \in Z_{n+1}(X),$$

and we have

$$\operatorname{pr}_*([\alpha-\gamma]_X) = [\overline{\alpha-\gamma}] = [\overline{\alpha}].$$

Therefore we have

$$\operatorname{Im}\operatorname{pr}_* \supset \ker\partial.$$

□

Another important tool when study relative homology is called the excision theorem. If one consider the homology of X relative to a subspace A, then naively, taking out a part of A should not effect on the relative homology groups. The excision theorem shows that it is indeed the case. There are two equivalent versions of this theorem.

We consider the subspace U of a space X. Let V be a subspace of X such that

$$\overline{V} \subset \mathring{U}.$$

Theorem 6.5.5 (Excision Theorem I)

For any $n \in \mathbb{N}$, we have

$$H_n(X,U) \cong H_n(X-V, U-V).$$

Equivalently, we can state it in the following way. Let A and B be two subspaces of X such that

$$X = \mathring{A} \cup \mathring{B}.$$

Theorem 6.5.6 (Excision Theorem II)

For any $n \in \mathbb{N}$, we have
$$H_n(A, A \cap B) \cong H_n(X, B).$$

Remark 6.5.7.
To see the equivalence, we can take take $U = B$ and $V = X \setminus A$.

We will proof the second version. Before giving the details, let us first have a look at the situation. Notice that $A \subset X$ and $B \subset X$ are both subspaces, hence any simplex in A or in B is also a simplex in X. Therefore for any n, the abelian group $C_n(A)$ and $C_n(B)$ are two subgroups of $C_n(X)$. We define

$$C_n(A+B) := \left\{ \sum_{i=1}^{m} k_i \sigma_i \mid \sigma_i \text{ is an } n\text{-simplex in } A \text{ or in } B \right\}.$$

In fact, this is just the subgroup of $C_n(X)$ generalized by $C_n(A)$ and $C_n(B)$. Since $C_n(X)$ is an abelian group, we can also rewrite it into

$$C_n(A+B) = C_n(A) + C_n(B).$$

Another subgroup of $C_n(X)$ is given by considering simplices in $A \cap B$. We define

$$C_n(A \cap B) := \left\{ \sum_{i=1}^{m} k_i \sigma_i \mid \sigma_i \text{ is an } n\text{-simplex in } A \cap B \right\}.$$

Since all subgroups of an abelian group are normal subgroup, by the fundamental theorem of group homomorphisms, for any $n \in \mathbb{N}$, we have

$$C_n(A)/C_n(A \cap B) \cong C_n(A+B)/C_n(B),$$

or equivalently, we have
$$C_n(A, A \cap B) \cong C_n(A+B, B).$$

Moreover, by the definition of the groups involved here, this isomorphism also induces isomorphism between the subgroup of relatives cycles and the subgroup of relative boundaries:

$$Z_n(A, A \cap B) \cong Z_n(A+B, B) \quad \text{and} \quad B_n(A, A \cap B) \cong B_n(A+B, B)$$

The details are left to readers to check. With these facts, we have the following observation:

Observation 6.5.8
For any $n \in \mathbb{N}$, we have
$$H_n(A, A \cap B) \cong H_n(A+B, B).$$

Hence it is enough to show for each $n \in \mathbb{N}$ the following isomorphism

$$H_n(A+B, B) \cong H_n(X, B).$$

The difference between two sides is this: for $H_n(X, B)$, we use simplices in X, and for $H_n(A+B, B)$ we use simplices in A or B. In the other words, if the above isomorphism holds, it means using "smaller" simplices, we obtain the same homology group for each $n \in \mathbb{N}$. This will be the key point in the proof of the excision theorem.

The main tool used here is the barycentric subdivision for an Euclidean simplex. We will apply such subdivision on Δ^n to subdivide simplex in X. Notice that by hypothesis, we have

$$X = \mathring{A} \cup \mathring{B}.$$

Hence for any $n \in \mathbb{N}$ and any n-simplex σ in X, if the preimage of $\sigma^{-1}(\mathring{A})$ and $\sigma^{-1}(\mathring{B})$ form an open cover of Δ^n. Since Δ^n is compact, there is an open cover of balls of Δ^n such that the image of each ball is either in $\mathring{A}$ or in $\mathring{B}$. Hence all we have to do in this step is to subdivide Δ^n into simplices with diameter uniformly small enough.

Let us first recall the varycentric subdivision. Let σ be a Euclidean n-simplex determined by $n+1$ points $v_0, ..., v_n \in \mathbb{R}^m$ $(m \geq n)$. Then the barycenter of σ is given by

$$v_\sigma = \frac{1}{n+1}(v_0 + \cdots + v_n).$$

If τ is a k-face of σ with $k < n$, we denote $\tau < \sigma$. Then in a barycentric subdivision, we first take the barycenters of all faces of σ and v_σ. Then we subdivide σ into the union of Euclidean n-simplices, each of with can be written as

$$[v_{\tau_0}, ..., v_{\tau_{n-1}}, v_\sigma],$$

such that the faces $\tau_0, ..., \tau_{n-1}$ satisfy

$$\tau_0 < \cdots < \tau_{n-1} < \sigma.$$

We denote by d the diameter of σ and

$$d' = \max\{\eta \mid \eta \text{ is an } n\text{-simplex obtained from the barycentric subdivision of } \sigma\}.$$

Then we have the following comparison.

Lemma 6.5.9

We have

$$d' \leq \frac{n}{n+1} d.$$

Proof. The diameter of an Euclidean simplex is given by its longest edge (1-face).

We use induction on n. Notice that for $n = 0$, we have $0 = 0$.

Assume that the inequality holds for k-simplicies with $0 \leq k \leq n-1$ for some $n > 0$, then let $\tau < \tau'$ be two faces of σ. Without loss of generality, we may assume that

$$\tau = [v_0, ..., v_k] < [v_0, ..., v_l] = \tau'$$

with $k < l \leq n$.

If $l < n$, then

$$||v_\tau - v'_\tau|| \leq \frac{l}{l+1}\mathrm{diam}(\tau') \leq \frac{n}{n+1}\mathrm{diam}(\sigma) = \frac{n}{n+1}d.$$

For the case $l = n$, notice that

$$\begin{aligned} ||v_\sigma - v_\tau|| &\leq \frac{k+1}{k+1}\max\{||v_\sigma - v_i|| \mid 0 \leq i \leq k\} \\ &= \max\{||v_\sigma - v_i|| \mid 0 \leq i \leq k\} \end{aligned}$$

On the other hand, for each i, we have

$$v_\sigma - v_i = \frac{1}{n+1}(v_0 + \cdots + v_n) - \frac{n+1}{n+1}v_i = \frac{1}{n+1}\sum_{j\neq i}(v_0 - v_i).$$

Hence we have

$$||v_\sigma - v_i|| \leq \frac{n}{n+1}\max\{||v_j - v_i|| \mid 0 \leq j \leq n, j \neq i\} = \frac{n}{n+1}\mathrm{diam}(\sigma) = \frac{n}{n+1}d.$$

Hence we have

$$|||v_\sigma - v_\tau|| \leq \frac{n}{n+1}d.$$

As a conclusion, we have

$$d' \leq \frac{n}{n+1}d.$$

□

We then have the following corollary.

Corollary 6.5.10

Given any Euclidean n-simplex σ, for any $\epsilon > 0$, there is an $m \in \mathbb{N}^*$, such that after taking m times barycentric subdivision, all simplices obtained have diameter smaller than ϵ.

Proof. Let d denote the diameter of σ. Let $m \in \mathbb{N}^*$ be such that

$$\left(\frac{n}{n+1}\right)^m d < \epsilon.$$

By the previous lemma, applying m times barycentric subdivision, all n-simplices obtained have diameter smaller than ϵ. □

For any $n \in \mathbb{N}$, we consider the barycentric subdivision of Δ^n, and denote the resulting Euclidean n-simplices by

$$\tau_0, \ldots, \tau_k.$$

There is a chosen orientation on Δ^n for the singular homology. For each τ_i, we consider a linear homeomorphism

$$f_i : \Delta^n \to \tau_i$$

such that the induced orientation on τ_i is the same as the one given by considering τ_i as subspace of Δ^n. We then define for any n-simplex in X, the following n-chain

$$S(\sigma) = \sum_{i=0}^{k} \sigma \circ f_i$$

Then S can be extended to a group homomorphism

$$S : C_n(X) \to C_n(X),$$

for any $n \in \mathbb{N}$. We then have the following proposition.

Proposition 6.5.11

For any $n \in \mathbb{N}$, for any σ an n-chain in $C_n(X)$, there is an $m \in \mathbb{N}^*$, such that

$$S^{(m)}(\sigma) \in C_n(A+B),$$

where $S^{(m)} = \underbrace{S \circ \cdots \circ S}_{m}$.

Remark 6.5.12.
It should be noticed that here the constant m depends on σ.

Another observation on S is that it commute with the boundary homomorphism.

Lemma 6.5.13

We have

$$S \circ \partial = \partial \circ S.$$

Now we would like to show that for any $n \in \mathbb{N}$, for any $\sigma \in Z_n(X)$, there is an n-cycle $\sigma' \in Z_n(A+B)$, such that

$$\sigma - \sigma' \in B_n(X).$$

In particular, we would like to show

$$\sigma - S^{(m)}(\sigma) \in B_n(X),$$

where m is given by the previous proposition. For this purpose, we start construct a chain homotopy T between the chain maps Id and S, where Id is the identity map.

The construction of T is inductive. To make it clear, we discuss what happens to the chains in Euclidean spaces given with simplices defined by linear maps from standard simplices to Euclidean spaces.

Consider the Euclidean space $\mathbb{R}^m$. Let $n \in \mathbb{N}$. we denote by $L_n(\mathbb{R}^m)$ the abelian group of singular n-chains given by singular linear n-simplices, i.e. maps from Δ^n to $\mathbb{R}^m$ which are restrictions of linear maps.

For any linear n-simplex σ in $\mathbb{R}^m$, denote

$$\Delta^n = [v_0, ..., v_n].$$

Choose $b \in \mathbb{R}^m$, then we have an linear $(n+1)$-simplex $b(\sigma)$ given by mapping

$$\Delta^{n+1} = [u_0, ..., u_{n+1}]$$

to $\mathbb{R}^m$ linearly and sending u_0 to b and u_i to $\sigma(v_{i-1})$ for $i > 0$.

This gives us a homomorphism

$$b : L_n(\mathbb{R}^m) \to L_{n+1}(\mathbb{R}^m)$$

Lemma 6.5.14

We have

$$b \circ \partial + \partial \circ b = \mathrm{Id}.$$

Proof. For any $\sigma \in L_n(\mathbb{R}^m)$ a linear n-simplex, we have

$$b(\partial\sigma) = \sum_{i=0}^{n}(-1)^i b(\sigma|_{[v_0,\ldots,\widehat{v_i},\ldots,v_n]}) = \sum_{i=0}^{n}(-1)^i b(\sigma)|_{[u_0,\ldots,\widehat{u_{i+1}},\ldots,u_n]}$$

On the other hand

$$\partial(b\sigma) = \sum_{i=1}^{n+1}(-1)^i b(\sigma)|_{u_0,\ldots,\widehat{u_i},\ldots,u_{n+1}} + b(\sigma)|_{[u_1,\ldots,u_{n+1}]}.$$

Since $b(\sigma)|_{[u_1,\ldots,u_{n+1}]} = \sigma$, the lemme is proved. □

Lemma 6.5.15

Let σ be an Euclidean n-simplex with b its barycenter, then we have

$$S(\sigma) = b(S(\partial\sigma)),$$

where S is the barycentric map for the chain complex $(L_n(\mathbb{R}^m))_{n\in\mathbb{N}}$.

Now we consider the following commutative diagram

$$\begin{array}{ccccccccccc}
\cdots \xrightarrow{\partial} & L_{n+1}(\mathbb{R}^m) & \xrightarrow{\partial} & L_n(\mathbb{R}^m) & \xrightarrow{\partial} \cdots \xrightarrow{\partial} & L_1(\mathbb{R}^m) & \xrightarrow{\partial} & L_0(\mathbb{R}^m) & \xrightarrow{\partial} 0 \\
& \downarrow S & & \downarrow S & & \downarrow S & & \downarrow S & \\
\cdots \xrightarrow{\partial} & L_{n+1}(\mathbb{R}^m) & \xrightarrow{\partial} & L_n(\mathbb{R}^m) & \xrightarrow{\partial} \cdots \xrightarrow{\partial} & L_1(\mathbb{R}^m) & \xrightarrow{\partial} & L_0(\mathbb{R}^m) & \xrightarrow{\partial} 0
\end{array}$$

Now we define the chain map

$$T_n : L_n(\mathbb{R}^m) \to L_{n+1}(\mathbb{R}^m)$$

inductively.

For $n = 0$, we define $T_0 : L_n(\mathbb{R}^m) \to L_{n+1}(\mathbb{R}^m)$ by sending each 0-simplex σ to $b_\sigma(\sigma)$ where b_σ is the barycenter of σ. For example, if $\sigma(v_0) = p \in \mathbb{R}^m$, we have $b_\sigma(\sigma)$ a linear map from $[v_0, v_1]$ to $\mathbb{R}^m$, such that $\sigma(v_0) = \sigma(v_1) = p$.

Then for any $n \in \mathbb{N}$, for any n-simplex $\sigma \in L_n(\mathbb{R}^m)$, we define

$$T(\sigma) = b_\sigma(\sigma - T(\partial(\sigma))),$$

where b_σ is the barycenter of σ. Then we have the following proposition

Proposition 6.5.16

We have

$$\mathrm{Id} - S = T \circ \partial + \partial \circ T.$$

Proof. We show it by induction on n.

When $n = 0$, for any 0-simplex σ, we have $\partial\sigma = 0$. Consider

$$(\partial \circ T)(\sigma) = \sigma - \sigma = \mathrm{Id}(\sigma) - S(\sigma) = 0.$$

Assume that $n \in \mathbb{N}$ and the identity holds for any k with $0 \leq k < n-1$. It is enough to show it for an n-simplex $\sigma \in L_n(\mathbb{R}^m)$, we have

$$\sigma - S(\sigma) = T(\partial(\sigma)) + \partial(T(\sigma)).$$

Notice that

$$\begin{aligned}\partial(T(\sigma)) &= \partial(b_\sigma(\sigma - T(\partial(\sigma)))) \\ &= (\mathrm{Id} - b_\sigma \circ \partial)(\sigma - T(\partial(\sigma))) \\ &= \sigma - T(\partial(\sigma)) - b_\sigma(\partial\sigma) + (b_\sigma \circ \partial \circ T \circ \partial)(\sigma) \\ &= \sigma - T(\partial(\sigma)) - b_\sigma(\partial\sigma) + (b_\sigma \circ (\mathrm{Id} - S - T \circ \partial))(\partial\sigma) \\ &= \sigma - T(\partial(\sigma)) - b_\sigma(\partial\sigma) + b_\sigma(\partial\sigma) - b_\sigma(S(\partial\sigma)) - T(\partial^{(2)}\sigma) \\ &= \sigma - T(\partial(\sigma)) - b_\sigma(S(\partial\sigma)) \\ &= \sigma - T(\partial(\sigma)) - S(\sigma).\end{aligned}$$

Here the second identity comes from

$$b \circ \partial + \partial \circ b = \mathrm{Id},$$

the fourth identity comes from the induction, the sixth identity comes from $\partial^{(2)} = 0$, and the last identity comes from

$$S(\sigma) = b_\sigma(S(\sigma)).$$

□

Now we consider simplices in X. Given any n-simplex σ in X, whatever happen to Δ^n (as an linear n-simplex) could be translate to σ by taking composition. Consider the following commutative diagram

$$\begin{array}{ccccccccccccc} \cdots & \xrightarrow{\partial} & C_{n+1}(X) & \xrightarrow{\partial} & C_n(X) & \xrightarrow{\partial} & \cdots & \xrightarrow{\partial} & C_1(X) & \xrightarrow{\partial} & C_0(X) & \xrightarrow{\partial} & 0 \\ & & \downarrow S & & \downarrow S & & & & \downarrow S & & \downarrow S & & \\ \cdots & \xrightarrow{\partial} & C_{n+1}(X) & \xrightarrow{\partial} & C_n(X) & \xrightarrow{\partial} & \cdots & \xrightarrow{\partial} & C_1(X) & \xrightarrow{\partial} & C_0(X) & \xrightarrow{\partial} & 0 \end{array}$$

For any $n \in \mathbb{N}$, we also have

$$T : C_n(X) \to C_{n+1}(X).$$

The relation

$$\mathrm{Id} - S = T \circ \partial + \partial \circ T,$$

still holds.

Proposition 6.5.17

For any $n \in \mathbb{N}$, for any $\alpha \in Z_n(X)$, for any $m \in \mathbb{N}^*$, we have

$$[S^{(m)}(\alpha)] = [\alpha] \in H_n(X),$$

where $S^{(m)} = \underbrace{S \circ \cdots \circ S}_{m}$.

Corollary 6.5.18

For any $n \in \mathbb{N}$, for any $\alpha \in Z_n(X)$, there is an n-cycle $\beta \in Z_n(A + B)$, such that

$$[\alpha] = [\beta] \in H_n(X).$$

Hence we have the following theorem

Theorem 6.5.19

For any $n \in \mathbb{N}$, the homomorphism

$$H_n(A+B) \to H_n(X),$$

induced by the inclusion $C_n(A+B) \to C_n(X)$ is an isomorphism.

Remark 6.5.20.
Roughly speaking this theorem says that to construct the homology group, we can use only small simplices. This theorem can be extended to one for an open cover $\{U_\alpha\}_{\alpha \in \Omega}$ of X

Now we consider the relative homology.

Lemma 6.5.21

For any $n \in \mathbb{N}$, for any $\overline{\alpha} \in Z_n(X,B)$, then for any $m \in \mathbb{N}^*$, we have

$$\overline{S^{(m)}(\alpha)} \in Z_n(X,B),$$

and $[\overline{S^{(m)}(\alpha)}] = [\overline{\alpha}]$ in $H_n(X,B)$ where $S^{(m)} = \underbrace{S \circ \cdots \circ S}_{m}$.

Proof. Let α be a relative n-cycle, then we have

$$\partial\alpha \in C_{n-1}(B).$$

Since $S \circ \partial = \partial \circ S$, we have

$$\partial(S^{(m)}(\alpha)) = S^{(m)}(\partial\alpha) \in C_{n-1}(B).$$

The previous discuss shows that

$$\alpha - S(\alpha) = \partial(T(\sigma)) + T(\partial\alpha).$$

By the definition of T, since $\partial\alpha \in C_{n-1}(B)$, we have

$$T(\partial\alpha) \in C_n(B).$$

Hence $\alpha - S(\alpha)$ is a relative boundary, and we have

$$[\overline{\alpha}] = [\overline{S(\alpha)}].$$

Using induction, we can show that for any $m \in \mathbb{N}^*$, we have

$$[\overline{\alpha}] = [\overline{S^{(m)}(\sigma)}].$$

□

Corollary 6.5.22

For any $n \in \mathbb{N}$, the embedding

$$i_* : H_n(A+B,B) \to H_n(X,B)$$

is surjective.

Lemma 6.5.23

The embedding i_* is also injective.

Proof. For any $n \in \mathbb{N}$, let $\alpha \in C_n(A+B)$ be a n-cycle relative to B. Hence we have

$$\partial\alpha \in C_{n-1}(B).$$

Assume that

$$i_*([\overline{\alpha}]_{A+B,B}) = [\overline{0}]_{X,B}.$$

Then $\alpha \in C_n(X)$ is an n-boundary relative to B. Hence there is an $(n+1)$-chain $\gamma \in C_{n+1}(X)$ and $\beta \in C_n(B)$, such that

$$\alpha = \partial\gamma + \beta.$$

Choose $m \in \mathbb{N}^*$ such that $S^{(m)}(\gamma) \in C_{n+1}(A+B)$, then we have

$$S^{(m)}(\alpha) = \partial(S^{(m)}(\gamma)) + S^{(m)}(\beta)$$

which is an n-boundary relative to B in $C_n(A+B)$. Notice that

$$[\overline{\alpha}]_{A+B,B} = [\overline{S^{(m)}(\sigma)}]_{A+B,B}$$

in $H_n(A+B,B)$. Hence

$$[\overline{\alpha}]_{A+B,B} = [\overline{0}]_{A+B,B} \in H_n(A+B,B).$$

Therefore the homomorphism i_* is injective. □

Combining all discussion above, we proved Theorem 6.5.6.

6.6 Homology of quotient spaces

Topologically, if we do not care about the information in some subspace, we could also consider the quotient space. In this part, we would like to consider a topological space X and its subspace A, and compare the following two homology groups for each $n \in \mathbb{N}$:

$$H_n(X,A) \quad \text{and} \quad H_n(X/A).$$

For technical reason, we assume that A admits a neighborhood U in X, such that A is a strong deformation retraction of U.

The relative homology groups are also homotopy invariant. More precisely, let X and Y be two spaces. Let A and B be subspaces of X and Y respectively. We consider the two pairs (X,A) and (Y,B). A ***morphism between the two pairs*** denoted by

$$f : (X,A) \to (Y,B),$$

is a continuous map

$$f : X \to Y,$$

such that $f(A) \subset B$. Two morphisms f, g between (X,A) and (Y,B) are said to be ***homotopic*** to each other if there is a continuous map

$$H : X \times [0,1] \to Y,$$

such that $H_0 = f$, $H_1 = g$, and for any $0 \leq t \leq 1$, $H_t(A) \subset B$.

A morphism between pairs (X, A) and (Y, B) induces the following commutative diagram for any $n \in \mathbb{N}$:

$$\begin{array}{ccc} C_n(X) & \xrightarrow{f_\#} & C_n(Y) \\ \downarrow & & \downarrow \\ C_n(X,A) & \xrightarrow{f_\#} & C_n(Y,B) \end{array}$$

where the two vertical arrows are given by the quotient maps. This homomorphism $f_\#$ between the two relative chain groups then induces a homomorphism between relative homology groups:

$$f_* : H_n(X, A) \to H_n(Y, B).$$

Proposition 6.6.1

If f and g are two homotopic morphisms between pairs (X, A) and (Y, B), then for any $n \in \mathbb{N}$, we have

$$f_* = g_*,$$

for the relative homology groups.

The proof is essentially the same as the one for the absolute homology groups. We leave it as an exercise.

Let B be a subspace of X containing A. The inclusions:

$$A \subset B \subset X$$

and the morphisms between pairs:

$$i : (B, A) \to (X, A) \quad \text{and} \quad j : (X, A) \to (X, B).$$

These maps induces an exact sequence for any $n \in \mathbb{N}$

$$0 \longrightarrow C_n(B, A) \xrightarrow{i_\#} C_n(X, A) \xrightarrow{j_\#} C_n(X, B) \longrightarrow 0.$$

The exactness comes from again the fundamental theorem of group homomorphism:

$$(C_n(X)/C_n(A))/(C_n(B)/C_n(A)) \cong C_n(X)/C_n(B).$$

Similar to the previous case, we have the following long exact sequence relating $H_n(X, B)$, $H_n(X, A)$ and $H_n(B, A)$.

Proposition 6.6.2

There is a long exact sequence

$$\cdots \longrightarrow H_{n+1}(B, A) \xrightarrow{i_*} H_{n+1}(X, A) \xrightarrow{j_*} H_{n+1}(X, B) \xrightarrow{\delta} H_n(B, A) \longrightarrow \cdots$$

$$\cdots \xrightarrow{j_*} H_0(X, B) \xrightarrow{\delta} 0$$

Then back to out situation, since A is a subspace in X admitting an open neighborhood U which has A as a strong deformation retraction. Using the homotopy invariance of the relative homology, we have

$$H_n(U, A) \cong H_n(A, A) \cong 0.$$

As a consequence of this fact with the above long exact sequence, we have the following proposition.

Proposition 6.6.3

With the same notation X, A, U as above, for any $n \in \mathbb{N}$, we have

$$H_n(X, A) \cong H_n(X, U).$$

Proof. Since for all $n \in \mathbb{N}$, we have

$$H_n(U, A) = 0,$$

from the above long exact sequence, for each $n \in \mathbb{N}$, we have

$$0 \xrightarrow{i_*} H_n(X, A) \xrightarrow{j_*} H_n(X, U) \xrightarrow{\delta} 0.$$

Therefore, we have the isomorphism:

$$H_n(X, A) \cong H_n(X, U).$$

□

The last ingredient is again a consequence of the long exact sequence of the relative homology.

Proposition 6.6.4

For any point $p \in Y$ in a topological space Y, for any $n \in \mathbb{N}^*$, we have

$$H_n(Y, p) \cong H_n(Y).$$

For $n = 0$, the group $H_0(Y, \{p\})$ is generated by simplices in the path connected components of Y not containing p.

Proof. By Theorem 6.5.4, we have the exact sequence for any $n > 2$

$$0 \cong H_n(\{p\}) \xrightarrow{i_*} H_n(Y) \xrightarrow{\mathrm{pr}_*} H_n(Y, \{p\}) \xrightarrow{\partial} H_{n-1}(\{p\}) \cong 0.$$

Hence we have

$$H_n(Y, p) \cong H_n(Y).$$

For $n = 0$ and 1, we consider

$$0 \cong H_1(\{p\}) \xrightarrow{i_*} H_1(Y) \xrightarrow{\mathrm{pr}_*} H_1(Y, \{p\}) \xrightarrow{\partial} H_0(\{p\}) \xrightarrow{i_*} H_0(Y) \xrightarrow{\mathrm{pr}_*} H_0(Y, \{p\}) \xrightarrow{\partial} 0.$$

Since

$$i_* : H_0(\{p\}) \to H_0(Y)$$

is injective, the ∂ on its left has trivial image, and we have

$$H_1(Y) \cong H_1(Y, \{p\}).$$

The last part of the statement for $H_0(Y, \{p\})$ is given by the injectivity of i_* in the following exact sequence

$$H_0(\{p\}) \xrightarrow{i_*} H_0(Y) \xrightarrow{\mathrm{pr}_*} H_0(Y, \{p\}) \xrightarrow{\partial} 0$$

□

Next, we would like to show the following relation.

Theorem 6.6.5

Let X, A and U be as above. For any $n \in \mathbb{N}$, we have

$$H_n(X, A) \cong H_n(X/A, A/A).$$

Proof. Let V be a subspace in U, such that

$$A \subset \mathring{V} \subset \overline{V} \subset U.$$

We use the following sequence of isomorphisms:

$$\begin{aligned} &H_n(X, A) \cong H_n(X, U) \cong H_n(X - V, U - V) \\ \cong &H_n(X/A - V/A, U/A - V/A) \cong H_n(X/A, U/A) \cong H_n(X/A, A/A). \end{aligned}$$

We explain these isomorphisms in order.

The first one is given by Proposition 6.6.3.

The second one is given by Excision Theorem 6.5.6.

The third one is given by Proposition 6.6.1 the homotopy invariance of relative homology, considering the following morphism

$$f : (X - V, U - V) \to (X/A - V/A, U/A - V/A).$$

which is given by a homeomorphism $f : (X - V) \to (X/A - V/A)$, hence is homotopic to the identity morphism between the pairs.

The fourth one is again given by Excision Theorem.

The last one is given by Proposition 6.6.3. □

Corollary 6.6.6

Let X, A and U be as above. For any $n > 0$, we have

$$H_n(X, A) \cong H_n(X/A).$$

Proof. The is given by considering the above theorem and Proposition 6.6.4. □

6.7 Mayer-Vietoris Sequences and some applications

Let X be a topological space. Let A and B be its subspaces, such that

$$\mathring{A} \cup \mathring{B} \cong X.$$

Excision Theorem tells us that the n-th homology group $H_n(X)$ of X can be define use only n-cycles in A or in B.

By considering inclusions of subspaces, we have the following commutative diagram

$$\begin{array}{ccc} A \cap B & \xrightarrow{j_1} & A \\ {\scriptstyle j_2}\downarrow & & \downarrow{\scriptstyle i_1} \\ B & \xrightarrow{i_1} & X \end{array}$$

We also have the inclusion of pairs

$$\iota : (A, A \cap B) \to (X, B).$$

From these, we have the following relation between the two long exact sequences

$$\begin{array}{ccccccccccc}
\cdots \longrightarrow & H_{n+1}(A, A \cap B) & \xrightarrow{\partial} & H_n(A \cap B) & \xrightarrow{(j_1)_*} & H_n(A) & \xrightarrow{\text{pr}_*} & H_n(A, A \cap B) & \xrightarrow{\partial} \cdots \\
& \downarrow{\scriptstyle \iota_*} & & \downarrow{\scriptstyle (j_2)_*} & & \downarrow{\scriptstyle (i_1)_*} & & \downarrow{\scriptstyle \iota_*} & \\
\cdots \longrightarrow & H_{n+1}(X, B) & \xrightarrow{\partial} & H_n(B) & \xrightarrow{(i_2)_*} & H_n(X) & \xrightarrow{\text{pr}_*} & H_n(X, B) & \xrightarrow{\partial} \cdots
\end{array}$$

Notice that the map ι_* is isomorphism by Excision Theorem.

We also have an short exact sequence fo chain complex

$$0 \longrightarrow C_n(A \cap B) \xrightarrow{\varphi} C_n(A) \oplus C_n(B) \xrightarrow{\psi} C_n(A + B) \longrightarrow 0.$$

where

$$\begin{array}{rl} \varphi : C_n(A \cap B) \to C_n(A) \oplus C_n(B) \\ \alpha \mapsto (\alpha, \alpha) \end{array} \quad \text{and} \quad \begin{array}{rl} \psi : C_n(A) \oplus C_n(B) \to C_n(A + B) \\ (\alpha, \beta) \mapsto \alpha - \beta \end{array}.$$

This induces a long exact sequence which is called the ***Mayer-Vietoris sequence***.

Theorem 6.7.1

We have a long exact sequence

$$\cdots \longrightarrow H_n(A \cap B) \xrightarrow{\varphi_*} H_n(A) \oplus H_n(B) \xrightarrow{\psi_*} H_n(A + B) \xrightarrow{\delta} H_{n-1}(A \cap B) \xrightarrow{\varphi_*}$$

$$\cdots \xrightarrow{\psi_*} H_0(A + B) \xrightarrow{\delta} 0$$

where

$$\delta : H_n(A + B) \to H_n(A + B, B) \to H_n(A, A \cap B) \to H_{n-1}(A \cap B),$$

and $\delta = \partial \circ \xi_* \circ \text{pr}_*$, with ξ the natural isomorphism between $C_n(A+B, B)$ and $C_n(A, A\cap B)$.

Proof. We first show that δ is well defined. For any $[z] \in H_n(A + B)$, we have an n-cycle

$$z = x + y \in Z_n(A + B),$$

with $x \in C_n(X)$ and $y \in C_n(Y)$. Hence we have

$$0 = \partial z = \partial x + \partial y.$$

Therefore, we have

$$\partial x = -\partial y \in C_{n-1}(A \cap B).$$

Since $\partial(\partial \alpha) = 0$, we have

$$\partial x \in Z_{n-1}(A \cap B).$$

Hence we have

$$\begin{array}{rcccccc} \delta : H_n(A + B) & \to & H_n(A + B, B) & \to & H_n(A, A \cap B) & \to & H_{n-1}(A \cap B), \\ [x + y]_{A+B} & \mapsto & [\overline{x}]_{A+B,B} & \mapsto & [\overline{x}]_{A,A\cap B} & \mapsto & [\partial x]_{A\cap B}. \end{array}$$

Now we try to show the following three equality

$$\begin{aligned} \operatorname{Im}\varphi_* &= \ker\psi_* \\ \operatorname{Im}\psi_* &= \ker\delta \\ \operatorname{Im}\delta &= \ker\varphi_* \end{aligned}$$

We first consider

$$H_n(A\cap B) \xrightarrow{\varphi_*} H_n(A)\oplus H_n(B) \xrightarrow{\psi_*} H_n(A+B).$$

$\operatorname{Im}\varphi_* \subset \ker\psi_*$: Since $\psi\circ\varphi = 0$, we have

$$\psi_*\circ\varphi_* = (\psi\circ\varphi)_* = 0,$$

hence

$$\operatorname{Im}\varphi_* \subset \ker\psi_*.$$

$\operatorname{Im}\varphi_* \supset \ker\psi_*$: For any $([x]_A, [y]_B) \in H_n(A)\oplus\ker\psi_*$, we have

$$[x-y]_{A+B} = [0]_{A+B}.$$

Hence

$$x - y \in B_n(A+B),$$

or equivalently, there is an $(n+1)$-chain $x_1 \in C_{n+1}(A)$ and $y_1 \in C_{n+1}(B)$, such that

$$x - y = \partial(x_1 + y_1).$$

Hence we have

$$\alpha = x - \partial x_1 = y + \partial y_1 \in C_n(A\cap B).$$

Since $x \in Z_n(A)$, we have

$$\partial\alpha = \partial x = 0.$$

Hence

$$\alpha \in Z_n(A\cap B).$$

We have

$$\varphi_*([\alpha]_{A\cap B}) = ([x]_A, [y]_B).$$

Hence

$$\operatorname{Im}\varphi_* \supset \ker\psi_*.$$

Next we consider

$$H_n(A)\oplus H_n(B) \xrightarrow{\psi_*} H_n(A+B) \xrightarrow{\delta} H_{n-1}(A\cap B)$$

$\operatorname{Im}\psi_* \supset \ker\delta$: Let $z \in \ker\delta$, hence we have $x \in C_n(A)$ and $y \in C_n(Y)$, such that

$$\begin{aligned} z &= x + y \\ \partial(x+y) &= 0 \\ [\delta(x+y)]_{A\cap B} &= [0]_{A\cap B} \end{aligned}$$

By the definition of δ, we have

$$[\partial y]_{A\cap B} = [\partial x]_{A\cap B} = [\delta(x+y)]_{A\cap B} = [0]_{A\cap B}.$$

Hence $x \in Z_n(A)$, $-y \in Z_n(B)$, and we have

$$\psi_*([x]_A, [-y]_B) = [x+y]_{A+B}.$$

Hence we have

$$\operatorname{Im}\psi_* \supset \ker\delta.$$

$\operatorname{Im}\psi_* \subset \ker\delta$: For any $([x]_A, [y]_B) \in H_n(A) \oplus H_n(B)$, we have

$$(\delta \circ \psi_*)([x]_A, [y]_B) = \delta([x-y]_{A+B}) = [\partial x]_{A\cap B} = [0]_{A\cap B}.$$

Hence

$$\operatorname{Im}\psi_* \subset \ker\delta.$$

Finally, we consider

$$H_n(A+B) \xrightarrow{\delta} H_{n-1}(A\cap B) \xrightarrow{\varphi_*} H_{n-1}(A) \oplus H_{n-1}(B)$$

$\operatorname{Im}\delta \subset \ker\varphi_*$: For any $[z]_{A+B} \in H_n(A+B)$, there is an n-chain $x \in C_n(A)$ and $y \in C_n(B)$, such that $z = x + y$. We have

$$\delta([x+y]_{A+B}) = [\partial x]_{A\cap B} = [-\partial y]_{A\cap B}$$

Hence

$$\varphi_*(\delta([x+y]_{A+B})) = ([\partial x]_A, [-\partial y]_B) = ([0]_A, [0]_B).$$

$\operatorname{Im}\delta \supset \ker\varphi_*$: Let $[z]_{A\cap B} \in \ker\varphi_*$, then

$$z \in B_{n-1}(A) \quad \text{and} \quad z \in B_{n-1}(B).$$

Hence there is an n-chain $x \in C_n(A)$ and $y \in C_n(B)$, such that

$$z = \partial x = \partial y.$$

Let $w = x - y \in C_n(A+B)$, we have

$$\partial w = z - z = 0.$$

Hence

$$w \in Z_n(A+B).$$

We consider

$$\delta([w]_{A+B}) = [\partial x]_{A\cap B} = [z]_{A\cap B}.$$

□

In the proof of Excision Theorem, we show the following two isomorphisms

$$\begin{aligned} \eta : H_n(A+B) &\to H_n(X) \\ [\alpha]_{A+B} &\mapsto [\alpha]_X \end{aligned}$$

and

$$\begin{aligned} \rho : H_n(A+B,B) &\to H_n(X,B) \\ [\overline{\alpha}]_{A+B,B} &\mapsto [\overline{\alpha}]_{X,B}. \end{aligned}$$

Therefore the previous exact sequence can be rewritten as the following one

$$\cdots \longrightarrow H_n(A\cap B) \xrightarrow{\varphi_*} H_n(A)\oplus H_n(B) \xrightarrow{\psi_*} H_n(X) \xrightarrow{\widetilde{\delta}} H_{n-1}(A\cap B) \xrightarrow{\varphi_*}$$

$$\cdots \xrightarrow{\psi_*} H_0(X) \xrightarrow{\widetilde{\delta}} 0$$

where

$$\widetilde{\delta} : H_n(X) \to H_n(X,B) \to H_n(A, A\cap B) \to H_{n-1}(A\cap B),$$

and $\delta = \partial \circ \iota_*^{-1} \circ \mathrm{pr}_*$.

Applications:

Next we give some applications of Mayer-Vietoris sequence. The first one is about the computation of the homology group of spheres.

Proposition 6.7.2

For any $k, n \in \mathbb{N}$, we have

$$H_k(S^n) \cong \begin{cases} \mathbb{Z} \oplus \mathbb{Z}, & k = n = 0, \\ \mathbb{Z}, & k = 0, n > 0, \\ \mathbb{Z}, & k = n > 0, \\ 0, & k \neq n, k \neq 0. \end{cases}$$

Proof. Since $S^0 = \{*, *\}$ has two points, hence we have

$$H_k(S^0) \cong H_k(\{*\}) \oplus H_k(\{*\}) \cong \begin{cases} \mathbb{Z} \oplus \mathbb{Z}, & k = 0 \\ 0, & k > 0 \end{cases}.$$

In the following we consider S^n with $n > 0$. Since S^n's for $n > 0$ are all path connected, we have

$$H_0(S^n) \cong \mathbb{Z},$$

for all $n > 0$. Hence in the following we also consider only $k > 0$.

Denote N and S the normal pole and south pole of S^n, and consider

$$A = S^n \setminus \{N\} \quad \text{and} \quad B = S^n \setminus \{S\}.$$

We have the following homotopy equivalence

$$A \sim B \sim D^n \quad \text{and} \quad A \cap B \sim S^{n-1}.$$

We consider the following part of the Mayer-Vietoris sequence

$$H_k(A) \oplus H_k(B) \xrightarrow{\psi_*} H_k(S^n) \xrightarrow{\widetilde{\delta}} H_{k-1}(A \cap B) \xrightarrow{\varphi_*} H_{k-1}(A) \oplus H_{k-1}(B)$$

Since $n > 0$, we have

$$H_k(A) \cong H_k(B) \cong H_k(D^n) \cong \begin{cases} \mathbb{Z}, & k = 0 \\ 0, & k > 0 \end{cases}.$$

For $k > 1$, the above sequence becomes

$$0 \xrightarrow{\psi_*} H_k(S^n) \xrightarrow{\widetilde{\delta}} H_{k-1}(S^{n-1}) \xrightarrow{\varphi_*} 0$$

By the exactness, we have the following isomorphism

$$H_k(S^n) \cong H_{k-1}(S^{n-1}).$$

For $k = 1$, we consider the following part of the long exact sequence

$$H_1(A) \oplus H_1(B) \xrightarrow{\psi_*} H_1(S^n) \xrightarrow{\widetilde{\delta}} H_0(A \cap B) \xrightarrow{\varphi_*} H_0(A) \oplus H_0(B) \xrightarrow{\psi_*} H_0(S^n) \longrightarrow 0$$

If $n = 1$, equivalently, we have

$$0 \xrightarrow{\psi_*} H_1(S^1) \xrightarrow{\widetilde{\delta}} \mathbb{Z} \oplus \mathbb{Z} \xrightarrow{\varphi_*} \mathbb{Z} \oplus \mathbb{Z} \xrightarrow{\psi_*} \mathbb{Z} \longrightarrow 0$$

Since the exactness at $H_0(S^1)$ (the $\mathbb{Z}$ on the most right) shows that

$$\psi_* : \mathbb{Z} \oplus \mathbb{Z} \to \mathbb{Z},$$

is surjective. Hence we have $\ker \psi_* \cong \mathbb{Z}$. This can be obtained by considering both $\mathbb{Z} \oplus \mathbb{Z}$ and $\mathbb{Z}$ as $\mathbb{Z}$-free modules. The kernal of ψ_* is a submodule of $\mathbb{Z} \oplus \mathbb{Z}$, hence a $\mathbb{Z}$-free module of rank at most 1. If it has rank 2, then the image should be 0 or a torsion module. If it has rank 0, then the image of ψ_* should be $\mathbb{Z} \oplus \mathbb{Z}$. Hence both are impossible. Hence it should be rank 1, i.e. isomorphic to $\mathbb{Z}$ as $\mathbb{Z}$-module and as group as well.

The exactness at $\mathbb{Z} \oplus \mathbb{Z}$ on the right shows that

$$\operatorname{Im} \varphi_* = \ker \psi_* \cong \mathbb{Z}.$$

A similar reason as above shows that

$$\ker \varphi_* \cong \mathbb{Z}.$$

The exactness at $\mathbb{Z} \oplus \mathbb{Z}$ on the left shows that

$$\operatorname{Im} \widetilde{\delta} \cong \ker \varphi_* \cong \mathbb{Z}.$$

The exactness at $H_1(S^1)$ shows that $\widetilde{\delta}$ is injective. Hence as a conclusion, we have

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

If $n > 1$, the exact sequence becomes

$$H_1(A) \oplus H_1(B) \xrightarrow{\psi_*} H_1(S^n) \xrightarrow{\widetilde{\delta}} \mathbb{Z} \xrightarrow{\varphi_*} \mathbb{Z} \oplus \mathbb{Z} \xrightarrow{\psi_*} \mathbb{Z} \longrightarrow 0$$

A similar discussion using the exactness at each position from the right to left shows that

$$H_1(S^n) \cong 0.$$

With all information obtained above, if $k > n$, we have

$$H_k(S^n) \cong H_{k-1}(S^{n-1}) \cong \cdots \cong H_{k-n}(S^0) \cong 0.$$

If $k < n$, we have

$$H_k(S^n) \cong H_{k-1}(S^{n-1}) \cong \cdots \cong H_1(S^{n-k}) \cong 0.$$

If $k = n$, we have

$$H_n(S^n) \cong \cdots \cong H_1(S^1) \cong \mathbb{Z}.$$

□

Using this result, we can get some results which are closer to our initial goal: classification of spaces in a topological way. This also shows that the homology groups can be used for this purpose.

Corollary 6.7.3

Let m and n be two distinct natural numbers.

1) The sphere S^m is not contractible.

2) The spheres S^m and S^n are not homotopy equivalent.

Proof. These are a consequence of invariance of homology groups.

For $m = 0$, the sphere S^0 is not path connected, hence is not contractible. For $m > 0$, we have

$$H_m(S^m) \cong \mathbb{Z},$$

which is different from the m-th homology group for a contractible space which is trivial. Hence S^m is not contractible for any m.

For $m \neq n$, without loss of generality, we may assume that $m < n$, then

$$H_n(S^m) \cong 0 \ncong \mathbb{Z} \cong H_n(S^n).$$

Hence S^m and S^n are not homotopy equivalent to each other. □

Remark 6.7.4.
Therefore spheres in different dimension are not homeomorphic to each other.

The above result may be not strange, but it seems impossible to prove it directly by simply constructing maps and playing with definition. With this result, we can also compare Euclidean spaces which seems also a mission impossible at the first glance.

Corollary 6.7.5

For two positive natural numbers $m \neq n$, the spaces $\mathbb{R}^m$ and $\mathbb{R}^n$ are not homeomorphic to each other.

Proof. Assume there is a homeomorphism

$$f : \mathbb{R}^m \to \mathbb{R}^n.$$

This induces a homeomorphism

$$f : \mathbb{R}^m \setminus \{O\} \to \mathbb{R}^n \setminus \{f(O)\},$$

where O is the origin of $\mathbb{R}^m$.

Notice that $\mathbb{R}^m \setminus \{O\}$ is homotopy equivalent to S^{m-1} while $\mathbb{R}^n \setminus \{f(O)\}$ is homotopy equivalent to S^{n-1}. Hence we have S^{m-1} and S^{n-1} homotopy equivalent to each other, which is a contradiction. □

In certain case, we can also show certain properties of continuous maps by comparing homology groups.

Corollary 6.7.6 (Brouwer Fixed Point Theorem)

Let $n \in \mathbb{N} \setminus \{0, 1\}$. Any continuous map from D^n to itself has a fixed point.

Proof. If not, there is a continuous map

$$f : D^n \to D^n,$$

has no fixed point.

Then for any $p \in D^2$, we have

$$f(p) \neq p.$$

We consider the ray

$$R(p) = \{f(p) + t(p - f(p)) \mid t \in \mathbb{R}_{\geq 0}\},$$

and denote

$$r(p) = R(p) \cap S^{n-1}.$$

This gives a map

$$r : D^n \to S^{n-1},$$

which is identity on S^{n-1} and can be used to show that S^{n-1} is a deformation retraction of D^n. Hence for each $m \in \mathbb{N}$, we have

$$H_m(D^n) \cong H_m(S^{n-1}),$$

which is impossible, since

$$H_{n-1}(D^n) \cong 0 \ncong \mathbb{Z} \cong H_{n-1}(S^{n-1}).$$

□

6.8 Equivalence between simplicial homology and singular homology

Let X be a topological space with a simplicial complex structure. For simplicity, we assume that X is finite dimensional, i.e. there is an upper bound on the dimension of simplices in the simplicial complex structure.

One could define the simplicial homology groups with respect to this simplicial complex structure and the singular homology groups. We denote by $(C_n^\Delta(X), \partial^\Delta)_{n_\mathbb{N}}$ the simplicial chain complex, and by $(C_n(X), \partial)_{n\in\mathbb{N}}$ be the singular chain complex. Notice that a simplex in the simplicial complex is also a singular simplex, hence for any $n \in \mathbb{N}$, we have a natural map

$$\varphi_n : C_n^\Delta(X) \to C_n(X).$$

Moreover taking the boundary of a simplex in a simplicial complex is defined in the same way as when considering it as a singular simplex. Hence these $(\varphi_n)_{n\in\mathbb{N}}$ are chain maps between the two chain complex:

$$\begin{array}{ccccccccccc}
\cdots \longrightarrow & C_{n+1}^\Delta(X) & \xrightarrow{\partial^\Delta} & C_n^\Delta(X) & \xrightarrow{\partial^\Delta} & C_{n-1}^\Delta(X) & \xrightarrow{\partial^\Delta} \cdots \xrightarrow{\partial^\Delta} & C_0^\Delta(X) & \xrightarrow{\partial^\Delta} & 0 \\
& \downarrow{\scriptstyle \varphi_{n+1}} & & \downarrow{\scriptstyle \varphi_n} & & \downarrow{\scriptstyle \varphi_{n-1}} & & \downarrow{\scriptstyle \varphi_0} & & \\
\cdots \longrightarrow & C_{n+1}(X) & \xrightarrow{\partial} & C_n(X) & \xrightarrow{\partial} & C_{n-1}(X) & \xrightarrow{\partial} \cdots \xrightarrow{\partial} & C_0(X) & \xrightarrow{\partial} & 0
\end{array}$$

Hence we have the homomorphisms in the homology groups level: for any $n \in \mathbb{N}$,

$$(\varphi_n)_* : H_n^\Delta(X) \to H_n(X).$$

To compare the two kinds of homology groups, it is enough to understand these homomorphism $(\varphi_n)_*$'s. In particular, we would like to show the following theorem.

Theorem 6.8.1

For any $n \in \mathbb{N}$, the group homomorphism defined above

$$(\varphi_n)_* : H_n^\Delta(X) \to H_n(X),$$

is an isomorphism.

Remark 6.8.2.
Before going into the proof, let us first make some remarks. First, in order to let this theorem make sense, the space X should be able to be equipped with a simplicial complex structure which is not true for any topological space. On the other hand, the definition of singular simplices are more flexible, since we only consider continuous maps.

Secondly, given a space with a simplicial complex structure, this theorem says that the simplicial homology and the singular homology are the same. Notice that the singular homology depends only on the topological structure on X, and is independent of the choice of any additional structure on it. Moreover, it is a immediate consequence of its definition that the singular homology groups are invariant under homeomorphism. Previously we have shown that it is even homotopy invariant. On the other hand, the simplicial homology groups depends on the given simplicial complex structure. From the previous discussion on the triangulations of surfaces, we have seen that the simplicial complex structure may be not unique. Hence the above theorem shows that the simplicial homology is independent of choice of simplicial complex structure.

Lastly, the singular homology groups is independent of choice of simplicial complex structure, but the cost is that we have to consider lots of singular simplices. The chain groups are too large, which makes the direct computation impossible. Although we have the homotopy invariance, the long exact sequence for relative singular homology and the Excision Theorem, it is still quite complicated to do computation. On the other hand, the simplicial homology are defined in a much simpler and geometric way, which makes the direct computation doable. Remember one reason that we would like to study the homology group is to use it as an invariant to tell different spaces.

The tool used in the proof is the relative homology. The definition of relative homology can be extends to the simplicial homology naturally. More precisely, for any $A \subset X$ a Δ-subcomplex, we can define for any $n \in \mathbb{N}$, the group of relative n-chains

$$C_n^\Delta(X,A) := C_n^\Delta(X)/C_n^\Delta(A).$$

Then the relative simplicial homology $H_n^\Delta(X,A)$ is the homology groups for the relative simplicial chain complex $(C_n^\Delta(X,A),\partial)_{n\in\mathbb{N}}$. We also have the long exact sequence for relative simplicial homology in this case. In particular, the above relation between the two homology groups induces the following commutative diagram relating the two long exact sequence together:

$$\begin{array}{ccccccccccccc}
\cdots \xrightarrow{\mathrm{pr}_*^\Delta} & H_{n+1}^\Delta(X,A) & \xrightarrow{\partial^\Delta} & H_n^\Delta(A) & \xrightarrow{i_*^\Delta} & H_n^\Delta(X) & \xrightarrow{\mathrm{pr}_*^\Delta} & H_n^\Delta(X,A) & \xrightarrow{\partial^\Delta} & H_{n-1}^\Delta(A) & \xrightarrow{i_*^\Delta} \cdots \\
& \downarrow{\scriptstyle(\varphi_{n+1})_*} & & \downarrow{\scriptstyle(\varphi_n)_*} & & \downarrow{\scriptstyle(\varphi_n)_*} & & \downarrow{\scriptstyle(\varphi_n)_*} & & \downarrow{\scriptstyle(\varphi_{n-1})_*} & \\
\cdots \xrightarrow{\mathrm{pr}_*} & H_{n+1}(X,A) & \xrightarrow{\partial} & H_n(A) & \xrightarrow{i_*} & H_n(X) & \xrightarrow{\mathrm{pr}_*} & H_n(X,A) & \xrightarrow{\partial} & H_{n-1}(A) & \xrightarrow{i_*} \cdots
\end{array}$$

For any $k \in \mathbb{N}$, we define the k-skeleton of X is the union of simplces in X with dimension k or less, and denote it by X^k. Hence it is a Δ-subcomplex of X. For any $n \in \mathbb{N}$, we define

$$C_n^\Delta(X^k,X^{k-1}) := C_n^\Delta(X^k)/C_n^\Delta(X^{k-1}),$$

with the convention that $X^{-1} = \emptyset$.

If $n \neq k, k-1$, then there is no n-simplex in X^k. Hence

$$C_n^\Delta(X^k,X^{k-1}) \cong 0.$$

If $n = k$, then there is no n-simplex in X^{k-1}, hence

$$C_k^\Delta(X^k,X^{k-1}) \cong \bigoplus_{\alpha\in\Omega} \mathbb{Z}_\alpha,$$

where Ω is the collection of k-simplices in X and $\mathbb{Z}_\alpha$ is the copy of $\mathbb{Z}$ associated to the k-simplex α.

If $n = k-1$, then X^k and X^{k-1} have the same collection of $(k-1)$-simplices. Hence

$$C^\Delta_{k-1}(X^k, X^{k-1}) \cong 0.$$

Hence we have

$$H^\Delta_n(X^k, X^{k-1}) \cong \begin{cases} \bigoplus_{\alpha\in\Omega} \mathbb{Z}_\alpha, & n = k; \\ 0, & n \neq k. \end{cases}$$

For each $n \in \mathbb{N}$, consider the singular relative homology groups $H_n(X^k, X^{k-1})$. Now we would like to discuss its relation with $H^\Delta_n(X^k, X^{k-1})$.

For any $k \in \mathbb{N}^*$, we consider the following commutative diagram for pairs of spaces

$$\begin{array}{ccccc} \left(\coprod_{\alpha\in\Omega}\Delta^k_\alpha, \coprod_{\alpha\in\Omega}\partial\Delta^k_\alpha\right) & \xrightarrow{f_1} & \left(\coprod_{\alpha\in\Omega}\Delta^k_\alpha, \coprod_{\alpha\in\Omega}(\Delta^k\setminus\{b_\alpha\})\right) & \xleftarrow{f_2} & \left(\coprod_{\alpha\in\Omega}\mathring{\Delta}^k_\alpha, \coprod_{\alpha\in\Omega}(\mathring{\Delta}^k\setminus\{b_\alpha\})\right) \\ \downarrow g_1 & & \downarrow g_2 & & \downarrow g_3 \\ (X^k, X^{k-1}) & \xrightarrow{h_1} & \left(X^k, X^k\setminus\coprod_{\alpha\in\Omega}\{\sigma_\alpha(b_\alpha)\}\right) & \xleftarrow{h_2} & \left(X^k\setminus X^{k-1}, (X^k\setminus X^{k-1})\setminus\coprod_{\alpha\in\Omega}\{\sigma_\alpha(b_\alpha)\}\right) \end{array}$$

Notice that g_3 is a homeomorphism between the two pairs, hence it induces isomorphism between the relative homology groups. The maps f_2 and h_2 also induces isomorphisms between the relative homology groups by Excision Theorem. Hence the map g_2 also induces an isomorphism between relative homology groups.

To see the maps f_1 and h_1 induce isomorphisms between relative homology groups, we consider the following fact.

Lemma 6.8.3

Let X be a topological space with subspaces A and U satisfying the inclusion relation:

$$U \subset A \subset X,$$

such that U is a deformation retraction of A, then for any $n \in \mathbb{N}$, we have

$$H_n(X, A) \cong H_n(X, U).$$

We first recall the five lemma for abelian groups.

Lemma 6.8.4 (Five Lemma for Abelian Groups)

Consider the following commutative diagram of abelian groups

$$\begin{array}{ccccccccc} A & \xrightarrow{i} & B & \xrightarrow{j} & C & \xrightarrow{k} & D & \xrightarrow{l} & E \\ \downarrow\alpha & & \downarrow\beta & & \downarrow\gamma & & \downarrow\delta & & \downarrow\epsilon \\ A' & \xrightarrow{i'} & B' & \xrightarrow{j'} & C' & \xrightarrow{k'} & D' & \xrightarrow{l'} & E' \end{array}$$

Assume that the first row is exact at B, C and D, the second row is exact at B', C' and E', and the homomorphisms α, β, δ and ϵ are isomorphisms. Then γ is an isomorphism.

Proof of Five Lemma. We first show that γ is surjective.

For any $c' \in C'$, denote
$$d' = k'(c') \in \operatorname{Im} k' = \ker l'.$$
Hence we have
$$l'(d') = 0 \in E'.$$
Since δ is an isomorphism, there is $d \in D$, such that
$$\delta(d) = d'.$$
Then since
$$\epsilon(l(d)) = l'(\delta(d)) = l'(d') = 0,$$
and ϵ is an isomorphism, we have
$$l(d) = 0,$$
hence
$$d \in \ker l = \operatorname{Im} k,$$
and there is an element $c \in C$, such that
$$k(c) = d.$$
Hence
$$k'(\gamma(c)) = \delta(k(c)) = d',$$
and
$$k'(c' - \gamma(c)) = 0.$$
Therefore, we have
$$c' - \gamma(c) \in \ker k' = \operatorname{Im} j'$$
and there is an element $b' \in B'$, such that
$$j'(b') = c' - \gamma(c).$$
Since β is isomorphism, there is an element $b \in B$, such that
$$\beta(b) = b'.$$
Hence we have
$$\gamma(j(b)) = j'(\beta(b)) = c' - \gamma(c).$$
Therefore
$$c' = \gamma(j(b) + c) \in \operatorname{Im} \gamma.$$
The homomorphism γ is surjective.

Now we would like to show that γ is injective. It is enough to show that
$$\ker \gamma = \{0\}.$$

Let $c \in \ker \gamma$. Then we have
$$\delta(k(c)) = k'(\gamma(c)) = k'(0) = 0.$$
Since δ is an isomorphism, we have
$$k(c) = 0.$$
Hence
$$c \in \ker k = \operatorname{Im} j,$$

and there is an element $b \in B$, such that

$$j(b) = c.$$

Now consider

$$j'(\beta(b)) = \gamma(j(c)) = 0.$$

We have

$$b' = \beta(b) \in \ker j' = \operatorname{Im} i',$$

and there is an element $a' \in A'$, such that

$$i'(a') = b'.$$

Since α is an isomorphism, there is an element $a \in A$, such that

$$a' = \alpha(a).$$

Notice that

$$i(a) = \beta^{-1}(i'(\alpha(a))) = b,$$

we have

$$b \in \operatorname{Im} i = \ker j.$$

Hence

$$c = j(b) = 0.$$

We therefore have

$$\ker \gamma = \{0\}.$$

□

Proof of Lemma 6.8.3. Let $\iota : U \to A$ be the inclusion map, and $r : A \to U$ be the retraction such that $\iota \circ r \cong \mathrm{id}_A$. By the homotopy invariant of singular homology, we have for any $n \in \mathbb{N}$

$$H_n(A) \cong H_n(U).$$

Let $X = \sqcup_{\alpha\in\Omega} X_\alpha$ be the path connected component decomposition of X. Let Ω_U (resp. Ω_A) be the collection of indices α such that $U \cap X_\alpha$ (resp. $A \cap X$) is not empty. Since U is a deformation retraction of A, we have $\Omega_U = \Omega_A$. Hence for $n = 0$, we have

$$H_0(X, U) \cong \bigoplus_{\alpha\in\Omega_U} \mathbb{Z}_\alpha \cong H_0(X, A).$$

Now let $n > 0$. For any n-chain $\alpha \in C_n(X)$, if $[\alpha]_U \in Z_n(X, U)$, then there is an $(n-1)$-chain $\beta \in C_{n-1}(U)$, such that $\partial\alpha = \beta$. Since $C_{n-1}(U) < C_{n-1}(A)$, we have

$$[\alpha]_A \in Z_n(X, A).$$

If $[\alpha]_U \in B_n(X, U)$, we have an $(n+1)$-chain $\beta \in C_{n+1}(X)$, such that

$$[\alpha]_U = \partial[\beta]_U = [\partial\beta]_U.$$

Hence there is an n-chain $\gamma \in C_n(U)$, such that $\alpha = \partial\beta + \gamma$. Since $\gamma \in C_n(A)$, we have $[\alpha]_A \in B_n(X, A)$. Hence the inclusions $U \subset A \subset X$ induce a group homomorphism

$$\varphi : H_n(X, U) \to H_n(X, A).$$

Now consider the long exact sequence for relative homology and have the following commutative diagram, we have

$$\begin{array}{ccccccccc}
H_{n+1}(U) & \xrightarrow{(j_U)_*} & H_n(X) & \xrightarrow{(\mathrm{pr}_U)_*} & H_n(X,U) & \xrightarrow{\delta_U} & H_n(U) & \xrightarrow{(j_U)_*} & H_{n-1}(X^{k-1}) \\
\downarrow \iota_* & & \downarrow (\mathrm{id}_X)_* & & \downarrow \varphi & & \downarrow \iota_* & & \downarrow (\mathrm{id}_X)_* \\
H_{n+1}(A) & \xrightarrow{(j_A)_*} & H_n(X) & \xrightarrow{(\mathrm{pr}_A)_*} & H_n(X,A) & \xrightarrow{\delta_A} & H_n(A) & \xrightarrow{(j_A)_*} & H_{n-1}(X)
\end{array}$$

By the Five lemma, we have φ an isomorphism. □

Using this lemma, the map f_1 and h_1 induce isomorphisms in between relative homology groups in each dimension. For any $n \in \mathbb{N}$, we have

$$H_n(X^k, X^{k-1}) \cong H_n\left(\coprod_{\alpha\in\Omega} \Delta^k_\alpha, \coprod_{\alpha\in\Omega} \partial\Delta^k_\alpha\right) \cong \bigoplus_{\alpha\in\Omega} H_n(\Delta^k_\alpha, \partial\Delta^k_\alpha)$$

We have the homeomorphism between the pairs $(\Delta^k, \partial\Delta^k)$ and $(D^k, \partial D^k)$. In the homework using the long exact sequence for relative homology groups, for any $k, n \in \mathbb{N}$, we have

$$H_n(D^{k+1}, S^k) \cong \begin{cases} H_{n-1}(S^k), & n \geq 2; \\ 0, & n = 1, k \geq 1; \\ \mathbb{Z}, & n = 1, k = 0; \\ 0, & n = 0. \end{cases}$$

Let $n = k+1$. If $k = 0$, then we have

$$H_0(D^1, S^0) \cong \mathbb{Z}.$$

If $k > 0$, then $n \geq 2$, and we have

$$H_{k+1}(D^{k+1}, S^k) \cong H_k(S^k) \cong \mathbb{Z}.$$

Hence for $k \geq 1$, we have

$$H_k(X^k, X^{k-1}) \cong H^\Delta_k(X^k, X^{-1}).$$

Moreover, from the above discussion, we can check that this isomorphism is given by

$$(\varphi_n)_* : H^\Delta_k(X^k, X^{-1}) \to H_k(X^k, X^{-1}).$$

Now we start to give the proof of Theorem 6.8.1.

Proof of Theorem 6.8.1. Notice that there is an inclusion relation among all k-skeleton's of X

$$X^0 \subset X^1 \subset \cdots \subset X^k \subset \cdots.$$

Since we assume that X is finite dimensional, there is $k \in \mathbb{N}$, such that

$$X = X^k.$$

The proof is by induction on k.

Consider $k = 0$, the 0-skeleton X^0 is a collection of points. Hence we have

$$H_n(X^0) \cong \begin{cases} \bigoplus_{p\in X^0} \mathbb{Z}[p]^\Delta, & n = 0; \\ 0, & n > 0. \end{cases}$$

On the simplicial side, by its definition, we have

$$H_n^\Delta(X^0) \cong \begin{cases} \bigoplus_{p\in X^0} \mathbb{Z}[p], & n = 0; \\ 0, & n > 0. \end{cases}$$

Hence we only have to check the case when $n = 0$. Notice that a singular 0-simplex is also a simplicial 0-simplex. Hence for any $p \in X^0$, the homomorphism $(\varphi_0)_*$ sends each $[p]^\Delta \in H_0^\Delta(X^0)$ to $[p] \in H_0(X^0)$. As a result, we have

$$H_0(X^0) \cong H_0^\Delta(X^0).$$

Now assume that $k > 0$ and for any $0 \le i \le k-1$, and for any $n \in \mathbb{N}$, the homomorphism

$$(\varphi_n)_* : H_n^\Delta(X^i) \to H_n(X^i),$$

is an isomorphism. We consider the commutative diagram

$$\begin{array}{ccccccccc} H_{n+1}^\Delta(X^k, X^{k-1}) & \xrightarrow{\partial^\Delta} & H_n^\Delta(X^{k-1}) & \xrightarrow{i_*^\Delta} & H_n^\Delta(X^k) & \xrightarrow{\mathrm{pr}_*^\Delta} & H_n^\Delta(X^k, X^{k-1}) & \xrightarrow{\partial^\Delta} & H_{n-1}^\Delta(X^{k-1}) \\ \downarrow{\scriptstyle (\varphi_{n+1})_*} & & \downarrow{\scriptstyle (\varphi_n)_*} & & \downarrow{\scriptstyle (\varphi_n)_*} & & \downarrow{\scriptstyle (\varphi_n)_*} & & \downarrow{\scriptstyle (\varphi_{n-1})_*} \\ H_{n+1}(X^k, X^{k-1}) & \xrightarrow{\partial} & H_n(X^{k-1}) & \xrightarrow{i_*} & H_n(X^k) & \xrightarrow{\mathrm{pr}_*} & H_n(X^k, X^{k-1}) & \xrightarrow{\partial} & H_{n-1}(X^{k-1}) \end{array}$$

Here the maps associated to vertical arrows are all isomorphism except the middle one. From left to right, the first and the fourth are because of the previous discussion on relative homology groups. The second and the fifth are because of the induction condition.

Using Five Lemma, we may conclude that

$$(\varphi_n)_* : H_n^\Delta(X^k) \to H_n(X^k),$$

is an isomorphism.

By induction, we have the theorem. □

From this, we have the following immediate corollaries.

Corollary 6.8.5

The simplicial homology groups of X given by different simplicial complex structures are isomorphic to each other. In other words, the isomorphism type of the homology groups is independent of choice of simplicial complex structure.

Corollary 6.8.6

For any $n \in \mathbb{N}$, if X has a simplicial complex structure with finitely many n-simplices, then $H_n(X)$ is finitely generated.

6.9 First homology and fundamental group

In the study of surface, we use the abelianizations of the fundamental groups of surfaces to give the classification of surfaces. We call it the first homology group of the surface. In fact the abelianization of the fundamental group of a surface is indeed isomorphic to the first homology group of the same surface. This is moreover a general fact.

More precisely, let X be a path connected space, and p be a base point. Recall that a path in X is a continuous map

$$\alpha : [0,1] \to X.$$

By definition, this is also a singular 1-simplex in X.

For any $\alpha \in \mathcal{L}(X,p)$, we have

$$\alpha(0) = \alpha(1),$$

hence $\partial\alpha = 0$ and we have

$$\alpha \in Z_1(X).$$

We have a natural map from $\mathcal{L}(X,p)$ to $H_1(X)$.

In order to have a map from $\pi_1(X,p)$ to $H_1(X)$, we should prove that the image is invariant under path homotopy. We consider α' be another loop in $\mathcal{L}(X,p)$ homotopic to α. Let us denote the homotopy by

$$H : [0,1] \times [0,1] \to X,$$

with $H_0 = \alpha$ and $H_1 = \alpha'$.

Consider the square $[0,1] \times [0,1]$ with

$$u_0 = (0,0),\ u_1 = (1,0),\ u_2 = (1,1),\ u_3 = (0,1).$$

There is a triangulation of $[0,1] \times [0,1]$ by adding the diagonal u_0u_2. The restriction to each triangle gives a singular 2-simplex in X, and we denote them by

$$\sigma_1 = H|_{[u_0,u_1,u_2]} \quad \text{and} \quad \sigma_2 = H|_{[u_0,u_2,u_3]}.$$

Let

$$\sigma = \sigma_1 + \sigma_2.$$

Then we have

$$\begin{aligned}\partial\sigma &= \partial H|_{[u_0,u_1,u_2]} + \partial H|_{[u_0,u_2,u_3]} \\ &= H|_{[u_1,u_2]} - H|_{[u_0,u_2]} + H|_{[u_0,u_1]} + H|_{[u_2,u_3]} - H|_{[u_0,u_3]} + H|_{[u_0,u_2]} \\ &= H|_{[u_1,u_2]} + \alpha - \alpha' - H|_{[u_0,u_3]}.\end{aligned}$$

Notice that both $H|_{[u_1,u_2]}$ and $H|_{[u_0,u_3]}$ are constant map by the definition of a path homotopy. We consider a singular 2-simplex

$$\tau : \Delta^2 \to X,$$

which is a constant map with image q. Let

$$\beta : \Delta^1 \to X,$$

be the singular 1-simplex which is a constant map with image q.

We still denote

$$\Delta^2 = [v_0, v_1, v_2].$$

Then we have

$$\partial\tau = \tau|_{[v_1,v_2]} - \tau|_{[v_0,v_2]} + \tau|_{[v_0,v_1]} = \beta - \beta + \beta = \beta.$$

Hence the singular 1-simplex in X given by constant path gives a 1-boundary in $C_1(X)$. Let τ_1 be the singular 2-simplex in X with image p and τ_2 be the singular 2-simplex in X with image p. We then have

$$\alpha - \alpha' = \partial\sigma - \partial\tau_1 + \partial\tau_2.$$

Hence

$$[\alpha] = [\alpha'].$$

From this discussion, we find a well-defined map

$$\begin{aligned}h : \pi_1(X,p) &\to H_1(X),\\ [\alpha]_{\pi_1} &\mapsto [\alpha]_X.\end{aligned}$$

The main goal of this section is to show the following theorem.

Theorem 6.9.1

The map h is a group homomorphism. Moreover it is surjective and its kernel is the commutator group of $\pi_1(X, p)$.

By the definition of the abelianization of a group, we have the following corollary

Corollary 6.9.2

We have
$$\pi_1(X, p)^{\mathrm{ab}} := \pi_1(X, p)/[\pi_1(X, p), \pi_1(X, p)] \cong H_1(X).$$

Proof. We first show that h is a group homomorphism.

For any α and α' loops in X based at p, we have

$$\beta = \alpha * \alpha' * \overline{\alpha * \alpha'}$$

is a loop based at p which is homotopically trivial. We identify S^1 with the quotient space $[0, 1]/0 \sim 1$. Since β is a loop, it descends to a map β' from S^1 to X, such that we have the following commutative diagram

$$\begin{array}{ccc} [0,1] & \xrightarrow{\beta} & X \\ {\scriptstyle \mathrm{pr}}\downarrow & \nearrow_{\beta'} & \\ S^1 & & \end{array}$$

Since β is homotopically trivial, the map β' can be extends to a map

$$\widetilde{\beta} : D^2 \to X.$$

We may identify Δ^2 with D^2 and get a singular 2-simplex σ, such that

$$\partial\sigma = \alpha + \alpha' - \alpha * \alpha'.$$

Hence we have
$$[\alpha]_X + [\alpha']_X = [\alpha * \alpha']_X,$$
and h is a homomorphism.

Now we would like to show that the map h is surjective. Let $z \in Z_1(X)$. It can be written as

$$z = \sum_{i=1}^{k} \sigma_i,$$

where σ_i's are singular 1-simplices in X. Here for different i and j, it is possible that $\sigma_i = \sigma_j$.

Since z is a singular 1 cycle, we have

$$\partial z = 0.$$

Hence
$$\sum_{i=1}^{k} (\sigma_i|_{v_0} - \sigma_i|_{v_1}) = 0.$$

This means each point in
$$\{\sigma_i(v_0), \sigma_{v_1} \mid 0 \le i \le k\}$$

appears even times. Moreover the times as starting points and the times as ending points of some σ_i are the same. Hence all σ_i's considered as paths in X form several loops by taking concatenation. Without loss of generality, we may assume that

$$\gamma = \sigma_1 * \cdots * \sigma_k$$

is a loop in X based at q.

First the concatenation

$$(\sigma_1 * \cdots * \sigma_k) * \overline{\gamma}$$

is a loop homotopically trivial, hence when consider the associated map from S^1 to X by identifying 0 and 1 of $[0, 1]$ together, it can be extends to a continuous map from D^2 to X. Moreover, the disk D^2 can be considered as an $(k+1)$-gon P, where sides are given by $\sigma_1, ..., \sigma_k, \overline{\gamma}$ following a given orientation of S^1. By taking the triangulation P, we can show that

$$\sigma_1 + \cdots \sigma_k + \overline{\gamma} = \sigma_1 + \cdots \sigma_k - \gamma \in B_1(X).$$

We have

$$[z] = [\gamma].$$

Now we consider a path α in X from p to q. The changes of base point yields a loop based at p:

$$\gamma' = \alpha * \gamma * \overline{\alpha}.$$

Notice that

$$\alpha * \gamma * \overline{\alpha} * \overline{\gamma'}$$

is again a loop homotopically trivial, by a similar argument as above, we have

$$(\alpha + \gamma - \alpha) - \gamma' \in B_1(X),$$

hence

$$[z] = [\gamma] = [\gamma'] \in \operatorname{Im} h.$$

Now we determine the kernel of h. Since $H_1(X)$ is abelian, we have

$$[\pi_1(X, p), \pi_1(X, p)] \subset \ker h.$$

Now let $[\alpha] \in \ker h$, hence we have

$$[\alpha]_X = [0]_X \in H_1(X),$$

or equivalently

$$\alpha \in B_1(X).$$

Let

$$\sigma = \sum_{i=0}^{k} \sigma_k,$$

where $\sigma_1, ..., \sigma_k \in C_2(X)$, such that

$$\sigma = \partial \sigma.$$

For each $1 \leq i \leq k$, we denote

$$\partial \sigma_i = \tau_{i0} - \tau_{i1} + \tau_{i2}$$

Hence

$$\alpha = \sum_{i=1}^{k} (\tau_{i0} - \tau_{i1} + \tau_{i2}).$$

Notice that on the left hand side there is one singular 1-simplex. Hence all τ_{ij}'s should be all canceled out with one left. We denote

$$\sigma_i : \Delta_i^2 \to X,$$

then the above observation induces a way to glue all Δ_i^2's together to a simplicial complex P and we have a map

$$\eta : P \to X,$$

such that $\eta|_{\Delta_i^2} = \sigma_i$. Here Δ_i^2 is identified with its image in P after gluing. In particular, it is connected. Hence the loop γ is homotopic to a concatenation of τ_{ij}'s considered as paths in X.

For each path τ_{ij}, we consider a path γ_{ij} going from p to the starting point of τ_{ij}. Hence we have loop

$$\widetilde{\tau}_{ij} = \gamma_{ij} * \tau_{ij} * \overline{\gamma_{i(j+1)}},$$

here j is taken up to mod 3. The loop γ is also changed to $\widetilde{\gamma}$ without changing its homotopy class since each change is given by adding $\overline{\gamma_{ij}} * \gamma_{ij}$ which is homotopic to a constant path.

Notice that the path $\widetilde{\tau}_{ij}$ is based at p now. And $\widetilde{\gamma}$ is a concatenation of $\widetilde{\tau}_{ij}$'s in certain order. Now we consider them in the abelianization of $\pi_1(X,p)$ (so that we can change their order) and have

$$[\widetilde{\gamma}]^{\text{ab}}_{\pi_1} = *_{i=1}^k([\widetilde{\tau}_{i0}]^{\text{ab}}_{\pi_1} * [\overline{\widetilde{\tau}_{i0}}]^{\text{ab}}_{\pi_1} * [\widetilde{\tau}_{i0}]^{\text{ab}}_{\pi_1}) = *_{i=1}^k[\widetilde{\tau}_{i0} * \overline{\widetilde{\tau}_{i0}} * \widetilde{\tau}_{i0}]^{\text{ab}}_{\pi_1}$$

Notice that for each i, the path $\widetilde{\tau}_{i0} * \overline{\widetilde{\tau}_{i0}} * \widetilde{\tau}_{i0}$ is homotopically trivial as a path in X, since it comes from the boundary map of a continuous map

$$\sigma_i : \Delta_i^2 \to X.$$

Hence we have

$$[\widetilde{\tau}_{i0} * \overline{\widetilde{\tau}_{i0}} * \widetilde{\tau}_{i0}]_{\pi_1} = [c_p]_{\pi_1},$$

and we have

$$[\gamma]^{\text{ab}}_{\pi_1} = [\widetilde{\gamma}]^{\text{ab}}_{\pi_1} = [c_p]^{\text{ab}}_{\pi_1}.$$

This shows that

$$[\gamma] \in [\pi_1(X,p), \pi_1(X,p)].$$

□

Chapter 7

CW-complex and cellular homology

We will introduce another homology group for a space using cells in it. The structure on the space to make this construction work is call a CW complex structure. The construction essentially follows the same idea as the one for simplicial homology by considering cells instead of simplices, although the first definition that we will give seems quite abstract.

7.1 CW-complex

Similar to simplicial complexes or simplicial complexes, the CW complex is also obtained by gluing spaces with certain topological type.

For any $n \in \mathbb{N}^*$, by a ***cell of dimension*** n, we mean a topological space homeomorphic to the Euclidean unit closed ball:

$$D^n := \{p \in \mathbb{R}^n \mid |p| \leq 1\},$$

where $|\cdot|$ stands for the Euclidean norm in $\mathbb{R}^n$.

An ***open cell of dimension*** n is a topological space homeomorphic to the Euclidean unit open ball

$$\mathring{D}^n := \{p \in \mathbb{R}^n \mid |p| < 1\}.$$

where $|\cdot|$ stands for the Euclidean norm in $\mathbb{R}^n$.

As a convention, a 0-cell consists of a single point.

Example 7.1.1 (Cells in dimension $0, 1, 2, 3$).
Figure 7.1.1 illustrates cells in dimension $0, 1, 2, 3$ respectively.

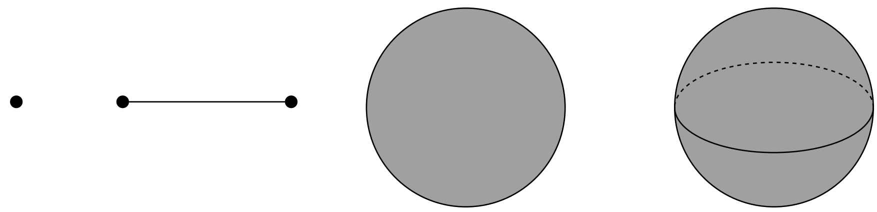


Figure 7.1.1: From left to right, we have D^0, D^1, D^2 and D^3.

Next we would like to glue cells possibly in different dimensions together. Let X and Y be two topological spaces, and A is a subspace of X. The space obtained by gluing X to Y along A is defined to be

$$X \cup_f Y := X \sqcup Y / x \sim f(x)$$

where $f : A \to Y$ is a continuous map, and $\sim$ is given by identifying $x \in A$ with its image $f(x) \in Y$.

Example 7.1.2 (Gluing a handle to a cup).
Topologically, We would like to glue a cylinder C to a disk D^2 by identifying the boundary of C to circles in D^2. Figure 7.1.2 illustrate how it works.

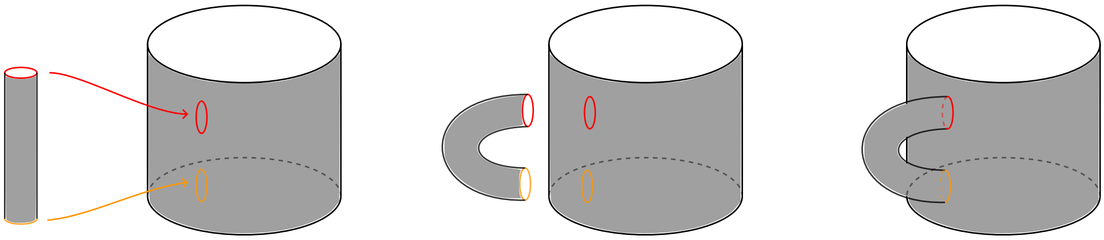


Figure 7.1.2: Glue a handle to a cup.

Definition of a CW-complex

A topological space X is called a CW-complex if it has a filtration of subspaces

$$X^{(0)} \subset X^{(1)} \subset \cdots \subset X^{(n)} \subset \cdots$$

such that

1) $X^{(0)}$ is a disjoint union of points;

2) for any $n \in \mathbb{N}^*$, we construct $X^{(n)}$ by

$$X^{(n)} := X^{(n-1)} \bigcup_{g_\alpha | \alpha \in \Omega} \left(\bigsqcup_{\alpha \in \Omega} D^n_\alpha \right)$$

where for each $\alpha \in \Omega$, D^n_α is a closed n-cell, and

$$g_\alpha : \partial D^n_\alpha \to X^{(n-1)},$$

is a continuous map.

3) we have

$$X = \bigcup_{n \in \mathbb{N}} X^{(n)},$$

and the topological on X is the weak topology given by $(X^{(n)})_{n \in \mathbb{N}}$. In the other words, a subset A of X is open if and only if $A \cap X^{(n)}$ is open in $X^{(n)}$.

For any $n \in \mathbb{N}^*$, for any $\beta \in \Omega$, we have a continuous map

$$f_\beta : D^n_\beta \to X^{(n-1)} \bigsqcup_{\alpha\in\Omega} D^n_\alpha \to X^{(n)} \to X,$$

where the first arrow stands for the inclusion of D^n_β in the disjoint union, the second arrow stands for the quotient map to glue all n-cells to $X^{(n-1)}$, and the last arrow stands for the inclusion of $X^{(n)}$ into X. Such a map f_β is called the ***characteristic map*** for D^n_α.

An easy observation shows the following two facts:

1) the restriction of f_α to ∂D^n_α is g_α;

2) the restriction of f_α to $\mathring{D}^n_\alpha$ is a homeomorphism to the image.

For any $n \in \mathbb{N}$, the subspace $X^{(n)}$ is called the n***-skeleton*** of X.

Example 7.1.3.

A ***sub-CW-complex*** Y of X is a subspace of X, such that it is a CW-complex and for any $n \in \mathbb{N}$, we have

$$Y^{(n)} = X^{(n)} \cap Y.$$

Hence Y is closed.

Example 7.1.4 (Some spaces considered as CW-complexes).
Sphere/Torus/$\mathbb{RP}^n$/Wedge sum of circles/Wedge sum of spheres

7.2 Properties of CW-complexes

Regarding the topological properties on X, since we construct X by gluing cells, we have the following proposition

Proposition 7.2.1

A CW-complex is Hausdorff and locally contractible.

Another property comes form the fact that when we cone off a subspace, the topology in that part becomes trivial. When making a cover over loop in a space, it is equivalent to glue a disk to along that circle. If this loop has non-trivial homotopy class, then after gluing the disk, this loop is homotopic to a constant path. A rigorous proof can be given by considering Seifert-Van-Kampen theorem.

By checking how gluing $n+1$-cells changes the topology of the n-skeleton $X^{(n)}$ of a CW-complex., we have the following proposition.

Proposition 7.2.2

Assume that the 2-skeleton $X^{(2)}$ of a CW-complex X is path connected. The inclusion

$$\iota : X^{(2)} \to X$$

induces an isomorphism

$$\iota_* : \pi_1(X^{(2)}, p) \to \pi_1(X, p),$$

where $p \in X^{(2)}$.

Proof. We consider the inclusion

$$j : X^{(n)} \to X^{(n+1)}$$

We try to show that the induced homomorphism

$$j_* : \pi_1(X^{(n)}, p) \to \pi_1(X^{(n+1)}, p),$$

is an isomorphism for $n \geq 2$.

Any point in any n-cell in $X^{(n)}$ can be connected to $X^{(n-1)}$ by a path. Then by induction, it can be connected to $X^{(2)}$ which is path connected. Hence $X^{(n)}$ is path connected, since any pair of points in $X^{(n)}$ can be connected to $X^{(2)}$ by paths, then by taking a concatenation of these two paths with a path in $X^{(2)}$, we have a path in $X^{(n)}$ to connect the two points.

Let D_α^{n+1} be a $(n+1)$-cell glued to $X^{(n)}$ by g_α with characteristic map f_α. Let γ_α be a path in $X^{(n)}$ with $\gamma_\alpha(0) = p$ and $\gamma_\alpha(1) \in g_\alpha(\partial D_\alpha^{n+1})$. We then glue a Euclidean band to $X^{(n+1)}$ along α.

Let $[0,1] \times [0,1]$ be the band. Then we consider the map

$$\widetilde{\gamma}_\alpha : [0,1] \times \{0\} \cup \{1\} \times [0,1] \to X^{(n+1)},$$

such that for any $t \in [0,1]$, we have

$$\widetilde{\gamma}_\alpha(t, 0) = \gamma_\alpha(t),$$

and

$$\widetilde{\gamma}_\alpha(\{1\} \times (0,1]) \subset f_\alpha(\mathring{D}_\alpha^{n+1}).$$

Then we glue $[0,1] \times [0,1]$ to $X^{(n+1)}$ along $[0,1] \times \{0\} \cup \{1\} \times [0,1]$.

For any other D_β^{n+1} if exists, we repeat the same construction. Denote a path γ_β in $X^{(n)}$ with $\gamma_\beta(0) = p$ and $\gamma_\beta(1) \in g_\beta(\partial D_\beta^{n+1})$. Then we glue a band $[0,1] \times [0,1]$ along

$$\{0\} \times [0,1] \cup [0,1] \times \{0\} \cup \{1\} \times [0,1].$$

by map $\widetilde{\gamma}_\beta$, such that

$$\widetilde{\gamma}_\beta|_{\{0\} \times [0,1]} = \widetilde{\gamma}_\alpha|_{\{0\} \times [0,1]},$$

and for any $t \in [0,1]$, we have

$$\widetilde{\gamma}_\beta(t, 0) = \gamma_\beta(t),$$

and

$$\widetilde{\gamma}_\beta(\{1\} \times (0,1]) \subset f_\beta(\mathring{D}_\beta^{n+1}).$$

We denote the resulting space by Z which has $X^{(n+1)}$ as a deformation retraction. Now we consider

$$U = Z - X^{(n)} \quad \text{and} \quad V = Z - \bigcup_{\beta \in \Omega} f_\beta(\mathring{D}_\beta^{n+1}).$$

Hence $U \cap V$ is a obtained by gluing as many copies as $(n+1)$-cells glued to $X^{(n)}$of $[0,1] \times [0,1]$ along $\{0\} \times [0,1]$, hence is contractible.

The space U is given by taking union of $U \cap V$ with all $f_\beta(\mathring{D}_\beta^{n+1})$, hence is also contractible. The space V has $X^{(n)}$ as a deformation retraction.

We choose a base point q in $\widetilde{\gamma}_\alpha(\{0\} \times (0,1))$, then by the Seifert-Van Kampen Theorem, we have

$$\pi_1(X^{(n+1)}, p) \cong \pi_1(Z, q) \cong \pi_1(U, q) \underset{\pi_1(U \cap V, q)}{*} \pi_1(V, q) \cong \pi_1(V, q) \cong \pi_1(X^{(n)}, p).$$

Hence we have the following sequence of isomorphism:

$$\pi_1(X^{(2)}, p) \cong \pi_1(X^{(3)}, p) \cong \cdots \cong \pi_1(X^{(n)}, p) \cong \cdots .$$

Now we consider a loop γ in $\mathcal{L}(X, p)$. If it is homotopically trivial, then there is a homotopy

$$H : [0,1] \times [0,1] \to X,$$

such that $H_0 = \gamma$ and $H_1 = c_p$. Notice that the image of H is compact in X.

Lemma 7.2.3

Given any K a compact subset in X, K only meets finitely many cells.

Proof. Suppose that K is compact and meets infinitely many open cells. Then we denote by

$$S = \{p_1, ..., p_m, ...\}$$

whose points are in K and meet different open cells.

Notice that S is closed subset of X, hence is compact, since it is a subset of a compact set K. A space with discrete topology which is compact must contain only finitely many points. Hence the contradiction. □

As a corollary, any compact subset of X contained in $X^{(n)}$ for some $n \in \mathbb{N}$. Since the image of γ and H are both compact, hence there is $n > 2$, such that

$$[\gamma] = [c_p] \in \pi_1(X^{(n)}, p).$$

Hence the homomorphism from $\pi_1(X^{(2)}, p)$ to $\pi_1(X, p)$ induced by the inclusion

$$\iota : X^{(2)} \to X,$$

is injective.

To see it is surjective, for any $\gamma \in \mathcal{L}(X, p)$, it is contained in $X^{(n)}$ for some $n > 2$. Since the inclusion from $X^{(2)}$ to $X^{(n)}$ induces an isomorphism between the fundamental group, we have a loop $\eta \in \mathcal{L}(X^{(2)}, p)$, such that

$$[\eta] = [\gamma] \in \pi_1(X^{(n)}, p).$$

Hence the homomorphism ι_* is surjective.

□

One corollary of this result is that

Corollary 7.2.4

Any group is a fundamental group of a CW-complex.

Proof. Any group G can have a presentation

$$\langle S \mid R \rangle,$$

where S is the set of generators and R is the set of relations. Notice that the cardinal of S and that of R could be arbitrary.

Notice that

$$G \cong \langle S \rangle / \langle\langle R \rangle\rangle.$$

Where $\langle S \rangle$ is the free group generated by S. To realize it as a fundamental group of a CW-complex, we consider $X^{(0)}$ be a single point and

$$X^{(1)} = \bigvee_{\alpha \in S} S^1_\alpha.$$

Now for any $w \in R$, it corresponds to a loop γ in $X^{(1)}$, by identifying 0 with 1, we can rewrite this loop as a map

$$\gamma' : S^1 \to X^{(1)}.$$

The by identify S^1_w with S^1, we consider this map γ' and use it to glue D^2_w to $X^{(1)}$ along $S^1_w = \partial D^2_w$. We repeat this for all relations $w \in R$ and obtain $X^{(2)}$. Then the fundamental group $X^{(2)}$ is isomorphic to G. □

Remark 7.2.5.
In the case where R is infinite, we can use the generalized version of Seifert-Van-Kampen theorem to see the final isomorphism.

7.3 Cellular homology group

Now we will give the construction of the cellular homology group for a CW-complex. For simplicity, we consider the case when X is of finite dimension, i.e. the dimension of cells glued has an upper bound.

We consider first the singular (relative) homology group of X. As a convention, let $X^{-1} = \emptyset$.

Proposition 7.3.1

For any $k, n \in \mathbb{N}$, we have

1) the relative homology groups

$$H_k(X^{(n)}, X^{(n-1)}) \cong \begin{cases} \bigoplus_{\alpha \in \Omega} \mathbb{Z}_\alpha, & k = n \\ 0, & k \neq n \end{cases}$$

where Ω is the index set of n-cells in X and $\mathbb{Z}_\alpha$ is isomorphic to $\mathbb{Z}$;

2) $H_k(X^{(n)}) = 0$ for $k > n$;

3) the inclusion

$$\iota : X^{(n)} \to X,$$

induces an isomorphisms

$$\iota_* : H_k(X^{(n)}) \to H_k(X),$$

for $k < n$.

Proof. The proof for 1) is similar to the one about simplicial complex when we try to show the equivalence between the simplicial homology and singular homology.

For any $n \in \mathbb{N}^*$, we consider the following commutative diagram for pairs of spaces

$$\begin{array}{ccccc}
\left(\coprod_{\alpha\in\Omega} D^n_\alpha, \coprod_{\alpha\in\Omega} \partial D^n_\alpha\right) & \xrightarrow{f_1} & \left(\coprod_{\alpha\in\Omega} D^n_\alpha, \coprod_{\alpha\in\Omega} (D^n \setminus \{b_\alpha\})\right) & \xleftarrow{f_2} & \left(\coprod_{\alpha\in\Omega} \mathring{D}^n_\alpha, \coprod_{\alpha\in\Omega} (\mathring{D}^n \setminus \{b_\alpha\})\right) \\
\downarrow{\scriptstyle g_1} & & \downarrow{\scriptstyle g_2} & & \downarrow{\scriptstyle g_3} \\
(X^{(n)}, X^{(n-1)}) & \xrightarrow{h_1} & \left(X^{(n)}, X^{(n-1)} \setminus \coprod_{\alpha\in\Omega} \{f_\alpha(b_\alpha)\}\right) & \xleftarrow{h_2} & \left(X^{(n)} \setminus X^{(n-1)}, (X^{(n)} \setminus X^{(n-1)}) \setminus \coprod_{\alpha\in\Omega} \{f_\alpha(b_\alpha)\}\right)
\end{array}$$

For 2), we consider part of the long exact sequence for relative homology

$$H_{k+1}(X^{(n)}, X^{(n-1)}) \longrightarrow H_k(X^{(n-1)}) \longrightarrow H_k(X^{(n)}) \longrightarrow H_k(X^{(n)}, X^{(n-1)}).$$

Since $k > n$, we have $k + 1 > 0$. By 1), we have

$$H_{k+1}(X^{(n)}, X^{(n-1)}) \cong H_k(X^{(n)}, X^{(n-1)}) \cong 0.$$

The exactness of the sequence shows that

$$H_k(X^{(n-1)}) \cong H_k(X^{(n)}).$$

Hence we have

$$H_k(X^{(n)}) \cong H_k(X^{(n-1)}) \cong \cdots \cong H_k(X^{(0)}) \cong 0.$$

For 3), we consider the same part of the long exact sequence

$$H_{k+1}(X^{(n+1)}, X^{(n)}) \longrightarrow H_k(X^{(n)}) \longrightarrow H_k(X^{(n+1)}) \longrightarrow H_k(X^{(n+1)}, X^{(n)}).$$

Since $k < n$, by 1) we have

$$H_{k+1}(X^{(n+1)}, X^{(n)}) \cong H_k(X^{(n+1)}, X^{(n)}) \cong 0.$$

By the exactness, we have

$$H_k(X^{(n)}) \cong H_k(X^{(n+1)}).$$

Hence we have for any $m \in \mathbb{N}$,

$$H_k(X^{(n)}) \cong H_k(X^{(n+1)}) \cong \cdots \cong H_k(X^{(n+m)}).$$

Since X is of finite dimension, we have $X = X^{(n+m)}$ for some $m \in \mathbb{N}$. Hence

$$H_k(X) = H_k(X^n).$$

□

We denote by

$$D_n := H_n(X^{(n)}, X^{(n-1)}).$$

For $n = 0$, we have

$$D_0 := H_0(X^{(0)}, \emptyset) = H_0(X^{(0)}).$$

Notice that for $n > 0$, the $(n-1)$-skeleton appears in two pairs $(X^{(n)}, X^{(n-1)})$ and $(X^{(n-1)}, X^{(n-2)})$, hence $H_{n-1}(X^{(n-1)})$ appears in the intersection between two long exact sequences:

$$\begin{array}{ccccccc}
 & & & & H_n(X^{(n)}, X^{(n-1)}) & & \\
 & & & & \downarrow \partial & & \\
H_n(X^{(n-1)}, X^{(n-2)}) & \xrightarrow{\partial} & H_{n-1}(X^{(n-1)}) & \xrightarrow{i_*} & H_{n-1}(X^{(n-1)}) & \xrightarrow{\mathrm{pr}_*} & H_{n-1}(X^{(n-1)}, X^{(n-2)}) \\
 & & & & \downarrow i_* & & \\
 & & & & H_{n-1}(X^{(n)}) & & \\
 & & & & \downarrow \mathrm{pr}_* & & \\
 & & & & H_{n-1}(X^{(n)}, X^{(n-1)}) & &
\end{array}$$

We then define

$$\delta : D_n \to D_{n-1},$$

by taking the composition $\delta = \mathrm{pr}_* \circ \partial$ in the above diagram from D_n to D_{n-1} at the upper right corner. As a convention for $n = 0$, we define δ to be the unique homomorphism from D_0 to 0 the trivial group.

Proposition 7.3.2

We have $\delta^2 = 0$.

Proof. For any $n > 0$, we can write the composition:

$$\delta^2 : H_{n+1}(X^{(n+1)}, X^{(n)}) \to H_n(X^{(n)}) \to H_n(X^{(n)}, X^{(n-1)}) \to H_{n-1}(X^{(n-1)}) \to H_{n-1}(X^{(n-1)}, X^{(n-2)})$$

Notice that this sequence is exact at $H_n(X^{(n)}, X^{(n-1)})$, hence $\delta^2 = 0$. □

From the above discussion, we conclude that $(D_n, \delta)_{n\in\mathbb{N}}$ is a chain complex. We call it the ***cellular chain complex*** for the CW-complex X. The n-th homology group associated to this chain complex is called the ***n-th cellular homology group***, we denote it by

$$H_n^{\mathrm{cw}}(X).$$

Theorem 7.3.3

For each $n \in \mathbb{N}$, we have

$$H_n^{\mathrm{cw}}(X) \cong H_n(X).$$

Proof. We review the above intersection between two exact sequences with the information given by Proposition 7.3.1:

$$\begin{array}{ccccccc}
 & & & & H_{n+1}(X^{(n+1)}, X^{(n)}) & \searrow^{\delta} & \\
 & & & & \downarrow \partial & & \\
0 & \longrightarrow & 0 & \longrightarrow & H_n(X^{(n)}) & \xrightarrow[(injective)]{\mathrm{pr}_*} & H_n(X^{(n)}, X^{(n-1)}) \\
 & & & & \downarrow i_* (surjective) & & \\
 & & & & H_n(X) & & \\
 & & & & \downarrow & & \\
 & & & & 0 & &
\end{array}$$

Since i_* is surjective, and the vertical sequence is exact at $H_n(X^{(n)})$, we have

$$H_n(X) \cong H_n(X^{(n)})/\mathrm{Im}\,\partial.$$

Since pr_* is surjective, we have

$$\mathrm{Im}\,\delta \cong \mathrm{Im}\,\partial,$$

and

$$\ker \delta = \ker \partial.$$

Consider the following diagram

$$\begin{array}{ccccc}
H_{n+1}(X^{(n+1)},X^{(n)}) & & & & \\
\downarrow \partial_{n+1} & \searrow^{\delta_{n+1}} & & & \\
H_n(X^{(n)}) & \xrightarrow[(injective)]{(\mathrm{pr}_{n+1})_*} & H_n(X^{(n)},X^{(n-1)}) & & \\
\downarrow (i_{n+1})_*(surjective) & & \downarrow \partial_n & \searrow^{\delta_n} & \\
H_n(X) & & H_{n-1}(X^{(n-1)}) & \xrightarrow[(injective)]{(\mathrm{pr}_n)_*} & H_{n-1}(X^{(n-1)},X^{(n-2)}) \\
 & & \downarrow (i_n)_*(surjective) & & \\
 & & H_{n-1}(X) & &
\end{array}$$

Notice that

$$H_n(X^{(n)}) \to H_n(X^{(n)},X^{(n-1)}) \to H_{n-1}(X^{(n-1)})$$

is exact at $H_n(X^{(n)},X^{(n-1)})$, hence

$$\mathrm{Im}\,(\mathrm{pr}_{n+1})_* = \ker \partial_n = \ker \delta_n.$$

Hence

$$\ker \delta_n/\mathrm{Im}\,\delta_{n+1} = \mathrm{Im}\,(\mathrm{pr}_{n+1})_*/\mathrm{Im}\,((\mathrm{pr}_{n+1})_* \circ \partial_{n+1}).$$

Since $(\mathrm{pr}_{n+1})_*$ is injective, we have

$$H_n^{\mathrm{cw}}(X) := \ker \delta_n/\mathrm{Im}\,\delta_{n+1} \cong H_n(X^{(n)})/\mathrm{Im}\,\partial_{n+1} \cong H_n(X).$$

□

Remark 7.3.4.
Given any CW-complex, for any $n \in \mathbb{N}$, if there is no n-cells, then $H_n(X^{(n)},X^{(n-1)})$ is trivial, hence

$$H_n(X) \cong H_n^{\mathrm{cw}}(X) \cong 0.$$

If there are finitely many n-cells, then the singular homology group $H_n(X)$ is finitely generated.

Cellular boundary formula

From the construction, we use $H_n(X^{(n)},X^{(n-1)})$ to define the n-complex. We would like to given another way to understand this chain complex which is more geometric and can be used to construct a way to compute the cellular homology and eventually compute the singular homology.

We first study the structure of $H_n(X^{(n)},X^{(n)})$ for each $n \in \mathbb{N}$. The case when $n = 0$ is clear. For any $\alpha \in \mathbb{N}$, we consider the morphism between pairs:

$$f_\alpha : (D^n_\alpha, \partial D^n_\alpha) \to (X^{(n)},X^{(n-1)}).$$

For any $n \in \mathbb{N}$, it induces a homomorphism

$$(f_\alpha)_* : H_n(D^n_\alpha, \partial D^n_\alpha) \to H_n(X^{(n)}, X^{(n-1)}).$$

To see the image, we use a discussion similar to the one used previously for studying $H_n(X^{(n)}, X^{(n-1)})$.

Consider the following commutative diagram for morphisms between pairs

$$\begin{array}{ccccc}
(D^n_\alpha, \partial D^n_\alpha) & \longrightarrow & (D^n_\alpha, D^n_\alpha \setminus \{b_\alpha\}) & \longleftarrow & (\mathring{D}^n_\alpha, \mathring{D}^n_\alpha \setminus \{b_\alpha\}) \\
\downarrow & & \downarrow & & \downarrow \\
\left(\coprod_{\alpha\in\Omega} D^n_\alpha, \coprod_{\alpha\in\Omega} \partial D^n_\alpha\right) & \xrightarrow{f_1} & \left(\coprod_{\alpha\in\Omega} D^n_\alpha, \coprod_{\alpha\in\Omega} (D^n \setminus \{b_\alpha\})\right) & \xleftarrow{f_2} & \left(\coprod_{\alpha\in\Omega} \mathring{D}^n_\alpha, \coprod_{\alpha\in\Omega} (\mathring{D}^n \setminus \{b_\alpha\})\right) \\
\downarrow g_1 & & \downarrow g_2 & & \downarrow g_3 \\
(X^{(n)}, X^{(n-1)}) & \xrightarrow{h_1} & \left(X^{(n)}, X^{(n-1)} \setminus \coprod_{\alpha\in\Omega} \{f_\alpha(b_\alpha)\}\right) & \xleftarrow{h_2} & \left(X^{(n)} \setminus X^{(n-1)}, (X^{(n)} \setminus X^{(n-1)}) \setminus \coprod_{\alpha\in\Omega} \{f_\alpha(b_\alpha)\}\right)
\end{array}$$

This commutative diagram gives a free generating set of $H_n(X^{(n)}, X^{(n-1)})$:

$$\{[e^n_\alpha] \mid \alpha \in \Omega\}.$$

Notice that the composition of the two vertical arrows on the left corresponds to f_α. Hence $(f_\alpha)_*$ sends the generator of $H_n(D^n_\alpha, \partial D^n_\alpha)$ to one free generator of $H_n(X^{(n)}, X^{(n-1)})$ associated to D^n_α.

As previously discussed (See Corollary 6.6.6), the homology groups of a space X relative to a subspace A are isomorphic to the homology group of the corresponding quotient space. Hence for $n > 0$, we have

$$H_n(X^{(n)}, X^{(n-1)}) \cong H_n(X^{(n)}/X^{(n-1)})$$

Although X and its skeletons of different dimensions could be quite complicated, the quotient space is quite simple. If

$$X^{(n)} = X^{(n-1)} \bigcup_{g_\alpha|\alpha\in\Omega} \left(\bigsqcup_{\alpha\in\Omega} D^n_\alpha\right),$$

the the quotient space

$$X^{(n)}/X^{(n-1)}$$

with quotient map denote by π is topologically can be considered as identifying the boundaries of all D^n_α's together:

$$X^{(n)}/X^{(n-1)} \cong \coprod_{\alpha\in\Omega} D^n_\alpha / \coprod_{\alpha\in\Omega} \partial D^n_\alpha \cong \bigvee_{\alpha\in\Omega} D^n_\alpha/\partial D^n_\alpha.$$

Here $\cong$ stands for being homeomorphic. We denote by

$$\varphi : X^{(n)}/X^{(n-1)} \to \coprod_{\alpha\in\Omega} D^n_\alpha / \coprod_{\alpha\in\Omega} \partial D^n_\alpha$$

the obvious homeomorphism.

Recall that when identifying all points on the boundary of an n-cell, we obtain an n-sphere. Hence we have

$$\psi : X^{(n)}/X^{(n-1)} \to \bigvee_{\alpha\in\Omega} S^n_\alpha,$$

the homeomorphism induced by φ.

To describe δ_n, it is enough to describe $\delta_n(e^n_\alpha)$ for all $\alpha \in \Omega$. In fact, by its definition, the information of this map is determined by g_α. All we have to study is the following composition

$$\partial D^n_\alpha \to X^{(n-1)} \to X^{(n-1)}/X^{(n-2)} \cong \bigvee_{\alpha\in\Omega} S^n_\alpha.$$

Notice that each S^n_α corresponds to a generator of $H_{n-1}(x^{(n-1)}, X^{(n-2)})$. Hence the whole study is boiled down to answer the following question: How the ∂D^n_α covers each $(n-1)$-sphere in

$$\bigvee_{\alpha\in\Omega} S^n_\alpha.$$

degree

Let $n \in \mathbb{N}^*$. We consider a continuous map

$$f : S^n \to S^n.$$

This map induces a homomorphism in the homology group level. In particular, we have

$$f_* H_n(S^n) \to H_n(S^n).$$

Since

$$H_n(S^n) \cong \mathbb{Z},$$

the homomorphism f_* is determined by the image of $[\alpha]$ a generator of $H_n(S^n)$. There is a integer $d \in \mathbb{Z}$, such that

$$f_*([\alpha]) = d[\alpha].$$

Definition 7.3.5

The integer d is defined to be the ***degree*** of f, denoted by $\deg f$

Here we list several properties of degree.

Proposition 7.3.6

For any $n \in \mathbb{N}^*$, the degree of continuous maps from S^n to S^n satisfies the following properties:

1) The identity map of S^n has degree 1.
2) If $f : S^n \to S^n$ is not surjective, then $\deg f = 0$.
3) If $f, g : S^n \to S^n$ are two homotopic continuous maps, then we have

$$\deg f = \deg g.$$

4) For any $f, g : S^n \to S^n$, we have

$$\deg f \circ g = \deg f \deg g.$$

5) A reflection of the sphere has degree -1.
6) The antipodal map has degree $(-1)^{n+1}$.
7) Any map $f : S^n \to S^n$ with no fixed point has degree $(-1)^{n+1}$.

Proof. The identity map of S^n induces an identity homomorphism of $H_n(S^n)$, hence the degree of identity map is 1.

Assume that f is not surjective, then there is a point $p \in S^n$ which is a not in the image of f. We may view f as a continuous map from S^n to $S^n \setminus \{p\}$. The latter is homotopy equivalent to a single point space. Hence $H_n(S^n \setminus \{p\})$ is trivial and the degree of f is 0.

By Theorem 6.4.4, since f and g are homotopic to each other, they induces a same homomorphism between homology groups. Hence we have $\deg f = \deg g$.

The fourth statement comes from the fact that $(f \circ g)_* = f_* \circ g_*$.

By the previous statement, to show the fifth one, it is enough to show that it holds for one reflection. Consider the map

$$f : S^n \to S^n,$$
$$(x_1, x_2 ..., x_{n+1}) \mapsto (-x_1, x_2 ..., x_{n+1}).$$

Let $H_+ = \{(x_1, x_2 ..., x_{n+1}) \in S^n \mid x_1 \geq 0\}$ and $H_- = \{(x_1, x_2 ..., x_{n+1}) \in S^n \mid x_1 \leq 0\}$ be the two hemispheres of S^n. It changes the orientation of S^n. Let $[\alpha]$ be a generator of $H_n(S^n)$, we then have

$$f_*([\alpha]) = -[\alpha].$$

Hence $\deg f = -1$.

The antipodal map is a composition of $n+1$ reflections, hence its degree is $(-1)^{n+1}$.

To see the last one, we will show that a self-map on S^n with no fixed point is homotopic to the antipodal map. Let $f : S^n \to S^n$ be a map with no fixed point. For any $p \in S^n$, we have $f(p) \neq p$. Hence the segment connecting $f(p)$ and $-p$ does not pass the origin, and the following map is well defined:

$$H : S^n \times I \to S^n$$
$$(p, t) \mapsto \frac{(1-t)f(p) + t(-p)}{|(1-t)f(p) + t(-p)|}.$$

Notice that $H(p, 0) = f(p)$ and $H(x, 1) = -x$. Hence f and the antipodal map are homotopic to each other. By 3), we have $\deg f = (-1)^{n+1}$. □

Here is one application of degree.

Proposition 7.3.7

If n is even, then any group acts on S^n freely is either trivial or isomorphic to $\mathbb{Z}_2$.

Proof. An action of a group G on a set X is free if for any $x \in X$, for any $g.x = x$, then $g = e$ which is the identity element of G. In the other words, the only element with at least one fixed point is the identity element.

Let G be a group acts freely on S^n. For any $g \in G$, we denote its action S^n still by g. Since any element $g \in G$ is invertible, the map $g : S^n \to S^n$ is a homeomorphism. Therefore, it induces an isomorphisms in the homology group level. Hence $\deg g \in \{1, -1\}$. Let $\{1, -1\}$ be the order 2 group. By considering the properties of the degree of maps on S^n, we have the group homomorphism

$$\varphi : G \to \{1, -1\}$$
$$g \mapsto \deg g.$$

Any $g \in G \setminus \{e\}$ has no fixed point, hence $\deg g = (-1)^{n+1}$. Since n is even, we have $\deg g = -1$. Therefore, the kernel is trivial $\{e\}$ and φ is injective. The group G is isomorphic to a subgroup of $\mathbb{Z}_2$ which is either trivial group or $\mathbb{Z}_2$. □

Cellular boundary formula

Now we are ready to give the cellular boundary formula.

Theorem 7.3.8 (Cellular boundary formula)

For any $n \in \mathbb{N}^*$, for any $\alpha \in \Omega$, we have

$$\delta_n([e_\alpha^n]) = \sum_{\beta\in\Lambda} d_{\alpha\beta}[e_\beta^{n-1}],$$

where Λ is the index set for $(n-1)$-cells in X, and for each $\beta \in \Lambda$, the integer $d_{\alpha\beta}$ is the degree of the following map

$$c_{\alpha\beta} : \partial D_\alpha^n \to X^{(n-1)} \to X^{(n-1)}/X^{(n-2)} \to X^{(n-1)}/(X^{(n-1)} \setminus f_\beta(\mathring{D}_\beta^{n-1})) \to D_\beta^{n-1}/\partial D_\beta^{n-1}.$$

Remark 7.3.9.
The last step can be considered as the projection to the $(n-1)$-sphere labeled by β:

$$\bigvee_{\eta\in\Lambda} S_\eta^{n-1} \to S_\beta^{n-1}.$$

Proof. For $n = 0$, by our convention, $D_{-1} = 0$.

For $n = 1$, we consider the map

$$\delta_1 = \partial : H_1(X^{(1)}, X^{(0)}) \to H_0(X^{(0)}).$$

The formula can be checked directly by considering the definition of $d_{\alpha\beta}$.

For $n > 2$, for each $\beta \in \Lambda$, we consider

$$\pi_\beta : X^{(n-1)}/X^{(n-2)} \to X^{(n-1)}/(X^{(n-1)} \setminus f_\beta(\mathring{D}_\beta^{n-1})) \to D_\beta^{n-1}/\partial D_\beta^{n-1},$$

and

$$\kappa_{n-1} : X^{(n-1)} \to X^{(n-1)}/X^{(n-2)}$$

Then we consider the following commutative diagram

$$\begin{array}{ccccc}
H_n(D_\alpha^n, \partial D_\alpha^n) & \xrightarrow[\cong]{\partial} & H_{n-1}(\partial D_\alpha^n) & \xrightarrow{(c_{\alpha\beta})_*} & H_{n-1}(D_\beta^{n-1}/\partial D_\beta^{n-1}) \\
\downarrow (f_\alpha)_* & & \downarrow (g_\alpha)_* & & \uparrow (\pi_\beta)_* \\
H_n(X^{(n)}, X^{(n-1)}) & \xrightarrow{\partial} & H_{n-1}(X^{(n-1)}) & \xrightarrow{(\kappa_{n-1})_*} & H_{n-1}(X^{(n-1)}/X^{(n-2)}) \\
& \searrow \delta_n & \downarrow \mathrm{pr}_* & \nearrow (\widetilde{\kappa}_{n-1})_*, \cong & \\
& & H_{n-1}(X^{(n-1)}, X^{(n-2)}) & &
\end{array}$$

Let $[D_\alpha^n]$ denote the generator in $H_n(D_\alpha^n, \partial D_\alpha^n)$, such that

$$[e_\alpha^n] = (f_\alpha)_*([D_\alpha^n])$$

We are interested in the expression of $\delta_n([e_\alpha^n])$ under the basis

$$\{[e_\beta^{n-1}] \mid \beta \in \Lambda\}.$$

The coordinate for label β is given by the degree associated to

$$\circ(\pi_\beta)_* \circ (\widetilde{\kappa}_{n-1})_* \circ \delta_n \circ (f_\alpha)_*.$$

Since

$$\partial : H_n(D^n_\alpha, \partial D^n_\alpha) \to H_{n-1}(\partial D^n_\alpha)$$

is an isomorphism, this is equivalently to compute the degree associated to

$$(\pi_\beta)_* \circ (\widetilde{\kappa}_{n-1})_* \circ \delta_n \circ (f_\alpha)_* \circ (\partial^{-1}) = (c_{\alpha\beta})_*.$$

□

Applications

Given any CW-complex X, we can now compute the cellular homology by using the cellular boundary formula. As a result, we obtain the singular homology of X.

Example 7.3.10 (Genus n surface).
Let $n \in \mathbb{N}^*$. We consider the oriented closed surface Σ_n of genus n. It can be considered as a connected sum of n torus.

From the classification of compact closed oriented surface (see Theorem 5.3.3 and Theorem 5.3.11), the surface Σ_n can be obtained by gluing paired sides of a $4n$-gon whose sides are labeled by $a_1, b_1, a_1^{-1}, b_1^{-1}, ..., a_n, b_n, a_n^{-1}, b_n^{-1}$ following a cyclic order. After the sides gluing all vertices are identified to a same point and the boundary of the polygon becomes an $2n$-rose.

This gives us a way to associated to Σ_n a CW-complex. In particular, there is one 0-cell, $2n$ 1-cell and one 2-cell. Moreover, the boundary of the 1-cells are all mapped to the only 0-cell, and the boundary of the 2-cell covers each 1-cell twice with different orientation. Hence the degrees of all maps induced by the gluing are 0. We summary the above information as follows. From the number of cells in each dimension, we have

$$\begin{aligned}
&H_0(\Sigma_n^{(0)}) \cong \mathbb{Z}, \\
&H_1(\Sigma_n^{(1)}, \Sigma_n^{(0)}) \cong \mathbb{Z}^{2n}, \\
&H_2(\Sigma_n^{(2)}, \Sigma_n^{(1)}) \cong \mathbb{Z}, \\
&H_k(\Sigma_n^{(k)}, \Sigma_n^{(k-1)}) \cong 0, \ \ k \geq 3.
\end{aligned}$$

Consider the chain complex

$$0 \to H_2(\Sigma_n^{(2)}, \Sigma_n^{(1)}) \to H_1(\Sigma_n^{(1)}, \Sigma_n^{(0)}) \to H_0(\Sigma_n^{(0)}) \to 0.$$

From the information about the degree of maps going from boundaries of k-cells to $(k-1)$-cells, the boundary maps involved above are all zero maps. Hence the cellular homology groups of Σ_n are as follows:

$$\begin{aligned}
&H_0^{\text{CW}}(\Sigma_n) \cong \mathbb{Z}, \\
&H_1^{\text{CW}}(\Sigma_n) \cong \mathbb{Z}^{2n}, \\
&H_2^{\text{CW}}(\Sigma_n) \cong \mathbb{Z}, \\
&H_k^{\text{CW}}(\Sigma_n) \cong 0, \ \ k \geq 3.
\end{aligned}$$

Example 7.3.11 (Projective space).
For each $n \in \mathbb{N}^*$, there is a degree 2 covering map from S^n to $\mathbb{RP}^n$. Meanwhile, we may consider the hemispheres in S^n defined as follows

$$\begin{aligned}
U^n &= \{(x_0, ..., x_n) \in S^n \mid x_0 \geq 0\}, \\
L^n &= \{(x_0, ..., x_n) \in S^n \mid x_0 \leq 0\}.
\end{aligned}$$

Notice that $U^n \cap L^n$ is homeomorphic to S^{n-1}. By identifying $(0, x_1 ..., x_n) \in U^n$ with $-(0, x_1 ..., x_n)$, we have a map from U^n to $\mathbb{RP}^n$. In this way, we may write $\mathbb{RP}^n$ as a disjoint union of an open ball $\mathring{D}^n$ with $\mathbb{RP}^{n-1}$. Repeating this process, we have

$$\mathbb{RP}^n = \mathring{D}^n \sqcup \mathring{D}^{n-1} \sqcup \cdots \sqcup \mathring{D}^2 \sqcup \mathring{D}^1 \sqcup \{*\}.$$

This gives us a CW-complex structure on $\mathbb{RP}^n$. From the number of cells in each dimension, we have

$$\begin{aligned}
&H_0((\mathbb{RP}^n)^{(0)}) \cong \mathbb{Z},\\
&H_1((\mathbb{RP}^n)^{(1)}, \mathbb{RP}^n)^{(0)}) \cong \mathbb{Z},\\
&\cdots\\
&H_n((\mathbb{RP}^n)^{(n)}, \mathbb{RP}^n)^{(n-1)}) \cong \mathbb{Z},\\
&H_k((\mathbb{RP}^n)^{(k)}, \mathbb{RP}^n)^{(k-1)}) \cong 0,\ k \geq n+1.
\end{aligned}$$

To compute the cellular homology group, we should find how cells of different dimension gluing together. The above discussion shows that it is actually given by the antipodal map. In particular, for any $n \in \mathbb{N}^*$, the boundary of D^n is S^{n-1} which is obtained by gluing two copies of D^{n-1} along its boundary. Hence the map from ∂D^n to $\mathring{D}^{n-1}$ is a 2 to 1 covering. If the map from $\mathring{U}^{n-1}$ to $\mathring{D}^{n-1}$ is denoted by f, then the map from $\mathring{L}^{n-1}$ to $\mathring{D}^{n-1}$ if given by $g \circ f$ where g is the antipodal map on S^{n-1}. We may take f to have degree 1, then $\deg g = (-1)^n$. Hence

$$\delta([e^n]) = (1 + (-1)^n)[e^{n-1}].$$

It is easy to understand this. The antipodal map on S^{n-1} sending one hemisphere to another hemisphere in an orientation preserving way if and only if n is even.

Now we consider the chain complex:

$$0 \to H_n((\mathbb{RP}^n)^{(n)}, (\mathbb{RP}^n)^{(n-1)}) \to \cdots \to H_1((\mathbb{RP}^n)^{(1)}, (\mathbb{RP}^n)^{(0)}) \to H_0((\mathbb{RP}^n)^{(0)}) \to 0.$$

By the above discussion, if n is even, we have

$$0 \xrightarrow{0} \mathbb{Z} \xrightarrow{f} \mathbb{Z} \xrightarrow{0} \cdots \xrightarrow{0} \mathbb{Z} \xrightarrow{0} 0$$

and if n is odd, we have

$$0 \xrightarrow{0} \mathbb{Z} \xrightarrow{0} \mathbb{Z} \xrightarrow{f} \cdots \xrightarrow{0} \mathbb{Z} \xrightarrow{0} 0$$

where $f : \mathbb{Z} \to \mathbb{Z}$ is defined by $f(1) = 2$. Therefore, if n is even, we have

$$H_k^{CW}(\mathbb{RP}^n) \cong \begin{cases} \mathbb{Z} & k = 0 \\ \mathbb{Z}_2 & 0 < k < n \text{ odd} \\ 0 & \text{otherwise} \end{cases}$$

If n is odd, we have

$$H_k^{CW}(\mathbb{RP}^n) \cong \begin{cases} \mathbb{Z} & k = 0, n \\ \mathbb{Z}_2 & 0 < k < n \text{ odd} \\ 0 & \text{otherwise} \end{cases}$$

7.4 Euler-Poincaré characteristic

In Chapter 5, we have discussed the Euler characteristic for a surface using triangulations. If we compare it with the definition of CW-complex, we can see that any triangulation of a surface is a special kind of CW-complex formed by cells of dimension 0, 1 and 2. In Chapter 5, the Euler characteristic is computed by counting numbers of vertices, edges and faces, i.e. the number of 0-cells, 1-cells and 2-cells. With this observation, we may try to generalize this notion for more general CW-complexes.

Let X be a CW-complex with finitely many cells of dimension at most $n \in \mathbb{N}$. We define its Euler-Poincaé characteristic to be

$$\chi(X) := \sum_{j=0}^{n} (-1)^j a_j,$$

where for each $0 \le j \le n$, $a_j \in \mathbb{N}$ is the number of j-cells in X. A surface may have different triangulations. Similarly a CW-complex could be decomposed into cells in different ways. Hence in order to have a topological invariant, we have to check if the quantity $\chi(X)$ depends on the cell structures on X. To answer this, we will relate the quantity to the cellular homology groups which are isomorphic to singular homology groups, hence the quantity depends only on the space X itself.

To be more precise, let us recall some background on finitely generated abelian groups. Let G be such a group. If it is torsion free, then there is $r \in \mathbb{N}$, such that $G \cong \mathbb{Z}^r$. The number r is called the rank of G. If it has torsion, then there are $r \in \mathbb{N}$ and $d_1, ..., d_s \in \mathbb{N} \setminus \{0, 1\}$ with $d_1 \mid \cdots \mid d_s$, such that

$$G \cong \mathbb{Z}^r \oplus \mathbb{Z}_{d_1} \oplus \cdots \oplus \mathbb{Z}_{d_s}.$$

Here r is called the rank of G and $d_1, ..., d_s$ are called the invariant factors of G.

We would like to compute the ranks of homology groups for X. To compute the cellular homology groups of X, we consider the chain complex

$$0 \to H_n(X^{(n)}, X^{(n-1)}) \to \cdots \to H_1(X^{(1)}, X^{(0)}) \to H_0(X^{(0)}) \to 0.$$

We still denote by δ the boundary map. Then we have the following short exact sequence

$$0 \to \ker \delta_k \to H_k(X^{(k)}, X^{(k-1)}) \to \operatorname{Im} \delta_k \to 0,$$

where the second homomorphism is the inclusion map and the third one is δ_k. By the definition of the cellular homology group, we have

$$0 \to \operatorname{Im} \delta_{k+1} \to \ker \delta_k \to H_k^{CW}(X) \to 0,$$

Then we have the relations among ranks of above groups

$$\begin{aligned} \operatorname{rank}(H_k(X^{(k)}, X^{(k-1)})) &= \operatorname{rank}(\ker \delta_k) + \operatorname{rank}(\operatorname{Im} \delta_k), \\ \operatorname{rank}(\ker \delta_k) &= \operatorname{rank}(\operatorname{Im} \delta_{k+1}) + \operatorname{rank}(H_k^{CW}(X)). \end{aligned}$$

For $H_0(X^{(0)})$, we have

$$\begin{aligned} \operatorname{rank}(H_0(X^{(0)}) &= \operatorname{rank}(\ker \delta_0) + \operatorname{rank}(\operatorname{Im} \delta_0), \\ \operatorname{rank}(\ker \delta_0) &= \operatorname{rank}(\operatorname{Im} \delta_1) + \operatorname{rank}(H_0^{CW}(X)). \end{aligned}$$

Notice that $\operatorname{Im} \delta_0 = \operatorname{Im} \delta_{n+1} = 0$. By summing up over k, we have

$$\chi(X) = \sum_{k=1}^{n} (-1)^n \operatorname{rank}(H_k(X^{(k)}, X^{(k-1)})) + \operatorname{rank} H_0(X^{(0)}) = \sum_{k=0}^{n} (-1)^k \operatorname{rank}(H_k^{CW}(X))$$

Hence the Euler-Poincaré characteristic is a topological invariant for X, independent of choice of cell structures for defining the cellular complex.

The rank of the k-th homology group $H_k(X)$ is called the *k-th Betti number* of X, which is usually denoted by b_k. Hence the above relation shows that

$$\chi(X) = \sum_{k=0}^{n} (-1)^k b_k.$$

Remark 7.4.1.
Historically, Betti numbers were defined in an combinatorical way which were are about decomposing a manifolds in the most efficient way with lower dimensional submanifolds. The relation among Betti numbers, ranks of homology groups and Euler-Poincaré characteristics are initially built based on series of work of Riemann, Betti, Poincaré, etc.

7.5 Lefschetz fixed point theorem

In the previous section, we discussed the degree of maps from a sphere S^n to itself, which helps us to study the cellular homology groups of a space. We saw that the degree of a map can tell us some information of the map. In particular, a map on S^n with no fixed point homotopic to the antipodal map. Hence any map of S^n with degree different from that of the antipodal map will have a fixed point. In this section, we would like to study similar questions for a general topological space.

Let X be a CW complex. Let f be a map from X to itself. As in the sphere case, the map f induces endomorphisms of homology groups of X. We would like to study the relation between this information and the existence of fixed points of f. For any $n \in \mathbb{N}$, we have the endomorphism of the n-th singular homology group

$$\begin{aligned} f_{*,n} : H_n(X) &\to H_n(X), \\ [\sigma] &\mapsto [f \circ \sigma]. \end{aligned}$$

We consider the trace of $f_{*,n}$ defined as follows. Let G be a finitely generated abelian group, and

$$\mathrm{Tor}G := \{g \in G \mid g \text{ has finite order}\}$$

denote its torsion subgroup. Given any endomorphism φ of G, it induces an endomorphism $\overline{\varphi}$ of $G/\mathrm{Tor}G$. Since $G/\mathrm{Tor}G$ is torsion free and is finitely generated, we have

$$G/\mathrm{Tor}G \cong \mathbb{Z}^r,$$

where r is the rank of G. Hence $\overline{\varphi}$ is an endomorphism of $\mathbb{Z}^r$. By choosing a basis of $\mathbb{Z}^r$, the endomorphism $\overline{\varphi}$ can be represented as a $\mathbb{Z}$-valued matrix M_φ. We define the trace of φ to be

$$\operatorname{tr}\varphi := \operatorname{tr} M_\varphi.$$

Notice that $\operatorname{tr} M_\varphi$ is independent of choice of basis of $\mathbb{Z}^r$, hence $\operatorname{tr}\varphi$ is well-defined.

Assume that X is of dimension n. We then define the *Lefschetz number of f* as follows:

$$\tau(f) := \sum_{k=0}^{n} \operatorname{tr} f_{*,k}.$$

Remark 7.5.1.
If f is homotopy equivalent to id_X, then $\tau(f) = \chi(X)$.

Theorem 7.5.2 (Lefschetz fixed point theorem)

Let X be a CW-complex with finitely many cells of dimension at most $n \in \mathbb{N}$. Any continuous map $f : X \to X$ with $\tau(f) \neq 0$ has a fixed point.

Before giving proof of the theorem, we first check some of its applications.

Example 7.5.3 (Ball).
Let D^n denote the ball of dimension n. It is homotopy equivalent to a point, hence its homology groups are as follows:

$$H_k(D^n) \cong \begin{cases} \mathbb{Z}, & k = 0 \\ 0, & k > 0 \end{cases}$$

Hence for any map $f : D^n \to D^n$, we have

$$\tau(f) = \operatorname{tr} f_{*,0} = 1.$$

By Lefschetz fixed point theorem, the map f has a fixed point.

Example 7.5.4 (Projective space).
Another slightly non trivial example is the even dimensional real projective space. Previously, we have compute the cellular homology groups of $\mathbb{RP}^n$ for any $n \in \mathbb{N}^*$. Let $n = 2k$ be even, then we have

$$H_j(\mathbb{RP}^{2k}) = \begin{cases} \mathbb{Z} & j = 0 \\ \mathbb{Z}_2 & 0 < j < n \text{ odd} \\ 0 & \text{otherwise} \end{cases}$$

Since to define the trace of a map from $\mathbb{RP}^n$ to itself, we have consider each homology group quotient by its torsion subgroup, we have

$$H_j(\mathbb{RP}^{2k})/\mathrm{Tor} H_j(\mathbb{RP}^{2k}) = \begin{cases} \mathbb{Z} & j = 0 \\ 0 & j > 0 \end{cases}$$

Hence for any $f : \mathbb{RP}^n \to \mathbb{RP}^n$, we have

$$\tau(f) = \operatorname{tr} f_{*,0} = 1.$$

By Lefschetz fixed point theorem, the map f has a fixed point.

Example 7.5.5 (Sphere).
The homology groups of sphere S^n for $n \in \mathbb{N}^*$ are as follows:

$$H_k(S^n) \cong \begin{cases} \mathbb{Z} & k = 0, n \\ 0 & \text{otherwise} \end{cases}$$

Let f be a map on S^n. The Lefschetz number of f is then

$$\tau(f) = \operatorname{tr} f_{*,0} + (-1)^n \operatorname{tr} f_{*,n} = 1 + (-1)^n \operatorname{tr} f_{*,n}.$$

Therefore, by Lefschetz fixed point theorem, if f has no fixed point, then $\operatorname{tr} f_{*,n} = (-1)^{n+1}$. Since $H_n(S^n) \cong \mathbb{Z}$, we have $\operatorname{tr} f_{*,n} = \deg f$.

In particular, if f is the antipodal map, then it has no fixed point, hence

$$\deg f = \operatorname{tr} f_{*,n} = (-1)^{n+1}.$$

Another consequence in the differential topology context is that there is a no non-zero vector field over S^2.

To prove the Lefschetz fixed point theorem, we would like to use the cellular approximation theorem.

Definition 7.5.6

Let X and Y be two cellular complex. A map $f : X \to Y$ is said to be *cellular* if for any $n \in \mathbb{N}$, we have $f(X^{(n)}) \subset Y^{(n)}$.

Definition 7.5.7

Let X and Y be two cellular complex. A map $f : X \to X$ is said to be *cellular approximated by* a map $g : X \to Y$ if g is cellular and f is homotopic to g.

Theorem 7.5.8 (Cellular approximation theorem)

Proof. content... □

Theorem 7.5.9 (Hopf trace formula)

content...

Proof. content... □

Proof of Lefschetz fixed point theorem. content... □

7.6 Homology with arbitrary coefficients

Chapter 8

Cohomology group

8.1 Cohomology of a chain complex

8.2 Cohomology and homology

8.3 Singular and cellular cohomology

8.4 Cup product

Chapter 9

Poincaré duality

Appendix A

Free groups

Let G be a finitely generated group. Let $n \in \mathbb{N}^*$, such that there is a subset

$$\{a_1, ..., a_n\} \subset G,$$

generating G, i.e.

$$G = \{b_1 \cdots b_m \mid m \in \mathbb{N}^*,\ b_1, ..., b_m \in \{a_1^{\pm 1}, ..., a_n^{\pm 1}\}\}.$$

Here all we know is that any element can be expressed as a products among finitely many elements in

$$\{a_1^{\pm 1}, ..., a_n^{\pm 1}\}.$$

We call these expressions *words* in letters

$$\{a_1^{\pm 1}, ..., a_n^{\pm 1}\}.$$

However, we have no idea if the expression is unique or not. An example is that

$$a_1 = a_1 a_2 a_2^{-1}.$$

The reason for this to happen is that the elements a_2 and a_2^{-1} satisfy a relation:

$$a_2 a_2^{-1} = e.$$

However, this is a little bit trivial, since this relation satisfied by any element and its inverse in any group. Here is a less trivial example. Consider the group $\mathbb{Z}_5^*$ of multiplication. By a direct computation, we know that $\mathbb{Z}_5^*$ can be generated by $\overline{2}$. Moreover by the Lagrange theorem, we know that

$$\overline{2}^4 = \overline{1}.$$

Of course

$$a^4 = e,$$

is not satisfied by all groups.

Another less trivial examples are abelian groups. Let G be an abelian group. For any $a, b \in G$, we have

$$ab = ba.$$

Same as before, this is not a property satisfied by all groups.

This raises some problems when we try to study groups using words of letters in generating sets:

- How do we know if two different words represent a same elements in the group?

- If we have two groups which both can be generated by n elements for some $n \in \mathbb{N}^*$, how do we know if they are isomorphic to each other?
- Is there a group with as less relation as possible?
- It seems that groups generated by n elements with more relations look smaller in sense than those with less relations. For example, for any $m \in \mathbb{N}^*$, the group $\mathbb{Z}_m$ can be considered as a quotient group of $\mathbb{Z}$. On the other hand, not every pair $\mathbb{Z}_m$ and $\mathbb{Z}_n$ can be compared in this way. Is there a way to make this kind of relation more clear?

A.1 Definitions of free groups

We will first try to construct a group with least possibly many relations satisfied by a given number of generators.

Let $n \in \mathbb{N}^*$. Consider the set of $2n$ distinct elements

$$A = \{a_1^{\pm 1}, ..., a_n^{\pm 1}\},$$

as $2n$ letters.

For any $k \in \mathbb{N}^*$, we call the expression

$$a_{i_1}^{\epsilon_1} \cdots a_{i_k}^{\epsilon_k},$$

where for each $1 \leq j \leq k$, $\epsilon_j \in \{1, -1\}$ a ***word*** of letters in A.

A word

$$a_{i_1}^{\epsilon_1} \cdots a_{i_k}^{\epsilon_k},$$

is said to be ***irreducible***, if for any $1 \leq j \leq k-1$, we have

$$a_{i_j}^{-\epsilon_j} \neq a_{i_{j+1}}^{\epsilon_{j+1}}.$$

We denote by F_n the following set

$$F_n := \{e\} \cup \{\text{irreducible word}\},$$

where distinct irreducible words are distinct elements and e is an elements distinct from all irreducible words and is called the ***empty word***.

Remark 1.1.1.
Another construction is using infinite sequences. We consider sequences in $A \cup \{e\}$ of the following form

$$(a_{i_1}^{\epsilon_{i_1}}, a_{i_2}^{\epsilon_{i_1}}, ..., a_{i_k}^{\epsilon_{i_1}}, e, e, ...),$$

i.e. only finitely many positions taking values in A, and all other entries are e.

An irreducible word is an infinite sequence such that if two adjacent elements are not e, then we have

$$a_{i_j}^{-\epsilon_{i_1}} \neq a_{i_{j+1}}^{\epsilon_{i_{j+1}}}.$$

In this setting, an empty set is the infinite sequence

$$(e, e, ...).$$

The set F_n still consists of irreducible words and empty words.

Next we will define a binary operator on F_n. It take two steps to get the result. Given any $w_1, w_2 \in F_n$, we first take their concatenation $\overline{w_1 w_2}$. Then we simplify $w_1 w_2$. The rule is as follows. If $w_1 w_2$ is not irreducible, then there exist two adjacent letters form $a_i a_i^{-1}$ or $a_i^{-1} a_i$. We remove them. Check the resulting word again to see if it is irreducible. If we get an irreducible word, then it is the result of the computation. If there is nothing left, then we set the result to be the empty word, i.e.

$$w_1 w_2 = e,$$

Example 1.1.2.
Let $n = 5$. Consider

$$w_1 = a_1 a_3 a_2 a_4, \quad w_2 = a_4^{-1} a_2^{-1} a_3 a_5,$$

in F_5, then

$$a_1 a_3 a_2 a_4 a_4^{-1} a_2^{-1} a_3 a_5 \longrightarrow a_1 a_3 a_2 a_2^{-1} a_3 a_5 \longrightarrow a_1 a_3 a_3 a_5$$

Hence we have

$$w_1 w_2 = a_1 a_3 a_3 a_5.$$

Proposition 1.1.3

The set F_n with the above binary operator form a group.

Definition 1.1.4

The group F_n is called the ***free group*** of n letters $\{a_1, ..., a_n\}$. We call n the ***rank*** of F_n.

Remark 1.1.5.
To distinguish with $\mathbb{Z}^n$, we also call F_n the rank n non abelian free group.

A.2 The universal property of free groups

The free group F_n has a so-called universal property, which can be stated as follows:

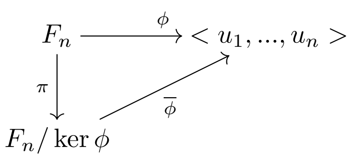
Theorem 1.2.1

For any group G, and n elements $u_1, ..., u_n$ in G, there is a unique group homomorphism

$$\phi : F_n \to G,$$

such that for any $1 \leq i \leq n$, we have $\phi(a_i) = u_i$.

The proof is a direct verification. Any group generated by n elements can be considered as a quotient group of F_n.



where $\overline{\phi}$ is a group isomorphism.

A.3 Free bases

Consider
$$A' = \{a_1a_2, a_2, ..., a_n\} \subset F_n.$$
Notice that
$$a_1 = (a_1a_2)a_2^{-1} \in \langle A' \rangle.$$
Hence we have
$$F_n = \langle A' \rangle.$$
If we consider
$$b_1 = a_1a_2, b_2 = a_2, ..., b_n = a_n,$$
as new letters, and we can get a free group of n letters $\{b_1, ..., b_n\}$. We denote it by F. Using the relation between $\{b_1, ..., b_n\}$ and $\{a_1, ..., a_n\}$, we have a group homomorphism
$$\psi : F \to F_n.$$
which is surjective ($a_1 = \psi(b_1b_2^{-1})$). To see the kernel of ψ, we should show that the image of any irreducible word in F is not identity in F_n. To see this, the rough idea is as follows. We consider a word
$$w(b_1, b_2..., b_n).$$
Then its image will be
$$w(a_1a_2, a_2, ..., a_n).$$
In order to cancel out everything, any b_1 should be followed by b_2^{-1}, then
$$b_1b_2^{-1} = a_1.$$
Then to cancel a_1, we need a_1^{-1}, but the only irreducible word for a_1^{-1} is $b_2b_1^{-1}$, then the word in $w(b_1, ..., b_n)$ is not irreducible.

Definition 1.3.1

Let $m \in \mathbb{N}^*$. We call any m elements $c_1, ..., c_m$ in F_n, satisfying

- $F_n = \langle c_1, ..., c_m \rangle$,
- irreducible words of $c_1, ..., c_m$ are not identity in F_n,

a ***free basis*** of F_n.

Example 1.3.2.
Hence A and A' are both free bases of F_n.

An immediate question is that do all free bases of F_n have n elements?

Theorem 1.3.3

Let X be a free basis of F_n. We have $|X| = n$.

Proof. We consider homomorphisms from F_n to $\mathbb{Z}_2$. Given any free basis $\{b_1, ..., b_m\}$, any homomorphism

$$\phi : F_n \to \mathbb{Z}_2,$$

can be determined by

$$\phi(b_1), ..., \phi(b_m) \in \mathbb{Z}_2.$$

On the other hand, by the universal property of F_n, given any $u_1, ..., u_n \in \mathbb{Z}_2$, we have a group homomorphism

$$\phi : F_n \to \mathbb{Z}_2,$$

such that

$$\phi(a_1) = u_1, ..., \phi(a_n) = u_n.$$

Hence the set of homomorphisms from F_n to $\mathbb{Z}_2$ has cardinality 2^n

$$|\mathrm{Hom}(F_n, \mathbb{Z}_2)| = 2^n.$$

By the definition of a free basis, the group F_n is isomorphic to F_X the free group with letters in X. By a similar argument, we have

$$|\mathrm{Hom}(F_X, \mathbb{Z}_2)| = 2^{|X|}.$$

Since a group homomorphism is a map which is independent of choice of free basis, hence we have

$$2^{|X|} = 2^n.$$

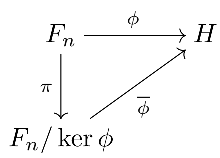
equivalently, we have $|X| = n$. □

A.4 Presentations of groups

Let us consider the universal property of F_n. Given any group G and its n elements $u_1, ..., u_n$, we consider the homomorphism

$$\phi : F_n \to H = \langle u_1, ..., u_n \rangle < G,$$

such that for any $1 \leq i \leq n$, we have $\phi(a_i) = u_i$. From the fundamental theorem of group homomorphism, we have the following commutative diagram



where $\overline{\phi}$ is an isomorphism. For any non identity element $w \in \ker \phi$, it is an irreudicble word of a_i's

$$w = a_{i_1}^{\epsilon_1} \cdots a_{i_k}^{\epsilon_k}.$$

By taking value u_i's, we have the following identity in G:

$$u_{i_1}^{\epsilon_1} \cdots u_{i_k}^{\epsilon_k} = e_G.$$

We call the left side is a ***relation*** among u_i's

Therefore, the construction of H can be considered as a two steps process. First, we consider the free group of letters $\{u_1, ..., u_n\}$, then taking its quotient by the relations satisfied by u_i's. The rough idea of a presentation of a group is that to describe a group by giving its generators and the relations satisfied among the generators.

Normal closure

Let G be a group.

Definition 1.4.1

Consider any non-empty subset $S \subset G$, we call

$$\langle\langle S\rangle\rangle := \bigcap\{N \lhd G \mid S \subset N\},$$

the ***normal subgroup generated by*** S, or the ***normal closure*** of S.

Remark 1.4.2.
Notice that given any normal subgroup $N_1, N_2 \lhd G$, we have $N_1 \cap N_2 \lhd G$. Moreover this holds for any intersections among normal subgroups of G. Hence the above definition is well defined.

There are two ways to understand the normal closure. The definition shows that it is the smallest normal subgroup of G containing S. The second way of understanding is that it is the biggest normal subgroup which could be "generated" by S.

To be more precise, for any $a \in G$, we consider $[a]$ the conjugacy class of a in G. For any non empty subset S of G, we denote

$$[S] := \cup\{[a] \mid a \in S\}.$$

Proposition 1.4.3

We have the following relation

$$\langle\langle S\rangle\rangle = \{a_1 \cdots a_k \mid k \in \mathbb{N}^*,\, a_1, ..., a_k \in [S] \cup [S^{-1}]\}.$$

Remark 1.4.4.
In the other words, $\langle\langle S\rangle\rangle$ is the subgroup generated by the union of conjugacy class of elements in S and S^{-1}.

Now back to our discussion on presentations of groups. Let H be a group generated by n distinct elements $\{u_1, ..., u_n\}$. Let F_n be the free group of letters $a_1, ..., a_n$. We consider the homomorphism

$$\phi : F_n \to H,$$

such that for each $1 \leq i \leq n$, we have $\phi(a_i) = u_i$. Denote N by its kernel. Let R' be a normal generating set of N. We denote by

$$R := \{w(u_1, ..., u_n) \mid w(a_1, ..., a_n) \in R'\}.$$

Then a presentation of H can be given as

$$H = \langle S|R\rangle.$$

Remark 1.4.5.
In general a group need not to be finitely generated, and the relation set R need not to be finite

either. The notion of presentation can be used for any group. When S is finite, we say it is finitely generated. When both S and R are finite, we say that it is finitely presented.

Example 1.4.6.
The quaternion Group Q_8 has the following presentation:

$$Q_8 = \langle a, b \mid a^4, b^4, a^2b^2, abab^{-1} \rangle.$$

Example 1.4.7.
The symmetry group S_3 has the following presentation:

$$S_3 = \langle a, b \mid a^2, b^3, abab \rangle.$$

Example 1.4.8.
The dihedral group D_4 has the following presentation:

$$D_4 = \langle r, s \mid r^4, s^2, srsr \rangle.$$

Example 1.4.9.
The free abelian group $\mathbb{Z}^2$ has the following presentation:

$$\mathbb{Z}^2 = \langle a, b \mid aba^{-1}b^{-1} \rangle.$$

A.5 Visualization of the presentations of groups

Cayley graph is an efficient tool to help us to see the structure of a group from the generating set and generating relations.

Let G be a group. Let S be a generating set of G, such that $S = S^{-1}$.

Definition 1.5.1

The Cayley graph of G with respect to the generating set S is a graph $\Gamma(G, S) = (V, E)$, where

- Vertices are elements in G ($V = G$),
- For any $w, w' \in G$, there exists an edge in E connecting w and w' if and onlt if $w' = ws$, for some $s \in S$.

By the cancellation rule in a group, if $w' = ws$, the element s is unique. Hence there is a unique way to associated to each orientation of an edge an element of S.

If we following a path in the graph from the identity element e to an element w, by writing down elements associated to each edge with the orientation induces by the orientation of the path from e to w, then we get a word for w of letters in S. On the other hand, any word of letters in S associated to the element w will correspond to a path from e to w.

For example, for the element

$$w = a_1 \cdots a_k.$$

we have

$$e \text{ ——— } a_1 \text{ ——— } a_1a_2 \text{ ——— } a_1a_2a_3 \text{ ——— } \cdots \text{ ——— } (a_1a_2 \cdots a_k).$$

Here are some examples

Example 1.5.2.
Let $G = \mathbb{Z}_2$, and $S = \{1\}$, then we have

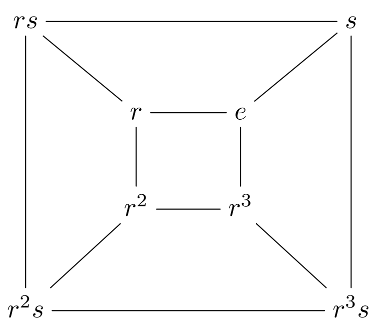
$$0 \text{ ——— } 1$$

Example 1.5.3.
Let $G = D_4$. Consider the four points in $\mathbb{R}^2$:

$$v_1 = (1,1),\ v_2 = (-1,1),\ v_3 = (-1,-1),\ v_4 = (1,-1).$$

Let s be the reflection of the plane fixing v_1 and v_3, and r be the rotation sending (v_1, v_2, v_3, v_4) to (v_2, v_3, v_4, v_1).

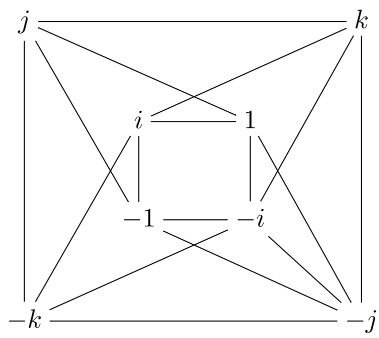
We denote $S = \{s, r, r^{-1} = r^3\}$, then have



Example 1.5.4.
Let $G = Q_8$. Consider the generating set

$$S = \{\pm i, \pm j\},$$

and we have



Now we assume that G is generated by n elements $a_1, ..., a_n$ and denote

$$S = \{a_1^{\pm}, ..., a_n^{\pm}\}.$$

As discussed previous, when different ways of writing $w \in G$ into words of letters in S correspond to different paths connecting e to w. In particular, we consider $w = e$, then each loop from e to e corresponds to a word of letters in S which is in the kernel of

$$\phi : F_n \to G,$$

if we view it as an elements in F_n.

The Cayley graph of $\mathbb{Z}^2$

Now we consider $\mathbb{Z}^2$ as an example to give a more precise description. Let

$$S = \{(1,0), (-1,0), (0,1), (0,-1)\},$$

then the Cayley graph of $\mathbb{Z}^2$ can be given as follows:

$$\begin{array}{ccccccccccccc}
 & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \\
 & & | & & | & & | & & | & & | & & \\
\cdots & \text{---} & (-1,3) & \text{---} & (0,3) & \text{---} & (1,3) & \text{---} & (2,3) & \text{---} & (3,3) & \text{---} & \cdots \\
 & & | & & | & & | & & | & & | & & \\
\cdots & \text{---} & (-1,2) & \text{---} & (0,2) & \text{---} & (1,2) & \text{---} & (2,2) & \text{---} & (3,2) & \text{---} & \cdots \\
 & & | & & | & & | & & | & & | & & \\
\cdots & \text{---} & (-1,1) & \text{---} & (0,1) & \text{---} & (1,1) & \text{---} & (2,1) & \text{---} & (3,1) & \text{---} & \cdots \\
 & & | & & | & & | & & | & & | & & \\
\cdots & \text{---} & (-1,0) & \text{---} & (0,0) & \text{---} & (1,0) & \text{---} & (2,0) & \text{---} & (3,0) & \text{---} & \cdots \\
 & & | & & | & & | & & | & & | & & \\
\cdots & \text{---} & (-1,-1) & \text{---} & (0,-1) & \text{---} & (1,-1) & \text{---} & (2,-1) & \text{---} & (3,-1) & \text{---} & \cdots \\
 & & | & & | & & | & & | & & | & & \\
 & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & &
\end{array}$$

Notice that this is an infinite graph. Now we consider the presentation of $\mathbb{Z}^2$.

$$F_2 = \langle a, b \rangle.$$

and the group homomorphism

$$\phi : F_2 \to \mathbb{Z}^2,$$

such that

$$\phi(a) = (1,0),\ \phi(b) = (0,1).$$

With the above graph, we would like to show that

$$\ker \phi = N(aba^{-1}b^{-1}).$$

Given any element $w \in F_2$, it can be written as a word of a, b, a^{-1}, b^{-1},

$$w = w(a, b).$$

Its image under ϕ is then

$$\phi(w) = w((1,0), (0,1)),$$

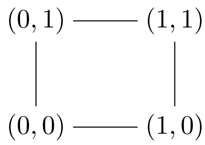
which corresponds to a path in the Cayley graph.

Moreover $w \in \ker \phi$ if and only if

$$w((1,0), (0,1)) = (0,0),$$

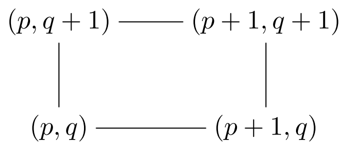
or equivalently, $w((1,0), (0,1))$ corresponds to a loop starting and ending at $(0,0)$.

First consider the loop corresponds to $aba^{-1}b^{-1}$:

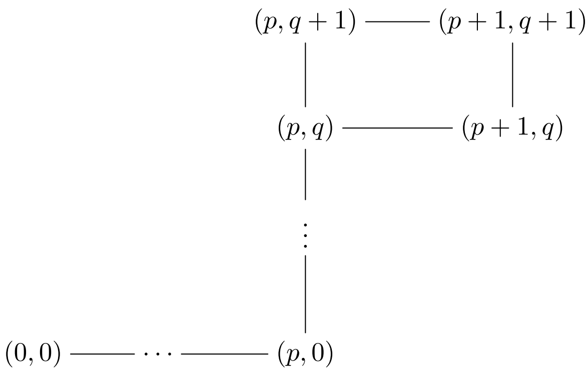


This relation tells us that the generators of $\mathbb{Z}^2$ commute with each other, hence $\mathbb{Z}^2$ is an abelian group.

Notice that any loop in $\mathbb{Z}^2$ can be decomposed into small squares of the form:



By connecting one vertex to $(0, 0)$, for example we take



then we have a loop starting and ending at $(0, 0)$ corresponding to the following word in F_2

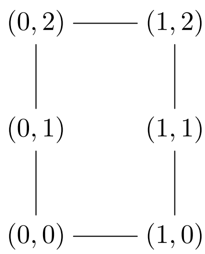
$$w = (a^p b^q)aba^{-1}b^{-1}(b^{-q}a^{-p}).$$

Notice that w is conjugate to $aba^{-1}b^{-1}$.

By this observation, any loop based starting and ending at $(0, 0)$ corresponds to a product among elements in $[aba^{-1}b^{-1}]$. Hence we have

$$\ker \phi = \langle\langle aba^{-1}b^{-1} \rangle\rangle.$$

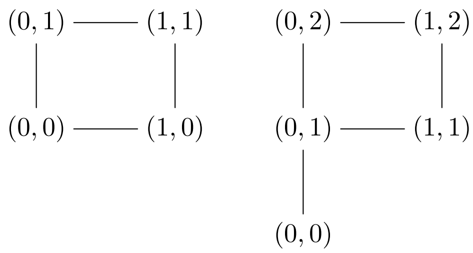
For example, we consider the following loop in the Cayley graph



If we walk around counterclockwise, the corresponding element in F_2 is

$$abba^{-1}b^{-1}b^{-1}.$$

If we go back to $(0,0)$ in the middle, we can decompose it into two loops



corresponds to $aba^{-1}b^{-1}$ and $baba^{-1}b^{-1}b^{-1}$ respectively.

Hence we have

$$abba^{-1}b^{-1}b^{-1} = (aba^{-1}b^{-1})(baba^{-1}b^{-1}b^{-1}) \in \langle\langle aba^{-1}b^{-1}\rangle\rangle.$$

Remark 1.5.5.

Another application of Cayley graphs is to allow us to define a metric on a group, so that we could study groups using geometric method. This leads us to the research area so called Geometric Group Theory.

Appendix B

Construction of groups

B.1 Free Product between groups

In the previous part, we introduce the notion of free group. We may follow this idea and define the so called free product between groups. Roughly speaking, a free group is a group generated by a collection of elements which have no relation among them. A free product between groups is a group generated by elements in two groups, such that elements in different groups has no relation among them.

Let G and H be two non trivial groups. We consider elements in $G^* \cup H^*$ as letters, for any $k \in \mathbb{N}^*$, we call any one of the following expressions a ***word***

$$\begin{aligned} & g_1 h_1 \cdots g_k h_k, \\ & g_1 h_1 \cdots h_{k-1} g_k, \\ & h_1 g_1 \cdots h_k g_k, \\ & h_1 g_1 \cdots g_{k-1} h_k. \end{aligned}$$

We denote

$$G * H := \{e\} \cup \{\text{words}\}.$$

To define an binary operator on $G * H$, we follow the same idea as what we have done for constructing free groups. For any $w_1, w_2 \in G*H$ different from e, we first take their concatenation, if the last letter of w_1 and the first letter of w_2 belong to different groups, then the concatenation is the result $w_1 w_2$. Otherwise, without loss of generality, we may assume that

$$w_1 = g_1 h_1 \cdots g_k h_k, \quad w_2 = h_1' g_1' \cdots h_k' g_k'.$$

Their concatenation is then

$$g_1 h_1 \cdots g_k h_k h_1' g_1' \cdots h_k' g_k'.$$

We do the computation in H to get the element $h_k h_1'$.

If $h_k h_1' \neq e_H$, then we obtain the result $w_1 w_2$.

If $h_k h_1' = e_H$, we cancel it and obtain

$$g_1 h_1 \cdots h_{k-1} g_k g_1' h_2' \cdots h_k' g_k'.$$

Now we consider the computation in G and get the element $g_k g_1'$.

If $g_k g_1' \neq e_G$, then we get the result $w_1 w_2$.

If $g_k g_1' = e_G$, we cancel

$$g_1 h_1 \cdots h_{k-1} h_2' \cdots h_k' g_k'.$$

Then we repeat the above process again.

Since there are only finitely many elements G and H involved, the process will stop at finite time.

If there are letters left after the cancellation, the word formed by them is the result w_1w_2.

Otherwise, all letters are canceled out, and we define in this case

$$w_1w_2 = e,$$

Proposition 2.1.1

The set $G * H$ with above binary operator is a group.

Definition 2.1.2

The group $G * H$ is called the ***free product*** between G and H.

Remark 2.1.3.
From its definition, we can see that $H * G = G * H$. By consider single letter words in $G * H$, both groups G and H can be considered as subgroups of $G * H$.

Remark 2.1.4.
By repeating this construction, we may define free product among several groups.

With this definition, we now review the notion of free group. Let $A = \{a\}$, denote

$$F(a) = \{a^n \mid n \in \mathbb{Z}\},$$

the free group of 1 letter.

For any $n \in \mathbb{N} \setminus \{0, 1\}$, let

$$A = \{a_1, ..., a_n\}$$

and denote F_n the free group of n letter $a_1, ..., a_n$.

Proposition 2.1.5

We then have

$$F_n \cong F(a_1) * \cdots * F(a_n).$$

Similar to free groups, the free products between groups also have certain universal property. Let G and H be two groups.

Theorem 2.1.6

For any group K and any group homomorphisms

$$\phi_G : G \to K, \quad \phi_H : H \to K,$$

there exists a unique group homomorphism

$$\phi : G * H \to K,$$

such that

$$\phi|_G = \phi_G, \quad \phi|_H = \phi_H.$$

B.2 Amalgamated free product between two groups

Let H, G_1 and G_2 be three groups. Assume that there exist group homomorphisms

$$\phi_1 : H \to G_1,$$
$$\phi_2 : H \to G_2.$$

We consider the free product between G_1 and G_2, and consider its subset

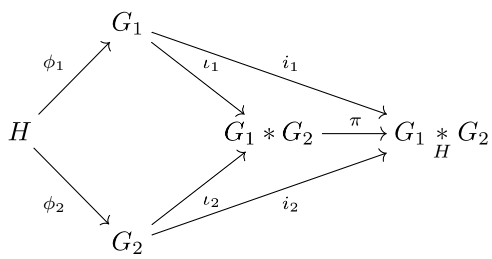
$$S = \{\phi_1(a)\phi_2(a)^{-1} \mid a \in H\}.$$

Definition 2.2.1

The amalgamated free product between G_1 and G_2 over H through ϕ_1 and ϕ_2 is the following quotient group

$$G_1 \underset{H}{*} G_2 := (G_1 * G_2)/\langle\langle S \rangle\rangle.$$

All groups mentioned above form the following commutative diagram:



where ι_1 and i_1 (resp. ι_2 and i_2) are inclusion of G_1 (resp. G_2) into $G_1 * G_2$ and $G_1 \underset{H}{*} G_2$ respectively.

Remark 2.2.2.
Notice that if $H = \{e\}$, the an amalgamated free product over H is a free product.

B.3 HNN extension

This method was first introduced by Graham Higman, Bernhard Neumann, and Hanna Neumann in 1949 in their paper "Embedding Theorems of Groups".

In stead of "gluing" different groups together along their subgroups, here we consider glue two parts of a group together.

More precisely, let H and G be two groups. Assume that there are two injective group homomorphisms

$$\phi_1 : H \to G, \quad \phi_2 : H \to G.$$

We denote by t a letter, and consider the free group generated by t denoted by $F(t)$. Let $G * F(t)$ be the free product between G and $F(t)$. Consider the subset

$$S = \{t\phi_1(a)t^{-1}\phi_2(a)^{-1} \mid a \in H\}.$$

Definition 2.3.1

The HNN extension of G over H through ϕ_1 and $\phi_2$is the following quotient group

$$\underset{H}{*} G := G * F(t)/\langle\langle S \rangle\rangle.$$

Bibliography

[1] M. A. Armstrong, *Basic Topology*, Springer-Verlag New York Inc., 1983.

[2] Jean Dieudonné, *A history of algebraic and differential topology, 1900-1960*, Birkhäuser, Boston, 2007.

[3] Alan Hatcher, *Algebraic Topology*, Cambridge University Press, 2002.

[4] James R. Munkres, *Topology*, Pearson College Div, 2000.

[5] Frédéric Paulin, *Topologie algébrique élémentaire*, lecture note.