NTNU – Trondheim
Norwegian University of
Science and Technology

The Fundamental Group of SO(3)

Eirik Andreas Mork

Master of Science
Submission date: December 2014
Supervisor: Petter Andreas Bergh, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

ABSTRACT

We study fundamental groups of topological spaces. In particular we will compute the fundamental group of $SO(3)$, the group of rotations in three dimensions, by studying covering spaces. We will see that the fundamental group is isomorphic to $\mathbb{Z}_2$. This is of interest because of its relation to physics.

Norsk sammendrag:
Vi studerer fundamentalgrupper av topologiske rom. Spesielt vil vi regne ut fundamentalgruppen til $SO(3)$, gruppen av rotasjoner i tre dimensjoner ved å studere overdekningsrom. Vi vil se at fundamental gruppen er isomorf med $\mathbb{Z}_2$. Dette er interessant på grunn av relasjonen til fysikk.

ACKNOWLEDGEMENTS

This thesis marks the end of my time as a student at the department of Mathematical Sciences and the LUR-programme at NTNU.

I would like to express my deepest gratitude to those who have helped me complete this thesis. First of all, I would like to thank my supervisors, Professor Petter Andreas Bergh and Marius Thaule. I would like to thank Petter for his great suggestion of topic, and for all his encouragement and help during this semester. I would like to thank Marius for our good discussions and for keeping me motivated. Without his guidance I would have had problems completing this thesis.

I would like to thank my friends for their support. Especially my fellow students, Marte Oldervoll and Therese Hagland at Matteland whom I have been working and sharing thoughts with. Finally, I would like to thank my mom and dad; without them I would never have been as hardworking as I am today, and I would probably not have been able to finish this tough education.

CONTENTS

1 Introduction **1**

2 Homotopy **3**
2.1 Paths . . . 3
2.2 Homotopy of paths . . . 4
2.3 Equivalence relations . . . 6
2.4 Groupoid properties . . . 9

3 The fundamental group **13**
3.1 Path connectedness and isomorphisms . . . 14
3.2 Simply connected space . . . 17
3.3 Induced homomorphisms . . . 17

4 Covering spaces **19**
4.1 Lifting properties . . . 23
4.2 The fundamental group of S^1 . . . 27

5 Homotopy equivalence **29**
5.1 Homotopy equivalence and homeomorphism . . . 31
5.2 Induced isomorphisms . . . 32
5.3 Contractible spaces . . . 34

6 Rotation groups **39**
6.1 Orthogonal groups . . . 39

6.2 Rotation groups 40
6.3 The SO(2) group 41
6.4 The SO(3) group 45

7 The fundamental group of SO(3) 49
7.1 Quaternions 49
7.2 Computing the fundamental group of SO(3) 53

CHAPTER 1

INTRODUCTION

The central idea behind algebraic topology is to associate a topological situation to an algebraic situation, and study the simpler algebraic setup. To each topological space a group can be associated, such that homeomorphic spaces give rise to isomorphic groups. To a map of spaces we can associate a homomorphism of groups such that compositions of maps yield compositions of homomorphisms of groups. Then anything we can say about a topological situation gives information about the algebraic one.

In this thesis we will study the first and simplest realization of this idea, the fundamental group of a space.

The fundamental group is a tool used for describing what a topological space looks like. It creates an algebraic "image" of the space using loops in the space. However, the group does not tell us everything about a space. So what does it actually tell us? It detects holes—it tells us if our space has any sort of holes.

Then how do we detect these holes? Imagine you are living in a topological space, e.g. a surface and you are totally blind. Your life and everything around is restricted to it, and you are only able to walk on the surface. Then suppose you want to get a feeling of what your world looks like. You cannot see, but you want to know the properties of the surface, so you are equipped with a lasso which serves as your detector. You have also been told that there may be some dangerous traps out there, so you decide to stay where you are. Then you start throwing your lasso in all directions, keeping track of where you might catch something. In the end you will have created some sort of map which gives us information about the holes of the surface, the algebraic ‘image” creating the fundamental group.

Our motivation for studying the fundamental group is to prove that the fundamental group of $SO(3)$, the group of rotations, is a cyclic group of order two. In physics applications this result is interesting to us because it is associated to spin and spinor representations in quantum mechanics. It explains why a rotation body can have a spin of half a quantum and no other fraction (see [6, p. 602]).

At the end of the thesis we will demonstrate this result practically by Dirac's scissors experiment.

In Chapter 2 we will define homotopy and present some group properties. Then we will describe the fundamental group and properties related in Chapter 3. In this chapter group homomorphisms will be presented, too. Chapter 4 describes covering spaces that play an important role in computing fundamental groups, and our main result will build upon this theory. In the end of the chapter we will compute the fundamental group of the circle. In Chapter 5 we will introduce the notion of homotopy equivalence, another tool used for computing fundamental groups. Chapter 6 presents orthogonal groups and especially rotation groups, which will lay the foundation for the last chapter. In this chapter we will compute the fundamental group of $SO(3)$. We will also introduce quaternions that we will need for proving our main result.

Throughout this thesis we will assume the knowledge of basic group theory and general topology. The exposure in Chapter 2 and 3 is closely related to [4].

CHAPTER 2

HOMOTOPY

2.1 Paths

We need to construct the group. All groups consists of elements and an operation that combines any two elements, and it turns out that the elements we are considering are constructed from paths. The paths will essentially be the basis for everything we talk about in this thesis.

Definition 2.1. Let X be a topological space. A ***path*** in X from x_0 to x_1 is a continuous map $f : I \to X$ such that $f(0) = x_0$ and $f(1) = x_1$. We say that x_0 is the ***initial*** point and x_1 the ***final point***.

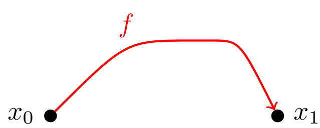


Figure 2.1: A path in X from x_0 to x_1.

See Figure 2.1 for an illustration. Paths are our building blocks, and we want to define an operation combining them. We can combine any two paths as long as they have a common endpoint. The operation is called ***concatenation*** of paths and defines a product given in the following definition:

Definition 2.2. Let $f : I \to X$ be a path in X from x_0 to x_1 and $g : I \to X$ be a path in X from x_1 to x_2. Then the ***product*** $f * g$ is defined to be the path $f * g : I \to X$ given by

$$f * g(s) = \begin{cases} f(2s) & \text{for } s \in [0, \frac{1}{2}], \\ g(2s-1) & \text{for } s \in [\frac{1}{2}, 1]. \end{cases}$$

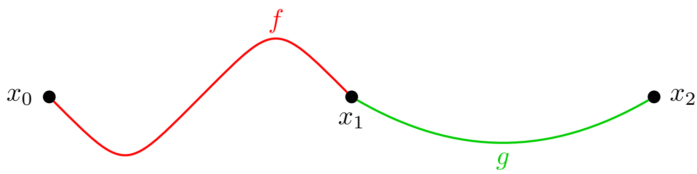


Figure 2.2: Concatenation of f and g.

The function $f * g : I \to X$ is well-defined and continuous (see Figure 2.2). The continuity of $f * g$ comes from the fact that a function defined on the union of two closed sets is continuous if it is continuous when restricted to each of the closed sets separately. See the ***pasting lemma***, [4, pp. 108–109].

Remark. $f * g$ is a path in X from x_0 to x_2, by going from x_0 to x_1 at twice the speed (i.e. in half the time) by f, and then from x_1 to x_2 via g (again in half the time).

Definition 2.3. Let $x \in X$. We define a ***constant path*** to be the path $e_x : I \to X$ carrying all of I to the point x.

Definition 2.4. Given a path f in X from x_0 to x_1, let $\bar{f}$ be the path defined by $\bar{f}(s) = f(1-s)$. It is called the ***reverse path*** of f.

In Theorem 2.14, we will see that the constant and reverse path will represent the identity and the inverse element, respectively, in the construction of the fundamental group.

2.2 Homotopy of paths

In general there are many paths on a topological space, in fact there can be too many of them to consider them all separately. We want to bring paths that are essentially the same together. If we can continuously deform one path into the other, we say they are essentially the same. Such a deformation is called a ***path homotopy*** between the two paths. This leads us to the following definition:

Definition 2.5. Let $f, f' : X \to Y$ be continuous maps. A ***homotopy*** between f and f' is a continuous map $F : X \times I \to Y$ such that

$$F(x,0) = f(x) \quad \text{and} \quad F(x,1) = f'(x)$$

for all x. Here $I = [0,1]$. Then f is ***homotopic*** to f'. We denote this by $f \simeq f'$.

Remark. We can think of F as a family of maps $\{f_t : X \to Y \mid t \in I\}$ connecting f and f', and $f_t(x) = F(x,t)$. Then $f_0 = f$ and $f_1 = f'$. If t represents time, then the homotopy F represents a continuous "deforming" of the maps, as t goes from 0 to 1.

Definition 2.6. A continuous map $f : X \to Y$ is said to be ***null homotopic*** if it is homotopic to a constant map.

Definition 2.7. Let $f, f' : I \to X$ be paths in X with same endpoints. A ***path homotopy*** between f and f' is a continuous map $F : I \times I \to X$ such that

(i) $F(s,0) = f(s) \quad \text{and} \quad F(s,1) = f'(s)$,

(ii) $F(0,t) = x_0 \quad \text{and} \quad F(1,t) = x_1$,

for each $s \in I$ and each $t \in I$. When two paths f and f' are connected in this way by a homotopy F, they are said to be ***path homotopic***. We denote this by $f \simeq_p f'$.

The first condition says that F represents a continuous way of deforming the path f to the path f', and the second condition says that the end points of the path remains fixed during the deformation. We can think of it as a "movie" or a sequence of slides going from f to f' (see Figure 2.3). If such a continuous deformation does not exists, we say they are ***non homotopic paths*** (see Figure 2.4).

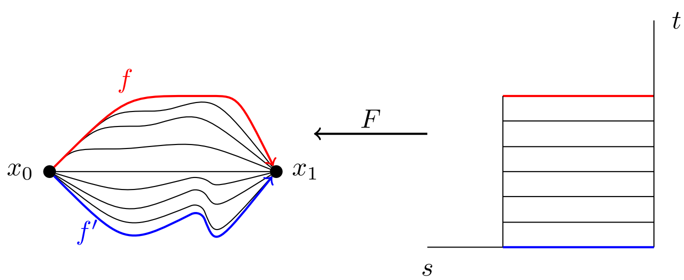


Figure 2.3: Homotopic paths: In this figure we have a valid homotopy, a continuous deformation of paths.

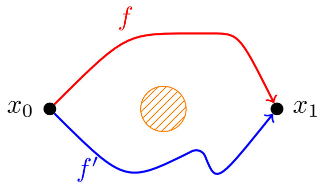


Figure 2.4: Non homotopic paths: In this figure the problem is the hole in the middle. Obviously we cannot have a continuous deformation.

Example 2.8. (**Linear homotopies.**) Let f and f' be any two paths in $\mathbb{R}^n$ having the same endpoints x_0 and x_1. Then $F(s,t) = (1-t)f(s) + tf'(s)$ is a homotopy between f and f'. We verify this:

(i) $F(s,0) = f(s) \quad \text{and} \quad F(s,1) = f'(s)$,

(ii) $F(0,t) = (1-t)x_0 + tx_0 = x_0 \quad \text{and} \quad F(1,t) = (1-t)x_1 + tx_1 = x_1$.

During the homotopy each point $f(s)$ travels along a line segment to $f'(s)$ at constant speed. It is called a ***straight line homotopy***.

In particular, if $U \subset \mathbb{R}^n$ is convex, then any two paths $f, g : I \to U$ with same endpoints are homotopic.

2.3 Equivalence relations

It is a well known fact that a congruence relation is an equivalence relation of an algebraic structure, such as a group or a ring. Being homotopic is an equivalence relation on the set of all continuous functions from X to Y. We will prove this in the following lemma.

Lemma 2.9. *The relations $\simeq$ and $\simeq_p$ of homotopy are equivalence relations.*

Proof. We need to verify the three properties of an equivalence relation, reflexibility, symmetry and transitivity.

(1) (**Reflexibility.**) Let $F(x,t) = f(x)$ be the constant homotopy. Then $f \simeq f$. Similarly, let $F(s,t) = f(s)$ be the constant path homotopy. Then $f \simeq_p f$.

(2) (**Symmetry.**) Let F be a homotopy between f and f'. Then we can find a homotopy G between f' and f, given by $G(x,t) = F(x, 1-t)$. If F is a path homotopy, G is also. Let G be this path homotopy. Then we have

(i) $G(s,0) = F(s,1) = f' \quad \text{and} \quad G(s,1) = F(s,0) = f$,

(ii) $G(0,t) = F(0,1-t) = x_0$ and $G(1,t) = F(1,1-t) = x_1$.

(3) (**Transitivity.**) Let F be a homotopy between f and f', and let G be a homotopy between f' and f''. Then there exists a homotopy H between f and f'' given by

$$H(x,t) = \begin{cases} F(x,2t) & \text{for } t \in [0,\frac{1}{2}], \\ G(x,2t-1) & \text{for } t \in [\frac{1}{2},1]. \end{cases}$$

These two definition agree for $t = 1/2$ since $F(x,2t) = f'(x) = G(x,2t-1)$, and the map H is well-defined. Continuity of H is evident by the pasting lemma. Since H is continuous on $X \times [0,\frac{1}{2}]$ and $X \times [\frac{1}{2},1]$, it is continuous on $X \times I$. If F and G are path homotopies, so is H. Let $H : I \times I \to X$ be this homotopy. Then

(i) $H(s,0) = F(s,0) = f$ and $H(s,1) = G(s,1) = f''$,
(ii) $H(0,t) = x_0$ and $H(1,t) = x_1$.

□

Definition 2.10. Given a space and a homotopy in X, the *homotopy class* of a path f in X, denoted $[f]$, is the subset of all paths in X which is path homotopic to f. We write

$$[f] = \{g \in X \mid g \simeq_p f\}.$$

As mentioned at the beginning of this section, we will find it useful to collect those paths that are essentially the same, and that is exactly what these homotopy classes do. In Chapter 3 we will see that it is exactly these collections of paths that will form the elements of the fundamental group.

So we consider the set of homotopy classes of paths. In order to have a group structure of these classes, we have to define a group operation. Recall that we defined a way to combine paths by an operation called concatenation of paths, and it turns out that this operator $*$ induces a well-defined operation on the path-homotopy classes, given by

$$[f] * [g] = [f * g].$$

We verify this. Let f and g be as defined in Definition 2.2. Let F be a path homotopy between f and f' and let G be a path homotopy between g and g'. Then we can define a new homotopy

$$H(s,t) = \begin{cases} F(2s,t) & \text{for } s \in [0,\frac{1}{2}], \\ G(2s-1,t) & \text{for } s \in [\frac{1}{2},1], \end{cases}$$

between $f * g$ and $f' * g'$. To see this, we have

(i) $H(0,t) = F(0,t) = x_0$,

(ii) $H(1,t) = G(1,t) = x_2$,

(iii) $H(s,0) = \begin{cases} F(2s,0) = f(2s) & \text{for } s \in [0, \frac{1}{2}], \\ G(2s-1,0) = g(2s-1) & \text{for } s \in [\frac{1}{2}, 1], \end{cases}$
which is equal to the product $f * g$,

(iv) $H(s,1) = \begin{cases} F(2s,1) = f'(2s) & \text{for } s \in [0, \frac{1}{2}], \\ G(2s-1,t) = g'(2s-1) & \text{for } s \in [\frac{1}{2}, 1], \end{cases}$
which is equal to the product $f' * g'$.

Since $F(1,t) = x_1 = G(0,t)$ for all t, the map H is well-defined; it is continuous by the pasting lemma. The homotopy H is pictured in Figure 2.5.

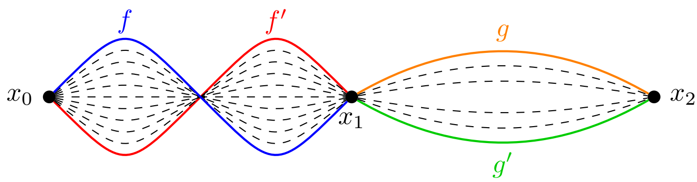


Figure 2.5: Illustration of homotopy between $f * g$ and $f' * g'$ given $f \simeq_p f'$ and $g \simeq_p g'$.

The homotopy class of $f * g$ depends only on the homotopy class of f and g, so the product $[f] * [g] = [f * g]$ is well defined.

Theorem 2.11. *If $f, f' : X \to Y$ are homotopic maps and $g, g' : Y \to Z$ are homotopic maps, then their compositions $g \circ f, g' \circ f' : X \to Z$ are also homotopic maps.*

Proof. Let $F : X \times I \to Y$ be a homotopy between f and f' and $G : Y \times I \to Z$ be a homotopy between g and g'. We define a map $H : X \times I \to Z$ by

$$H(x,t) = G(F(x,t),t).$$

Clearly, H is continuous. Moreover,

(i) $H(x,0) = G(F(x,0),0) = G(f(x),0) = g(f(x))$,

(ii) $H(x,1) = G(F(x,1),1) = G(f'(x),1) = g'(f'(x))$.

Thus, H is the required homotopy between $g \circ f$ and $g' \circ f'$. □

2.4 Groupoid properties

In this section we will study the induced operation $*$ on the homotopy classes, and its properties. However, first we need to make some definitions that we will make use of in the upcoming theorem.

Definition 2.12. Let $k : X \to Y$ be a continuous map, and let F be a path homotopy between the paths f and f'. Then $k \circ F$ is a path homotopy in Y between the paths $k \circ f$ and $k \circ f'$.

Definition 2.13. Let $k : X \to Y$ be a continuous map and let f and g be paths in X such that $f(1) = g(0)$. Then

$$k \circ (f * g) = (k \circ f) * (k \circ g).$$

This equality follows at once from the definition of the product operation $*$.

Theorem 2.14. *The operation $*$ has the following properties:*

*(1) (**Associativity.**) If both relations are defined, then*

$$[f] * ([g] * [h]) = ([f] * [g]) * [h].$$

*(2) (**Right and left identities.**) If f is a path in X from x_0 to x_1, then*

$$[f] * [e_{x_1}] = [f] \quad \text{and} \quad [e_{x_0}] * [f] = [f].$$

*(3) (**Inverses.**) Let f be a path in X from x_0 to x_1. Then*

$$[f] * [\bar{f}] = [e_{x_0}] \quad \text{and} \quad [\bar{f}] * [f] = [e_{x_1}].$$

Before we start proving anything, we need to be clear on how we approach the given properties. What does it mean for two homotopy classes to be equal? Consider the classes $[f]$ and $[g]$. Since $[f]$ consists of all paths that are homotopic to f, and $[g]$ consists of all paths homotopic to g, then solving $[f] = [g]$ must be the same as to show that f is homotopic to g.

Proof. To verify (1), we need to find a homotopy between $[f] * ([g] * [h])$ and $([f] * [g]) * [h]$. For this proof we will find it convenient to use another notation for the product $f * g$ than the one we are used to. Let $[a, b]$ and $[c, d]$ be two intervals in $\mathbb{R}$. Then we can construct a positive linear map $p : [a, b] \to [c, d]$ that is given by $p(x) = mx + k$ and maps a to c and b to d.

Then the product $f * g$ can be described as follows: On $[0, \frac{1}{2}]$ it equals the positive linear map of $[0, \frac{1}{2}]$ to $[0, 1]$, followed by f; and on $[\frac{1}{2}, 1]$ it equals the positive linear map of $[\frac{1}{2}, 1]$ to $[0, 1]$, followed by g.

Let f, g, and h be paths in X. The products $f * (g * h)$ and $(f * g) * h$ are defined precisely when $f(1) = g(0)$ and $g(1) = h(0)$. Further we define a "triple product" of the paths f, g, and h: Choose points a and b of I so that $0 < a < b < 1$. Define a path $k_{a,b}$ in X as follows: On $[0, a]$ it equals the positive linear map of $[0, a]$ to I followed by f; on $[a, b]$ it equals the positive linear map of $[a, b]$ to I followed by g; and on $[b, 1]$ it equals the positive linear map of $[b, 1]$ to I followed by h. The path $k_{a,b}$ depends of course on the choice of the points a and b. But its path homotopy class does not!

If we let c and d be another points of I with $0 < c < d < 1$, and we manage to show that $k_{c,d}$ is path homotopic to $k_{a,b}$, we are done. The product $f * (g * h)$ is equal to $k_{a,b}$ in the case $a = 1/2$ and $b = 3/4$, while the product $(f * g) * h$ equals $k_{c,d}$ in the case $c = 1/4$ and $d = 1/2$.

Let $p : I \to I$ be a map, and restrict p to the intervals $[0, a]$, $[a, b]$ and $[b, 1]$. Then it equals the positive linear maps of these onto the intervals $[0, c]$, $[c, d]$ and $[d, 1]$, respectively (see Figure 2.6). It follows at once that $k_{c,d} \circ p$ equals $k_{a,b}$. But p is a path in I from 0 to 1, and so is the identity map $i : I \to I$. Hence, there is a path homotopy P in I between p and i. Then $k_{c,d} \circ P$ is a path homotopy in X between $k_{a,b}$ and $k_{c,d}$.

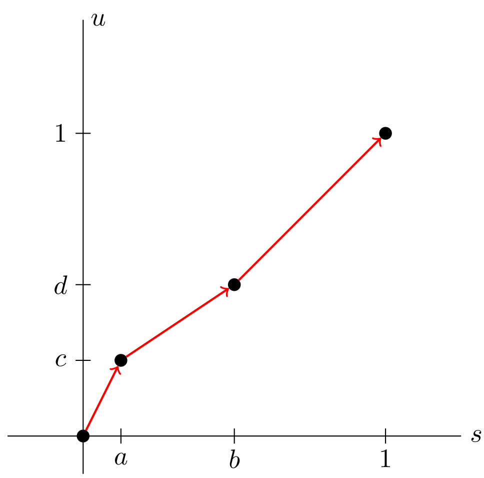


Figure 2.6: Positive linear maps of intervals.

We verify (2). Let e_0 denote the constant path in I at 0, and let $i : I \to I$ denote the identity map, which is a path in I from 0 to 1. Then $e_0 * i$ is also a path in I from 0 to 1 (see Figure 2.7).

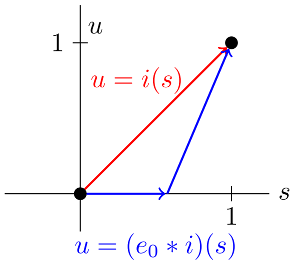


Figure 2.7: Observe $e_0 * i$ going twice as "fast" relative to i.

Using Example 2.8 and the fact that I is convex we know there is a path homotopy G in I between i and $e_0 * i$. Then we know that $f \circ G$ is a path homotopy in X between the paths $f \circ i = f$ and

$$f \circ (e_0 * i) = (f \circ e_0) * (f \circ i) = e_{x_0} * f.$$

We apply the same method using the fact that e_1 denotes the constant path at 1. Then $i*e_1$ is path homotopic in I to the path i, which gives us that $[f]*[e_{x_1}] = [f]$.

Finally, we verify (3). Let i be defined as before. We know that the reverse of i is $\bar{i}(s) = i(1-s) = 1-s$. Then $i * \bar{i}$ is a path in I with endpoints at 0, and so is the constant path e_0 (see Figure 2.8).

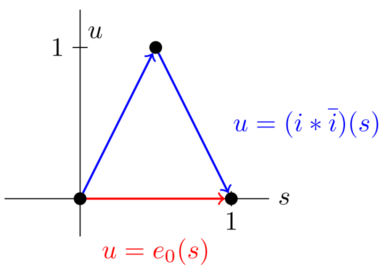


Figure 2.8: e_0 and $i * \bar{i}$ are homotopic paths in I.

Again, using the fact that I is convex, there is a path homotopy H in I between e_0 and $i * \bar{i}$. Then $f \circ H$ is a path homotopy between $f \circ e_0 = e_{x_0}$ and

$$(f \circ i) * (f \circ \bar{i}) = f * \bar{f}.$$

An entirely similar argument, using the fact that $\bar{i} * i$ is path homotopic in I to e_1, shows that $[\bar{f}] * [f] = [e_{x_1}]$. □

We see that the preceding properties are very similar to the axioms of a group. There is only one problem; the product $[f] * [g]$ is not defined for every pair of classes, but only for those pairs $[f]$, $[g]$ for which $f(1) = g(0)$. For every group all the group axioms, namely closure, identity, associativity and invertibility must be satisfied. However, in this case the closure property fails. These properties are called the ***groupoid properties*** of $*$. To ensure closure, we need to make a generalization.

CHAPTER 3

THE FUNDAMENTAL GROUP

Suppose we pick out a point x_0 of X to serve as a "base point" and restrict ourselves to those paths that begin and end at x_0 (just choose $e_{x_0} = e_{x_1}$). Then we can multiply any path with each other because one path will start where the last one ended, and we will automatically have a closed system. Associativity, the existence of an identity element $[e_{x_0}]$, and the existence of an inverse $[\bar{f}]$ for $[f]$ are immediate. Then all the group axioms are satisfied, and the set of these path homotopy classes form a group under $*$.

Definition 3.1. A ***loop*** in X is a continuous map $f : I \to X$ such that $f(0) = f(1)$.

Then two loops can be combined together in an obvious way; first travel along the first loop, then along the second.

Definition 3.2. Let X be a topological space, and x_0 a point in X. The ***fundamental group*** of X is the set of path homotopy classes $[f]$ of loops $f : I \to X$ based at x_0, together with the operation $*$. We denote it by $\pi_1(X, x_0)$.

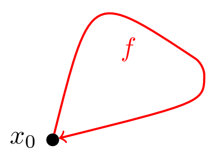


Figure 3.1: A loop based at x_0.

The fundamental group is also called the ***first homotopy group***. It has a generalization to homotopy groups $\pi_n(X, x_0)$, defined for all $n \in \mathbb{Z}_+$, but we will not study them in this thesis.

Definition 3.3. Given two loop classes $[f]$ and $[g]$ we define:

(i) $[f] * [g] = [f * g]$.

(ii) The inverse of $[f]$ is given by $[f^{-1}]$, that is $[f]^{-1} = [f^{-1}]$, where $f^{-1}(t) = \bar{f}(t) = f(1-t)$.

Definition 3.4. Let $x_0 \in X$. We define a ***constant loop*** to be the loop $e_{x_0} : I \to X$ carrying all of I to the point x_0.

Definition 3.5. A loop f is called ***nullhomotopic*** if it is homotopic to the constant loop.

Example 3.6. Let $\mathbb{R}^n$ denote the Euclidean n-space. Then $\pi_1(\mathbb{R}^n, x_0)$ is the trivial group. For if f is a loop in $\mathbb{R}^n$ based at x_0, the straight line homotopy is a path homotopy between f and the constant path at x_0. Hence there is only one homotopy class of loops. More generally, if X is any convex subset of $\mathbb{R}^n$, $\pi_1(X, x_0)$ is the trivial group. In particular, a disk

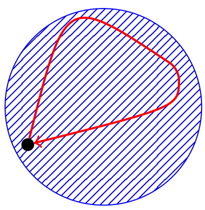


Figure 3.2: Loop based at x_0 in a disk.

has trivial fundamental group (see Figure 3.2). Lets choose some point x_0 inside it and look at loops based at this point. Then all loops at x_0 are nullhomotopic.

How can we make a "movie" that continuously deforms these loops to the constant loop? Imagine we can see every point from x_0, and imagine a spider at x_0 collecting all the threads of its web to x_0, at a constant rate.

3.1 Path connectedness and isomorphisms

Throughout the thesis, we only deal with path connected spaces when studying fundamental groups. We therefore find it useful to define path connectedness.

Definition 3.7. A space X is said to be ***path connected*** if there is a path joining any two points in X.

Let X be path connected. Since X is path connected there exist a path α in X connecting two point x_0 and x_1 in X. Then we can define a map between the two fundamental groups whose loops are based at the respective points x_0 and x_1 in X:

Definition 3.8. Let α be a path in X from x_0 to x_1. Then we have a map

$$\hat{\alpha} : \pi_1(X, x_0) \to \pi_1(X, x_1)$$

given by

$$\hat{\alpha}([f]) = [\bar{\alpha}] * [f] * [\alpha],$$

mapping a loop f based at x_0 to a loop $\bar{\alpha} * (f * \alpha)$ based at x_1. Hence $\hat{\alpha}$ maps $\pi_1(X, x_0)$ into $\pi_1(X, x_1)$ as desired. The map $\hat{\alpha}$ is well defined since the operation $*$ is well defined.

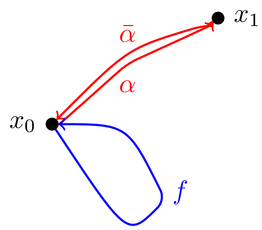


Figure 3.3

We will show that the map is an isomorphism. Then using the fact that X is path connected and that x_0 and x_1 are two arbitrary points in X, we know that any path in X give rise to an isomorphism.

Theorem 3.9. *The map $\hat{\alpha}$ is a group isomorphism.*

Proof. The map $\hat{\alpha}$ is a group homomorphism, as is seen from

$$\begin{aligned}\hat{\alpha}([f]) * \hat{\alpha}([g]) &= ([\bar{\alpha}] * [f] * [\alpha]) * ([\bar{\alpha}] * [g] * [\alpha]) \\ &= [\bar{\alpha}] * [f] * [g] * [\alpha] \\ &= \hat{\alpha}([f] * [g]).\end{aligned}$$

For $\hat{\alpha}$ to be an isomorphism we need to find an inverse of $\hat{\alpha}$. Let β denote the reverse path $\bar{\alpha}$, a path from x_1 to x_0. Then $\hat{\beta}$ is an inverse for $\hat{\alpha}$. Then for each $[h]$ of $\pi_1(X, x_1)$, we have

$$[\hat{\beta}]([h]) = [\bar{\beta}] * [h] * [\beta] = [\alpha] * [h] * [\bar{\alpha}],$$

$$\hat{\alpha}(\hat{\beta}([h])) = [\bar{\alpha}] * ([\alpha] * [h] * [\bar{\alpha}]) * [\alpha] = [h].$$

Similarly, for each $[f]$ of $\pi_1(X, x_0)$ we have

$$[\hat{\alpha}]([h]) = [\bar{\alpha}] * [f] * [\alpha] = [\beta] * [f] * [\bar{\beta}],$$

$$\hat{\beta}(\hat{\alpha}([f])) = [\bar{\beta}] * ([\beta] * [f] * [\bar{\beta}]) * [\beta] = [f].$$

Hence, $\hat{\alpha}$ is an isomorphism. ☐

Corollary 3.10. *if X is path connected and x_0 and x_1 are two points of X, then $\pi_1(X, x_0)$ is isomorphic to $\pi_1(X, x_1)$.*

Hence, letting X be path connected ensures us that the fundamental group is independent of the base point. This is an important fact we will use throughout the thesis when computing fundamental groups. Then for a path connected space X we will denote its corresponding fundamental group by $\pi_1(X)$, omitting the base point x_0.

Example 3.11. Consider the ***punctured euclidean space***, $\mathbb{R}^n \setminus (\mathbf{0})$. If $n \geq 1$, the space is path connected. Let $\mathbf{x}$ and $\mathbf{y}$ be two points in $\mathbb{R}^n$ different from $\mathbf{0}$. Then we can join them by the path $f : I \to \mathbb{R}^n \setminus (\mathbf{0})$ given by $f(t) = (1-t)\mathbf{x} + t\mathbf{y}$ if the path does not go through the origin. Otherwise, we can join $\mathbf{x}$ and $\mathbf{y}$ by two paths through a third point z.

Lemma 3.12. *Let $F : X \to Y$ be continuous and onto. If X is path connected, then Y is also.*

Proof. Since f is onto there exist x_1, $x_2 \in X$ such that $f(x_1) = y_1$ and $f(x_2) = y_2$ for any y_1, $y_2 \in Y$. If X is path connected there is a path $g : I \to X$ from x_1 to x_2. Then the composition $f \circ g : I \to Y$ is a path from y_1 to y_2, so Y is path connected. ☐

The map $f : \mathbb{R}^n \setminus (\mathbf{0}) \to S^1$ given by $f(x) = \frac{x}{\|x\|}$ is continuous is surjective; hence by Lemma 3.12 we know that S^1 is path connected.

3.2 Simply connected space

In a simply connected space we can continuously shrink any closed curve to a point while remaining in the domain, so any loop based at x_0 can be deformed to the constant map at that point, to the identity element. Then the homotopy class of each element is the same as the homotopy class of the identity element, so there is only one homotopy class.

Definition 3.13. A space is called ***simply connected*** if it is path connected and if $\pi_1(X, x_0)$ is the trivial group. We denote this by writing $\pi_1(X, x_0) = 0$.

So the fundamental group can measure the extent to which a space fails to be simply connected. Intuitively, the group gives us information about the holes in the space; if no holes exist, the group is trivial and the space is simply connected.

Lemma 3.14. *Let X be simply connected. Then any two paths in X having the same initial and final points are path homotopic.*

Proof. Let α and γ be two paths from x_0 to x_1. Then $\alpha * \bar{\gamma}$ is a loop in X based at x_0. Since X is simply connected, the loop $\alpha * \bar{\gamma}$ is path homotopic to the constant loop e_{x_0}. In particular, we get

$$[\alpha * \bar{\gamma}] * [\gamma] = [e_{x_0}] * [\gamma] = [\alpha].$$

Hence, it follows that $[\alpha] = [\gamma]$ and $\alpha \simeq_p \gamma$. □

Example 3.15. The sphere S^2 is simply connected because every loop on the surface can be contracted to a point.

In fact, the spheres, S^n, for $n \geq 2$ are simply connected spaces.

3.3 Induced homomorphisms

Group homomorphisms are maps that preserve group structure, and as with any two groups we can find a homomorphism between them. Suppose we have a continuous function $h : X \to Y$. Then associated to h is a an algebraic analog, h_*, which is a function between the corresponding fundamental groups. We will take a look at the definition.

Definition 3.16. Let $h : (X, x_0) \to (Y, y_0)$ be a continuous map such that $h(x_0) = y_0$. Then we define a map $h_* : \pi_1(X, x_0) \to \pi_1(Y, y_0)$ given by $h_*([f]) = [h \circ f]$. Then h_* is a homomorphism of fundamental groups called the ***homomorphism induced by h.***

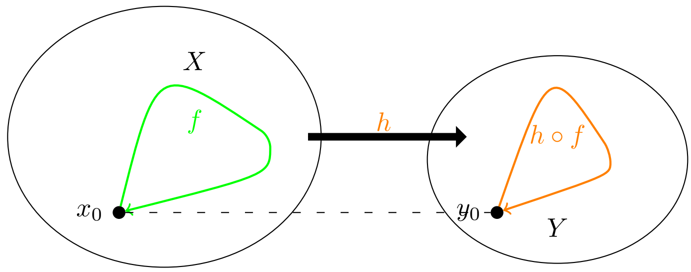


Figure 3.4: Basepoint x_0 in X mapped to y_0, basepoint in Y.

As Figure 3.4 illustrates, if f is a loop in $\pi_1(X, x_0)$, then $h \circ f$ is a loop in $\pi_1(Y, y_0)$.

The map h_* is well-defined, for if F is a path homotopy between the paths f and f' in X, then we can "transfer" this homotopy over to a homotopy in Y, $h \circ F$, that will define a path homotopy between the paths $h \circ f$ and $h \circ f'$. We check that h_* actually is a homomorphism:

$$h_*[f * g] = h \circ (f * g),$$

$$h_*[f] * h_*[g] = (h \circ f) * (h \circ g).$$

We already know that the identity

$$(h \circ f) * (h \circ g) = h \circ (f * g)$$

holds, so h_* is a homomorphism.

It is important to be aware of that the homomorphism h_* not only depends on the map $h : X \to Y$ but also on the choice of the base point x_0. We may have to consider different base points of X. Then we can not use the same symbol h_* to stand for different homomorphisms. To distinguish them from each other, we find it natural to use the notation $(h_{x_0})_*$ for base point x_0.

The induced homomorphism has two important properties that are given in the following theorem:

Theorem 3.17. *If $h : (X, x_0) \to (Y, y_0)$ and $k : (Y.y_0) \to (Z, z_0)$ are continuous, then $(k \circ h)_* = k_* \circ h_*$. If $id_X : (X, x_0) \to (X, x_0)$ is the identity map, then id_{X*} is the identity homomorphism.*

Proof. By definition,

$$(k \circ h)_*([f]) = [(k \circ h) \circ f],$$

$$(k_* \circ h_*)([f]) = k_*(h_*([f])) = k_*([h \circ f]) = [k \circ (h \circ f)].$$

Similarly, $id_{X*}([f]) = [id_X \circ f] = [f]$. □

CHAPTER 4

COVERING SPACES

Covering spaces have many uses, especially in topology. Given a space X, we are interested in spaces that "cover" X in a nice way. Our immediate goal is to use them as a general tool for calculating fundamental groups of topological spaces because there is an intimate connection between them. At the end of the chapter we will use our knowledge of covering spaces to compute the fundamental group of the circle S^1.

Let $p : C \to B$ denote a covering map, where C and B denotes the covering space and base space, respectively. Before we start looking at the definition and what makes p a covering map, we will give a concrete example. Consider the map $p : \mathbb{R} \to S^1$ of the real line onto a circle, in which we think of the real line as an "infinite" spiral floating in the air. We imagine a projection map from $\mathbb{R}^3$ to $\mathbb{R}^2$ given by $(x, y, z) \mapsto (x, y)$, mapping an infinite number of points to the same point on the circle in the plane.

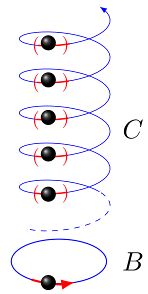


Figure 4.1: Covering map of S^1.

We can also imagine a ball circling down a spiral, and for every time the ball reaches a new red neighbourhood, its shadow in the plane reaches the same neighbourhood over and over again (see Figure 4.1).

Arithmetically, we can define this more precisely. Let us think of $\mathbb{R}$ as the real line, and the circle being the interval I with 0 identified with 1 (see Figure 4.2).

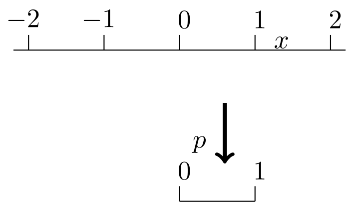


Figure 4.2: Covering map of S^1.

We define $p(x) = x \pmod 1$. Then we have

$$p(3.2) = 3.2 \pmod 1 = 0.2 \quad \text{and} \quad p\left(\frac{7}{5}\right) = \frac{7}{5} \pmod 1 = \frac{2}{5}.$$

There is another way of thinking of this map which is useful and uses some abelian group theory. If we think of $\mathbb{R}$ as an additive abelian group, we can consider the quotient $\mathbb{R}/\mathbb{Z}$, the space of cosets of $\mathbb{Z}$ in $\mathbb{R}$, in which the elements can be expressed as $a + \mathbb{Z}$, $a \in \mathbb{R}$. Then the space of all cosets, $\mathbb{R}/\mathbb{Z}$ is isomorphic to S^1, because the cosets themselves are parametrised by elements belonging to the interval I. We then have

$$\mathbb{R}/\mathbb{Z} \cong S^1 \cong I \text{ with 0 identified with 1.}$$

Now, with that set up we can describe the covering map in an algebraic way:

$$p : \mathbb{R} \to \mathbb{R}/\mathbb{Z},$$

a projection map given by $p(a) = a + \mathbb{Z}$. Then

$$p\left(\frac{3}{2}\right) = \frac{3}{2} + \mathbb{Z} = \frac{1}{2} + \mathbb{Z}.$$

We are reducing the number $(\bmod 1)$, removing an integer part. This is a typical situation of a covering space mapping from a space to a quotient of the space. The question is; what defines it to be a covering map, and not just any map? We will take a look at the definition.

Definition 4.1. A ***covering space*** or cover of a space B is a space C together with a map $p : C \to B$ such that the following hold:

(i) $p : C \to B$ is surjective.

(ii) Every point $b_0 \in B$ has an open neighbourhood $U \subset B$ such that $p^{-1}(U)$ is a disjoint union of open sets, V_α, each of which is mapped by p homeomorphically onto U.

(iii) $p^{-1}(b_0)$ has the discrete topology.

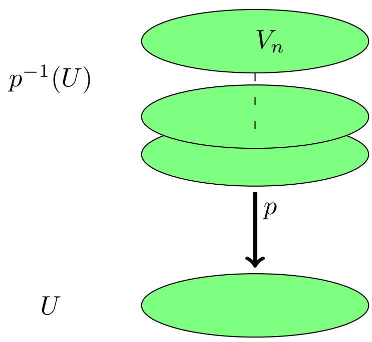


Figure 4.3: Covering map.

We often visualize the pre-image of U as a stack of slices, each slice being a copy of U, floating in the air above U; the map p squashes them all down onto U (see Figure 4.3).

Looking back at Figure 4.1, we can clearly see that it satisfies the conditions for a covering space. If we let x be any point in the base B, representing the circle, there is a discrete number of elements in $p^{-1}(x)$. We can also find an open interval U around x in the base B such that $p^{-1}(U)$ is a disjoint union of open sets in the covering space C, each of which a copy of U (the red neighbourhoods). We will prove this more formally:

Theorem 4.2. *The map $p : \mathbb{R} \to S^1$ given by*

$$p(x) = (\cos 2\pi x, \sin 2\pi x)$$

is a covering map.

Proof. In this case we picture p as a function wrapping the real line $\mathbb{R}$ around S^1.

Let us consider the subset U of S^1 consisting of those points having positive x-coordinates. The pre-image $p^{-1}(U)$ consists of those points $x \in \mathbb{R}$ for which $\cos 2\pi x$ is positive; that is, it is the union of the intervals

$$V_n = \left(n - \frac{1}{4}, n + \frac{1}{4}\right),$$

for all $n \in \mathbb{Z}$ (see Figure 4.4).

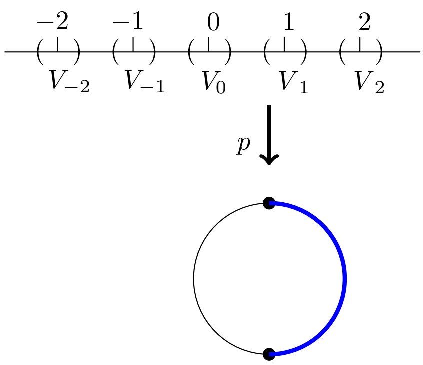


Figure 4.4: Covering map of S^1.

If we restrict ourselves to the closed intervals $\bar{V}_n$, the map is injective, because $\sin 2\pi x$ is strictly monotonic on such an interval (see Figure 4.5).

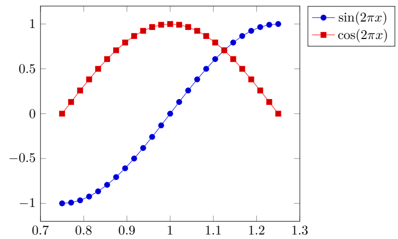


Figure 4.5: $\sin 2\pi x$ and $\cos 2\pi x$ for $x \in \bar{V}_1$.

Moreover, $p : \bar{V}_n \to \bar{U}$ is surjective, and since $\bar{V}_n$ is compact, $p|\bar{V}_n$ is a homeomorphism of $\bar{V}_n$ with $\bar{U}$. In particular, $p|V_n$ is a homeomorphism of V_n with U.

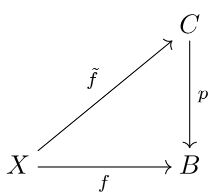
We can easily apply the same arguments to the intersection of S^1 with the upper and lower open half-planes, and with the open left-hand half-plane. Hence, $p : \mathbb{R} \to S^1$ is a covering map. □

Corollary 4.3. *If $p : C \to B$ is a covering map, then p is a local homeomorphism of C with B.*

That is, each point $c \in C$ has a neighbourhood that is mapped homeomorphically by p onto an open subset of B.

4.1 Lifting properties

If $p : C \to B$ is a covering map, then what is the relationship between $\pi_1(C)$ and $\pi_1(B)$? To answer this, we will need to consider paths, and connect paths in B to paths in C. We establish two important results related to this, called the ***Path lifting property*** and the ***Homotopy lifting property***.



The "lifting problem" is to decide when we can "lift" a map $f : X \to B$ to a map $\tilde{f} : X \to C$, where $p : C \to B$ is given. What are the conditions for the diagram to commute? We take a look at the following lemma:

Lemma 4.4. *(**Path lifting property.**) Let $p : C \to B$ be a covering map with $b_0 \in B$, $c_0 \in C$ such that $p(c_0) = b_0$. Given any path $f : I \to B$ beginning at b_0 there exists a unique path $\tilde{f} : I \to C$ starting at c_0 with the lifting property $p \circ \tilde{f} = f$.*

For a proof of Lemma 4.4 , see [4, pp. 342–343]. An illustration is pictured in Figure 4.6:

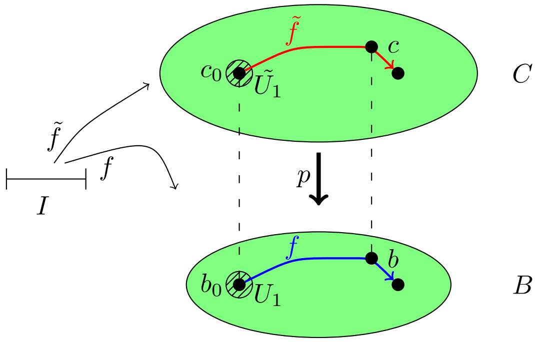


Figure 4.6: Illustration of Lemma 4.4.

Let b_0 be any point in B, and c_0 lying right above b_0 in C such that $p(c_0) = b_0$. Let f be a path in B beginning at b_0. Then Lemma 4.4 says we can lift f to a path $\tilde{f}$ in the covering space C such that is starts at c_0. In other words, going by $\tilde{f}$ to the point c, and then going down via p to b, will be the same as going to b with f. In addition, we want c to be right above b.

How can we lift such a map? Let U_1 be a neighbourhood of b_0. Then $p^{-1}(U_1) = \tilde{U}_1$ is a neighbourhood containing c_0. In fact, the covering property ensures that $\tilde{U}_1$ is an isomorphic copy of U_1. Then consider the path obtained by the intersection of U_1 and f. We can lift it up to a path in C in a unique way because $p|\tilde{U}_1$ is locally an isomorphism (go with $(p|U)^{-1}$).

We repeat the process by considering neighbourhoods $U_2, U_3, \ldots, U_n$, until we have gone through the whole path. We look at an example:

Example 4.5. Let $p : \mathbb{R} \to S^1$ be the covering of Theorem 4.2. We want to lift paths in S^1 to paths in $\mathbb{R}$. Consider the path $f : I \to S^1$ beginning at $b_0 = (1.0)$ given by $f(s) = (\cos \pi s, \sin \pi s)$. Then f lifts to a path $\tilde{f}(s) = s/2$ beginning at 0 and ending at 1/2 (see Figure 4.7).

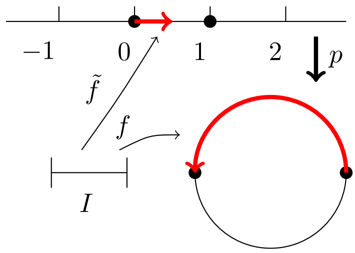


Figure 4.7: Lifting of a path in S^1.

The second lemma we consider is constructed to support the picture of C lying above B, by letting a homotopy in B to be moved "upstairs" to C:

Lemma 4.6. *(**Homotopy lifting property.**) Let $p : C \to B$ be a covering map such that $p(c_0) = b_0$. Let the map $F : I \times I \to B$ be continuous, with $F(0,0) = b_0$. Then there is a unique lifting of F to a continuous map*

$$\tilde{F} : I \times I \to C$$

such that $\tilde{F}(0,0) = c_0$. If F is a path homotopy, then $\tilde{F}$ is a path homotopy.

For a proof of Lemma 4.6, see [4, pp. 343–344]. An illustration is pictured in Figure 4.8.

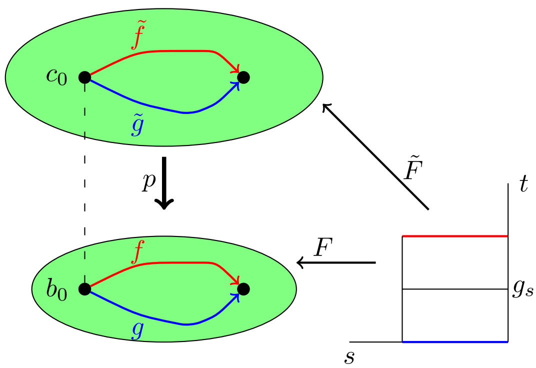


Figure 4.8: An illustration of Lemma 4.6.

Let $p : C \to B$ be a covering space, and let $p(c_0) = b_0$. Let f and g be homotopic paths in B, starting at b_0 with homotopy F.

Then from Lemma 4.6 we know that there is a unique homotopy $\tilde{F}$ from the lifts of f and g to c_0 starting at c_0.

We want to show that $\tilde{f}$ and $\tilde{g}$ are homotopic. To do this we consider the intermediate path g_s which is a part of the "movie" going from g to f. Then lift g_s for each $s \in [0,1]$ to $\tilde{g}_s$ in C. In total we are lifting the homotopy F.

Theorem 4.7. *Let $p : C \to B$ be a covering map such that $p(c_0) = b_0$. Let f and g be two paths in B beginning at b_0 and ending at b_1. Let $\tilde{f}$ and $\tilde{g}$ be their respective liftings to paths in C beginning at c_0. If f and g are path homotopic, then $\tilde{f}$ and $\tilde{g}$ are path homotopic and end at the same point of C.*

Proof. Let f and g be path homotopic and let $F : I \times I \to B$ be the homotopy between them. Since the paths begin at b_0 we know that $F(0,0) = b_0$. Let $\tilde{F} : I \times I \to C$ be the lifting of F to C. Then we know that $\tilde{F}(0,0) = c_0$. By Lemma 4.6, $\tilde{F}$ is a path homotopy, such that $\tilde{F}(0 \times I) = \{c_0\}$ and $\tilde{F}(1 \times I)$ is a one-point set $\{c_1\}$.

Consider the restriction map $\tilde{F}|I \times 0$ of $\tilde{F}$ to the bottom edge of $I \times I$. This is a path in C beginning at c_0 that is a lifting of $F|I \times 0$. Since lifting of paths are unique, we must have $\tilde{F}(s,0) = \tilde{f}(s)$. Similarly, $\tilde{F}|I \times 1$ is a path on C that is a lifting of $F|I \times 1$, and it begins at c_0 because $\tilde{F}(0 \times I) = \{c_0\}$. Again, since liftings are unique, $\tilde{F}(s,1) = \tilde{g}(s)$. Hence, both $\tilde{f}$ and $\tilde{g}$ end at c_1, and $\tilde{F}$ is a path homotopy between them. □

Definition 4.8. Let $p : C \to B$ be a covering map such that $p(c_0) = b_0$. Given an element $[f]$ of $\pi_1(B, b_0)$, let $\tilde{f}$ be the lifting of f to a path in C that begins at c_0. Let the map

$$\phi : \pi_1(B, b_0) \to p^{-1}(b_0)$$

be such that $\phi([f])$ denotes the end point $\tilde{f}(1)$ of $\tilde{f}$. Then ϕ is a well-defined set map. We call ϕ the ***lifting correspondence*** derived from the covering map p.

Theorem 4.9. *Let $p : C \to B$ be covering map such that $p(c_0) = b_0$. If C is path connected, then the lifting correspondence*

$$\phi : \pi_1(B, b_0) \to p^{-1}(b_0)$$

is surjective. If C is simply connected, it is bijective.

Proof. Let C be path connected, and let $c_1 \in p^{-1}(b_0)$. Then there is a path $\tilde{f}$ in C from c_0 to c_1. Then $f = p \circ \tilde{f}$ is a loop in B at b_0 and $\phi([f]) = c_1$ by Definition 4.8.

Suppose C is simply connected. Let $[f]$ and $[g]$ be two elements of $\pi_1(B, b_0)$ such that $\phi([f]) = \phi([g])$. Let $\tilde{f}$ and $\tilde{g}$ be the liftings of f and g, respectively, to paths in C that begin at c_0. Then, by Theorem 4.7, $\tilde{f}(1) = \tilde{g}(1)$. Then by Lemma 3.14 there is a path homotopy $\tilde{F}$ in C between $\tilde{f}$ and $\tilde{g}$. Then $p \circ \tilde{F}$ is a path homotopy in B between f and g. Since simply connectedness implies path connectedness we have that ϕ is both surjective and injective; hence we have a bijective correspondence. □

4.2 The fundamental group of S^1

We have already seen that the the map $\mathbb{R} \to S^1$ defines a covering map of S^1. To visualize this, we think of the real line as an "infinite" spiral (see Figure 4.1) as we did before, covering the circle and mapping an infinite number of points to the same point on the circle. Using Theorem 4.9, and the fact that $\mathbb{R}$ is simply connected we might suggest that the fundamental group is the integers, $\mathbb{Z}$.

Consider S^1 after we have identified 0 with 1, and parametrize from 0 to 1. We want to describe loops on S^1 and we find it natural to choose $0 = 1$ as our base point (see Figure 4.9).

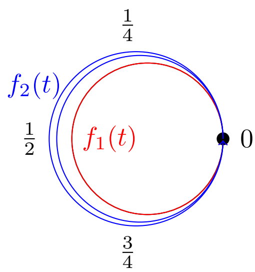


Figure 4.9: Paths f_1 and f_2 going around the circle once, and twice, respectively.

We define the loop f_1 such that $f_1(t) = t$. As t ranges from 0 to 1, f_1 goes once around the circle in the positive direction. We then define f_2 such that $f_2(t) = 2t \pmod 1$ going around the circle twice. More generally, we have $f_n(t) = nt \pmod 1$. Further there is no need to restrict ourselves to positive multiplies. Then $f_{-1}(t) = -t \pmod 1$ is a path going once around the circle in the negative direction.

What does the multiplication look like? The product $f_1 * f_{-2}$ is a loop going once around the circle in the positive direction, and then it goes around twice in the negative direction. In total the path has gone once in the negative direction; hence the loop must be homotopic to f_{-1}. Remember, in turns of the fundamental group we are not multiplying the loops themselves, but the homotopy classes of the loops. In general we have:

$$[f_m] * [f_n] = [f_{m+n}].$$

The operation on the elements acts as to adding the integer indices.

Then all loops in S^1 are characterized by the number of times they wind around the origin. A positive integer i is isomorphic to a loop winding i times counterclockwise; similarly for negative integers winding clockwise. The concatenation of loops in $\pi_1(S^1)$ is equivalent to addition of integers in $\mathbb{Z}$.

Theorem 4.10. *The fundamental group of S^1 is isomorphic to the additive group of integers, $\mathbb{Z}$.*

Proof. We need to show that the map $\theta : \mathbb{Z} \to \pi_1(S^1)$ mapping an integer n to the homotopy class of the loop $f_n(s)$ is an isomorphism.

We have already seen in Theorem 4.2 that the map $p : \mathbb{R} \to S^1$ given by $p(s) = (\cos 2\pi s, \sin 2\pi s)$ is a map covering S^1.

Using the fact that S^1 is path connected, we can choose the base point to be any point on S^1. We choose $f_n(s) = (\cos 2\pi ns, \sin 2\pi ns)$ be a loop in S^1 based at $(1, 0)$, winding around the circle n times. Let $\tilde{f_n} : I \to \mathbb{R}$ be the path starting at 0 and ending at n, defined by $\tilde{f_n}(s) = ns$. Then $\tilde{f_n}$ is a lifting of $f_n = p \circ \tilde{f_n}$.

Set $\theta(n) = [p \circ \tilde{g}]$, $\tilde{g}$ a path in $\mathbb{R}$ from 0 to n. Then $\tilde{g}$ is homotopic to $\tilde{f_n}$ by the straight-line homotopy defined in Example 2.8. Hence $p \circ \tilde{g}$ is homotopic to $p \circ \tilde{f_n} = f_n$ by Theorem 4.7 and the definition of $\theta(n)$ holds.

We verify that θ is an isomorphism. Let $\pi_m : \mathbb{R} \to \mathbb{R}$ be defined by $\pi_m(x) = x+m$. Then $\tilde{f_m} * (\pi_m \circ \tilde{f_n})$ is a path in $\mathbb{R}$ from 0 to $n+m$. The image of this path under p is $\theta(n+m)$, the homotopy class of the loop in S^1 going around the circle $n+m$ times. The image is in fact $f_n * f_m$, so $\theta(m + n) = \theta(m) * \theta(n)$.

What is left is to show that θ is a bijection. Let f_n be as before, and let $[f_n]$ represent an element of $\pi_1(S^1)$. By Lemma 4.4 we have a lift $\tilde{f_n}$ starting at 0 and ending at some integer n, since $p \circ \tilde{f_n}(1) = f_n(1) = (1, 0)$ and $p^{-1}(1, 0) = \mathbb{Z}$. Then we have $\theta(n) = [p \circ \tilde{f_n}] = [f_n]$.

We show that θ is injective. Let $\theta(m) = \theta(n)$. Then $f_n \simeq_p f_m$. Let F be the homotopy between f_n and f_m. By Lemma 4.6 the homotopy lifts to a homotopy $\tilde{F}$ starting at 0, and from Theorem 4.7, the lifted paths end at the same point, so $m = n$. □

CHAPTER 5

HOMOTOPY EQUIVALENCE

We have now seen that working with covering spaces is a useful tool for studying the fundamental group of a space. Another method that we will use for studying fundamental groups is called ***homotopy equivalence***. It provides a way for reducing the problem of computing the fundamental group of a space to that of computing fundamental group of some other space—preferably one that is more familiar.

To take an everyday example, there are many ways to draw the letters of the alphabet. One way is to draw them either thick or thin. The thin letter X will obviously be a subspace of the thick letter **X**, and we can continuously shrink the thick letter to the thin one. We will think of this shrinking process as taking place during a time interval $t \in I$, and it will define a family of functions $f_t : \mathbf{X} \to \mathbf{X}$, letting $f_t(x)$ be the point to which a point $x \in \mathbf{X}$ has moved at time t. Examples like these will lead us to the following definitions:

Definition 5.1. If $A \subset X$, a ***retraction*** of X onto A is a continuous map $r : X \to A$ such that $r(a) = a$ for all $a \in A$ (i.e. $r|A$ is the identity map of A).

We think of it as a continuous map of X onto a subspace A, leaving each point of the subspace fixed.

Definition 5.2. Let $A \subset X$. A continuous map

$$H : X \times I \to X$$

is a ***deformation retraction*** of X onto a subspace A if the following holds:

(i) $H(x, 0) = x$ and $H(x, 1) \in A$ for all $x \in X$,

(ii) $H(a, t) = a$ for all $a \in A$.

The subspace A is called a ***deformation retract*** of X. The map $r : X \to A$ defined by $H(x, 1)$ is a retraction of X onto A, and H is a homotopy between id_X and the map $j \circ r$, where $j : A \to X$ is the inclusion map.

A deformation retraction is a homotopy between the identity map on X and a retraction. It captures the idea of continuously shrinking a space X to a subspace A (see Figure 5.1). A retraction however, does not need to be a deformation retraction. We take a look at an example.

Example 5.3. Let $x_0 \in X$, and $\{x_0\}$ be a retraction of X. Then the map $r : X \to \{x_0\}$ satisfies $r(x_0) = x_0$.

For $\{x_0\}$ to be a deformation retract of X there has to be a map $H : X \times I \to X$ such that $H(x, 0) = x$, $H(x, 1) = x_0$ and $H(x_0, t) = x_0$ for all $t \in I$. This gives us a homotopy between id_X and a constant map at x_0. A space X with a homotopy satisfying these properties is called a ***contractible space*** as we will see later in the chapter.

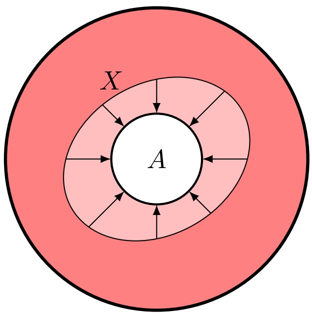


Figure 5.1: Deformation retraction: X retracts to subset A

Definition 5.4. Given two spaces X and Y, we say they are ***homotopy equivalent*** if there exist continuous maps $f : X \to Y$, $g : Y \to X$ such that $g \circ f$ is homotopic to id_X and $f \circ g$ is homotopic to id_Y. The maps f and g are called ***homotopy equivalences***, and each is said to be a ***homotopy inverse*** of the other.

Looking closer at these definitions, we observe that if A is a deformation retract of X, then A must be homotopy equivalent to X. Let $j : A \to X$ be the inclusion map and $r : X \to A$ be a retraction map. Then the composite map $r \circ j$ equals the identity map of A, and $j \circ r$ is homotopic to the identity map of X by hypothesis. This is an important fact, and we will use it later in the chapter (see Example 5.12).

Definition 5.5. Two spaces, X and Y that are homotopy equivalent are said to have the same ***homotopy type***. These are spaces that can be deformed continuously into one another. We write $X \simeq Y$.

5.1 Homotopy equivalence and homeomorphism

A ***homeomorphism*** is related to homotopy equivalence, but it is important to be able to distinguish them from each other. We look closer at the definition.

Definition 5.6. Let X and Y be topologically spaces and let $f : X \to Y$ be a bijection. If both the function f and the inverse function $f^{-1} : Y \to X$ are continuous, then f is called a ***homeomorphism***.

In fact, every homeomorphism $f : X \to Y$ is a homotopy equivalence: simply take $g = f^{-1}$. However, the converse is not true in general. Take a solid disk and the single point as an example. They are homotopy equivalent (deform the disk along radial lines continuously to a point). On the other hand, they are not homeomorphic (there is no bijection between them).

A homeomorphism preserves the topological structure involved and is “stronger” in the sense that the deformation is bijective. Homotopy allows squashing and does not require the spaces to have the same dimension.

Example 5.7. Let X be the letter “X” and Y be the letter “Y”. Then X and Y are homotopy equivalent, but they are not homeomorphic.

(Sketch proof.) Let $f : X \to Y$ map the three segments of “X” onto the “Y” letter in a obvious way, and let it map the fourth segment to the centre point. Let $g : Y \to X$ be the mapping of “Y” onto three segments of “X”. Then f and g are both continuous. The map f is surjective, but not injective, while the map g is injective, but not surjective. Now, the compositions $g \circ f$ and $f \circ g$ are both easily seen to be homotopic to id_X and id_Y, respectively, so “X” and “Y” are homotopy equivalent.

However, “X” and “Y” are not homeomorphic. To see this, removing the point at the centre of the “X” gives us a space with four connected components, while removing any point from “Y” gives at most three connected components.

Theorem 5.8. *Homotopy equivalence is an equivalence relation on topological spaces.*

Proof. We need to verify the three properties, reflexibility, symmetry and transitivity.

(1) (**Reflexibility.**) Given X, we must show $X \simeq X$. The identity map $id_X : X \to X$ is a homeomorphism, and thus a homotopy equivalence.

(2) (**Symmetry.**) Given $X \simeq Y$, we must show $Y \simeq X$. Suppose $f : X \to Y$ is a homotopy equivalence. Then from Definition 5.4, f has a homotopy inverse g. Then $g : Y \to X$ is a homotopy equivalence with homotopy inverse f.

(3) (**Transitivity.**) Given $X \simeq Y$ and $Y \simeq Z$, we must show $X \simeq Z$. Suppose $f : X \to Y$ is a homotopy equivalence with homotopy inverse g, and $h : Y \to Z$ is a homotopy equivalence with homotopy inverse k. Then, using Theorem 2.11 (and the associativity of compositions) it follows that $h \circ f : X \to Z$ is a homotopy equivalence with homotopy inverse $g \circ k$.

□

5.2 Induced isomorphisms

We studied homomorphisms of groups in Chapter 3. There we only considered maps $f : X \to Y$ that were continuous. Let us now extend their properties and consider the two cases in which we both have a homeomorphism and a homotopy equivalence. Then the maps will induce an isomorphism of fundamental groups.

Theorem 5.9. *Let $h : (X, x_0) \to (Y, y_0)$ be a homeomorphism. Then the map $h_* : \pi_1(X, x_0) \to \pi_1(Y, y_0)$ is an isomorphism.*

Proof. We know that h_* is a homomorphism. What is left is to check that h_* is bijective. Let $h^{-1} : (Y, y_0) \to (X, x_0)$ be the inverse of h. From Theorem 3.17 we have $h_*^{-1} \circ h_* = (h^{-1} \circ h)_* = id_{X*}$. Similarly, $h_* \circ h_*^{-1} = (h \circ h^{-1})_* = id_{Y*}$, where id_Y is the identity map of (Y, y_0). Since id_{X*} and id_{Y*} are the identity homomorphisms of the group $\pi_1(X, x_0)$ and $\pi_1(Y, y_0)$, respectively, h_*^{-1} is the inverse of h_*. Hence, h_* is an isomorphism. □

Example 5.10. The fundamental group of the torus, $\pi_1(S^1 \times S^1) = \mathbb{Z} \times \mathbb{Z}$. For a proof, see [4, p. 372]. Since we already have seen that $\pi_1(S^1) = \mathbb{Z}$ we conclude that the torus is not homeomorphic to the circle.

Theorem 5.11. *Let $f : X \to Y$ be a continuous map such that $f(x_0) = y_0$. If f is a homotopy equivalence, then $f_* : \pi_1(X, x_0) \to \pi_1(Y, y_0)$ is an isomorphism.*

Proof. Let $g : Y \to X$ be a homotopy inverse for f. Consider the maps

$$(X, x_0) \xrightarrow{f} (Y, y_0) \xrightarrow{g} (X, x_1) \xrightarrow{f} (Y, y_1)$$

where $g(y_0) = x_1$ and $f(x_1) = y_1$. Then we can study the corresponding induced homomorphisms of groups:

$$\pi_1(X, x_0) \xrightarrow{(f_{x_0})_*} \pi_1(Y, y_0) \xrightarrow{g_*} \pi_1(X, x_1) \xrightarrow{(f_{x_1})_*} \pi_1(Y, y_1).$$

Since f and g are homotopy equivalences, we have that the map

$$g \circ f : (X, x_0) \longrightarrow (X, x_1)$$

is homotopic to the identity map id_X. Then there is a path α in X such that

$$(g \circ f)_* = \hat{\alpha} \circ (id_X)_* = \hat{\alpha}, \tag{5.1}$$

and it follows that the homomorphism

$$(g \circ f)_* = g_* \circ (f_{x_0})_* \tag{5.2}$$

is an isomorphism.

Similarly, because $f \circ g$ is homotopic to the identity map id_Y, the homomorphism

$$(f \circ g)_* = (f_{x_1})_* \circ g_* \tag{5.3}$$

is an isomorphism.

Then (5.2) implies that g_* is surjective, and (5.3) implies that g_* is injective, so g_* is an isomorphism. By combining (5.2) and (5.1) we conclude that

$$(f_{x_0})_* = (g_*)^{-1} \circ \hat{\alpha},$$

so that $(f_{x_0})_*$ is an isomorphism. □

Example 5.12. A solid ball deformation retracts to a point, so it must also be homotopy equivalent to a point. Since the ball is a subset of $\mathbb{R}^n$, we know from Example 3.6 that its fundamental group is trivial. Then, from Theorem 5.11 the same must be for the one-point space. We observe that $\mathbb{R}^2 \setminus (\mathbf{0})$ deformation retracts to S^1, so it must also be homotopy equivalent to S^1. Hence, the fundamental group of $\mathbb{R}^2 \setminus (\mathbf{0})$ must be isomorphic to $\mathbb{Z}$.

5.3 Contractible spaces

In Example 5.12 we stated that the solid ball was homotopy equivalent to a point. We will prove that any space having these properties, is a ***contractible*** space. Intuitively, a contractible space is one that can be continuously shrunk to a point. It is a stronger condition than simply connectedness and does not allow holes of any dimension. Remember that the sphere is simply connected, even though it has a "hole" in the hollow center. We start looking at the definition.

Definition 5.13. A space X is said to be ***contractible*** if the identity map $id_X : X \to X$ is nullhomotopic.

Example 5.14. $\mathbb{R}$ is contractible. We have to show there exist a homotopy F between the identity map $id_{\mathbb{R}}$ and the constant map $e_0(x) = 0$. We choose our homotopy to be $F : \mathbb{R} \times I \to \mathbb{R}$ such that

$$F(x, 0) = x \quad \text{and} \quad F(x, 1) = 0$$

for each $x \in \mathbb{R}$. The map $F(x,t) = (1-t) \cdot id_{\mathbb{R}}(x) = (1-t)x$ is well-defined, continuous, and satisfies the conditions.

In fact, $\mathbb{R}^n$ is contractible, and any convex subset $U \subset \mathbb{R}^n$ is contractible. In particular, the disk in Example 3.6 is contractible.

Theorem 5.15. *A space is contractible if and only if it has the homotopy type of a point.*

Proof. $\Rightarrow$ Let X be a contractible space. Then we have that id_X is homotopic to a constant map $f : X \to X$ such that $f(x) = x_0$. Consider a one-point space $\{x_0\}$. Then we have a map from X to $\{x_0\}$. If $j : \{x_0\} \to X$ is the inclusion map, then $f \circ j = id_{\{x_0\}}$, so they are trivially homotopic. Further, $j \circ f = f$ and since $f \simeq id_X$, then $j \circ f \simeq id_X$ and X is homotopy equivalent to a one-point space.

$\Leftarrow$ Let X be homotopy equivalent to a one-point space $Y = \{y_0\}$. Then there exist maps $f : X \to Y$ and $g : Y \to X$ such that $f \circ g \simeq id_Y$ and $g \circ f \simeq id_X$. Since Y only consists of one point, $g(y_0) = x_0$ for some $x_0 \in X$. For all $x \in X$ we have $g \circ f(x) = g(y_0) = x_0$. Then $g \circ f$ must be the constant map. Since $g \circ f \simeq id_X$, id_X must be homotopic to the constant map. □

From Theorem 5.11, we know that the fundamental groups of a contractible space and a one-point space are isomorphic. Since the fundamental group of a one-point space is trivial (see Example 5.12), it follows immediately that the same must hold for a contractible space.

Now, what do we mean by saying that the fundamental group of a space X is trivial? Given an element $[f] \in X$ there has to exist a homotopy between f and

e_{x_0}, the identity element. We will prove that such a homotopy exists, but to do that we need to make use of a result that we will prove in the following theorem:

Theorem 5.16. *A contractible space is path connected.*

Proof. Let a homotopy $G : X \times I \to X$ such that $G(x, 0) = id_X$ and $G(x, 1) = x_0$ for some fixed $x_0 \in X$. For any point $x \in X$, the function $f(t) = G(x, t)$ takes I into X and satisfies $f(0) = x$ and $f(1) = x_0$. Consequently every $x \in X$ is connected by a path to x_0, so it must be path connected. □

This is obvious since every point will be contracted to a fixed point x_0, and we therefore have a path from every point to the fixed point. Let f and g be two such paths. If we then go f, followed by the reverse $\bar{g}$, we have a path between any two points.

Theorem 5.17. *The fundamental group of a contractible space X is trivial.*

Proof. Since X is contractible, we know that X is path connected. Then all groups are isomorphic, independent of the based point. Consider the group $\pi_1(X)$, and choose the basepoint x_0 to be the contraction point of X. Since X is contractible, we know that a homotopy between id_X and e_{x_0} exists. Let $H : X \times I \to X$ be this homotopy such that

$$H(x, 0) = id_X(x) = x \quad \text{and} \quad H(x, 1) = e_{x_0}(x)$$

for each x. Let f be a loop, and let $[f] \in \pi_1(X)$. If we show $[f] = [e_{x_0}]$ we are done.

We need to find a homotopy between f and e_{x_0}. Let $F : I \times I \to X$ be a homotopy given by

$$F(s, t) = H(f(s), t).$$

Then we have

(i) $F(s, 0) = H(f(s), 0) = f(s)$ for all s,

(ii) $F(s, 1) = H(f(s), 1) = e_{x_0}(f(s)) = x_0$ for all s,

(iii) $F(0, t) = H(f(0), t) = H(x_0, t)$ for all t,

(iv) $F(1, t) = H(f(1), t) = H(x_0, t)$ for all t.

There is no reason for $H(x_0, t) = x_0$. Hence, we need to modify F:

Let $n : I \to I \times I$ be a path in $I \times I$ such that $n(0) = (0, 0)$ and $n(1) = (1, 0)$ (travelling along the bottom edge). Further, let $\phi : I \to I \times I$ be another path such that $\phi(0) = (0,0)$ and $\phi(1) = (1,0)$ (travelling first up the left edge, then along the top edge, and down the right edge). Since $I \times I$ is convex, we have a homotopy $G : I \times I \to I \times I$ between them. It has the following properties:

(i) $G(0,t) = (0,0)$ and $G(1,t) = (1,0)$ for all t,

(ii) $G(s,0) = (s,0)$ and $G(s,1) = \phi(s)$ for all s.

Let us now define $F' : I \times I \to X$ to be the homotopy given by

$$F'(s,t) = F(G(s,t)).$$

It has the following properties:

(i) $F'(0,t) = F(G(0,t)) = F(0,0) = H(f(0),0) = f(0) = x_0$ for all t,

(ii) $F'(1,t) = F(1,0) = H(f(1),0) = f(1) = x_0$ for all t,

(iii) $F'(s,0) = F(s,0) = H(f(s),0) = f(s)$ for all s,

(iv) $F'(s,1) = F(\phi(s))$.

We are still not there. If we let $\Phi : I \to X$ be a path such that $\Phi(t) = H(x_0,t)$ we get that $F(\phi(s)) = (\Phi * e_{x_0}) * \bar{\Phi})$. We verify this:

We need to choose a bracketing for $\phi(s)$: let us suppose that we concatenate the left hand side and top first, and then concatenate the result with the right and side. Then we get

$$\phi(s) = \begin{cases} (0,4s) & \text{for } s \in [0,\frac{1}{4}], \\ (4s-1,1) & \text{for } s \in [\frac{1}{4}\frac{1}{2}], \\ (1,2-2s) & \text{for } s \in [\frac{1}{2},1]. \end{cases}$$

Then from the definition of F we know that

$$F(\phi(s)) = \begin{cases} F(0,4s) = H(f(0),4s) = H(x_0,4s) = \Phi(4s) & \text{for } s \in [0,\frac{1}{4}], \\ F(4s-1,1) = H(f(4s-1),1) = x_0 & \text{for } s \in [\frac{1}{4},\frac{1}{2}], \\ F(1,2-2s) = \Phi(2-2s) = \Phi^{-1}(2s-1) & \text{for } s \in [\frac{1}{2},1]. \end{cases}$$

which gives us that $F(\phi(s)) = (\Phi * e_{x_0}) * \bar{\Phi}$.

From the construction of the fundamental groups we know that the concatenation $(\Phi * e_{x_0}) * \bar{\Phi}$ is homotopic to e_{x_0}. We denote this homotopy by $H^{'}$. Then we can ensure that $H^{'}(0,t) = x_0$ and $H^{'}(1,t) = x_0$ for all t.

By concatenating $F^{'}$ and $H^{'}$, we obtain a homotopy

$$F''(s,t) = \begin{cases} F'(s,2t) & \text{for } t \in [0, \frac{1}{2}], \\ H'(s,2t-1) & \text{for } t \in [\frac{1}{2}, 1]. \end{cases}$$

Then we get

(i) $F^{''}(0,t) = x_0$ for all t,

(ii) $F''(1,t) = x_0$ for all t,

(iii) $F^{''}(s,0) = F^{'}(s,0) = f(s)$ for all s,

(iv) $F^{''}(s,1) = H'(s,1) = e_{x_0}(s) = x_0$ for all s.

Hence, we have a homotopy between the loops f and e_{x_0} and we conclude that $\pi_1(X) = 0$. □

CHAPTER 6

ROTATION GROUPS

When we think of rotation, there are many associations that come to mind. A child would probably associate it to a carousel or roller coaster, while others probably think of wheels running or the earth rotating. These are all rotations in three dimensional space that we will study in the following chapters.

6.1 Orthogonal groups

We start by defining the orthogonal group. We recall that an orthogonal matrix is a matrix whose column vectors are mutually orthonormal. Let $\mathbb{R}_{n\times n}$ be the set of all real $n \times n$ matrices. Then we denote

$$O(n) = \{R \in \mathbb{R}_{n\times n} \mid R^T R = RR^T = I \text{ and } \det(R) = \pm 1\}$$

where the group operation is given by matrix multiplication.

Theorem 6.1. *The set $O(n)$ is a group under matrix multiplication.*

Proof. It is obvious that we have an identity element I since the inverse is equal to the transpose, so every element has an inverse. Since matrix multiplication is associative, it is clear that $O(n)$ is associative. What is left is to show that $O(n)$ is closed and has determinant ± 1.

To prove that $O(n)$ is closed, we consider two arbitrary elements, $R, P \in O(n)$. Then we have the following:

$$\begin{aligned}(RP)(RP)^T &= RPP^TR^T\\ &= RR^T\\ &= I.\end{aligned}$$

Hence, $(RP)^T = (RP)^{-1}$ and $O(n)$ is closed under matrix multiplication.

From the properties of determinants, we know that $\det(R) = \det(R^T)$ and $\det(RP) = \det(R)\det(P)$. Let $R \in O(n)$ and observe the following:

$$\begin{aligned}\det(R)^2 &= \det(R)\det(R^T)\\ &= \det(RR^T)\\ &= \det(I)\\ &= 1.\end{aligned}$$

Hence, $\det(R)^2 = 1$, so all matrices must have determinant ± 1. □

6.2 Rotation groups

If we further restrict ourselves to those matrices having determinant 1 we get a subgroup $SO(n) \subset O(n)$. This is obvious, since $\det(R) = \det(P) = 1$ implies that $\det(RP) = 1$, so it is closed. The group is also referred to as the rotation group, since in dimensions two and three, its elements are rotations around a point and a line, respectively. In general the group acts on $\mathbb{R}^n$ and is the group of all rotations of $\mathbb{R}^n$.

Definition 6.2. The group $SO(n) \subset O(n)$ is called the ***rotation group***.

Given two rotations, a composition results in another rotation, and given a rotation we can always find the inverse by rotating in the opposite direction. The identity matrix is indeed a rotation, the trivial one. Since matrix multiplication is associative, rotations obey this property, too.

The algebraic structure of $SO(n)$ is coupled with a topological structure. Performing operations as multiplication and finding inverses are continuous functions, which give us a topological group. The group is in fact a *Lie group* since operations also are smooth.

In this thesis, we will look at rotation groups, and especially we will study the $SO(3)$ group. In the next chapter we will find its corresponding fundamental group.

However, to get a good grip of what we are working with, we turn to lower dimensions and start by looking at 2×2 matrices, and the $SO(2)$ group.

6.3 The SO(2) group

We start by looking at an example:

Example 6.3. Let A be a symmetric 2×2 matrix. Using linear algebra, we can find a basis of eigenvectors, $\beta = \{e_1, e_2\}$ such that the matrix $P = [e_1, e_2]$ is orthogonal. Consider the following system:

$$3x^2 + 10xy + 3y^2 = 8 \Rightarrow \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 8. \tag{6.1}$$

A function of the form (6.1) is called a ***quadratic form*** and represents a conic section. The equation can be written in matrix form as $\mathbf{x}^T A \mathbf{x} = 8$, A being our symmetric matrix.

Since A is symmetric, A is also orthogonal diagonalizable, which means that $A = PDP^{-1} = PDP^T$, where D is a diagonal matrix, and P is an orthogonal eigenvalue matrix.

Having done some linear algebra we get the matrix

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

What kind of curve are we studying? Solving problems like these are based on simplifying the quadratic form $\mathbf{x}^T A \mathbf{x}$ by making a substitution $\mathbf{x} = P\mathbf{X}$, expressing the variables (x, y) in terms of new variables (X, Y). We obtain a new quadratic form which is easier to recognize and sketch;

$$8X^2 - 2Y^2 = 8 \Rightarrow X^2 - \left(\frac{Y}{2}\right)^2 = 1.$$

We recognize this as the equation of a hyberbola. However, by the substitution we also changed the coordinate system from (x, y) to (X, Y). We have pictured the hyperbola in the originally coordinate system (x, y) in Figure 6.1.

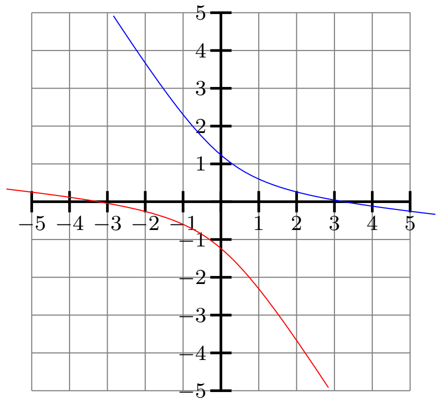


Figure 6.1: Hyperbola formed by a conic section.

To summarize, we started with an equation that was hard to interpret, and after some linear algebra and a substitution we obtained a generalized equation that we could recognize. But what has this to do with the $SO(2)$ group? The matrix P used in the substitution acts as to rotate the coordinate system. The orthonormal vectors in P form a basis for $\mathbb{R}^2$, and by multiplying with P we obtain a new axis system which is rotated relatively to the original xy-axes. We know that every rotation maps an orthonormal basis to another orthonormal basis, and this is exactly what happens.

In Example 6.1 the axis system is rotated 45° relative to the original one:

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \cos 45^\circ & -\sin 45^\circ \\ \sin 45^\circ & \cos 45^\circ \end{bmatrix}.$$

We are ready to construct our 2×2 rotation matrices in $SO(2)$.

Definition 6.4. If P is a matrix such that $P \in SO(2)$, then P can be written on the following from:

$$P(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

P has the effect of rotating the xy-axes of a coordinate system counterclockwise through positive angles θ.

Remark. We observe that all matrices on the form P are orthogonal, just multiply the column vectors. We also see they have determinant 1 by the trigonometric identity $\cos^2\theta + \sin^2\theta = 1$.

Note. Although we in Example 6.1 observe that the rotation matrices represent arbitrary rotations of the axes about the origin, it is worth mentioning that a rotation P also acts as to rotate points in the plane counterclockwise through positive angles θ about the origin. It is also this last approach we choose when describing rotations in this thesis.

Definition 6.5. Let $P \in SO(2)$. Then P rotates column vectors by means of the following matrix multiplication,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

where the coordinates (x', y') of the point (x, y) after rotation are

$$x' = x\cos\theta - y\sin\theta,$$

$$y' = x\sin\theta + y\cos\theta.$$

Example 6.6. Rotate the point $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$ $3\pi/2$ around the unit circle. Then we have

$$P\left(\frac{3\pi}{2}\right) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

and the coordinates after rotation are

$$x' = \frac{1}{2}\cdot 0 - \frac{\sqrt{3}}{2}\cdot -1 = \frac{\sqrt{3}}{2},$$

$$y' = \frac{1}{2}\cdot -1 + \frac{\sqrt{3}}{2}\cdot 0 = -\frac{1}{2}.$$

6.3.1 Commutativity of SO(2)

Is $SO(2)$ abelian? We consider two elements of $SO(2)$:

$$P = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad R = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

We have to check commutativity of P and R:

$$
\begin{aligned}
PR &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta\cos\phi - \sin\theta\sin\phi & -\cos\theta\sin\phi - \sin\theta\cos\phi \\ \sin\theta\cos\phi + \cos\theta\sin\phi & -\sin\theta\sin\phi + \cos\theta\cos\phi \end{bmatrix} \\
&= \begin{bmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{bmatrix}.
\end{aligned}
$$

We swap the elements, and get:

$$
\begin{aligned}
RP &= \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{bmatrix}.
\end{aligned}
$$

The elements in $SO(2)$ commute; hence $SO(2)$ is abelian.

6.3.2 Fundamental group of SO(2)

From complex analysis we learn that every complex number $z = x + iy$ can be represented by a 2×2 matrix,

$$
A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.
$$

Specially, we have that every complex number $e^{i\theta} = \cos\theta + i\sin\theta$, also known as the elements of $U(1)$, can be represented by the 2×2 matrix,

$$
A' = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.
$$

We also know from complex analysis that the geometric description of multiplication of unit complex numbers is rotation around the unit circle. In fact we have a bijective correspondence,

$$
e^{i\theta} \leftrightarrow \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},
$$

between $U(1)$ and the $SO(2)$ group.

Geometrically, we observe that $U(1)$ is the unit circle

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

which makes it clear that the circle group is, by the correspondence above, isomorphic to the group $SO(2)$, and hence,

$$\pi_1(SO(2)) \cong \pi_1(S^1) = \mathbb{Z}.$$

6.4 The SO(3) group

As we saw in the previous section, it was quite easy to construct elements for $SO(2)$. On the contrary, for $SO(n)$ as n grows larger, determinants get quite harder to compute and thus it is more difficult to check whether a matrix is in $SO(n)$ or not. However, since we already know what elements in $SO(2)$ look like, we try to insert one into a 3×3 matrix.

Consider the matrices,

$$P(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad R(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

and

$$Q(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}.$$

Since $P^T P = R^T R = Q^T Q = I$ and $\det(P) = \det(R) = \det(Q) = 1$ the matrices are elements in $SO(3)$.

In fact, many copies of n-dimensional rotations are found within $(n+1)$-dimensional rotations, as subgroups. Each embedding leaves one direction fixed, as in case of $\mathbb{R}^3$ which fixes a unique one-dimensional linear subspace of $\mathbb{R}^3$. This is called the ***axis of rotation***. Any rotation in three dimensions can be represented by a pair (u, θ), consisting of a unit vector u indicating the direction of the axis of rotation, together with an ***angle of rotation*** θ about the axis.

Note. The matrix R defined above is a rotation about the positive x-axis by angle θ, P is the rotation about the positive z-axis and Q is the rotation about the positive y-axis.

Note. Rotation of vectors in $\mathbb{R}^3$ appears counterclockwise when u points towards the observer (right-hand rule), and the angle θ is positive.

6.4.1 Determining the angle and the axis

Given a matrix R it can be difficult to find the pair (u, θ).

We know that a vector u parallel to the axis of rotation has to satisfy

$$Ru = u$$

since u is fixed by the rotation. We may rewrite this equation as

$$Ru = Iu \Rightarrow (R - I)u = 0.$$

Hence, to find u is the same as to find the nullspace of $R - I$.

To find θ, pick a vector v perpendicular to the rotation axis. Then θ is the angle between v and Rv.

Of course, there are other ways as well. We take a look at an example.

Example 6.7. Consider the matrix,

$$M = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We can check that $M^{-1} = M^T$ and $\det(M) = 1$, so $M \in SO(3)$, and we also observe that it leaves the z-axis fixed.

To find θ, we observe what M does to an arbitrary vector $v = (1, 1, 0)$ in the xy-plane:

$$Mv = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

We know that θ is the angle between $(1, 1, 0)$ and $(-1, 1, 0)$. Since the dot product of the two vectors is 0, the vectors are orthogonal, which means that θ is equal to $\pi/2$.

6.4.2 Non-commutativity of SO(3)

We have already seen that $SO(2)$ is abelian. Does the same hold for $SO(3)$? We consider P, $R \in SO(3)$ with

$$PR = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$
$$= \begin{bmatrix} \cos\theta & -\sin\theta\cos\theta & \sin^2\theta \\ \sin\theta & \cos^2\theta & -\sin\theta\cos\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}.$$

We swap the elements, and get:

$$RP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta\cos\theta & \cos^2\theta & -\sin\theta \\ \sin^2\theta & \sin\theta\cos\theta & \cos\theta \end{bmatrix}.$$

So elements in $SO(3)$ do not in general commute and $SO(3)$ is not abelian. However, it is worth mentioning that there do exist commutative subgroups of $SO(n)$ for all n. We look at an example:

Example 6.8. The subgroup of $SO(3)$ consisting of the matrices of the form

$$P = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is commutative. They are rotations in the xy-plane which we have already seen are commutative.

CHAPTER 7

THE FUNDAMENTAL GROUP OF SO(3)

We want to compute the fundamental group of $SO(3)$, and for this we want to use our knowledge of covering spaces. We will in particular study the map $p : S^3 \to SO(3)$. We also assume that it is well known that $SO(3)$ is path connected.

To help us on the way, we will find it useful to define something called quaternions that are used for representing elements in $SO(3)$.

7.1 Quaternions

We all know from complex analysis that complex numbers can be used to represent rotations in the plane. The mathematician William Hamilton found a way to represent rotations in $\mathbb{R}^3$. He already knew that complex numbers could be interpreted as points in the plane, so he wanted to find a way to do the same for points in $\mathbb{R}^3$. He defined a number system that extended the complex numbers called ***quaternions***.

Although we shall use them to represent matrix rotations in this thesis, they also find uses in applied mathematics, such as three dimensional computer graphics, robots, navigation, orbital mechanics of satellites etc. They are easier to deal with, and hence often preferred in real-world applications.

Definition 7.1. Quaternions form a four-dimensional associative normed division algebra over the real numbers, denoted $\mathbb{H}$, and represents a four dimensional real vector space $\mathbb{R}^4$ with basis $1, i, j$ and k.

Then every quaternion may be expressed as a linear combination

$$q = t1 + xi + yj + zk,$$

for $t, x, y, z \in \mathbb{R}$. It is easy to check that the basis element 1 is the identity element of $\mathbb{H}$, so multiplication by 1 leaves the element fixed.

Definition 7.2. A quaternion $q \in \mathbb{H}$ is written as

$$q = t + xi + yj + zk,$$

for $t, x, y, z \in \mathbb{R}$.

Definition 7.3. We define the ***conjugate*** of q to be

$$\bar{q} = t - xi - yj - zk,$$

for $t, x, y, z \in \mathbb{R}$. Conjugation is an involution, meaning that conjugating the element twice, returns the original element.

Remark. We can write the elements as quadruples:

$$\mathbb{H} = \{(t, x, y, z) \mid t, x, y, z \in \mathbb{R}\}.$$

Then the basis elements are:

$$1 = (1, 0, 0, 0),$$

$$i = (0, 1, 0, 0),$$

$$j = (0, 0, 1, 0),$$

$$k = (0, 0, 0, 1).$$

7.1.1 Multiplication of basis elements

We all know from complex analysis that a complex number of the form $z = t + xi$, $t, x \in \mathbb{R}$ satisfies the equation $i^2 = -1$. How does this relate to quaternions? Hamilton studied this for many years, and finally he came up with an answer that turned out to be the great breakthrough in quaternion theory.

Definition 7.4. Every quaternion $z = t + xi + yj + zk$ satisfies the identity

$$i^2 = j^2 = k^2 = ijk = -1.$$

Then it follows immediately that the products of the basis elements will satisfy

$$ij = k, ji = -k,$$

$$jk = i, kj = -i,$$

$$ki = j, ik = -j,$$

and multiplication of quaternions is ***non-commutative***.

7.1.2 Multiplication and addition of quaternions

Let $q_1 = t_1 + x_1 i + y_1 j + z_1 k$ and $q_2 = t_2 + x_2 i + y_2 j + z_2 k$. The product $q_1 q_2$, called the ***Hamilton product*** is defined as follows:

$$\begin{aligned} q_1 q_2 &= (t_1 + x_1 i + y_1 j + z_1 k)(t_2 + x_2 i + y_2 j + z_2 k) \\ &= t_1 t_2 + t_1 x_2 i + t_1 y_2 j + t_1 z_2 k \\ &\quad + x_1 t_2 i + x_1 x_2 i^2 + x_1 y_2 ij + x_1 z_2 ik \\ &\quad + y_1 t_2 j + y_1 x_2 ji + y_1 y_2 j^2 + y_1 z_2 jk \\ &\quad + z_1 t_2 k + z_1 x_2 ki + z_1 y_2 kj + z_1 z_2 k^2. \end{aligned}$$

It is a product of basis elements. From Definition 7.4 and using the distributive law, we get:

$$\begin{aligned} q_1 q_2 &= (t_1 + x_1 i + y_1 j + z_1 k)(t_2 + x_2 i + y_2 j + z_2 k) \\ &= t_1 t_2 - x_1 x_2 - y_1 y_2 - z_1 z_2 \\ &\quad + (t_1 x_2 + x_1 t_2 + y_1 z_2 - z_1 y_2) i \\ &\quad + (t_1 y_2 - x_1 z_2 + y_1 t_2 + z_1 x_2) j \\ &\quad + (t_1 z_2 + x_1 y_2 - y_1 x_2 + z_1 t_2) k. \end{aligned}$$

Let $q_1 = t_1 + x_1 i + y_1 j + z_1 k$ and $q_2 = t_2 + x_1 i + y_1 j + z_1 k$. Then addition of quaternions is defined as follows:

$$\begin{aligned} q_1 + q_2 &= (t_1 + x_1 i + y_1 j + z_1 k) + (t_2 + x_2 i + y_2 j + z_2) k \\ &= (t_1 + t_2) + (x_1 + x_2) i + (y_1 + y_2) j + (z_1 + z_2) k. \end{aligned}$$

7.1.3 Scalar and vector parts

Having quaternions written on the form above, can make the computations very long and complicated, so we want to give another way to represent them.

Definition 7.5. An element of the form $t + 0i + 0j + 0k$ is called a ***real quaternion***, and an element of the form $0 + xi + yj + zk$, is called a ***pure quaternion***. If $t + xi + yj + zk$ is any quaternion, then t is called the ***scalar part***, and $xi + yj + zk$ is called the ***vector part***.

Then we can write any quaternion as

$$q = (t, \vec{v}),$$

for $q \in \mathbb{H}, t \in \mathbb{R}, \vec{v} \in \mathbb{R}^3$.

Then addition and multiplication are as follows:

$$\begin{aligned} q_1 + q_2 &= (t_1, \vec{v_1}) + (t_2, \vec{v_2}) \\ &= (t_1 + t_2, \vec{v_1} + \vec{v_2}), \end{aligned}$$

$$\begin{aligned} q_1 q_2 &= (t_1, \vec{v_1})(t_2, \vec{v_2}) \\ &= (t_1 t_2 - \vec{v_1} \cdot \vec{v_2}, t_1 \vec{v_2} + t_2 \vec{v_1} + \vec{v_1} \times \vec{v_2}). \end{aligned}$$

where " $\cdot$ " is the dot product, and "$\times$" is the cross product.

7.1.4 Unit quaternions

As mentioned in the beginning of the chapter we will use quaternions to represent rotations, but there are only some special quaternions that are used for this purpose. They are called ***unit quaternions*** and these are the only quaternions we will consider in this thesis.

Definition 7.6. We define the ***norm*** of q to be

$$N(q) = \|q\| = \sqrt{t^2 + x^2 + y^2 + z^2} \in \mathbb{R}.$$

Definition 7.7. A ***unit quaternion*** is a quaternion of norm 1.

Using conjugations and norms of quaternions makes it possible to define the ***reciprocal*** of a quaternion:

Definition 7.8. The ***reciprocal*** of q is defined to be

$$q^{-1} = \frac{\bar{q}}{\|q\|^2}$$

Then for unit quaternions we know that

$$q^{-1} = \frac{\bar{q}}{\|q\|^2} = \bar{q}.$$

This will we make use of later when we present a ***conjugation*** by a quaternion q.

Recall that the sphere S^3 is the set of points in $\mathbb{R}^4$ such that

$$t^2 + y^2 + z^2 + t^2 = 1.$$

This gives us the opportunity to identify the set of all unit quaternions with S^3. In particular, S^3 forms a group (a Lie group) under multiplication:

$$S^3 = \{q \in \mathbb{H}^* \mid q\bar{q} = 1\},$$

where $\mathbb{H}^* = \mathbb{H} \setminus (\mathbf{0})$ is the multiplicative group of all non-zero quaternions.

By representing a unit quaternion on the form $q = (t, \vec{v})$, we can think of the vector part as being the axis about which a rotation occurs, and the scalar part as the amount of rotation that occurs about the given axis.

In fact, there is an explicit expression of a unit quaternions, representing a rotation of an angle θ around a unit vector u, given by

$$q = \cos\frac{\theta}{2} + u\sin\frac{\theta}{2}. \tag{7.1}$$

7.2 Computing the fundamental group of SO(3)

As we have seen earlier, any rotation in $\mathbb{R}^3$ can be represented as a combination of a unit vector u, indicating the direction of axis of rotation, and a scalar θ, the angle describing the magnitude of the rotation about the axis.

Until now we have only looked at matrices whose axis of rotation is around the $x-, y-$ and z-axes. We want to construct a rotation matrix whose axis of rotation is around an arbitrary unit vector u.

Let $u = (u_1, u_2, u_3)$ be this unit vector. Then choose two unit orthonormal vectors $v = (v_1, v_2, v_3)$ and $w = (w_1, w_2, w_3)$ such that $u = v \times w$. The set

(v, w, u) forms an oriented orthonormal basis of $\mathbb{R}^3$. If we combine them into a matrix

$$S_u = \begin{bmatrix} v_1 & w_1 & u_1 \\ v_2 & w_2 & u_2 \\ v_3 & w_3 & u_3 \end{bmatrix}$$

we obtain an element $S_u \in SO(3)$:

$$\begin{aligned} S_u^T S_u &= \begin{bmatrix} v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \end{bmatrix} \begin{bmatrix} v_1 & w_1 & u_1 \\ v_2 & w_2 & u_2 \\ v_3 & w_3 & u_3 \end{bmatrix} \\ &= \begin{bmatrix} v \cdot v & v \cdot w & v \cdot u \\ w \cdot v & w \cdot w & w \cdot u \\ u \cdot v & u \cdot w & u \cdot u \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I. \end{aligned}$$

So the orthonormal columns of S_u imply that $S_u^T S_u = I$. Then $S_u^T = S_u^{-1}$, so $S_u S_u^T = (S_u^T)^T (S_u^T) = I$, so the columns of S_u^T are also an orthonormal basis. The columns of S_u^T is equal to the rows of S_u. Hence, the rows of S_u are also orthogonal, and we get in particular:

$$(\text{row}\, i) \cdot (\text{row}\, j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Then we compute $S_u S_u^T$:

$$\begin{aligned} S_u S_u^T &= \begin{bmatrix} v_1 & w_1 & u_1 \\ v_2 & w_2 & u_3 \\ v_3 & w_3 & u_3 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \end{bmatrix} \\ &= \begin{bmatrix} (\text{row}\,1)_{S_u} \cdot (\text{row}\,1)_{S_u} & (\text{row}\,1)_{S_u} \cdot (\text{row}\,2)_{S_u} & (\text{row}\,1)_{S_u} \cdot (\text{row}\,3)_{S_u} \\ (\text{row}\,2)_{S_u} \cdot (\text{row}\,1)_{S_u} & (\text{row}\,2)_{S_u} \cdot (\text{row}\,2)_{S_u} & (\text{row}\,2)_{S_u} \cdot (\text{row}\,3)_{S_u} \\ (\text{row}\,3)_{S_u} \cdot (\text{row}\,1)_{S_u} & (\text{row}\,3)_{S_u} \cdot (\text{row}\,2)_{S_u} & (\text{row}\,3)_{S_u} \cdot (\text{row}\,3)_{S_u} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I. \end{aligned}$$

What is left is to check that $\det(S_u) = 1$. From the crossproduct $u = v \times w$ we know that we can write u as

$$u = (v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_2 w_1).$$

Similarly, for $v = w \times u$ and $w = u \times v$ we know that

$$v = (u_3w_2 - u_2w_3, w_3u_1 - w_1u_3, u_2w_1, u_1w_2),$$
$$w = (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1).$$

Then by the evaluating the determinant, we get

$$\det(S_u) = \begin{vmatrix} v_1 & w_1 & u_1 \\ v_2 & w_2 & u_2 \\ v_3 & w_3 & u_3 \end{vmatrix} = v_1 \cdot \begin{vmatrix} w_2 & u_2 \\ w_3 & u_3 \end{vmatrix} - w_1 \cdot \begin{vmatrix} v_2 & u_2 \\ v_3 & u_3 \end{vmatrix} + u_1 \cdot \begin{vmatrix} v_2 & w_2 \\ v_3 & w_3 \end{vmatrix}$$

$$= v_1 \cdot v_1 - w_1 \cdot (-w_1) + u_1 \cdot u_1$$
$$= v_1^2 + w_1^2 + u_1^2 = 1.$$

The last equality comes the fact that S_u^T is orthogonal.

Further we consider the map: $h : \mathbb{R}^3 \to SO(3)$ that maps $u \in \mathbb{R}^3$ to an element in $SO(3)$, given by

$$u \mapsto S_u P_\theta S_u^T,$$

where

$$P_\theta = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is the rotation matrix around the z-axis.

We set $R = S_u P_\theta S_u^T$. After some long computation, we get

$$R = \begin{bmatrix} v_1 & w_1 & u_1 \\ v_2 & w_2 & u_3 \\ v_3 & w_3 & u_3 \end{bmatrix} \begin{bmatrix} c_\theta & -s_\theta & 0 \\ s_\theta & c_\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \end{bmatrix} \tag{7.2}$$
$$= \begin{bmatrix} c_\theta + u_1^2(1-c_\theta) & u_1u_2(1-c_\theta) - u_3s_\theta & u_1u_3(1-c_\theta) + u_2s_\theta \\ u_1u_2(1-c_\theta) + u_3s_\theta & c_\theta + u_2^2(1-c_\theta) & u_2u_3(1-c_\theta) - u_1s_\theta \\ u_1u_3(1-c_\theta) - u_2s_\theta & u_2u_3(1-c_\theta) + u_1s_\theta & c_\theta + u_3^2(1-c_\theta) \end{bmatrix},$$

where $c_\theta = \cos\theta$, $s_\theta = \sin\theta$, a matrix whose rotation is by angle θ leaving axis u fixed.

We see that the resultant matrix is independent of the vectors v and w; their coordinates vanishes by the crossproduct terms earlier and the fact that S_u is orthogonal.

7.2.1 Quaternion representation

Recall from Section 7.1.4 that unit quaternions consist of a rotation vector u and a rotation angle θ. Therefore we also call them ***rotation quaternions***, and we shall show that they give another way to represent the rotation matrix (7.2) leaving u fixed.

We identify $S^3 \subseteq \mathbb{R}^4$ with the unit quaternions and identify $\mathbb{R}^3$ with the pure quaternions. Such a unit quaternion q induces a rotation of $\mathbb{R}^3$, though not simply by multiplication as with matrices. The product of q and a member of $\mathbb{R}^3$ may not belong to $\mathbb{R}^3$. A rotation in $SO(3)$ by q requires a ***conjugation*** with q, given by

$$r \mapsto qrq^{-1},$$

for its representation.

In Section 7.1.4 we mentioned that there is an explicit formula of q given by (7.1). But why by the argument $\theta/2$? We conjugate by q and see what happens. To make computations easier, we set $u = i$. Then using the Hamilton product, we get

$$\begin{aligned} p_q(r) &= qrq^{-1} \\ &= \left(\cos\frac{\theta}{2} + i\sin\frac{\theta}{2}\right)(xi + yj + zk)\left(\cos\frac{\theta}{2} - i\sin\frac{\theta}{2}\right) \\ &= i\left(\left(\sin^2\frac{\theta}{2} + \cos^2\frac{\theta}{2}\right)x\right) \\ &\quad + j\left(y\cos^2\frac{\theta}{2} - y\sin^2\frac{\theta}{2} - 2z\sin\frac{\theta}{2}\cos\frac{\theta}{2}\right) \\ &\quad + k\left(z\cos^2\frac{\theta}{2} - z\sin^2\frac{\theta}{2} + 2y\sin\frac{\theta}{2}\cos\frac{\theta}{2}\right) \\ &= ix + j(y\cos\theta - z\sin\theta) + k(z\cos\theta + y\sin\theta). \end{aligned}$$

The result is a quaternion that fixes the x-axis and acts as to rotate through an angle θ, which is exactly what we want. Hence $\theta/2$ seems to be a good choice. We take a look at an example:

Example 7.9. Use quaternions to rotate the point $x = (1, 0, 0)$ $\pi/6$ around the z-axis.

We need to determine the rotation quaternion. Since $\theta = \frac{\pi}{6}$ and $u = (0, 0, 1)$ we get $q = \cos\frac{\pi}{12} + (0, 0, 1)\sin\frac{\pi}{12} = \frac{\sqrt{6}+\sqrt{2}}{4} + \left(0, 0, \frac{\sqrt{6}-\sqrt{2}}{4}\right)$. The desired rotation can be applied to the vector $r = (1, 0, 0)$, considered as a pure quaternion in

$\mathbb{R}^3$. A computation show that $p_q(r) = \left(\frac{\sqrt{3}}{2}, \frac{1}{2}, 0\right)$, the new position vector after rotation, which also is a member of $\mathbb{R}^3$.

Consider the map $p_q : \mathbb{H} \to \mathbb{H}$, given by $r \mapsto qrq^{-1}$ sending $r \in \mathbb{R}^3$ to qrq^{-1}. Then qrq^{-1} will also be a member of $\mathbb{R}^3$. We prove this:

Let $p_q(r) = qrq^{-1}$. Set $r = xi + yj + zk$. Then consider the rotation represented by the rotation vector $\vec{u} = i + j + k$ and angle $\theta = \frac{2\pi}{3}$. Using (7.1) for q we get

$$q = \frac{1 + i + j + k}{2}.$$

We conjugate r by q:

$$\begin{aligned} p_q(r) &= qrq^{-1} \\ &= \frac{1 + i + j + k}{2}(xi + yj + zk)\frac{1 - i - j - k}{2}. \end{aligned}$$

After a lengthy computation we get

$$p_q(xi + yj + zk) = zi + xj + yk,$$

an element that obviously belongs to $\mathbb{R}^3$.

Therefore $p_q : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation. With the standard $\mathbb{R}^3$ basis, p_q can be regarded as a 3×3 matrix.

To find p_q, we observe what the transformation does to each of the the basis elements of $\mathbb{R}^3$. For the first basis element we get:

$$\begin{aligned} p_q(i) &= (t + xi + yj + zk)(i)(t - xi - yj - zk) \\ &= i(x^2 + t^2 - y^2 - z^2) \\ &\quad + j(xy + tz + zt + yx) \\ &\quad + k(xz - ty + zx - yt). \end{aligned}$$

Similarly, for the second basis element:

$$\begin{aligned} p_q(j) &= (t + xi + yj + zk)(j)(t - xi - yj - zk) \\ &= i(yx - zt - tz + xy) \\ &\quad + j(y^2 - z^2 + t^2 - x^2) \\ &\quad + k(yz + zy + tx + xt). \end{aligned}$$

And finally, for the third basis element:

$$\begin{aligned} p_q(k) &= (t + xi + yj + zk)(k)(t - xi - yj - zk) \\ &= i(zx + yt + xz + ty) \\ &\quad + j(zy + yz - xt - tx) \\ &\quad + k(z^2 - y^2 - x^2 + t^2). \end{aligned}$$

In total we get the matrix with respect to the basis,

$$p_q = \begin{bmatrix} t^2 + x^2 - y^2 - z^2 & 2(xy - tz) & 2(xz + ty) \\ 2(xy + tz) & t^2 - x^2 + y^2 - z^2 & 2(yz - tx) \\ 2(xz - ty) & 2(yz + tx) & t^2 - x^2 - y^2 + z^2 \end{bmatrix}.$$

What does p_q represent? By (7.1) we set $t = \cos\theta/2$, $x = u_1 \sin\theta/2$, $y = u_2 \sin\theta/2$ and $z = u_3 \sin\theta/2$, and insert the change of variables into p_q.

Given that the entries in p_q can be enumerated as in the matrix below,

$$\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

then by the change of variables we get a new matrix p'_q whose entries are as follows:

(1') $t^2 + x^2 - y^2 - z^2 = \cos^2\frac{\theta}{2} + u_1^2 \sin^2\frac{\theta}{2} - u_2^2 \sin^2\frac{\theta}{2} - u_3^2 \sin^2\frac{\theta}{2} = \frac{1+\cos\theta}{2} + \frac{1-\cos\theta}{2}(2u_1^2 - 1) = \cos\theta + u_1^2(1 - \cos\theta)$.

(2') $2(xy + tz) = 2(u_1u_2 \sin^2\frac{\theta}{2} + \cos\frac{\theta}{2} u_3 \sin\frac{\theta}{2}) = 2(u_1u_2\frac{1-\cos\theta}{2} + u_3\frac{\sin\theta}{2}) = u_1u_2(1 - \cos\theta) + u_3 \sin\theta$.

(3') $2(xz - ty) = 2(u_1u_3 \sin^2\frac{\theta}{2} - \cos\frac{\theta}{2} u_2 \sin\frac{\theta}{2}) = 2(u_1u_3\frac{1-\cos\theta}{2} - u_2\frac{\sin\theta}{2}) = u_1u_3(1 - \cos\theta) + u_2 \sin\theta$.

(4') $2(xy - tz) = 2(u_1u_2 \sin^2\frac{\theta}{2} - \cos\frac{\theta}{2} u_3 \sin\frac{\theta}{2}) = 2(u_1u_2\frac{1-\cos\theta}{2} - u_3\frac{\sin\theta}{2}) = u_1u_2(1 - \cos\theta) - u_3 \sin\theta$.

(5') $t^2 - x^2 + y^2 - z^2 = \cos^2\frac{\theta}{2} - u_1^2 \sin^2\frac{\theta}{2} + u_2^2 \sin^2\frac{\theta}{2} - u_3^2 \sin^2\frac{\theta}{2} = \frac{1+\cos\theta}{2} + \frac{1-\cos\theta}{2}(2u_2^2 - 1) = \cos\theta + u_2^2(1 - \cos\theta)$.

(6') $2(yz + tx) = 2(u_2u_3 \sin^2\frac{\theta}{2} + \cos\frac{\theta}{2} u_1 \sin\frac{\theta}{2}) = 2(u_2u_3\frac{1-\cos\theta}{2} + u_1\frac{\sin\theta}{2}) = u_2u_3(1 - \cos\theta) + u_1 \sin\theta$.

(7') $2(xz + ty) = 2(u_1u_3 \sin\frac{\theta}{2}^2 + \cos\frac{\theta}{2} u_2 \sin\frac{\theta}{2}) = 2(u_1u_3\frac{1-\cos\theta}{2} + u_2\frac{\sin\theta}{2}) = u_1u_3(1 - \cos\theta) + u_2 \sin\theta$.

(8') $2(yz - tx) = 2(u_2u_3 \sin\frac{\theta}{2}^2 - \cos\frac{\theta}{2}u_1 \sin\frac{\theta}{2}) = 2(u_2u_3\frac{1-\cos\theta}{2} - u_1\frac{\sin\theta}{2}) = u_2u_3(1-\cos\theta) - u_1\sin\theta$.

(9') $t^2 - x^2 - y^2 + z^2 = \cos\frac{\theta}{2}^2 - u_1^2 \sin\frac{\theta}{2}^2 - u_2^2 \sin\frac{\theta}{2}^2 + u_3^2 \sin\frac{\theta}{2}^2 = \frac{1+\cos\theta}{2} + \frac{1-\cos\theta}{2}(2u_3^2 - 1) = \cos\theta + u_3^2(1-\cos\theta)$.

Then by the change of variables we get

$$
\begin{aligned}
p'_q &= \begin{bmatrix} 1' & 4' & 7' \\ 2' & 5' & 8' \\ 3' & 6' & 9' \end{bmatrix} \\
&= \begin{bmatrix} c_\theta + u_1^2(1-c_\theta) & u_1u_2(1-c_\theta) - u_3s_\theta & u_1u_3(1-c_\theta) + u_2s_\theta \\ u_1u_2(1-c_\theta) + u_3s_\theta & c_\theta + u_2^2(1-c_\theta) & u_2u_3(1-c_\theta) - u_1s_\theta \\ u_1u_3(1-c_\theta) - u_2s_\theta & u_2u_3(1-c_\theta) + u_1s_\theta & c_\theta + u_3^2(1-c_\theta) \end{bmatrix},
\end{aligned}
$$

where $c_\theta = \cos\theta$ and $s_\theta = \sin\theta$. This is exactly the same matrix R as we derived earlier from the product $S_uR_\theta S_u^T$, a matrix whose rotation is by angle θ leaving axis u fixed. We call p_q the ***quaternion derived rotation matrix*** and $p_q \in SO(3)$.

Remark. A rotation through angles between 0 and π corresponds to the point on the opposite side of the origin, at the same distance from the origin.

Let us consider the quaternion $-q$ on the opposite side of the origin given by

$$-q = \cos\left(\frac{\theta}{2} + \pi\right) + u\sin\left(\frac{\theta}{2} + \pi\right). \tag{7.3}$$

If we insert (7.3) for q in the matrix p_q we will actually get the same rotation matrix as we did when we inserted (7.1) .

This can also be seen from the fact that the rotation is expressed as $p_q(r) = qrq^{-1}$ and we have $p_{-q}(r) = (-q)r(-q)^{-1} = qrq^{-1}$ which implies that $-q$ is mapped to the same rotation. Then, either we rotate through an angle of π, or we rotate through an angle of $-\pi$, we still get the same rotation.

Then we have a two-to-one map and a continuous function

$$p : S^3 \to SO(3) \tag{7.4}$$

given by $p(q) = p_q$.

Further we want to convince ourselves that p is surjective. Let $p_q \in SO(3)$, and let us claim that q maps to p_q. From (7.4) it is clear that we have such a function satisfying $p(q) = p_q$.

We also know that for every rotation in $\mathbb{R}^3$ there is a quaternion that represents it, given by (7.1). Hence, we are convinced, and p is surjective.

Moreover, p is a group homomorphism:

$$p_{q_1 q_2}(r) = q_1 q_2 r (q_1 q_2)^{-1} = q_1 q_2 r q_2^{-1} q_1^{-1},$$

$$p_{q_1} \circ p_{q_2}(r) = p_{q_1}(p_{q_2}(r)) = p_{q_1}(q_2 r q_2^{-1}) = q_1 q_2 r q_2^{-1} q_1^{-1}.$$

Hence, $p_{q_1 q_2} = p_{q_1} p_{q_2}$ and the structure is preserved.

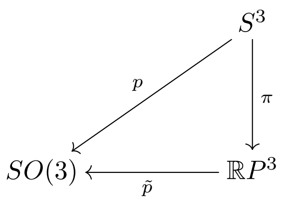
By the map $p_q(r) = qrq^{-1}$ we easily observe that $p(q) = I \Leftrightarrow q = \pm 1$. Then by the first isomorphism theorem, the induced map

$$\tilde{p} : S^3/\{1, -1\} \xrightarrow{\sim} SO(3)$$

is a group isomorphism.

Definition 7.10. The projective space $\mathbb{R}P^3$ is the quotient space obtained from S^3 by identifying each point $q \in S^3$ with its antipodal point $-q$.

We know that the projective map $\pi : S^3 \to RP^3 \cong S^3/\{1, -1\}$ is a double covering map identifying antipodal points of S^3. Hence, by the commutative diagram,



$p = \tilde{p} \circ \pi$ is a double covering map, identifying antipodal points on the surface of S^3.

We now use Theorem 4.9. Since S^3 is simply connected, the lifting correspondence

$$\phi : \pi_1(SO(3)) \to p^{-1}(p_q)$$

is a bijective correspondence.

Since $p^{-1}(p_q)$ is a two element set, we know that $\pi_1(SO(3))$ is a group of order two.

Let $G = \{e, a\}$ be a group of order two, and let e be the identity element. Then we have the following: Since e is the identity element, we have $e * x = x * e = x$ for all $x \in \{e, a\}$. Then we get the multiplication table

$*$	e	a
e	e	a
a	a	

Since a is a nontrivial element, a has an inverse a' such that $a' * a = a * a' = e$. In our case, a' must be either e or a. Since $a' = e$ does not work, we must have $a' = a$, and we complete the table,

$*$	e	a
e	e	a
a	a	e

We know that $\mathbb{Z}_2 = \{0, 1\}$ under addition modulo 2 is a group. By our arguments, its table must be the one above with e replaced by 0, and a by 1.

This shows that the fundamental group of $SO(3)$ is $\mathbb{Z}_2$.

As we mentioned in the introduction, this result is related to physics; more specifically that a rotating object only can have spin equal to a half. An illustration of this, called ***Dirac's scissors experiment***, was performed by the Nobel laureate in physics P.A.M. Dirac at lectures in the 1930's (see [3, p. 39]). In fact, there are several ways to demonstrate this practically (see [2, pp. 166–167]). We will explain Dirac's scissors experiment, but first we need to make some considerations.

We want to consider an element of the group. A generator for $\pi_1(SO(3))$ is any closed curve obtained by projecting any curve connecting antipodal points in S^3 down to $SO(3)$. Consider the semicircle

$$\{\cos\theta + i\sin\theta \in S^3 \mid 0 \le \theta \le \pi\},$$

the intersection of S^3 with the upper half of the xy-plane. This maps by p to a loop in $SO(3)$ and represents a nontrivial element of $\pi_1(SO(3))$. We look at what it does to an arbitrary element in $\mathbb{R}^3$. Let $r = xi + yj + zk$ and $q = \cos\theta + i\sin\theta$. Then we have

$$
\begin{aligned}
p_q(r) &= qrq^{-1}\\
&= (\cos\theta + i\sin\theta)(xi + yj + zk)(\cos\theta - i\sin\theta)\\
&= i((\sin^2\theta + \cos^2\theta)x)\\
&\quad + j(y\cos^2\theta - y\sin^2\theta - 2z\sin\theta\cos\theta)\\
&\quad + k(z\cos^2\theta - z\sin^2\theta + 2y\sin\theta\cos\theta)\\
&= ix + j(y\cos 2\theta - z\sin 2\theta) + k(z\cos 2\theta + y\sin 2\theta)
\end{aligned}
$$

Hence, it fixes the complex plane on so only acts on the jk-plane. For $q = yj + zk$ we have

$$
\begin{aligned}
p_q(r) &= qrq^{-1}\\
&= (\cos\theta + i\sin\theta)(yj + zk)(\cos\theta - i\sin\theta)\\
&= j(x\cos^2\theta - x\sin^2\theta - 2y\sin\theta\cos\theta)\\
&\quad + k(y\cos^2\theta - y\sin^2\theta + 2x\sin\theta\cos\theta)\\
&= j(x\cos 2\theta - y\sin 2\theta) + k(y\cos 2\theta + x\sin 2\theta),
\end{aligned}
$$

which tells us that $q = \cos\theta + i\sin\theta$ acts on the jk-plane as to rotate through 2θ. Since $\theta \in [0, \pi]$, it follows that the nontrivial element of $\pi_1(SO(3))$ is given by a path of rotations about a given axis whose angle $\in [0, 2\pi]$.

Further we let the rotation to be around the z-axis and let $r(t)$ be counter-clockwise rotation in the xy-plane through the angle $2\pi t$, r a loop representing the nontrivial element of $\pi_1(SO(3))$.

Since $\pi_1(SO(3))$ is a group of order two, it follows that the loop $f = r * r$ must be homotopic to the constant loop $e \in SO(3)$, f a loop whose value $f(t)$ is the rotation of angle $4\pi t$.

Let f be this loop, and let $[f] \in \pi_1(SO(3))$. Since $[f] = [e]$ we have a homotopy

$$F : I \times I \to SO(3)$$

such that

(i) $F(s, 0) = e \quad \text{and} \quad F(s, 1) = f(s)$,

(ii) $F(0, t) = e \quad \text{and} \quad F(1, t) = e$.

With this set up we are ready to demonstrate the experiment.

Dirac's scissor experiment. Take two threads and attach one end of them to the two inside rings of a pair of scissors. Attach the other ends to some fixed objects in the room, e.g. two chairs (see Figure 7.1). Then rotate the scissors one complete turn, i.e. 360°. Then, while keeping the object stationary, try to untangle the threads. Although patience is a good thing, it will not help you, you will either way be unable to do this no matter how long you try. However, if you rotate the scissors one more time in the same direction, that means two complete turns in total, you will actually be able to untangle them, keeping in mind to hold the scissors stationary as you are doing it. Then what does this tell us? Well, by rotation not once, but two complete turns around the same axis, we got back to where we started, and hence we have a cyclic group of order two, whose generator is a loop of rotation around an axis of angle 2π.

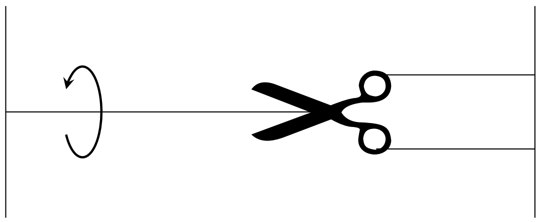


Figure 7.1: Dirac's scissors experiment.

BIBLIOGRAPHY

[1] M. Arkowitz. *Introduction to homotopy theory.* Springer-Verlag, 2011.

[2] G. E. Bredon. *Topology and geometry.* Springer-Verlag, New York, 1997.

[3] V. L. Hansen. *Braids and coverings: selected topics.* Cambridge University Press, 1989.

[4] J. R. Munkres. *Topology.* Featured Titles for Topology Series. Prentice Hall, Incorporated, 2000.

[5] J. Stuelpnagel. On the parametrization of the three-dimensional rotation group. *SIAM review*, 6(4):422–430, 1964.

[6] G. L. Trigg. *Mathematical tools for physicists.* John Wiley & Sons, 2006.