What is algebraic topology?

Algebraic topology studies 'geometric' shapes, spaces and maps between them by *algebraic means*.

An example of a space is a circle, or a doughnut-shaped figure, or a Möbius band. A little more precisely, the objects we want to study belong to a certain geometric 'category' of topological spaces (the appropriate definition will be given in due course). This category is hard to study directly in all but the simplest cases. The objects involved could be multidimensional, or even have infinitely many dimensions and our everyday life intuition is of little help. To make any progress we consider a certain 'algebraic' category and a 'functor' or a 'transformation' from the geometric category to the algebraic one. I say 'algebraic category' because its objects have algebraic nature, like natural numbers, vector spaces, groups etc. This algebraic category is more under our control. The idea is to obtain information about geometric objects by studying their image under this functor. For example, we have two geometric objects, say, a circle S^1 and a two-dimensional disc D^2 and we want to somehow distinguish one from the other. A more precise formulation of such a problem is this: given two topological spaces we ask whether one could be continuously 'deformed' into the other. It is intuitively clear that a two-dimensional square could be deformed into D^2, however S^1 cannot be. The reason is that S^1 has a hole in it which must be preserved under continuous deformation. However D^2 is solid, and therefore, S^1 cannot be deformed into it.

We will make these consideration a little more precise by looking at the image of the corresponding functor. In the case at hand this functor associates to a geometric figure the number of holes in it. This is an invariant under deformation, that is, this number does not change as the object is being deformed. This invariant equals 1 for S^1 and 0 for D^2, therefore one is not deformable into the other.
However this 'invariant' is not quite sufficient yet. Look at S^1 and S^2, the circle and the two-dimensional sphere. Each has one hole, but it is intuitively clear the 'natures' of these holes are different, and that one still cannot be deformed into the other.
So the basic problem of algebraic topology is to find a system of algebraic invariants of topological spaces which would be powerful enough to distinguish different shapes. On the other hand these invariants should be computable. Over the decades people have come up with lots of invariants of this sort. In this course we will consider the most basic, but in some sense, also the most important ones, the so-called *homotopy* and *homology* groups.

Here we will discuss one of the famous results in algebraic topology which is proved using the ideas explained above. This is only a very rough outline and much of our course will be spent trying to fill in the details of this proof.
Let D^n be the n-dimensional disc. You could think of it as a solid ball of radius one having its center at the origin of $\mathbb{R}^n$, the n-dimensional real space. Simpler yet, you could take n to be equal to 3 or 2, or even 1.

Theorem

Let $f : D^n \longrightarrow D^n$ be a continuous map. Then f has at least one fixed point, i.e. there exists a point $x \in D^n$ for which $f(x) = x$.

This is the celebrated *Brouwer Fixed Point Theorem*.
To get some idea why it should be true consider the almost trivial case $n = 1$. In this case we have a continuous map $f : [-1, 1] \longrightarrow [-1, 1]$. We can assume that $f(-1) > -1$ and $f(1) < 1$, since otherwise there is nothing to prove. To say that f has a fixed point is equivalent to saying that the function $g(x) := f(x) - x$ is zero at some point $c \in [-1, 1]$. Our assumption on f implies that $g(-1) > 0$ and $g(1) < 0$. By the Intermediate Value Theorem from calculus we then conclude that there is a point $c \in [-1, 1]$ for which $g(c) = 0$ and the theorem is proved.

Unfortunately, this elementary proof does not generalize to higher dimensions, so we need a new idea.
Suppose that there exists a continuous map $f : D^n \longrightarrow D^n$ without fixed points. Take any $x \in D^n$ and draw a line between x and $f(x)$; such a line is unique since $f(x) \neq x$ by our assumption. This line intersects the boundary S^{n-1} of D^n in precisely two points, take the one that's closer to x than to $f(x)$ and denote it by $l(x)$. Then the map $x \longrightarrow l(x)$ is a continuous map from D^n to its boundary S^{n-1} and $l(x)$ restricted to S^{n-1} is the identity map on S^{n-1}.
We postpone for a moment to introduce the relevant

Definition

Let Y be a subset of X. A map $f : X \longrightarrow Y$ is called a retraction of X onto$Y$ *if f restricted to Y is the identity map on Y. Then Y is called a* retract of X.

Proof of Brouwer fixed point theorem

Now return to the proof of the our theorem. Note, that assuming that $f: D^n \longrightarrow D^n$ has no fixed points we constructed a retraction of D^n onto S^{n-1}. We will show that this is impossible. For this we need the following facts to be proved later on:
Associated to any topological space X (of which D^n or its boundary S^{n-1} are examples) is a sequence of abelian groups $H_n(X), n = 0, 1, 2, \dots$ called *homology groups* such that:

- to any continuous map $X \longrightarrow Y$ there corresponds a homomorphism of abelian groups $H_n(X) \longrightarrow H_n(Y)$ and to the composition of continuous maps there corresponds the composition of homomorphisms
- $H_i(D^n) = 0$ for $i > 0$ and any n,
- $H_n(S^n) = \mathbb{Z}$, the group of integers.

Taking this for granted we will now deduce the Brouwer fixed point theorem. Note that the correspondence $X \longrightarrow H_i(X)$ is an example of a functor or, rather a collection of functors from the geometric category of spaces to the algebraic category of abelian groups and group homomorphisms.

Continuation of proof

Denote by $i : S^{n-1} \longrightarrow D^n$ the obvious inclusion map. Then the composition

$$S^{n-1} \xrightarrow{i} D^n \xrightarrow{l} S^{n-1}$$

is the identity map.
Associated to this sequence of maps is the sequence of group homomorphisms

$$H_{n-1}(S^{n-1}) \longrightarrow H_{n-1}(D^n) \longrightarrow H_{n-1}(S^{n-1}) \ .$$

whose composition should also be the identity homomorphism on $H_{n-1}(S^{n-1}) = \mathbb{Z}$.
But remember that $H_{n-1}(D^n) = 0$ if $n > 1$. Therefore our sequence of homomorphisms has the form

$$\mathbb{Z} \longrightarrow 0 \longrightarrow \mathbb{Z}$$

and clearly the composition must be zero, not the identity on $H_{n-1}(S^{n-1}) = \mathbb{Z}$.
This contradiction proves our theorem.

Review of background material: algebra, topology and category theory

We will review first some of the preliminary material which will be needed later on. Some of it you have hopefully seen before, the rest will be developed here from scratch.
We begin with some basic definitions and facts.

Definition

A group *is a set G together with a map $G \times G \longrightarrow G : (g, h) \longrightarrow gh \in G$ called* multiplication *such that*

- $(gh)k = g(hk)$ *for any $g, h, k \in G$ (*associativity*).*
- *There exists an element $e \in G$ for which $eg = ge = g$ for any $g \in G$ (*existence of two-sided unit*).*
- *For any $g \in G$ there exists $g^{-1} \in G$ for which $gg^{-1} = g^{-1}g = e$ (*existence of a two-sided inverse*).*

If for any $g, h \in G$ $gh = hg$ then the group G is called abelian*. For an abelian group G we will usually use the additive notation $g + h$ to denote the product of g and h.*
A subgroup *H of G is a subset $H \subseteq G$ which contains the unit, together with any element contains its inverse and is closed under multiplication in G. A subgroup $H \subseteq G$ is* normal *if for any $g \in G$ and $h \in H$ the element ghg^{-1} also belongs to H.*

Definition

For a group G, its element $g \in G$ and its subgroup H a left coset *gH is the collection of elements of the form gh with $h \in H$.*
Similarly a right coset *is the collection of elements of the form hg with $h \in H$.*
If the subgroup H is normal then the collections of left and right cosets coincide and both are called the quotient *of G by H, denoted by G/H.*
For two groups G and H a homomorphism *$f : G \longrightarrow H$ is such a map that $f(gh) = f(g)f(h)$ for any $g, h \in G$.*
The kernel *of a homomorphism $f : G \longrightarrow H$, denoted $\operatorname{Ker} f$ is the set of elements in G mapping to $e \in H$. $\operatorname{Ker} f$ is a normal subgroup in G.*
The image *of a homomorphism $f : G \longrightarrow H$, denoted $\operatorname{Im} f$ is the set of elements in H having a nonempty preimage under f. $\operatorname{Im} f$ is a subgroup of H.*
A homomorphism $f : G \longrightarrow H$ is called an epimorphism *or* onto *if $\operatorname{Im} f = H$.*
A homomorphism $f : G \longrightarrow H$ is called a monomorphism *if $\operatorname{Ker} f = \{e\}$.*
A homomorphism that is both a monomorphism and an epimorphism is called an isomorphism.
An isomorphism $f : G \longrightarrow H$ admits an inverse isomorphism $f^{-1} : H \longrightarrow G$ so that $f \circ f^{-1} = id_H$ and $f^{-1} \circ f = id_G$.
The basic theorem about homomorphisms says that $\operatorname{Im} f$ is isomorphic to the quotient group $G/\operatorname{Ker} f$, i.e. $\operatorname{Im} f \cong G/\operatorname{Ker} f$.

Exact sequences

We will now introduce a concept which may be new to you, that of an *exact sequence*. This is one of the most important algebraic working tools in algebraic topology.

Definition

A sequence of abelian groups and homomorphisms

$$\ldots \xleftarrow{d_{-2}} A_{-2} \xleftarrow{d_{-1}} A_{-1} \xleftarrow{d_0} A_0 \xleftarrow{d_1} A_1 \xleftarrow{d_2} \ldots \xleftarrow{d_n} A_n \xleftarrow{d_{n+1}} \ldots$$

is called exact if $\operatorname{Ker} d_n = \operatorname{Im} d_{n+1}$ *for any* $n \in \mathbb{Z}$.

Let us consider special cases of this definition.
Suppose that all groups A_i are trivial save A_n. In that case the exactness of the sequence

$$\ldots \longleftarrow 0 \longleftarrow A_n \longleftarrow 0 \longleftarrow \ldots$$

clearly means that $A_n = 0$, the trivial group.
If our sequence consists of trivial groups except for the two neighboring ones:

$$\ldots \longleftarrow 0 \longleftarrow A_{n-1} \longleftarrow A_n \longleftarrow 0 \longleftarrow \ldots$$

then it is easy to see that the homomorphism $A_n \longrightarrow A_{n-1}$ is an isomorphism (check this!)

Short exact sequences

Further consider the case when *three* consecutive groups are nonzero and the rest is zero.
In that case our sequence takes the form

$$0 \longleftarrow A_{n-1} \longleftarrow A_n \longleftarrow A_{n+1} \longleftarrow 0\ . \tag{0.1}$$

This sort of exact sequence is called a *short exact sequence*. The exactness in this case amount to the condition that

- The homomorphism $A_n \longrightarrow A_{n-1}$ is an epimorphism;
- the homomorphism $A_{n+1} \longrightarrow A_n$ is a monomorphism so A_{n+1} could be considered as a subgroup in A_n;
- the kernel of the homomorphism $A_n \longrightarrow A_{n-1}$ is precisely the subgroup A_{n+1} in A_n.

So we see that the short exact sequence (0.1) gives rise to an isomorphism

$$A_{n-1} \cong A_n/A_{n+1}.$$

A generalization of the notion of an exact sequence is that of a *complex*:

Definition

The sequence of abelian groups and homomorphisms

$$\dots \xleftarrow{d_{-2}} A_{-2} \xleftarrow{d_{-1}} A_{-1} \xleftarrow{d_0} A_0 \xleftarrow{d_1} A_1 \xleftarrow{d_2} \dots \xleftarrow{d_n} A_n \xleftarrow{d_{n+1}} \dots \tag{0.2}$$

is called a chain complex (A, d) if the composition of any two consecutive homorphisms is zero: $d_n \circ d_{n+1} = 0$.

Exercise

Show that an exact sequence is a complex.

A complex which is exact is considered trivial in some sense. We will see later on, why. The characteristic that measures the 'nontriviality' of a complex is its *homology*:

Definition

The nth homology group $H_n(A, d)$ of the complex (0.2) (A, d) is the quotient group $\operatorname{Ker} d_n / \operatorname{Im} d_{n+1}$*.*

We see, therefore, that an exact sequence has its homology equal to zero in all degrees $i \in \mathbb{Z}$.

Definition

A topology on a set X is a collection τ of subsets of X satisfying:

1. $\varnothing, X \in \tau$.
2. *For any collection of sets $U_i \in \tau$ their union $\bigcup_i U_i$ also belongs to τ.*
3. *If two sets U, V belong to τ, then $U \bigcap V \in \tau$.*

A set X with a topology τ on it is called a topological space, denoted (X, τ).

Definition

1. *A subset $U \in \tau$ will be called an open set in X.*
2. *A set $Y \subseteq X$ is closed iff its complement $X - Y$ is open.*
3. *A* neighborhood *of a point x in X is an open set $U \subseteq X$ such that $x \in U$.*
4. *An interior point of a set $Y \subseteq X$ is a point $y \in Y$ such that Y contains a neighborhood of y.*

Example

An example is the real line $\mathbb{R}^1$, the topology being specified by the collection of open sets (in the usual sense) of $\mathbb{R}^1$, that is countable unions of open intervals. Another example is $\mathbb{R}^n$, the n-dimensional real space. Again, the topology is given by the collection of usual open sets in $\mathbb{R}^n$.

Any set X has a *discrete topology* which is defined by declaring *all* subsets to be open. The opposite extreme is the *antidiscrete topology* in which the open sets are $X, \varnothing$ and nothing else.
We will now consider how to build new topological spaces out of a given one.

Definition

Let X be a topological space and $Y \subseteq X$. Then Y becomes a topological space with the subspace topology *defined by declaring the open sets the sets of the form $U \bigcap Y$, where U is open in X.*

Another example is given by taking quotients by an equivalence relation. Recall that an equivalence relation on a set X is a subset $R \subseteq X \times X$ such that

1. reflexive: if $(x, x) \in R$ for any $x \in X$,
2. symmetric: if $(x, y) \in R$ then $(y, x) \in R$,
3. transitive: if $(x, y), (y, z) \in R$, then $(x, z) \in R$.

We will usually write $x \sim y$ if $(x, y) \in R$ and read it as *'x is equivalent to y'.* The *equivalence class* $[x]$ of $x \in X$ is the set of all elements $y \in X$ which are equivalent to x. If two equivalence classes are not equal then they are disjoint, and any element of X belongs to a unique equivalence class, namely $[x]$. The set of equivalence classes is written as X/R, the *quotient* of X by the equivalence relation R. There is a surjective map $p : X \longrightarrow X/R$ given by $x \longrightarrow [x]$.

Now we can make the following

Definition

If τ is a topology on X then the quotient topology *τ/R on X/R is given by*

$$\tau/R = \{U \subseteq X/R : p^{-1}(U) \in \tau.\}$$

Example

An example of a quotient topological space which is frequently encountered is the contraction of a subspace. *Let X be a topological space and $Y \subseteq X$. We define an equivalence relation on X by declaring $x_1 \sim x_2$ iff $x_1, x_2 \in Y$. The resulting set of equivalence classes is denoted by X/Y.*
Let $X = [0, 1]$, the unit interval with its usual topology and Y be its boundary (consisting of two endpoints). Then, clearly, X/Y can be identified with the unit circle S^1.

Perhaps the simplest way to build a new topological space is by taking the *disjoint union* $\coprod_{i \in I} X_i$ of a collection of topological spaces X_i indexed by a set I. The open sets in $\coprod_{i \in I} X_i$ are just the disjoint unions of open sets in X_i.

Another important construction is the *product* of two topological spaces. For two spaces X and Y consider its cartesian product

$$X \times Y := \{(x, y) : x \in X, y \in Y\}.$$

Certainly, if U is an open set in X and V is an open set in Y then we want $U \times V$ to be an open set in $X \times Y$. But this is not enough, as an example of $[0,1] \times [0,1]$ makes clear. We say that a subset $W \subseteq X \times Y$ is open if W is a union, possibly infinite, of subsets in $X \times Y$ of the form $U \times V$ where $U \subseteq X$ and $V \subseteq Y$.
Similarly we could define a product of a collection, possibly infinite, of topological spaces X_i. The relevant notation is $\prod_i X_i$.

We now come to the central notion of the point-set topology, that of a *continuous map*.

Definition

If X and Y are two topological spaces then a map of sets $f : X \longrightarrow Y$ is continuous *if for any open set $U \subseteq Y$ the set $f^{-1}(U)$ is open in X.*

Exercise

Recall that a real valued function $\mathbb{R} \longrightarrow \mathbb{R}$ is called continuous at a point a if for any $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon.$$

Show that for the topological space $\mathbb{R}$ with its usual topology the two notions of continuity are equivalent.

Example

Examples of continuous maps.

1. *If $X \subseteq Y$ is a subspace of a topological space Y with the subspace topology then the inclusion map $X \longrightarrow Y$ is continuous.*
2. *If X/R is a quotient of a topological space X by the equivalence relation R with the quotient topology then the projection map $X \longrightarrow X/R$ is continuous.*
3. *The inclusion maps $i_i : X_i \longrightarrow \coprod_i X_i$ are all continuous.*
4. *The projection maps $p_i : \prod_i X_i \longrightarrow X_i$ are continuous. Here for $(x_1, x_2, \dots,) \in \prod X_i$ we define $p_i(x_1, x_2, \dots,) = x_i$.*

Definition

A map $f : X \longrightarrow Y$ between topological spaces is called a homeomorphism *if f is continuous, bijective and the inverse map $f^{-1} : Y \longrightarrow X$ is also continuous.*

Note that it is possible for a map to be continuous and bijective, but not a homeomorphism. Indeed, consider the semi-open segment $X = [0, 1)$ on the real line. Clearly there is a continuous map $X \longrightarrow S^1$ from X to the circle. This map could be visualized by bringing the two ends of $[0, 1]$ closer to each other until they coalesce. This is clearly a bijective map, but the inverse map would involve tearing the circle and is, therefore, not continuous.
Another example is given by the map $X^\delta \longrightarrow X$ where X^δ coincides with X as a set but is supplied with the discrete topology. The map is just the tautological identity. It is clear that it is continuous and one-to-one but it cannot be a homeomorphism unless the topology on X is discrete.
Informally speaking the topological spaces X and Y are homeomorphic if they have *'the same number'* of open sets. In the examples above the space $[0, 1)$ has *'more'* open sets then S^1 and X^δ has more open sets then X.

Definition

A topological space X is called connected *if it cannot be represented as a union $V \bigcup U$ of two nonempty open sets V and U which have empty intersection $U \bigcap V = \varnothing$.*

A related notion is that of *path connectedness*. A path in a topological space X is a continuous map $\gamma : [0,1] \longrightarrow X$. If $a = \gamma(0) \in X$ and $b = \gamma(1) \in X$ we say that a and b are connected by the path γ.

Definition

A space X is path connected *if any two points in X can be connected by a path.*

Proposition

If X is a (path-)connected topological space and $f : X \to Y$ a continuous map, then $f(X) = \operatorname{Im} f$ is a (path-)connected subspace of Y.

Proof: exercise!

Proposition

A path-connected topological space X is connected.

Proof: Let X be path-connected and suppose that $X = U \bigcup V$, where both U and V are nonempty disjoint open sets in X.

Let $\gamma : [0,1] \longrightarrow X$ be any path in X. Being the continuous image of the connected set $[0,1]$, the image $\gamma([0,1]) \subset Y$ is connected, and so it lies entirely in either U or V.
Therefore, there is no path in X joining a point $a \in U$ with a point $b \in V$, which contradicts the assumption that X is path-connected. QED.

On the other hand a topologogical space X could be connected but *not* be path-connected.

Example

Consider the topological space X which is the union of the graph of the function $f = \sin \frac{1}{x}$ and the segment $[-1,1]$ on the y-axis. Then X is connected but not path-connected. (Exercise!)

No matter what kind of mathematics you are doing it is useful to get acquainted with *category theory* since the latter gives you a sort of 'big picture' from which you can gain various patterns and insights. The term category was mentioned already in the introduction, but now we will be more precise.

Definition

A category $\mathcal{C}$ *consists of*

1. *A class of* objects $Ob(\mathcal{C})$*.*
2. *A set of* morphisms $Hom(X, Y)$ *for every pair of objects* X *and* Y*. If* $f \in Hom(X, Y)$ *we will write* $f : X \longrightarrow Y$*.*
3. *A* composition law*. In more detail, for any ordered triple* (X, Y, Z) *of objects of* $\mathcal{C}$ *there is a map*

$$Hom(X, Y) \times Hom(Y, Z) \longrightarrow Hom(X, Z).$$

If $f \in Hom(X, Y)$ *and* $g \in Hom(Y, Z)$ *then the image of the pair* (f, g) *in* $Hom(X, Z)$ *is called the* composition *of* f *and* g *and is denoted by* $g \circ f$*.*

Moreover the following axioms are supposed to hold:

- Associativity*:*

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

  for any morphisms f, g and h for which the above compositions make sense.
- *For any object X in $\mathcal{C}$ there exists a morphism $1_X \in Hom(X, X)$ such that for arbitrary morphisms $g \in Hom(X, Y)$ and $f \in Hom(Y, X)$ we have $1_X \circ f = f$ and $g \circ 1_X = g$.*

The notion of a category is somewhat similar to the notion of a group. Indeed, if a category $\mathcal{C}$ consists of only one object X and all its morphisms are invertible then clearly the set $Hom(X, X)$ forms a group. This analogy is important and useful, however we will not pursue it further.
For us the notion of a category encodes the collection of sets with structure and maps which preserve this structure.

Remark

We will frequently use the notion of a commutative diagram *in a category $\mathcal{C}$. The latter is a directed graph whose vertices are objects of $\mathcal{C}$, the edges are morphisms in $\mathcal{C}$ and any two paths from one vertex to another determine the same morphism. A great many formulas in mathematics can be conveniently expressed as the commutativity of a suitable diagram.*

Example

Examples of categories abound. We can talk about

1. *the category $\mathcal{S}$ of sets and maps between sets;*
2. *the category $Vect_k$ of vector spaces over a field k and linear transformations;*
3. *the category $\mathcal{G}r$ of groups and group homomorphisms;*
4. *the category $\mathcal{A}b$ of abelian groups and group homomorphisms;*
5. *the category $\mathcal{R}ings$ of rings and ring homomorphisms;*
6. *the category $\mathcal{T}op$ of topological spaces and* continuous *maps.*

In the list above the category $\mathcal{S}$ is the most basic but for us not very interesting. The categories (2)-(5) are familiar, have algebraic nature and are more or less easy to work with. The category $\mathcal{T}op$ and its variations is what we are really interested in.

Definition

A morphism $f \in Hom(X, Y)$ is called an isomorphism *if it admits a two-sided inverse, i.e. a morphism $g \in Hom(Y, X)$ such that $f \circ g = 1_Y$ and $g \circ f = 1_X$.*

For example in the category of groups the categorical notion of isomorphism specializes to the usual group isomorphism whereas in the category $\mathcal{T}op$ it is a homeomorphism.

Exercise

Let $f : X \longrightarrow Y$ be an isomorphism in a category $\mathcal{C}$ and Z be an arbitrary object in $\mathcal{C}$. Then f determines by composition a map of sets $f_* : Hom(Z, X) \longrightarrow Hom(Z, Y)$. Likewise there is a map of sets $f^* : Hom(Y, Z) \longrightarrow Hom(X, Z)$. Show that both f^* and f_* are bijections of sets.

The next notion we want to discuss is that of a *functor* between two categories.

Definition

A functor *F from the category $\mathcal{C}$ into the category $\mathcal{D}$ is a correspondence which*

1. *associates the object $F(X) \in Ob(\mathcal{D})$ to any $X \in Ob(\mathcal{C})$;*
2. *associates a morphism $F(f) \in Hom_{\mathcal{D}}(F(X), F(Y))$ to any morphism $f \in Hom_{\mathcal{C}}(X, Y)$.*

Moreover the following axioms have to hold:

- *For any object* $X \in \mathcal{C}$ *we have* $F(1_X) = 1_{F(X)}$.
- *For any* $f \in Hom_{\mathcal{C}}(X, Y)$ *and* $g \in Hom_{\mathcal{C}}(Y, Z)$ *we have:*

$$F(g \circ f) = F(g) \circ F(f) \in Hom_{\mathcal{D}}(F(X), F(Z)).$$

Remark

Sometimes a functor as it was defined above is referred to as a covariant *functor to emphasize that it respects the direction of arrows. There is also the notion of a* contravariant *functor. The most economical definition of it uses the notion of an* opposite category $\mathcal{C}^{op}$ *which has the same objects as* $\mathcal{C}$ *and for every arrow (morphism) from* A *to* B *in* $\mathcal{C}$ *there is precisely one arrow from* B *to* A *in* $\mathcal{C}^{op}$. *Then a contravariant functor* $\mathcal{C} \to \mathcal{D}$ *is by definition a (covariant) functor* $\mathcal{C}^{op} \to \mathcal{D}$.

Example

There are very many examples of functors and you could think of some more. Take $\mathcal{C} = \mathcal{A}b$, *the category of abelian groups and* $\mathcal{D} = \mathcal{G}r$ *be the category of groups. Then there is an obvious functor which takes an abelian group and considers it as an object in* $\mathcal{G}r$. *Functors of this sort are called* forgetful *functors for obvious reasons.*

Example

Another example: take a set I and consider a real vector space $\mathbb{R}\langle I\rangle$ whose basis is indexed by the set I. This gives a functor $\mathcal{S} \longrightarrow Vect_{\mathbb{R}}$.

Exercise

Show that if $F : \mathcal{C} \longrightarrow \mathcal{D}$ is a functor and $f \in Hom(X, Y)$ is an isomorphism in $\mathcal{C}$ then $F(f) \in Hom(F(X), F(Y))$ is an isomorphism in $\mathcal{D}$.

Exercise

An example of a contravariant functor: let $\mathcal{C}$ be the category of vector spaces over a field k and associate to any vector space V its k-linear dual V^*. Show that this correspondence gives a contravariant functor.

Exercise

Let X be a topological space and introduce an equivalence relation on X by declaring $x \sim y$ for $x, y \in X$ if there is a path in X connecting x and y. Show that the above is indeed an equivalence relation.

Definition

The equivalence classes of X under the equivalence relation introduced above are called the path components *of X. The set of equivalence classes is denoted by $\pi_0 X$.*

We see, that every space is the disjoint union of path connected subspaces, its path components.

The set $\pi_0 X$ gives in fact rise to a functor π_0 from $\mathcal{T}op$ to $\mathcal{S}$. Indeed, let $f : X \longrightarrow Y$ be a map. Let $[x] \in \pi_0 X$, the connected component containing $x \in X$. Then $f[x] := [f(x)]$. It is an easy exercise to check that π_0 preserves compositions and identities, therefore it is indeed a functor $\pi_0 : \mathcal{T}op \to \mathcal{S}$.

If you think of categories as of something like groups then functors are like homomorphisms between groups. We are most interested in the category $\mathcal{T}op$. However this category is hard to study directly. We will proceed by constructing various functors from $\mathcal{T}op$ into more algebraically manageable categories like $\mathcal{A}b$ and studying the images of these functors.

Let us now consider the next level of abstraction – the *category of functors*. We stress that this is not some arcane notion but is indispensable in many concrete questions.

Definition

Let $\mathcal{C}$ and $\mathcal{D}$ be two categories and $F, G : \mathcal{C} \to \mathcal{D}$ be two functors. A morphism *of functors (also called a* natural transformation*) $f : F \to G$ is a family of morphisms in $\mathcal{D}$:*

$$f(X) : F(X) \to G(X),$$

one for each object X in $\mathcal{C}$ such that for any morphism $\phi : X \to Y$ the following diagram is commutative:

$$\begin{tikzcd} F(X) \arrow[r, "f(X)"] \arrow[d, "F(\phi)"'] & G(X) \arrow[d, "G(\phi)"] \\ F(Y) \arrow[r, "f(Y)"] & G(Y) \end{tikzcd}$$

If, for every object X in $\mathcal{C}$, the morphism $f(X)$ is an isomorphism in $\mathcal{D}$, then f is called a natural isomorphism *(or* isomorphism of functors*).*

The composition of morphisms of functors as well as the identity morphism are defined in an obvious way. We see, therefore, that functors from one category to another themselves form a category, the *category of functors*.
The next notion we will consider is that of an *equivalence of categories*. If we view a category as an analogue of a group then this is analogues to the notion of an isomorphism. However we will see that there is important subtlety in the definition of an equivalence of categories.

Definition

Let $\mathcal{C}$ and $\mathcal{D}$ be two categories. They are said to be equivalent *if there exist two functors $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{C}$ such that the composition $F \circ G$ is isomorphic to the identity functor on $\mathcal{D}$ and the composition $G \circ F$ is isomorphic to the identity functor on $\mathcal{C}$. In this situation the functors F and G are called* quasi-inverse equivalences *between $\mathcal{C}$ and $\mathcal{D}$.*

Remark

There is also a notion of an isomorphism between categories which is obtained if one requires that the compositions of F and G be equal to the identity functors on $\mathcal{C}$ and $\mathcal{D}$ (as opposed to isomorphic). This notion, surprisingly, turns out to be more or less useless since a natural construction hardly ever determines an isomorphism of categories. The following example is instructive.

Example

Consider the category $Vect_k$ *of finite-dimensional vector spaces over a field* k *and a functor* $Vect_k^{op} \to Vect_k$ *given by associating to a vector space* V *its dual* V^**. We claim that this functor establishes an equivalence of categories* $Vect_k$ *and* $Vect_k^{op}$ *where the quasi-inverse functor is likewise given by associating to a vector space its dual. Indeed, the composition of the two functors associates to a vector space* V *it double dual* V^{**}*. There is then a natural (i.e. functorial) isomorphism* $V \to V^{**}$*: a vector* $v \in V$ *determines a linear function* $\tilde{v}$ *given for* $\alpha \in V^*$ *by the formula* $\tilde{v}(\alpha) = \alpha(v)$*. Note that* V^{**} *is* not *equal to* V*, only canonically* isomorphic *to it.*

Exercise

Fill in the details in the above proof that the dualization functor is an equivalence of categories.

Remark

Many of the theorems in mathematics, particularly in algebraic topology, could be interpreted as statements that certain categories are equivalent. We will see some examples later on. For now let us formulate a useful criterion for a functor to be an equivalence which does not require constructing a quasi-inverse explicitly. It is somewhat analogous to the statement that a map of sets is an isomorphism (bijection) if and only if it is a surjection and an injection.

Definition

A functor $F : \mathcal{C} \to \mathcal{D}$ is called full *if for any $X, Y \in Ob(\mathcal{C})$ the map $Hom_{\mathcal{C}}(X, Y) \to Hom_{\mathcal{D}}(F(X), F(Y))$ is surjective. If the latter map is injective then F is said to be* faithful*.*

Theorem

A functor $F : \mathcal{C} \to \mathcal{D}$ is an equivalence if and only if:

1. *F is full and faithful.*
2. *Every object in $\mathcal{D}$ is isomorphic to an object of the form $F(X)$ for some object X in $\mathcal{C}$.*

Proof: Let $F : \mathcal{C} \to \mathcal{D}$ be an equivalence and $G : \mathcal{D} \to \mathcal{C}$ be the quasi-inverse functor. Let

$$f(X) : GFX \to X, X \in Ob(\mathcal{C}),$$

$$g(Y) : FGY \to Y, Y \in Ob(\mathcal{D})$$

be the given isomorphisms of functors $GF \to Id_{\mathcal{C}}$ and $FG \to Id_{\mathcal{D}}$. Note that an object Y of $\mathcal{D}$ is isomorphic to $F(GY)$ which proves that F is surjective on isomorphism classes of objects.

Further, let $\phi \in Hom_{\mathcal{C}}(X, X')$ be a morphism in $\mathcal{C}$ and consider the commutative diagram

$$\begin{tikzcd} GFX \arrow[r, "f(X)"] \arrow[d, "GF(\phi)"'] & X \arrow[d, "\phi"] \\ GFX' \arrow[r, "f(X')"] & X' \end{tikzcd}$$

We see that ϕ can be recovered from $F(\phi)$ by the formula

$$\phi = f(X') \circ GF(\phi) \circ f(X)^{-1}$$

Which shows that F is a faithful functor. Similarly, G is likewise faithful. Now consider a morphism $\psi \in Hom_{\mathcal{D}}(FX, FX')$ and set

$$\phi = f(X') \circ G(\psi) \circ f(X)^{-1} \in Hom_{\mathcal{C}}(X, X').$$

Then (as has just been proved) $\phi = f(X') \circ GF(\phi) \circ f(X)^{-1}$ and $G(\psi) = GF(\phi)$ because $f(X), f(X')$ are isomorphisms. Since G is faithful, $\psi = F(\phi)$ so F is fully faithful as required.
Conversely suppose that the conditions (1) and (2) hold. For any $Y \in Ob(\mathcal{D})$ fix $X_Y \in Ob(\mathcal{C})$ so that there exists an isomorphism $g(Y) : FX_Y \to Y$. We define the functor G quasi-inverse to F by $GY = X_Y$ and for $\psi \in Hom_{\mathcal{D}}(Y, Y')$ define $G(\psi) \in Hom_{\mathcal{C}}(GY, GY')$ as the unique morphism $GY \to GY'$ such that:

$$FG(\psi) = g(Y')^{-1} \circ \psi \circ g(Y) \in Hom_{\mathcal{D}}(FGY, FGY').$$

Such a morphism exists, because F is assumed to be full and it is unique, because F is also assumed to be faithful. It is easy to check that G defines a functor and that $g = \{g(Y)\} : FG \to Id_{\mathcal{D}}$ is an isomorphism of functors.
Finally $g(FX) : FGFX \to FX$ is an isomorphism for all $X \in Ob(\mathcal{C})$. Therefore by (1) $g(FX) = F(f(X))$ for a unique isomorphism $f(X) : GFX \to X$. An easy inspection shows that $f = \{f(X)\}$ is an isomorphism of functors $GF \to Id_{\mathcal{C}}$.
Therefore G is indeed quasi-inverse to F. QED.

Let $I = [0, 1]$ be the unit interval with its usual topology induced from the real line $\mathbb{R}$. It will play the role of a parameter space.

Definition

Let X and Y be two topological spaces and $f, g : X \longrightarrow Y$ be two (continuous) maps. Then f is said to be homotopic *to g if there exists a map $F : X \times I \longrightarrow Y$, called a* homotopy *such that $F(x, 0) = f(x)$ and $F(x, 1) = g(x)$. In that case we will write $f \sim g$.*

A homotopy F can be considered as a *continuous* family of maps $f_t : X \longrightarrow Y$ indexed by the points $t \in I$. Then $f_0 = f$ and $f_1 = g$. In other words the homotopy F *continuously deforms* the map f into the map g.

Proposition

The relation $\sim$ is an equivalence relation on the set of maps from X to Y.

Proof.

1. *Reflexivity.* Let $f : X \longrightarrow Y$ be a map and define $F : X \times I \longrightarrow Y$ by the formula $F(x, t) = f(x)$. Then F is a homotopy between f and itself.
2. *Symmetry.* Assume that $f \sim g$. Then there exists a homotopy $F : X \times I \longrightarrow Y$ such that $F(x, 0) = f(x)$ and $F(x, 1) = g(x)$. Define the homotopy $G : X \times I \longrightarrow Y$ by the formula $G(x, t) = F(x, 1 - t)$. Then, clearly $G(x, 0) = g(x)$ and $G(x, 1) = f(x)$ so $g \sim f$.
3. *Transitivity.* Suppose that $f \sim g$ and $g \sim h$. Let $F : X \times I \longrightarrow Y$ be the homotopy relating f and g and $G : X \times I \longrightarrow Y$ be the homotopy relating g and h. Define the homotopy $H : X \times I \longrightarrow Y$ by the formula

$$H(x, t) = \begin{cases} F(x, 2t) & \text{if } t \leq \frac{1}{2}, \\ G(x, 2t - 1) & \text{if } t \geq \frac{1}{2}. \end{cases}$$

Then clearly $f \sim h$ via the homotopy H. $\square$

Since $\sim$ is an equivalence relation the set of maps from X to Y is partitioned into equivalence classes.These classes are called *the homotopy classes of maps from X into Y*. The set of all homotopy classes is denoted by $[X, Y]$.

Proposition

Let $f, f' : X \longrightarrow Y$ and $g, g' : Y \longrightarrow Z$ be the continuous maps. Suppose that $f \sim f'$ and $g \sim g'$. Then $g \circ f \sim g' \circ f'$.

Proof.

We'll first prove that $g \circ f \sim g \circ f'$. Let $F : X \times I \longrightarrow Y$ be the homotopy connecting f and f'. Define the homotopy $F' : X \times I \longrightarrow Z$ as the composition

$$X \times I \xrightarrow{F} Y \xrightarrow{g} Z \ .$$

Now we'll show that $g \circ f' \sim g' \circ f'$. Let $G : Y \times I \longrightarrow Z$ be the homotopy connecting g and g'. Define the homotopy $G' : X \times I \longrightarrow Z$ as the composition

$$X \times I \xrightarrow{f' \times id} Y \times I \xrightarrow{G} Z \ .$$

Therefore $g \circ f \sim g \circ f' \sim g' \circ f'$ and we are done. $\square$

Definition

The homotopy category *$h\mathcal{T}op$ is the category whose objects are topological spaces and the set of morphisms between two objects X and Y is the set of homotopy classes of maps $X \longrightarrow Y$.*

Exercise

Check that $h\mathcal{T}op$ is indeed a category. Note that there is a tautological functor from the category $\mathcal{T}op$ to $h\mathcal{T}op$.

Definition

Two topological spaces X and Y are called homotopy equivalent *if there are maps $f : X \longrightarrow Y$ and $g : Y \longrightarrow X$ such that $f \circ g \sim 1_Y$ and $g \circ f \sim 1_X$. In that case f and g are called* homotopy equivalences.

In other words a homotopy equivalence is a map which admits a two-sided inverse *up to homotopy*. 'Homotopy equivalence' is just the categorical isomorphism in $h\mathcal{T}op$. Note that a homeomorphism is a special case of a homotopy equivalence. However there are many examples where a homotopy equivalence is not a homeomorphism as we'll see shortly.

Definition

A topological space is called contractible *if it is homotopy equivalent to a point $\{pt\}$.*

Definition

Let X, Y be spaces and $y \in Y$. Then the map $f : X \longrightarrow Y$ is called nullhomotopic *if it is homotopic to the constant map taking every point x in X into $y \in Y$.*

Note that a map $\{pt\} \longrightarrow X$ is nothing but picking a point in X. We see that a topological space X is contractible if and only if the identity map $1_X : X \longrightarrow X$ is nullhomotopic.
Recall that a subset X of $\mathbb{R}^n$ is *convex* if for each pair of points $x, y \in X$ the line segment joining x and y is contained in X, i. e. $ty + (1-t)x \in X$ for all $t \in I$.

Proposition

Every convex set $X \subseteq \mathbb{R}^n$ is contractible.

Proof.

Choose $x_0 \in X$ and define $f : X \longrightarrow X$ by $f(x) = x_0$ for all $x \in X$. Then define a homotopy $F : X \times I \longrightarrow X$ between f and 1_X by $F(x,t) = tx_0 + (1-t)x$. □

This shows that there are many contractible spaces which are *not* points. In other words the notion of homotopy equivalence is strictly weaker then that of homeomorphism. On the other hand note that a contractible set need not necessarily be convex. (Show that a hemispere is contractible!).

It is easy to construct nullhomotopic maps and contractible spaces. In fact you could deduce (Exercise!) that for a contractible space X and any space Y all maps $X \longrightarrow Y$ are nullhomotopic (and homotopic to each other) so that the set $[X, Y]$ consists of just one element. Similarly the set $[Y, X]$ consists of only one element.
It is much harder to show that a given map is *essential* that is, not null-homotopic or that a given space is *not* contractible. We will now give an example of an essential map.

Let $\mathbb{C}$ denote the field of complex numbers and $\Sigma_\rho \subset \mathbb{C}$ denote the circle with center at the origin and radius ρ. Consider the function $z \longrightarrow z^n$ and denote by f_ρ^n its restriction to Σ_ρ. Thus $f_\rho^n : \Sigma_\rho \longrightarrow \mathbb{C} \setminus \{0\}$.

Theorem

For any $n > 0$ and any $\rho > 0$ the map f_ρ^n is essential.

The proof of this theorem will be given later. For now we will deduce from it the *fundamental theorem of algebra*.

Theorem (Fundamental Theorem of Algebra)

Any non constant polynomial $g(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n$ with complex coefficients has at least one complex root.

Proof: Consider the polynomial with complex coefficients:

$$g(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n.$$

Choose $\rho > \max\{1, \Sigma_{i=0}^{n-1} |a_i|\}$ and define $F : \Sigma_\rho \times I \longrightarrow \mathbb{C}$ by

$$F(z,t) = z^n + \Sigma_{i=0}^{n-1} (1-t) a_i z^i.$$

It would be clear that F is a homotopy between f_ρ^n and $g|_{\Sigma_\rho}$ if we can show that the image of F is contained in $\mathbb{C} \setminus \{0\}$. In other words we need to show that $F(z,t) \neq 0$. Indeed, if $F(z,t) = 0$ for some $t \in I$ and some z with $|z| = \rho$ then

$$z^n = -\Sigma_{i=0}^{n-1} (1-t) a_i z^i.$$

By the triangle inequality

$$\rho^n \leq \Sigma_{i=0}^{n-1} (1-t) |a_i| \rho^i \leq \Sigma_{i=0}^{n-1} |a_i| \rho^i \leq \Sigma_{i=0}^{n-1} |a_i| \rho^{n-1},$$

because $\rho > 1$ implies that $\rho^i \leq \rho^{n-1}$. Canceling ρ^{n-1} gives $\rho \leq \Sigma_{i=0}^{n-1} |a_i|$ which contradicts our choice of ρ.

Now suppose that g has no complex roots. Define $G : \Sigma_\rho \times I \longrightarrow \mathbb{C} \setminus \{0\}$ by $G(z,t) = g((1-t)z)$.
Note that since g has no roots the values of G do lie in $\mathbb{C} \setminus \{0\}$. Then G is a homotopy between g restricted to Σ_ρ and the constant function $z \longrightarrow g(0) = a_0$.
Therefore $g|\Sigma_\rho$ is nullhomotopic and since f_ρ^n is homotopic to g it is also nullhomotopic. This contradicts the preceding theorem about essential maps.
QED.

For technical reasons it is often more convenient to work in the category of *pointed* topological spaces. Here's the definition:

Definition

A pointed space *is a pair* (X, x_0) *where* X *is a space and* $x_0 \in X$. *Then* x_0 *is called the* base point *of* X. *A map of based spaces* $(X, x_0) \longrightarrow (Y, y_0)$ *is just a continuous map* $f : X \longrightarrow Y$ *such that* $f(x_0) = y_0$. *The category of pointed topological spaces and their maps is denoted by* $\mathcal{T}op_*$.

Question

What is a homotopy in the category $\mathcal{T}op_*$*?*

For this, let us introduce a slightly more general notion of a *relative* homotopy.

Definition

Let X *be a topological space,* $A \subseteq X$ *and* Y *is another space. Let* $f, g : X \longrightarrow Y$ *be two maps such that their restrictions to* A *coincide. Then* f *is homotopic to* g relative to A *if there exists a map* $F : X \times I \longrightarrow Y$ *such that* $F(a, t) = f(a) = g(a)$ *for all* $a \in A$ *and* $t \in I$. *We will say that* $f \sim g$ *rel* A.

Now let (X, x_0) and (Y, y_0) be pointed spaces and $f, g : X \longrightarrow Y$ be two pointed maps.

Definition

The maps f and g are called homotopic as pointed maps if they are homotopic rel x_0. The set of pointed homotopy classes of maps from X to Y is denoted by $[X, Y]_*$.

In other words the homotopy between f and g goes through pointed maps where the cross-section $X \times t \subset X \times I$ has (x_0, t) for its base point. Similarly to the unpointed case one shows that pointed homotopy is an equivalence relation on the set of pointed maps from one space to another. Moreover the composition of pointed homotopy classes of maps is well-defined and we are entitled to the following

Definition

The homotopy category of pointed spaces $h\mathcal{T}op_*$ is the category whose objects are pointed spaces and morphisms are homotopy classes of pointed maps.

More often than not we will work with pointed spaces pointed maps, homotopies etc.

Question

What is the relevant notion of homotopy equivalence for pointed spaces?

We are now preparing to define the *fundamental group* of a spaces. As the name suggests this is one of the most important invariants that a space has. Recall that a path γ in X is just a continuous map $\gamma : I \longrightarrow X$.

Definition

We say that two paths δ and γ in X are homotopic if they are homotopic as maps $I \longrightarrow X$ rel $\{0, 1\}$. We denote the homotopy class of the path γ by $[\gamma]$.

Note our abuse of language here; the notion of homotopy of paths differs from usual homotopy. So two paths are homotopic if they could be continuously deformed into each other in such a way that in the process of deformation their endpoints don't move.

Given two paths δ, γ in X such that $\delta(1) = \gamma(0)$ their product $\delta \cdot \gamma$ is the path that travels first along δ then along γ. More formally,

$$\delta \cdot \gamma(t) = \begin{cases} \delta(2t) & \text{if } 0 \leq t \leq \frac{1}{2}, \\ \gamma(2t-1) & \text{if } \frac{1}{2} \leq t \leq 1. \end{cases}$$

Note that the product operation respects homotopy classes in the sense that $[\delta \cdot \gamma] = [\delta' \cdot \gamma']$ if $[\delta] = [\delta']$ and $[\gamma] = [\gamma']$ (prove that!).
Further for a path γ define its inverse γ^{-1} by the formula $\gamma^{-1}(t) = \gamma(1-t)$.

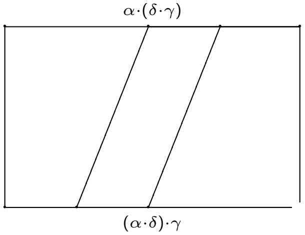
Let us now assume that the space X is pointed and restrict attention to those paths γ for which $\gamma(0) = \gamma(1) = x_0$, the base point of X. Such a path is called a *loop* in X (based at x_0) and the set of homotopy classes of loops based at x_0 is denoted by $\pi_1(X, x_0)$. The *product* of two loops is again a loop. Let us define the *constant loop* $e : I \to X$ by the formula $e(t) = x_0$. Then we have

Proposition

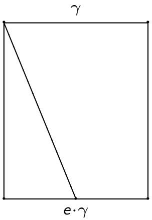
The set $\pi_1(X, x_0)$ *is a group with respect to the product* $[\delta] \cdot [\gamma] = [\delta \cdot \gamma]$.

Proof: The following slightly stylized pictures represent relevant homotopies. You could try to translate them into formulas if you like. (Check it!)

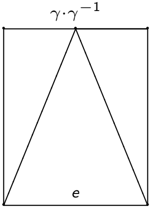
Associativity:



Existence of a unit:



Existence of inverse:



QED.

The group $\pi_1(X, x_0)$ is called the *fundamental group* of the pointed space (X, x_0). Note that a loop $\gamma : I \longrightarrow X$ could be considered as a map $S^1 \longrightarrow X$ which takes the base point $1 \in S^1$ to $x_0 \in X$. The homotopies of paths correspond to homotopies of *based* maps $S^1 \longrightarrow X$ and therefore the fundamental group $\pi_1(X, x_0)$ is the same as $[S^1, X]_*$, the set of homotopy classes of pointed maps from S^1 to X. Note also that the correspondence $(X, x_0) \mapsto \pi_1(X, x_0)$ is a functor $h\mathcal{T}op_* \mapsto \mathcal{G}r$.
It is natural to ask how the fundamental group of X depends on the choice of the base point. It is clear that if we choose base points lying in different connected components of X then there is no connection whatever between the corresponding fundamental groups. We assume, therefore that X is connected. Let x_0, x_1 be two points in X and choose a path $h : I \longrightarrow X$ connecting x_0 and x_1. The inverse path $h^{-1} : I \longrightarrow X$ then connects x_1 and x_0. Then we can associate to any loop γ of X based at x_1 the loop $h \cdot \gamma \cdot h^{-1}$ based at x_0. Strictly speaking we should choose an order of forming the product $h \cdot \gamma \cdot h^{-1}$, either $(h \cdot \gamma) \cdot h^{-1}$ or $h \cdot (\gamma \cdot h^{-1})$ but the two choices are homotopic and we are only interested in homotopy classes here.

Proposition

The map $\beta_h : \pi_1(X, x_1) \longrightarrow \pi_1(X, x_0)$ *defined by* $\beta_h[\gamma] = [h \cdot \gamma \cdot h^{-1}]$ *is an isomorphism of groups.*

Proof.

If $f_t : I \longrightarrow X$ is a homotopy of loops based at x_1 then $h \cdot f_t \cdot h^{-1}$ is a homotopy of loops based at x_0 so β_h is well-defined. Further β_h is a group homomorphism since

$$\beta_h[\gamma_1 \cdot \gamma_2] = [h \cdot \gamma_1 \cdot \gamma_2 \cdot h^{-1}] = [h \cdot \gamma_1 \cdot h^{-1} \cdot h \cdot \gamma_2 \cdot h^{-1}] = \beta_h[\gamma_1]\beta_h[\gamma_2].$$

Finally β_h is an isomorphism with inverse $\beta_{h^{-1}}$ since

$$\beta_h \beta_{h^{-1}}[\gamma] = \beta_h[h^{-1} \cdot \gamma \cdot h] = [h \cdot h^{-1} \cdot \gamma \cdot h \cdot h^{-1}] = [\gamma]$$

and similarly $\beta_{h^{-1}}\beta_h[\gamma] = [\gamma]$. $\square$

So we see that if X is (path-)connected then the fundamental group of X is independent, up to an isomorphism, of the choice of the base point in X. In that case the notation $\pi_1(X, x_0)$ is often abbreviated to $\pi_1(X)$ or $\pi_1 X$.

Exercise

Show that two homotopic paths h_1 *and* h_2 *connecting* x_0 *and* x_1 *determine the same isomorphism between* $\pi_1(X, x_0)$ *and* $\pi_1(X, x_1)$*. That is,* $\beta_{h_1} = \beta_{h_2}$*.*

Definition

A space X is called simply-connected *if it is (path-)connected and has a trivial fundamental group.*

Exercise

Show that a space X is simply-connected iff there is a unique homotopy class of paths connecting any two points in X.

Proposition

$\pi_1(X \times Y, (x_0, y_0))$ *is isomorphic to* $\pi_1(X, x_0) \times \pi_1(Y, y_0)$ *if the pointed spaces* (X, x_0) *and* (Y, y_0) *are connected.*

Proof.

A basic property of the product topology is that a map $f : Z \longrightarrow X \times Y$ is continuous iff the maps $g : Z \longrightarrow X$ and $h : Z \longrightarrow Y$ defined by $f(z) = (g(z), h(z))$ are both continuous. Therefore a loop γ in $X \times Y$ based at (x_0, y_0) is the same as a pair of loops γ_1 in X and γ_2 in Y based at x_0 and y_0 respectively. Similarly a homotopy f_t of a loop in $X \times Y$ is the same as a pair of homotopies g_t and h_t of the corresponding loops in X and Y. Thus we obtain a bijection $\pi_1(X \times Y, (x_0, y_0)) \mapsto \pi_1(X, x_0) \times \pi_1(Y, y_0)$ so that $[\gamma] \mapsto [\gamma_1] \times [\gamma_2]$. This clearly is a group homomorphism and we are done. $\square$

We consider S^1 embedded into $\mathbb{R}^2$ as a unit circle having its center at the origin. The point $(1,0)$ will be the base point.

Theorem

The fundamental group of S^1 based at $(1,0)$ is isomorphic to $\mathbb{Z}$, the group of integers, i. e.

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

In particular it is abelian.

Proof. To any point $x \in S^1$ we associate in the usual way a real number defined up to a summand of the form $2\pi k, k \in \mathbb{Z}$. For example, the base point $(1,0)$ is associated with the collection $\{2\pi k\}$, the point $(0,1)$ with the collection $\{\frac{\pi}{2}+2\pi k\}$. Then any loop $\omega : I \longrightarrow S^1$ corresponds to a multivalued function ω' on I whose value at any point is defined up to a summand $2\pi k$ and the values of which at 0 and 1 is the collection of numbers $\{2\pi k\}$.
Let us call a function $\omega'' : I \longrightarrow \mathbb{R}$ a *single-valued branch* of ω' if ω'' is continuous and its (single) value at any point $x \in I$ belongs to the set of values at x assumed by ω'. We claim that ω' has a single-valued branch ω'' which is determined uniquely by the condition $\omega''(0) = 0$. Indeed, let n be a positive integer such that if $|x_1 - x_2| \leq \frac{1}{n}$ then the points $\omega(x_1), \omega(x_2) \in S^1$ are *not* diametrically opposite.

Set $\omega''(0) = 0$. Further for $0 \leq x \leq \frac{1}{n}$ we choose for $\omega''(x)$ that value of $\omega'(x)$ for which $\omega'(x) < \pi$. Then for $\frac{1}{n} \leq x \leq \frac{2}{n}$ we take for $\omega''(x)$ the value of $\omega'(x)$ for which $\omega'(x) < \omega''(\frac{1}{n})$. And so forth.
Note the following properties of the function $\omega'' : I \longrightarrow \mathbb{R}$:

- $\omega''(1)$ is an integer multiple of 2π.
- A homotopy ω_t of the loop ω determines a homotopy ω''_t of ω''.

Note that the integer $k = \frac{\omega''(1)}{2\pi}$ does not change under any homotopy because it can only assume a discrete set of values. So this integer only depends on the homotopy class of ω, that is, on the element in $\pi_1(S^1, (1,0))$ which ω represents.
Next, for any given k there exists a loop ω for which $\frac{\omega''(1)}{2\pi} = k$. Indeed, it suffices to set $\omega = h_k = 2\pi kx$.
Finally if ω and λ are two loops for which $\omega''(1) = \lambda''(1) = k$ then ω'' and λ'' are homotopic in the class of functions $I \longrightarrow \mathbb{R}$ having fixed values at 0 and 1 and both are homotopic to h_k. (Check it!)
That shows that the correspondence $\omega \mapsto \frac{\omega''(1)}{2\pi}$ determines a bijection of sets $\pi_1(S^1, (1,0)) \to \mathbb{Z}$. To see that this is in fact an isomorphism of groups, check

$$(h_k \cdot h_l)''(1) = h''_{k+l}(1).$$

QED.

The map $S^1 \longrightarrow S^1$ corresponding to the loop having invariant n is called a *degree n map*. Thus, a degree n map from S^1 into itself wraps S^1 around itself n times.
The correspondence $(X, x_0) \mapsto \pi_1(X, x_0)$ is a functor

$$\mathcal{T}op_* \longrightarrow \mathcal{G}r.$$

For a map of pointed spaces $f : (X, x_0) \longrightarrow (Y, y_0)$ we have a map

$$f_* : \pi_1(X, x_0) \longrightarrow \pi_1(Y, y_0), f_*([\omega]) := [f \circ \omega]$$

called the *induced map* of fundamental groups of (X, x_0) and (Y, y_0). (Check it!)

Exercise

Consider a map $f : S^1 \longrightarrow S^1$ of degree n. Then, clearly, the induced map $\pi_1(S^1, (1, 0)) \cong \mathbb{Z} \longrightarrow \pi_1(S^1, (1, 0)) \cong \mathbb{Z}$ is just multiplication by n.

Exercise

If two spaces X and Y are homotopically equivalent through a basepoint preserving homotopy then $\pi_1 X \cong \pi_1 Y$

Proposition

If $f : X \longrightarrow Y$ is an (unpointed) homotopy equivalence then the induced homomorphism $\pi_1(X, x_0) \longrightarrow \pi_1(Y, f(x_0))$ is an isomorphism for all $x_0 \in X$.

Proof. The proof will use a simple fact about homotopies that do not fix the basepoint:

Lemma

Let $f_t : X \longrightarrow Y$ be a homotopy between $f_0, f_1 : X \longrightarrow Y$ and h be the path $f_t(x_0)$ formed by the images of the basepoint $x_0 \in X$. Then $f_{0_*} = \beta_h f_{1*}$. In other words the following diagram of groups is commutative:

$$
\begin{tikzcd}
& \pi_1(X, x_0) \arrow[dl, "f_{0*}"'] \arrow[dr, "f_{1*}"] & \\
\pi_1(Y, f_0(x_0)) & & \pi_1(Y, f_1(x_0)) \arrow[ll, "\beta_h"']
\end{tikzcd}
$$

(Recall that $\beta_h : \pi_1(Y, f_1(x_0)) \longrightarrow \pi_1(Y, f_0(x_0))$ is a homomorphism induced by the path h.)

This lemma is almost obvious after you draw the picture (Check it!). Let h_t be the restriction of h to the interval $[0, t]$ rescaled so that its domain is still $[0, 1]$. Then if ω is a loop in X based at x_0 the product $h_t \cdot (f \circ \omega) h_t^{-1}$ gives a homotopy of loops at $f_0(x_0)$. Restricting this homotopy to $t = 0$ and $t = 1$ we see that $f_{0*}[\omega] = \beta_h(f_{1_*}[\omega])$ so our lemma is proved.

Let us now return to the proof of the Proposition. Let $g : Y \longrightarrow X$ be a homotopy inverse for f so that $f \circ g \sim 1_Y$ and $g \circ f \sim 1_X$. Consider the maps

$$\pi_1(X, x_0) \xrightarrow{f_*} \pi_1(Y, f(x_0)) \xrightarrow{g_*} \pi_1(X, g \circ f(x_0)) \xrightarrow{f_*} \pi_1(Y, f \circ g \circ f(x_0)) .$$

The composition of the first two maps is an isomorphism since $g \circ f \sim 1_X$ implies that $g_* \circ f_* = \beta_h$ for some h by the previous lemma. In particular since $g_* \circ f_*$ is an isomorphism, f_* must be injective. The same reasoning with the second and third map shows that g_* is injective. Thus the first two of the three maps are injections and their composition is an isomorphism, so the first map f_* must be surjective as well as injective. QED.

Even though most of the time we work in the pointed context occasionally we use unpointed maps and homotopies. For two spaces X and Y and maps $f, g : X \longrightarrow Y$ we say that f and g are *freely* homotopic to emphasize that they are homotopic through a non-basepoint-preserving homotopy.

Exercise

If $f : X \longrightarrow Y$ is freely nullhomotopic then the induced homomorphism $f_* : \pi_1(X) \longrightarrow \pi_1(Y)$ is trivial. (Hint: use the Lemma of the last proof!)

Remember the theorem claiming that the map $f_\rho^n : \Sigma_\rho \longrightarrow \mathbb{C} \setminus \{0\}$ given by $z \to z^n$ is *not* nullhomotopic. We have the following commutative diagram of spaces:

$$\begin{tikzcd} \Sigma_\rho \arrow[r, "f_\rho^n"] \arrow[d, "\sim"] & \mathbb{C} \setminus \{0\} \arrow[d, "\sim"] \\ S^1 \arrow[r] & S^1 \end{tikzcd}$$

Here the downward arrows are homotopy equivalences and the lower horizontal map is a map of degree n. (Check this!) Applying the functor $\pi_1(?)$ to the above diagram we would get a commutative diagram of abelian groups

$$\begin{tikzcd} \mathbb{Z} = \pi_1\Sigma_\rho \arrow[r] \arrow[d, equal] & \mathbb{Z} = \pi_1(\mathbb{C} \setminus \{0\}) \arrow[d, equal] \\ \mathbb{Z} = \pi_1 S^1 \arrow[r, "n"] & \mathbb{Z} = \pi_1 S^1 \end{tikzcd}$$

Now if f_ρ^n were nullhomotopic then the upper horizontal map in the above diagram would be the zero map which is impossible. Therefore this theorem is also proved. QED.

Exercise

Following the ideas in the introduction prove the Brouwer fixed point theorem for a two-dimensional disk using the fundamental group functor! Assuming more generally that there are functors $\pi_n : \mathcal{T}op_* \mapsto \mathcal{G}r$ with $\pi_n S^n = \mathbb{Z}$ prove it in the general case!

Remark

We will eventually give a proof of the Brouwer fixed point theorem in the general case using homology groups rather than homotopy groups.

Sometimes using homotopy groups we could prove that spaces are not *homeomorphic* to each other. Here's an example

Corollary

The two-dimensional sphere S^2 is not *homeomorphic to $\mathbb{R}^2$.*

Proof.

Suppose such a homeomorphism $f : S^2 \longrightarrow \mathbb{R}^2$ exists. Let $p = f^{-1}\{0\}$. Then the punctured sphere $S^2 \setminus \{p\}$ is homeomorphic to $\mathbb{R}^2 \setminus \{0\}$. However $S^2 \setminus \{p\}$ is contractible, in particular $\pi_1 S^2 \setminus \{p\}$ is trivial. On the other hand $\mathbb{R}^2 \setminus \{0\}$ is homotopically equivalent to S^1 and therefore has a nontrivial fundamental group, a contradiction. $\square$

Question

What about fundamental groups of higher-dimensional spheres?

It turns out that they are all trivial.

Proposition

$\pi_1 S^n = 0$ *for* $n > 1$*.*

Proof. Let ω be a loop in S^n at a chosen basepoint x_0. If the image of ω is disjoint from some other point $x \in S^n$ then ω is actually a map $S^1 \longrightarrow S^n \setminus \{x\}$. Note that $S^n \setminus \{x\}$ could be collapsed to the one-point space along the meridians. Therefore $S^n \setminus \{x\}$ is homotopically equivalent to the point, in particular it is simply-connected. Therefore in that case ω is null-homotopic. So it suffices to show that ω is homotopic to a map that is non-surjective. To this end consider a small open ball in S^n about any point $x \neq x_0$. Note that the number of times ω enters B, passes through x and leaves B is finite (why?) so each of the portions of ω can be pushed off x without changing the rest of ω. More precisely, we consider ω as a map $I \longrightarrow S^n$. Then the set $\omega^{-1}(B)$ is open in $(0, 1)$ and hence is the union of a possibly infinite collection of disjoint open intervals (a_i, b_i). The compact set $\omega^{-1}(x)$ is contained in the union of these intervals, so it must be contained in the union of finitely many of them.

Consider one of the intervals (a_i, b_i) meeting $\omega^{-1}(x)$. The path ω_i obtained by restricting ω to the interval $[a_i, b_i]$ lies in the closure of B and its endpoints $\omega(a_i), \omega(b_i)$ lie in the boundary of B. Since $n \geq 2$ we can choose a path γ_i from $\omega(a_i)$ to $\omega(b_i)$ inside the closure of B but disjoint from x. (For example, we could choose γ_i to lie in the boundary of B which is a sphere of dimension $n-1$ which is connected if $n \geq 2$). Since the closure of B is simply-connected the path ω_i is homotopic to γ_i so we may deform ω by deforming ω_i to γ_i. After repeating this process for each of the intervals (a_i, b_i) that meet $\omega^{-1}(x)$ we obtain a loop γ homotopic to the original ω and with $\gamma(I)$ disjoint from x. QED.

Corollary

$\mathbb{R}^2$ is not homeomorphic to $\mathbb{R}^n$ for $n \neq 2$.

Proof.

Suppose that $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^n$ is a homeomorphism. Let $n > 2$ (the case $n = 1$ is an excercise!). Then $\mathbb{R}^2 \setminus \{0\}$ is homotopy equivalent to S^2 whereas $\mathbb{R}^n \setminus \{f(0)\}$ is homotopically equivalent to S^n (Check this!). Therefore by the Proposition $\mathbb{R}^n \setminus \{f(0)\}$ cannot be homotopy equivalent to $\mathbb{R}^2 \setminus \{0\}$, let alone homeomorphic to it. □

The group $\pi_1(X, x_0)$ is the first of the infinite series of *homotopy invariants* of pointed spaces called *homotopy groups*. Let us sketch the definition and basic properties of these invariants!
Let I^n denote the *n-dimensional cube*, i.e. the product of unit intervals $[0,1]^n$.

Definition

For a pointed space (X, x_0) *define the* n-th homotopy group of X *(denoted by* $\pi_n(X, x_0)$*) as the set of homotopy classes of maps* $f : (I^n, \partial I^n) \to (X, x_0)$ *(i.e. such that the boundary of* I^n *goes to the basepoint* x_0*) where the homotopy* f_t *is required to satisfy* $f_t(\partial I^n) = x_0$ *for all* $t \in [0,1]$.

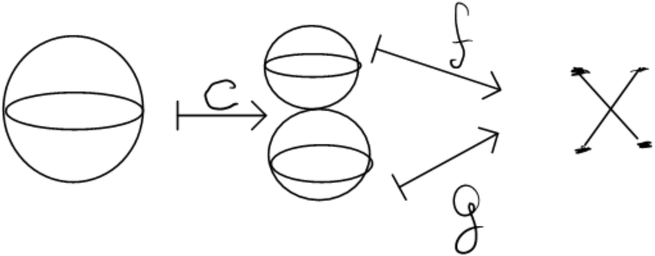
Remark

Note that $I^n/(\partial I^n)$ *is homeomorphic to an n-dimensional sphere* S^n*. It is furthermore clear that* $\pi_n(X, x_0)$ *could alternatively be defined as the set of homotopy classes of based maps* $(S^n, s_0) \to (X, x_0)$*. When* $n = 1$ *we recover the definition of* $\pi_1(X, x_0)$*.*

The set $\pi_n(X, x_0)$ has a group structure defined as follows:
For $f, g : I^n \to X$ set

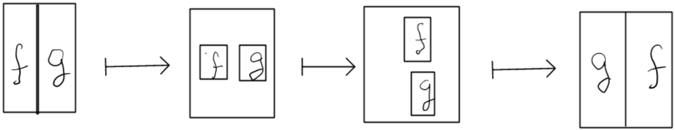
$$(f+g)(s_1, \ldots, s_n) = \begin{cases} f(2s_1, s_2, \ldots, s_n), s_1 \in [0, 1/2] \\ g(2s_1 - 1, s_2, \ldots, s_n), s_1 \in [1/2, 1]. \end{cases}$$

In other words, we cut I^n in half and define the map $f + g$ on each half separately. On the left half this map is (a suitably rescaled) f, on the right half it is (a suitably rescaled) g. When we view the maps f and g as maps of spheres the following picture illustrates the situation:

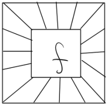


It is clear that this sum is well-defined on homotopy classes. Since only the first coordinate is involved in the sum operation, the same arguments as for π_1 show that $\pi_n(X, x_0)$ is a group. The identity element is the constant map sending the whole of I^n into x_0 and the inverse (or negative) to the element given by a map $f : I^n \to X$ is given by the formula $(-f)(s_1, \ldots, s_n) = f(1 - s_1, s_2, \ldots, s_n)$.

We use the additive notation because (contrary to the case of π_1) the group $\pi_n(X, x_0)$ is always abelian for $n > 1$. The following picture illustrates the homotopy $f + g \sim g + f$:



Just as in the π_1-case different choices of a basepoint x_0 lead to isomorphic groups $\pi_n(X, x_0)$ when X is path-connected. Indeed, given a path $\gamma : I \to X$ from $x_0 = \gamma(0)$ to $x_1 = \gamma(1)$ we may associate to each map $f : (I^n, \partial I^n) \to (X, x_1)$ another map $\gamma f : (I^n, \partial I^n) \to (X, x_0)$ by shrinking the domain of f to a smaller concentric cube in I^n, then inserting the path γ on each radial segment on the region between the smaller cube and ∂I^n as the following picture illustrates:



Let us assume that all our spaces are now *locally path-connected* and *locally simply-connected* (meaning any point possesses a path-connected simply-connected neighborhood). This is not necessary for developing much of the theory but in practice all spaces of interest will even be *locally contractible* and so we will not strive for maximum generality here.

Definition

A covering space *of a connected space X is a connected space $\tilde{X}$ together with a map $p : \tilde{X} \longrightarrow X$ satisfying the following condition:*
There exists an open cover $\{U_\alpha\}$ of X such that for each α the set $p^{-1}(U_\alpha)$ is a disjoint union of open sets in $\tilde{X}$ each of which maps homeomorphically onto U_α.
Sometimes we will refer to the map p as a covering*. The open sets U_α will be called* elementary*.*

Example

An example of a covering is the map $p : \mathbb{R} \longrightarrow S^1$ with $p(t) = (\cos 2\pi t, \sin 2\pi t)$.
Another example is the map $p : S^1 \longrightarrow S^1$ with $p(z) = z^n$ where we view points of S^1 as complex numbers having modulus 1.

The function $x \mapsto \{\text{the number of preimages of } x\}$ is a locally constant function on X; since we assume that X is connected it is actually constant. This number is sometimes called the *number of sheets* of the covering p.

Theorem (Homotopy Lifting property)

Let $p : \tilde{X} \longrightarrow X$ be a covering and $f_t : Y \longrightarrow X$ a homotopy. Suppose that the map $f_0 : Y \longrightarrow X$ lifts *to a map $\tilde{f}_0 : Y \longrightarrow \tilde{X}$. In other words we assume that there exists $\tilde{f}_0 : Y \longrightarrow \tilde{X}$ such that the following diagram is commutative:*

$$\begin{tikzcd} & \tilde{X} \arrow[d, "p"] \\ Y \arrow[ur, "\tilde{f}_0"] \arrow[r, "f_0"] & X \end{tikzcd} .$$

Then there exists a unique homotopy $\tilde{f}_t : Y \longrightarrow \tilde{X}$ lifting the homotopy f_t. That means that the following diagram is commutative for any $t \in I$:

$$\begin{tikzcd} & \tilde{X} \arrow[d, "p"] \\ Y \arrow[ur, "\tilde{f}_t"] \arrow[r, "f_t"] & X \end{tikzcd} .$$

Equivalently if we replace the family $f_t : Y \longrightarrow X$ by a single map $F : Y \times I \longrightarrow X$ and the family $\tilde{f}_t$ by a map $\tilde{F} : Y \times I \longrightarrow \tilde{X}$ then the following diagram should be commutative:

$$\begin{tikzcd} & \tilde{X} \arrow[d, "p"] \\ Y \times I \arrow[ur, "\tilde{F}"] \arrow[r, "F"] & X \end{tikzcd} .$$

Proof. We need a special case of our theorem to prove the general case:

Lemma

For any path $s : I \longrightarrow X$ and any point $\tilde{x}_0$ such that $p(\tilde{x}_0) = s(0) = x_0$ there is a unique path $\tilde{s} : I \longrightarrow \tilde{X}$ such that $\tilde{s}(0) = \tilde{x}_0$ and $\tilde{s}$ lifts s, i.e. the following diagram is commutative:

$$\begin{tikzcd} & \tilde{X} \arrow[d, "p"] \\ I \arrow[ur, "\tilde{s}"] \arrow[r, "s"] & X \end{tikzcd} .$$

(Note that this is a special case of the theorem when Y is a one-point space.)

Proof.

For any $t \in I$ denote by $U(t)$ an elementary neighborhood of the point $s(t)$. Since the unit interval I is compact we can choose a finite collection $U_1, \ldots, U_N$ among $\{U(t)\}$ such that $U_i \supset s(t_i, t_{i+1})$ where $0 = t_1 < t_2 < \ldots < t_{N+1} = 1$. The preimage of U_1 is a disjoint union of open sets in $\tilde{X}$ each of which is homeomorphic to U_1. Among this union we will choose the one which contains $\tilde{x}_0$ and denote it by $\tilde{U}_1$. As a partial lift $\tilde{s}$ of s take the preimage in $\tilde{U}_1$ of the path $s(t)$ restricted to $[t_1, t_2]$ (draw a picture!). Then do the same thing with the neighborhood U_2, the point $s(t_2)$ and the path $s(t)$ restricted to $[t_2, t_3]$ and so one. Since there are only finitely many neighborhoods covering $s(t)$ this process will end. Also since the lift is unique at each neighborhood the resulting path $\tilde{s}$ lifting s will also be unique. □

Let's go back to the proof of the theorem:
Let $y \in Y$ be an arbitrary point. Define a path s_y in X by the formula $s_y(t) = f_t(y)$. This path could be uniquely lifted to $\tilde{s}_y : I \longrightarrow \tilde{X}$ so that $\tilde{s}_y(0) = \tilde{f}(y)$. Letting y vary we obtain the map $\tilde{F}(y, t) := \tilde{s}_y(t)$ so $\tilde{F}$ is the homotopy $Y \times I \longrightarrow \tilde{X}$ which lifts the homotopy $F : Y \times I \longrightarrow X$. QED.

Now we will start making connections with fundamental groups.

Proposition

If $p : \tilde{X} \longrightarrow X$ with $p(\tilde{x}_0) = x_0$ is a covering then $p_* : \pi_1(\tilde{X}, \tilde{x}_0) \longrightarrow \pi_1(X, x_0)$ is a monomorphism.

Proof.

We need to prove that if the loop $\tilde{s} : I \longrightarrow \tilde{X}$ projects onto the loop $s : I \longrightarrow X$ which is nullhomotopic then $\tilde{s}$ itself is nullhomotopic. Consider a homotopy $s_t : I \longrightarrow X$ such that $s_0 = s$, $s_t(1) = s_t(0) = x_0$ and $s_1(I) = x_0$. (The homotopy s_t deforms s into the constant loop in X.) By the homotopy lifting property there exists a homotopy $\tilde{s}_t : I \longrightarrow \tilde{X}$ such that $\tilde{s}_0 = \tilde{s}$ and $p \circ \tilde{s} = s$. Since the preimage of x_0 in $\tilde{X}$ is discrete we have $\tilde{s}_t(0) = \tilde{s}(0) = \tilde{x}_0$ and $\tilde{s}_t(1) = \tilde{s}(1) = \tilde{x}_0$. Furthermore $\tilde{s}_1(t) = \tilde{x}_0$. Therefore the loop $\tilde{s}$ is nullhomotopic. □

We will call the group $p_*\pi_1(\tilde{X}, \tilde{x}_0) \subseteq \pi_1(X, x_0)$ the *group of the covering p*. The group of the covering depends on the choice of the point $\tilde{x}_0$ in $p^{-1}(x_0)$ and also on the point $x_0 \in X$.

Proposition

Let $\tilde{x}_0' \in \tilde{X}$ *be such that* $p(\tilde{x}_0') = x_0$*. Then the subgroups* $p_*\pi_1(\tilde{X}, \tilde{x}_0)$ *and* $p_*\pi_1(\tilde{X}, \tilde{x}_0')$ *inside* $\pi_1(X, x_0)$ *are conjugate.*

Proof.

Let $s : I \longrightarrow X$ be a loop in X which represents an element in $p_*\pi_1(\tilde{X}, \tilde{x}_0)$. That means that there exists a loop $\tilde{s} : I \longrightarrow \tilde{X}$ based at $\tilde{x}_0$ which projects down to s under the map p. Let $\tilde{h}$ be a path from $\tilde{x}_0'$ to $\tilde{x}_0$ and consider the loop $\tilde{s}' := \tilde{h} \cdot \tilde{s} \cdot \tilde{h}^{-1}$. This loop is now based at $\tilde{x}_0'$. Let $s' := p(\tilde{s}')$ and $h := p(\tilde{h})$ be the loops in X obtained by projecting $\tilde{s}'$ and $\tilde{h}$ down to X. Note that $[s'] \in p_*\pi_1(\tilde{X}, \tilde{x}_0')$. It follows that in $\pi_1(X, x_0)$ we have

$$[s'] = [h][s][h]^{-1}.$$

We showed that any element $[s]$ in $p_*\pi_1(\tilde{X}, \tilde{x}_0)$ is conjugate to some element $[s']$ in $p_*\pi_1(\tilde{X}, \tilde{x}_0')$. Symmetrically any element in $p_*\pi_1(\tilde{X}, \tilde{x}_0')$ is conjugate to some element in $p_*\pi_1(\tilde{X}, \tilde{x}_0)$. $\square$

Question

What happens if we change the point $x_0 \in X$?

Take a point $x_1 \in X$ and consider the group $\pi_1(X, x_1)$. There is a collection of subgroups in $\pi_1(X, x_1)$ corresponding to the various choices of the point in $p^{-1}(x_1)$. There is also a collection of subgroups in $\pi_1(X, x_0)$ corresponding to the various choices of the point in $p^{-1}(x_0)$.

Exercise

These two collections correspond to each other under an isomorphism between $\pi_1(X, x_0)$ and $\pi_1(X, x_1)$ given by a path h in X connecting x_0 and x_1. (Hint: use the lifting homotopy property to lift the path h to $\tilde{X}$.)

It turns out that the difference between $\pi_1(\tilde{X}, \tilde{x}_0)$ and $\pi_1(X, x_0)$ is measured by the number of preimages of the point x_0 (which is the number of sheets of the covering p). More precisely:

Proposition

There is a bijective correspondence between the collection of cosets $\pi_1(X, x_0)/p_*\pi_1(\tilde{X}, \tilde{x}_0)$ and the set $p^{-1}(x_0)$.

Proof.

Consider a loop s based at x_0 in X. Using the homotopy lifting property we could lift it to $\tilde{X}$ as a path $h_s : I \longrightarrow \tilde{X}$ with $h_s(0) = \tilde{x}_0$. Consider the correspondence $s \mapsto h_s(1)$. If s is being deformed then $h_s(1)$ could vary only within a discrete set. Therefore it does not change. Therefore our correspondence is a map $\pi_1(X, x_0) \longrightarrow p^{-1}(x_0)$. Furthermore two loops s_1 and s_2 determine the same element in $p^{-1}(x_0)$ iff the loop $s_1^{-1}s_2$ lifts to $\tilde{X}$ as a *loop* (apriori it could lift as as a path with two different endpoints). Therefore our correspondence gives in fact an injective map

$$\pi_1(X, x_0)/p_*(\pi_1(\tilde{X}, \tilde{x}_0)) \longrightarrow p^{-1}(x_0).$$

It remains to see that the last map is surjective but this follows from the connectedness of $\tilde{X}$: any point in $p^{-1}(x_0)$ can be connected with $\tilde{x}_0$ by a path in $\tilde{X}$ an the projection of this path is a loop in X based at x_0. □

We will now formulate and proof a criterion for lifting arbitrary maps.

Proposition

Suppose that $p : (\tilde{X}, \tilde{x}_0) \to (X, x_0)$ is a covering and $f : (Y, y_0) \to (X, x_0)$ is a (based) map with Y path-connected. Then the lift $\tilde{f} : (Y, y_0) \to (\tilde{X}, \tilde{x}_0)$ exists if and only if $f_*(\pi_1(Y, y_0)) \subset p_*(\pi_1(\tilde{X}, \tilde{x}_0))$. Such a lift is then unique.

Proof.

Note first that the 'only' statement is obvious. For the converse let $y \in Y$ and let γ be a path from y_0 to y. The path $f\gamma$ in X starting at x_0 has a unique lift in $\tilde{X}$ starting at $\tilde{x}_0$. Now set $\tilde{f}(y) := \widetilde{f\gamma}(1)$. Let us show that $\tilde{f}$ is independent of the choice of γ. Indeed, let δ be another path from y_0 to y. Then $(f\gamma)(f\delta)^{-1}$ is a loop h_0 at x_0 whose homotopy class belongs (by the original assumption) to $p_*(\pi_1(\tilde{X}, \tilde{x}_0))$. Therefore there is a loop h_1 at x_0 lifting to a loop $\tilde{h}_1$ at $\tilde{x}_0$ and homotopic to h_0 through a family h_t. Note that as h_0 was itself liftable the statements of independence were obvious, but what we have also suffices.

Indeed, we can lift the homotopy h_t to $\tilde{h}_t$ in $\tilde{X}$. Since $\tilde{h}_1$ is a loop then so is $\tilde{h}_0$ which shows that the loop h_0 does lift, after all, and we are done.

What remains is to show that $\tilde{f}$ is continuous which is left as an exercise. The uniqueness is likewise clear. □

Definition

A covering $p : \tilde{X} \longrightarrow X$ is called regular *if the group $p_*\pi_1(\tilde{X}, \tilde{x}_0)$ is a normal subgroup in $\pi_1(X, x_0)$.*

Remark

The notion of a regular covering is independent of the choice of $x_0 \in X$ and $\tilde{x}_0 \in p^{-1}(x_0)$ (Check it!).

Let us consider a loop s based at the point $x_0 \in X$ and lift it to the path $\tilde{s}$ in $\tilde{X}$ so that $\tilde{s}(0) = \tilde{x}_0$. Then $\tilde{s}$ could be a loop in $\tilde{X}$ (in which case $\tilde{s}(1) = \tilde{x}_0$) or else $\tilde{s}(1) \neq \tilde{x}_0$.

In the latter case $\tilde{s}$ is a path with two different endpoints in $\tilde{X}$. We will call such a path a *nonclosed* path to distinguish it from the *closed* path, i.e. a loop.

Proposition

A covering $p : \tilde{X} \longrightarrow X$ is regular if and only if no loop in X can be the image of both a closed and a nonclosed path in $\tilde{X}$.

Proof.

Suppose that a loop s based at x_0 in X lifts to a closed path $\tilde{s}$ based at $\tilde{x}_0 \in \tilde{X}$ and also to a nonclosed path $\tilde{s}'$ based at $\tilde{x}_0' \in \tilde{X}$. Then clearly s represents an element in $p_*\pi_1(\tilde{X}, \tilde{x}_0)$, but the subgroup $\pi_1(\tilde{X}, \tilde{x}_0')$ in $\pi_1(X, x_0)$ *does not* contain the loop s. Therefore the subgroups $p_*\pi_1(\tilde{X}, \tilde{x}_0')$ and $p_*\pi_1(\tilde{X}, \tilde{x}_0)$ inside $\pi_1(X, x_0)$ are different and p cannot be a regular covering.

Conversely, suppose that any lifting of a loop in X is either a loop or a nonclosed path. Any loop s liftable to a loop based at $\tilde{x}_0$ is also liftable to a loop base at $\tilde{x}_0'$. That shows that the subgroups $p_*\pi_1(\tilde{X}, \tilde{x}_0')$ and $p_*\pi_1(\tilde{X}, \tilde{x}_0)$ inside $\pi_1(X, x_0)$ coincide. In other words the group $p_*\pi_1(\tilde{X}, \tilde{x}_0')$ does not depend on the choice of $\tilde{x}_0 \in p^{-1}(x_0)$. When $\tilde{x}_0$ varies the subgroup $p_*\pi_1(\tilde{X}, \tilde{x}_0')$ gets replaced with its conjugate. We see, that the conjugating does not have effect on $p_*\pi_1(\tilde{X}, \tilde{x}_0')$. In other words $p_*\pi_1(\tilde{X}, \tilde{x}_0')$ is a normal subgroup of $\pi_1(X, x_0)$. $\square$

We are going to study regular coverings more closely. To do it properly we need to discuss *group actions*.

Definition

Let G be a group. We say that G acts on the left on the set X *if there is given a map of sets $f : G \times X \longrightarrow X$. We will denote $f(g,x) \in X$ simply by gx. Moreover the following axioms must be satisfied:*

- *$ex = x$ for any $x \in X$. Here e is the identity element in G.*
- *$(gh)x = g(hx)$ for any $g, h \in G$ and any $x \in X$.*

Exercise

Consider the group $Aut(X)$ consisting of all *permutations of the set X. Show that $Aut(X)$ acts on X. Moreover show that the action of any group G on X is equivalent to a group homomorphism $G \longrightarrow Aut(X)$.*

Remark

One can also define a right action of G on X *as a map $f : X \times G \longrightarrow X$ so that $(x,g) \mapsto xg \in X$. The corresponding axioms are:*

- *$xe = x$ for any $x \in X$ and*
- *$x(gh) = (xg)h$ for any $g, h \in G$ and any $x \in X$.*

Formulate and prove an analogue of the exercise above!

Remark

Furthermore for any left action of G on X there is an associated right action defined by the formula $xg := g^{-1}x$. (Show that this is indeed a right action!). Likewise for any right action the formula $gx := xg^{-1}$ defines a left action. Thus, we can switch back and forth between left and right actions if needed.

Example

Let G be a group. Then G acts on itself by left translations: $(g, h) \mapsto gh$. (Show that this is indeed a left action!) Similarly G acts on itself by left conjugations: $(g, h) \mapsto hgh^{-1}$. (Show that this is a left action!) Similarly we can define the action of G on itself by right translations and right conjugations.

Example

The group $GL(n, k)$ of invertible matrices whose entries belong to the field k acts on the left on the set (actually, a vector space) of vector-columns. Similarly $GL(n, k)$ acts on the right on the set of vector-rows (check this!)

Definition

Suppose that the set X is supplied with an action of the group G. Let us introduce the equivalence relation on X by $x_1 \sim x_2$ if $x_1 = gx_2$ for some $g \in G$. The equivalence class of $x \in X$ is called the orbit *of the element x and will be denoted by $O(x)$. The set of all orbits is called the* quotient of X by the group G*, denoted by X_G or X/G. Clearly there is a map $X \longrightarrow X/G$ which associates to a point $x \in X$ its equivalence class. If there is only one orbit of the action of G on X then the action is called* transitive*.*

Definition

Fix a point $x \in X$. The collection G_x of elements $g \in G$ for which $gx = x$ is called the stabilizer *of x.*

Proposition

Let the action of G on X be transitive. Then there is a bijective correspondence between X and the collections of left cosets G/G_x for any $x \in X$.

Proof.

Let $x' \in X$. Since the action is transitive there exists $g \in G$ such that $gx = x'$. We associate the coset gG_x to the element x. Conversely, we associate to a given coset gG_x the element $gx \in X$. It is straightforward to check that this correspondence is well-defined and one-to-one. □

Question

What is the relationship between stabilizers of different points in X?

It is not hard to see that they are conjugate in G. We have:

Proposition

Suppose that G acts on X transitively and x, x' are elements in X. Let $g \in G$ be such that $gx = x'$. Then $gG_xg^{-1} = G_{x'}$. In other words the subgroups G_x and $G_{x'}$ are conjugate in G via g.

Proof.

Note that $g^{-1}x' = x$. Let $h \in G_x$. Then $ghg^{-1}x' = ghx = gx = x'$. Therefore $ghg^{-1} \in G_{x'}$. Similarly for $h' \in G_{x'}$ we have $g^{-1}h'g \in G_x$. But $h' = gg^{-1}h'gg^{-1}$. Therefore every element in $G_{x'}$ is of the form ghg^{-1} for some $h \in G_x$ $\square$

Remark

By analogy with the theory of coverings we can call an action of G on X regular *if the stabilizer of some point $x \in X$ is a normal subgroup in G, In that case the proposition tells us that stabilizers of all points in the orbit of x will be normal and will coincide. Then there is a one-to one correspondence between $O(x)$ and the quotient group G/G_x.*

We will now make a connection between group actions and the theory of covering spaces. The set X on which a group G acts will now assumed to be a topological space and the action will be continuous in the sense that the action map $f : G \times X \longrightarrow X$ is supposed to be a continuous map. Equivalently any element g acts on the space X by *homeomorphisms*.

Definition

The action of a group G on a topological space X is called free if any point $x \in X$ possesses a neighborhood $U_x \supset \{x\}$ such that $gU_x \bigcap g' U_x = \varnothing$ for $g \neq g'$.

Example

An example of a free action is the action of the group $\mathbb{Z}$ on $\mathbb{R}$ by translations: for $n \in \mathbb{Z}$ and $x \in \mathbb{R}$ we define $nx := x + n$. In that case the set of orbits $\mathbb{R}/\mathbb{Z}$ is clearly homeomorphic to the circle S^1. Another example is the action of the group $\mathbb{Z}/2\mathbb{Z}$ on S^2 by reflections about the center. The corresponding quotient space is called the real projective plane $\mathbb{RP}^2$.

Theorem

Let G act freely on $\tilde{X}$. Then the natural map $\tilde{X} \longrightarrow \tilde{X}/G$ is a regular covering. Conversely every regular covering $\tilde{X} \longrightarrow X$ is of the form $\tilde{X} \longrightarrow \tilde{X}/G$ where G is some group acting freely on $\tilde{X}$.

Proof. Suppose first that G acts freely on $\tilde{X}$. We show that the projection $p : \tilde{X} \longrightarrow \tilde{X}/G$ is a covering. For any point $\tilde{x} \in \tilde{X}$ we choose a neighborhood $U_{\tilde{x}}$ as in the definition of the free group action. Consider the set $V_{\tilde{x}} := p(U_{\tilde{x}}) \in \tilde{X}/G$. Then $p^{-1}(V_{\tilde{x}})$ is by definition the disjoint union of open sets $\{gU_{\tilde{x}}\}, g \in G$. We see that $V_{\tilde{x}}$ is open in $\tilde{X}$ (recall the definition of the quotient topology). Moreover $V_{\tilde{x}}$ is exactly an elementary neighborhood of the point $p(x) \in \tilde{X}/G$ as in the definition of a covering space.
We still need to prove that p is a regular covering. But this is obvious: suppose that a closed and a nonclosed paths are in the preimage of some loop in $\tilde{X}/G$. Then there exists an element of G which maps a closed path into a nonclosed path in $\tilde{X}$. This cannot happen since any element g in G determines a *homeomorphism* of $\tilde{X}$ and therefore closed paths should go to closed paths; the nonclosed paths - to nonclosed paths under the transformation of $\tilde{X}$ determined by g.

Conversely, let us assume that $\tilde{X} \longrightarrow X$ is a regular covering. Take any point $x_0 \in X$ and $\tilde{x}_0 \in \tilde{X}$ for which $p(\tilde{x}_0) = x_0$ and consider the quotient group $G := \pi_1(X, x_0)/p_*\pi_1(\tilde{X}, \tilde{x}_0)$. We claim that G acts on $\tilde{X}$ so that $\tilde{X}/G = X$.

To see that take a loop s in X based at x_0 which represents a coset in $\pi_1(X,x_0)/p_*\pi_1(\tilde{X},\tilde{x}_0)$ and lift it to a path $\tilde{s}$ starting at $\tilde{x}_0$. Let $\tilde{x}_0'$ be the ending point of $\tilde{s}$. Now consider a point $\tilde{x}_1 \in \tilde{X}$ and a path $\tilde{h}$ in $\tilde{X}$ from $\tilde{x}_1$ to $\tilde{x}_0$. The path $\tilde{h}$ projects to the path h in X connecting $x_1 = p(\tilde{x}_1)$ and x_0. Let $\tilde{h}'$ be the path lifting h and starting at $\tilde{x}_0'$ and consider the composite path $\tilde{h}\cdot\tilde{s}\cdot\tilde{h}'$ in $\tilde{X}$. (Draw the picture!)

We define the action of G on $\tilde{x}_1$ by the formula $s\tilde{x}_1 := \tilde{h}\cdot\tilde{s}\cdot\tilde{h}'(1)$. (In other words $s\tilde{x}_1$ is the ending point of the path $\tilde{h}\cdot\tilde{s}\cdot\tilde{h}'$.)

To see that this action does not depend on the choice of the path $\tilde{h}$ consider another path $\tilde{l}$ connecting $\tilde{x}_1$ and $\tilde{x}_0$ and let $\tilde{l}'$ be the path lifting l and starting at $\tilde{x}_0'$. We claim that the ending point of that paths $\tilde{l}'$ and $\tilde{h}'$ in $\tilde{X}$ coincide. Indeed, denoting by l the projection of $\tilde{l}$ to X we see that the loop $h\cdot l$ lifts to a *closed path* $\tilde{h}\cdot\tilde{l}$ in $\tilde{X}$. Since p is a regular covering all liftings of $h\cdot l$ are closed paths; in particular $\tilde{h}\cdot\tilde{l}'$ is a closed path. Our claim is proved.

To finish the proof we need to show that the action of G is free. Take an elementary neighborhood U_x of any point $x\in X$ and consider $p^{-1}(U_x)$. Then $p^{-1}(U_x) = \coprod_{\tilde{x}\in p^{-1}(x)} V_{\tilde{x}}$. Clearly G permutes the neighborhoods $V_{\tilde{x}} \subset \tilde{X}$ and since these are disjoint we see that G indeed acts freely. QED.

Definition

A covering $p : \tilde{X} \longrightarrow X$ *is called* universal *if* $\tilde{X}$ *is simply-connected.*

Remark

The universal covering is always regular. (Check it!).

Corollary

Suppose that a group G *acts freely on a simply-connected space* X*. Then* $\pi_1(X/G) \cong G$*.*

Thus, in order to find the fundamental group of a space X it suffices to find a universal covering of X. This covering is always determined by a free action of some group G, and this group is isomorphic to $\pi_1(X)$. Thus we have a method for computing fundamental groups of spaces.

Example

To illustrate the force of this method consider again the case $X = S^1$*. The group* $\mathbb{Z}$ *acts on* $\mathbb{R}$ *by translations and the canonical map* $\mathbb{R} \longrightarrow \mathbb{R}/\mathbb{Z} \cong S^1$ *is clearly a universal covering. Therefore* $\pi_1(S^1) \cong \mathbb{Z}$*. Another example: the group* $\mathbb{Z}/2\mathbb{Z}$ *acts freely on* S^2 *and the quotient* $S^2/(\mathbb{Z}/2\mathbb{Z}) \cong \mathbb{RP}^2$*, the real projective space. Since* S^2 *is simply-connected we conclude for the real projective plane* $\pi_1(\mathbb{RP}^2) \cong \mathbb{Z}/2\mathbb{Z}$*.*

Exercise

Construct a universal covering over a two-dimensional torus $T^2 = S^1 \times S^1$ and compute $\pi_1(T^2)$.

Definition

Let S be a set. The free group *$F(S)$ on S is the group whose elements are the formal symbols of the form $s_1^{i_1} s_2^{i_2} \dots s_n^{i_n}$. Here i_k are integers, possibly negative. These symbols are called* words *in the alphabet $\{s_i\}, s_i \in S$. The formal symbols s_i are called the* generators*. The multiplication of two words $s_1^{i_1} s_2^{i_2} \dots s_n^{i_n}$ and $h_1^{j_1} h_2^{j_2} \dots h_k^{j_k}$ is the word $s_1^{i_1} s_2^{i_2} \dots s_n^{i_n} h_1^{j_1} h_2^{j_2} \dots h_k^{j_k}$ obtained by concatenation of $s_1^{i_1} s_2^{i_2} \dots s_n^{i_n}$ and $h_1^{j_1} h_2^{j_2} \dots h_k^{j_k}$. The unit e is by definition, the empty word. The* cancellation rule *associates to a word of the form $s_1^{i_1} s_2^{i_2} \dots s_n^{i_n} s_n^{k} s_{n+1}^{i_{n+1}} \dots s_m^{i_m}$ the word of the form $s_1^{i_1} s_2^{i_2} \dots s_n^{i_n+k} s_{n+1}^{i_{n+1}} \dots s_m^{i_m}$. Two words are considered equal if one could be obtained from the other by a finite number of cancellations. (Thus $ss^{-1} = e$, for example.)*

Example

If $S = \{s_1\}$ consists of one element, the corresponding free group is just the group consisting of symbols $s_1^n, n \in \mathbb{Z}$. Clearly this is just the group of integers $\mathbb{Z}$, in particular, it is abelian. The free group on two generators $S = \{s_1, s_2\}$ is already highly nontrivial and nonabelian.

Remark

Why is $F(S)$ a group? The multiplication is clearly associative, and the empty word is the left and right unit for the multiplication. The inverse for the word $s_1^{i_1} s_2^{i_2} \dots s_n^{i_n}$ is the word $s_n^{-i_n} s_{n-1}^{-i_{n-1}} \dots s_1^{-i_1}$. (Check this!)

Proposition

Let G be a group. Then there exists a free group F and an epimorphism $F \longrightarrow G$.

Proof.

Let S be a set whose elements are in one-to-one correspondence with the elements of G. (In other words, S is just G, only we forget that G is a group and consider it as just a set.) The element of S corresponding to $g \in G$ will be denoted by s_g. Consider the free group $F(S)$ and let $f : F(S) \longrightarrow G$ be the map that associates to a word $s_{g_1}^{i_1} s_{g_2}^{i_2} \dots s_{g_n}^{i_n}$ the element $g_1^{i_1} g_2^{i_2} \dots g_n^{i_n} \in G$ (the multiplication in the last term is taken in the group G). Then f is clearly a surjective homomorphism of $F(S)$ onto G. $\square$

Remark

The homomorphism f constructed in the previous proposition is 'universal', that is it works for all groups G uniformly. However, it is very 'wasteful' in the sense that typically there exists a free group with a much smaller set of generators which surjects on to a given group G. For example, if $G = \mathbb{Z}$, then G itself is free, so we could just take for f the identity homomorphism $\mathbb{Z} \longrightarrow \mathbb{Z}$. By contrast, the 'universal' homomorphism constructed before involves a free group on countably many generators.

Definition

Let G be a group and $f : F \longrightarrow G$ be a surjective homomorphism where $F = F(S)$ is a free group on the set S. Let H be the kernel of f. Then S is called the set of generators *for G and H the subgroup of* relations*. In that case we say that G is* defined *by a set of generators and relations. Note that $G \cong F(S)/H$.*

Remark

It is very hard to determine, in general, whether two sets of generators and relations determine the same group. When working with generators and relations one usually tries to find a 'small' presentation, i.e. such that the number of generators and the size of the subgroup of relations are as small as possible.

Exercise

Show that the group with two generators f, g subject to the relation $fg = gf$ is isomorphic to the free abelian group $\mathbb{Z} \times \mathbb{Z}$. More precisely, we consider the free group $F(2)$ with two generators f and g and the normal subgroup H in $F(2)$ generated by the element $fgf^{-1}g^{-1}$. You need to show that the quotient $F(2)/H$ is isomorphic to $\mathbb{Z} \times \mathbb{Z}$.
Hint: using the relation $fg = gf$ any word in f and g can be reduced to a canonical form $f^i g^j$. Show that the canonical form is unique, i.e. if $f^i g^j = f^k g^l$ then $i = k$ and $j = l$. On the other hand the group $\mathbb{Z} \times \mathbb{Z}$ is none other then the set of pairs $(i, j), i, j \in \mathbb{Z}$ with the multiplication law $(i, j)(k, l) = (i + k, j + l)$. It follows that $F(2)/H \cong \mathbb{Z} \times \mathbb{Z}$.

Exercise

Find the set of generators and relations for the cyclic group $\mathbb{Z}/n\mathbb{Z}$.

Using the developed technology we will compute fundamental groups of some more spaces.

Construction. Consider the two dimensional plane $\mathbb{R}^2$, we will identify it with the complex plane $\mathbb{C}$ whenever convenient. (Recall that the point $z = x + iy \in \mathbb{C}$ corresponds to the point $(x, y) \in \mathbb{R}^2$.) Let X be the square in $\mathbb{R}^2$ bounded by the lines $x = 0, y = 0, x = 1, y = 1$. Let us introduce the equivalence relation on X by declaring

1. $(0, y) \sim (1, y)$ for $0 \leq y \leq 1$;
2. $(0, y) \sim (1, y)$ for $0 \leq y \leq 1$ and $(x, 0) \sim (x, 1)$ for $0 \leq x \leq 1$;
3. $(0, y) \sim (1, y - 1)$ for $0 \leq y \leq 1$;
4. $(x, 0) \sim (x - 1, 1)$ for $0 \leq x \leq 1$ and $(0, y) \sim (1, y - 1)$ for $0 \leq y \leq 1$;
5. $(x, 0) \sim (x, 1)$ for $0 \leq x \leq 1$ and $(0, y) \sim (1, y - 1)$ for $0 \leq y \leq 1$.

Now consider the space $X/\sim$. It is clear that in the case
(1) $X/\sim$ is a cylinder
(2) $X/\sim$ is a torus
(3) $X/\sim$ is a Möbius strip
(4) $X/\sim$ is homeomorphic to $\mathbb{RP}^2$ (why?)
(5) $X/\sim$ is called the *Klein bottle*. It will be denoted by K.

We are interested in $\pi_1(X/\sim)$. In fact we already know the answer in all cases save (5) (why?). So let us work out case (5).

Consider the following transformation of $\mathbb{R}^2 = \mathbb{C}$:

$$\phi(z) = z + i;$$
$$\psi(z) = \bar{z} + 1.$$

Here $\bar{z}$ denotes complex conjugation. Let G denote the subgroup generated by ϕ, ψ in the group of all transformations of $\mathbb{C}$. We claim that

1. the only relation in G is of the form $\phi\psi\phi = \psi$. (More precisely, G is isomorphic to the quotient of the free group on generators ϕ, ψ by the normal subgroup generated by the element $\phi\psi\phi\psi^{-1}$).
2. G acts freely on $\mathbb{C}$ and
3. $\mathbb{C}/G$ is homeomorphic to K.

This will give us the complete description of $\pi_1 K$. Let us prove the claims (1)-(3).

1. We have

   $$\phi\psi\phi(z) = \phi\psi(z+i) = \phi(\overline{z+i}+1) = \phi(\bar{z}-i+1) = \bar{z}-i+1+i = \bar{z}+1 = \psi(z).$$

   Next we need to check that *all relations in G are consequences of* $\phi\psi\phi = \psi$. Rewriting this relation as $\phi\psi = \psi\phi^{-1}$ and using it to permute ϕ past ψ we see that any word $\phi^{i_1}\psi^{j_1}\phi^{i_2}\psi^{j_2}\ldots\phi^{i_k}\psi^{j_k}$ could be reduced to the form $\phi^i\psi^j$. We will call such a form *canonical*.

We need to check that two canonical forms are equal in G iff they are identical, i.e. if $\phi^i\psi^j = \phi^k\psi^l$ then $i = k$ and $j = l$. This is reduced to showing that if $\phi^i\psi^j = e$ then $i = j = 0$. Next, if $j \neq 0$ then applying the transformation $\phi^i\psi^j$ to any $z \in \mathbb{C}$ we see that $Re(\phi^i\psi^j(z)) = Re(z) + j$ and $\phi^i\psi^j$ could be the identity transformation iff $j = 0$ (ϕ does not change $Re(z)$!). Further, clearly, $\phi^i = e$ iff $i = 0$.

1. For any $z \in \mathbb{C}$ we need to choose a neighborhood $U_z \supset \{z\}$ such that $\phi^i\psi^j(U_z) \cap \phi^k\psi^l(U_z) = \varnothing$ if $(i,j) \neq (k,l)$. Take U_z to be the ball around z of radius $\frac{1}{4}$. If $j \neq l$ then using the fact that $Re(\phi^i\psi^j(z)) = Re(z) + j$ and $Re(\phi^k\psi^l(z)) = Re(z) + l$ we see that $\phi^i\psi^j(U_z) \cap \phi^k\psi^l(U_z) = \varnothing$. Complete the argument in the case $j = l$.
2. Note that the orbit of any point $z \in \mathbb{C}$ under the action of G has a representative inside the square X. Moreover the points in the interior of X never lie in the same orbit and the points on the boundary of X are identified precisely as in the definition of K. That completes our calculation of $\pi_1 K$.

Our next example is concerned with the fundamental group of the *wedge* of two copies of S^1, i.e. the figure eight "8".

Remark

A wedge *is the analogue in $\mathcal{T}op_*$ of the disjoint union construction in $\mathcal{T}op$. Namely, for two pointed spaces (X, x_0) and (Y, y_0) their* wedge *(or* bouquet*) is the space $X \vee Y$ obtained from $X \coprod Y$ by identifying the points $x_0 \in X$ with $y_0 \in Y$. Thus, $X \vee Y$ is a pointed space whose basepoint is $x_0 = y_0$.*

Proposition

$\pi_1(S^1 \vee S^1)$ is the free group on two generators.

Proof. Let $F = F(g, h)$ be the free group with generators g and h. We will construct a free action of F on a contractible space so that the quotient is homeomorphic to $S^1 \vee S^1$. This will prove the proposition. The construction will be done step by step. First, we consider the space T_1 which is by definition the figure $+$ (a cross). Next we attach to each outer vertex of $+$ (there are four of them) three new edges so that these vertices become centers of four new crosses. Denote the obtained figure by T_2. All edges of T_2 are either horizontal or vertical. (Draw the picture!)

Note that $T_1 \subset T_2$. Repeating this procedure we construct the sequence of graphs $T_1 \subset T_2 \subset T_3 \subset \dots$
Denote by T the union of all T_n's. Thus, T is an infinite graph with a marking on the edges. We make T into a metric space by requiring each edge of T to have length 1. In particular, T is a topological space. We claim

1. T is contractible as a topological space.
2. $F(g, h)$ acts freely on T.

To see (1) it suffices to construct a homotopy $f_t : T \longrightarrow T$ connecting the identity map on T with the map collapsing T onto its center. Take $x \in T$ and consider a path of minimal length connecting x with the center of T.
Considering it as a map $\gamma_x : I \longrightarrow T$ denote by $f_t : T \longrightarrow T$ the map given by $f_t(x) = \gamma_x(t)$. Then clearly f_0 is the identity map whereas f_1 is mapping T onto its center.
For (2) define the action of $F(g, h)$ on T as follows. The element g acts as shift upwards by the length 1 whereas h acts by a unit shift to the right. (Then, necessarily, g^{-1} acts as a downward shift while h^{-1} is a shift to the left). This is clearly a group action.

To see that this action is free take a point $x \in T$. Suppose first that x is a vertex of T. Taking a ball U_x in T of radius $\frac{1}{4}$ around x we see that the images of U_x under the action of any word $g^i h^k g^l \dots$ are balls of radius $\frac{1}{4}$ around vertices of T. In particular they are disjoint. The case when x is an internal point of an edge is considered similarly. Therefore the action is free.

What is the quotient of T with respect to the action of $F(g, h)$?
The quotient is, by definition, the set of orbits. Now, any point $x \in T$ has a representative inside T_1 and, furthermore, the internal points of T_1 never lie in the same orbit (why?). The outer vertices of T_1 do lie in the same orbit and, therefore, $T/F(g, h)$ is just the quotient of T_1 by the equivalence relation identifying the outer edges of T_1. The resulting space is clearly homeomorphic to $S^1 \vee S^1$. QED.

There is a rich family of coverings over the space $S^1 \vee S^1$. We will describe some of them.

Example

Consider the space X obtained by attaching circles to integer points of the real line $\mathbb{R}^1$. Formally, X is the union of $\mathbb{R}^1$ and an infinite number of S^1's modulo the equivalence relation identifying the point $n \in \mathbb{R}^1$ with the basepoint of the nth copy of S^1.
Consider the covering $p : X \longrightarrow S^1 \vee S^1$ which maps every copy of S^1 in X onto the first wedge summand of $S^1 \vee S^1$ and every interval $[n, n+1]$ in X onto the second wedge summand of X. Then, clearly, p is indeed a covering with infinitely many sheets.
Moreover, it is easy to see that X is homotopy equivalent to the wedge of countably many copies of S^1 (why?). From this we deduce that $\pi_1(X) = F(\infty)$, the free group on countably many generators. Since $p_* : \pi_1(X) \longrightarrow \pi_1(S^1 \vee S^1)$ is an injective homomorphism we see that $F(\infty)$ can be embedded as a subgroup in $F(2)$, the free group on two generators (recall that $\pi_1(S^1 \vee S^1) = F(2)$).

Example

Consider the unit circle $S^1 \in \mathbb{R}^2$ and attach a copy of the circle to the points with coordinates $(0, 1), (\cos 2\pi/3, \sin 2\pi/3), (\cos 4\pi/3, \sin 4\pi/3)$. The resulting space denoted by X has the homotopy type of $S^1 \vee S^1 \vee S^1$ (why?).

Exercise

Show that $\pi_1(X) = F(3)$, the free group on three generators.

The group $\mathbb{Z}/3\mathbb{Z}$ acts on X by rotations by $2\pi/3$ and the quotient space is clearly homotopy equivalent to $S^1 \vee S^1$. On the level of fundamental groups this gives a monomorphism of $F(3)$ into $F(2)$.
Note that all of the covering spaces considered so far were *regular*. These are the most important ones and also easiest to construct. There exist, however, coverings which are not regular.

Exercise

Construct an example of a nonregular covering. Hint: take $S^1 \vee S^1$ for the base space of the covering.

Remark

Note that if $p : \tilde{X} \longrightarrow X$ is a two-sheeted covering then the subgroup $p_*\pi_1(\tilde{X})$ inside $\pi_1(X)$ has index two and, therefore, is normal. That means that a nonregular covering has to be at least three-sheeted.

Question

How to classify *covering spaces over a given base space X?*

Definition

Let $p_1 : (\tilde{X}, \tilde{x}_0) \to (X, x_0)$ and $p_2 : (\overline{X}, \overline{x}_0) \to (X, x_0)$ be two coverings of (X, x_0). A basepoint preserving map between these two coverings is a (based) map $f : (\tilde{X}, \tilde{x}_0) \to (\overline{X}, \overline{x}_0)$ such that the following diagram is commutative:

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

If the map f is not required to be based then we simply have a map between two coverings. Finally, if f is a homeomorphism (basepointed or not depending on the basepointedness of f) the two covering are said to be isomorphic.

Remark

We see that there are two categories associated with a spaces X; the objects in both categories are coverings over X and objects are maps of coverings, based or not. We will mostly concentrate on the category of based covering maps from now on; it will be denoted by $Cov(X)$.

Definition

Let G be a group and let $\mathcal{C}(G)$ be the category whose objects are subgroups of G and morphisms are inclusions of one subgroup into another. It is clear how to compose morphisms and that we indeed obtain a category.

Now we construct a functor $F : Cov(X) \to \mathcal{C}(\pi_1(X, x_0))$ where X is a path-connected space; it is clear that for different choices of the basepoint x_0 the categories $\mathcal{C}(\pi_1(X, x_0))$ are equivalent (even isomorphic) and we will leave as an exercise the analysis of the dependence of F on x_0.

Namely, F sends a covering $p : (\tilde{X}, \tilde{x}_0) \to (X, x_0)$ to the subgroup $p_*(\pi_1(\tilde{X}_0, \tilde{x}_0)) \subset \pi_1(X, x_0)$. It is immediate to define F on morphisms and check that F is indeed a functor.

Theorem

The functor F is an equivalence of categories.

The proof of this theorem will rely on several results and constructions.

Proposition

The functor F is fully faithful.

Proof.

Let $p_1 : (\tilde{X}, \tilde{x}_0) \to (X, x_0)$ and $p_2 : (\overline{X}, \overline{x}_0) \to (X, x_0)$ be two coverings of (X, x_0) and $f : {p_1}_*(\pi_1(\tilde{X}, \tilde{x}_0)) \to {p_2}_*(\pi_1(\overline{X}, \overline{x}_0))$ be the corresponding inclusion. Then by the lifting criterion there exists a map of coverings $\tilde{f} : (\tilde{X}, \tilde{x}_0) \to (\overline{X}, \overline{x}_0)$ which induces f. Such a map is unique which shows that F induces an 1-1 map on the set of morphisms of the corresponding category as required. □

In order to show that F is surjective on morphisms we have to build, for any subgroup of $\pi_1(X, x_0)$ a corresponding covering. Let us start with the trivial subgroup. Recall that in that case the corresponding covering is called *universal*.

Proposition

For any (path-connected, locally simply connected) space (X, x_0) a universal covering $(\tilde{X}, \tilde{x}_0)$ exists.

Proof. Let $\tilde{X}$ be defined as the set of homotopy classes of paths γ in X starting at x_0. As usual, we consider homotopies fixing the endpoint $\gamma(0)$ and $\gamma(1)$. The map $p : \tilde{X} \to X$ associates to a path γ the point $\gamma(1)$. It is clear that p is surjective. To define a topology on $\tilde{X}$ consider the collection of open sets $U \subset X$ such that U is simply-connected. (Clearly such collection forms a base for the topology in X).

Now let

$$U_{[\gamma]} := \{[\gamma\eta] | \eta \text{ is a path in } U \text{ with } \eta(0) = \gamma(1).\}$$

As the notation indicates, $U_{[\gamma]}$ only depends on $[\gamma]$. Note also that $p : U_{[\gamma]} \to U$ is surjective since U is path-connected and injective since different choices of η joining $\gamma(1)$ to a fixed $x \in U$ are all homotopic, the set U being simply-connected.

Next, it is possible to show that the collection $U_{[\gamma]}$ forms a base of a topology in $\tilde{X}$ (we skip this verification) and that a map p is a local homeomorphism (this is more or less clear). Thus, p is a covering.

It remains to show that $\tilde{X}$ is simply-connected. For a point $[\gamma] \in \tilde{X}$ let γ_t be the path in X that equals γ on $[0, t]$ and is constant on $[t, 1]$. Then the function $t \mapsto [\gamma_t]$ is a path in $\tilde{X}$ lifting γ that start at $[x_0]$, the homotopy class of a constant path at x_0, and ends at $[\gamma]$. Since $[\gamma]$ was arbitrary, this shows that $\tilde{X}$ is path-connected. To show that it is simply-connected it suffices to show that $p_*(\pi_1(\tilde{X}, [x_0])$ is trivial inside $\pi_1(X, x_0)$. Elements in the image of p_* are represented by loops γ at x_0 that lift to loops in $\tilde{X}$ starting at $[x_0]$. We saw that the path $t \mapsto [\gamma_t]$ lifts γ starting at $[x_0]$ and for this lifted path to be a loop means that $[\gamma_1] = [x_0]$. Since $\gamma_1 = \gamma$ this means that $[\gamma] = [x_0]$ so γ is nullhomotopic as required. QED.

Finally, the general case (from which the surjectivity of F on isomorphism classes of objects follows).

Proposition

Let X be a path-connected and locally simply-connected space. Then for every subgroup $H \subset \pi_1(X, x_0)$ there is a covering $p : X_H \to X$ such that $p_*(\pi_1(X_H, x)) \cong H$ for a suitably chosen basepoint $x \in X_H$.

Proof.

For points $[\gamma], [\delta]$ in the universal covering $\tilde{X}$ constructed above, set $[\gamma] \sim [\delta]$ if $\gamma(1) = \delta(1)$ and $[\gamma\delta^{-1}] \in H$. It is easy to see that this is an equivalence relation: it is reflexive since H contains the identity element, symmetric since H is closed under inverses and transitive since H is closed under multiplication. Now let $X_H := \tilde{X}/\sim$. Note that if $\gamma(1) = \delta(1)$ then $[\gamma] \sim [\delta]$ if and only if $[\gamma\eta] \sim [\delta\eta]$. That means that if any two points in basic neighborhoods $U_{[\gamma]}$ and $U_{[\delta]}$ are identified then their whole neighborhoods are identified. It follows that the natural projection $X_H \to X$ is a covering.
If we choose for the basepoint x in X_H the equivalence class of the constant path at x_0 then the image of $p_* : \pi_1(X_H, x) \to \pi_1(X, x_0)$ is exactly H.This is because for a loop γ in X based at x_0 its lift in $\tilde{X}$ starting at x ends at $[\gamma]$, so the image of this lifted path in X_H is a loop if and only if $[\gamma] \in H$. $\square$

There are essentially two methods of computing the fundamental group of a topological space:

- the method of *covering spaces*.
- the method of *decomposing* a topological space into unions of simpler subspaces for which the fundamental groups are known.

The second method is described in *Van Kampen's Theorem*. We will first need some preparations from group theory:

Definition

1. *Let G and H be groups; then their* free product *is the group $G * H$ whose elements are words $g_1 h_1 g_2 h_2 \dots$ (or $h_1 g_1 h_2 g_2 \dots$) of arbitrary finite length modulo the relations already present in G and H. The group operation is concatenation of words and the empty word is the identity element e for the group operation.*
2. *More generally, let K be a third group together with group homomorphisms $i : K \to G$ and $j : K \to H$. Then the* free product of G and H over K *(or their* amalgamated product over K*) is the quotient of $G * H$ by the following relation: for any two words A and B in $G * H$ and $k \in K$ we have*
$$Ai(k)B = Aj(k)B.$$
We will denote this group by $G *_K H$.

Remark

An amalgamated product of groups can also be considered as a pushout *in the category of groups. It is clear that the notion of amalgamated product of groups could be defined for a collection of groups.*

Now suppose that a space X is the union of path-connected open sets A_α containing the basepoint $x_0 \in X$. We denote by $i_{\alpha\beta}$ the group homomorphism $i_{\alpha\beta} : \pi_1(A_\alpha \bigcap A_\beta) \to \pi_1(A_\alpha)$ induced by the inclusion $A_\alpha \bigcap A_\beta \hookrightarrow A_\alpha$. Then we can formulate the main result of this section.

Theorem (Van Kampen Theorem)

Adopting the above notation suppose that each intersection $A_\alpha \bigcap A_\beta$ is path-connected. Then the natural homomorphism $\Phi : *_\alpha \pi_1(A_\alpha) \to \pi_1(X)$ is surjective. If, further, every triple intersection $A_\alpha \bigcap A_\beta \bigcap A_\gamma$ is also path-connected then, the kernel of Φ is the normal subgroup N generated by all elements of the form $i_{\alpha\beta}(\omega) i_{\beta\alpha}(\omega)^{-1}$ for $\omega \in \pi_1(A_\alpha \bigcap A_\beta)$ and so Φ induces an isomorphism $\pi_1(X) \cong *_\alpha \pi_1(A_\alpha)/N$.

In particular, if X is a union of only two such sets A and B as above with $C := A \bigcap B$ then $\pi_1(X) \cong \pi_1(A) *_{\pi_1(C)} \pi_1(B)$. This is the most useful special case of the van Kampen theorem.

Proof. We first prove the surjectivity statement. Given a loop $f : I \to X$ based at $x_0 \in X$ we choose a partition $0 = s_0 < s_1 < \ldots < s_m = 1$ of I such that every subinterval $[s_i, s_{i+1}]$ is mapped to a single A_α by f.
Denote the A_α containing $f([s_i, s_{i+1}])$ by A_i and let f_i be the corresponding restriction of the path f. It follows that f is the composition $f_1 \ldots f_m$. Since $A_i \bigcap A_{i+1}$ is path-connected we can choose a path g_i from x_0 to $f(s_i)$ lying in $A_i \bigcap A_{i+1}$. Consider the loop

$$(f_1 g_1^{-1})(g_1 f_2 g_2^{-1}) \ldots (g_{m-1} f_m).$$

It is clear that this loop is homotopic to f and is a composition of loops lying in separate A_i. Hence $[f]$ is in the image of Φ as required.
Now the identification of the kernel. For an element $[f] \in \pi_1(X)$ consider its representation as a product of loops $[f_1] \ldots [f_k]$ such that every loop f_i is a loop in some A_α.
We will call this a *factorization* of f. It is, thus, a word in the free product of $\pi_1(A_\alpha)$'s that is mapped to $[f]$ via Φ. We showed above that each homotopy class of a loop in X has a factorization. To describe the kernel of Φ is tantamount to describing possible factorizations of a given loop of X.

We will call two factorizations *equivalent* if they are related by two sorts of moves or their inverses:

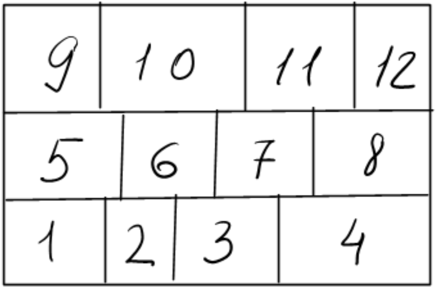
- Combine adjacent terms $[f_i][f_{i+1}]$ into a single term $[f_i f_{i+1}]$ if f_i and f_{i+1} lie in the same space A_α.
- regard the term $[f_i] \in \pi_1(A_\alpha)$ as lying in $\pi_1(A_\beta)$ if f_i is a loop in $A_\alpha \bigcap A_\beta$.

It is clear that two factorizations are equivalent if and only if they determine the same element in $*_\alpha \pi_1(A_\alpha)/N$. Therefore, we are reduced to showing that any two factorizations of a loop f in X are equivalent.

So let $[f_1] \dots [f_k]$ and $[f_1'] \dots [f_l']$ be two factorizations of f. Then the corresponding compositions of paths are homotopic via some homotopy $F : I \times I \to X$.

Consider partitions $0 = s_0 < s_1 < \dots < s_m = 1$ and $0 = t_0 < \dots < t_n = 1$ such that each rectangle $R_{ij} = [s_{i-1}, s_i] \times [t_{j-1}, t_j]$ is mapped by F to a single A_α which we relabel A_{ij}.

We may also assume there are at least three rows of rectangles so we can do this perturbation just on the rectangles in the intermediate rows, not on the top and bottom ones. We further relabel the small rectangles as $R_1, \ldots, R_{mn}$ as on the following picture.



We will represent loops in X as paths in $I \times I$ running from left to right edges. Let γ_r be the path separating the first r rectangles from the remaining rectangles. Thus, γ_0 is the bottom edge of $I \times I$ while γ_{mn} is its top edge.

The idea is that as we push from γ_r to γ_{r+1} we obtain equivalent factorizations. This idea needs to be massaged, however, since each γ_r does not quite determine a factorization; it needs to be further refined.

We will call the corners of the R_i's *vertices*. For each vertex v with $F(v) \neq x_0$ let g_v be the path from x_0 to $F(v)$. We can choose g_v to lie in the intersection of the two or three A_{ij}'s corresponding to the R_i's containing v since we assumed that the intersections of any two or three of our open sets in the cover are path-connected. Let us insert the paths $g_v^{-1} g_v$ into $F|_{\gamma_r}$ at appropriate vertices as in the proof of surjectivity of Φ. This will give a factorization of $F|_{\gamma_r}$.

The factorizations associated to successive paths γ_r and γ_{r+1} are equivalent since pushing γ_r across R_{r+1} to γ_{r+1} changes the path $F|_{\gamma_r}$ to $F|_{\gamma_{r+1}}$ by a homotopy within A_{ij} corresponding to R_{r+1} and we can choose this A_{ij} for all the segments of γ_{r+1} in R_{r+1}.

We can arrange that the factorization associated to γ_0 is equivalent to $[f_1] \dots [f_k]$ by choosing the path g_v for each vertex v in the lower edge of $I \times I$ to lie not only in the two A_{ij}'s corresponding to the two adjacent small rectangles containing v, but also in the open set A_α for the map f_i containing v in its domain.

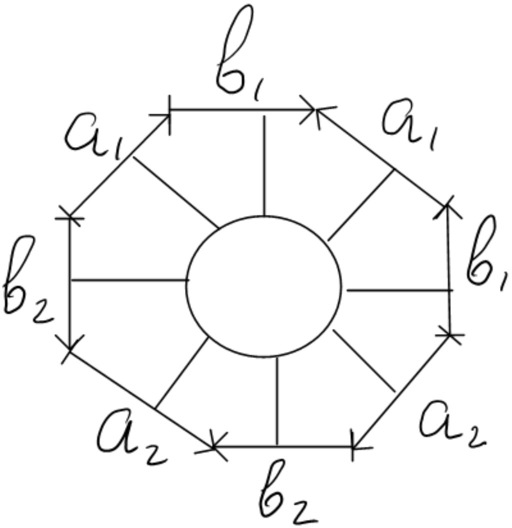
In the case v is the common endpoint of two such domains we have $F(v) = x_0$ and there is no need to insert g_v.

In a similar fashion we show that the factorization associated with the restriction of F onto the upper edge of $I \times I$ is equivalent to that of the factorization associated to γ_0 is equivalent to $[f_1'] \dots [f_l']$.

This concludes the proof of the second part of the van Kampen theorem. QED.

Example

- It follows immediately from the van Kampen theorem that $\pi_1(X \bigvee Y) \cong \pi_1(X) * \pi_1(Y)$ for any two pointed spaces X and Y; in particular we recover the result that the fundamental group of the wedge of two circles is the free group on two generators, i. e. $\pi_1(S^1 \bigvee S^1) \cong \mathbb{Z} * \mathbb{Z}$
- Let X be the *oriented surface of genus g* which is homeomorphic to a sphere with g handles. Recall that X is obtained from $4g$-gon by identifying pairs of edges according to the following picture. Choose the basepoint to be one of the vertices (independent on the choice of the basepoint).


b_1
a_1
a_1
b_2
b_1
a_2
a_2
b_2

Let us cut a disc D in the center of the $4g$-gon. The resulting surface with boundary will be homotopy equivalent to the wedge of $2g$ spheres. The required homotopy is obtained by flowing the boundary of the disc by the radial rays onto the boundary of the $4g$-gon. It follows that $\pi_1(X \setminus D) \cong \langle a_1, b_1, \dots a_g, b_g \rangle$, the free group on the boundary loops $a_1, b_1, \dots a_g, b_g$. It is, furthermore, clear, that the boundary of D viewed as a loop in $X \setminus D$ is homotopic to $a_1 b_1 a_1^{-1} b_1^{-1} \dots a_g b_g a_g^{-1} b_g^{-1}$. (Strictly speaking, this loop is not based at the boundary of the $4g$-gon but it is clear that it does not matter – we could have started with D touching the boundary of the $4g$-gon).

We are now in a position to apply the van Kampen theorem: one open set is $X \setminus D$, the other is a slight thickening of D, their intersection is an annulus (having the same fundamental group as S^1, i.e. $\mathbb{Z}$.) It follows that $\pi_1(X) \cong \langle a_1, b_1, \dots a_g, b_g \rangle / a_1 b_1 a_1^{-1} b_1^{-1} \dots a_g b_g a_g^{-1} b_g^{-1}$; i.e. the free group on $2g$ generators subject to one relation $a_1 b_1 a_1^{-1} b_1^{-1} \dots a_g b_g a_g^{-1} b_g^{-1} = 1$.

Exercise

Use the van Kampen theorem to compute the fundamental group of an unorientable surface obtained by cutting a small disc in an oriented surface of genus g and glueing back in the Möbius strip (whose boundary is a circle S^1).

We will now introduce and study another extremely important algebraic invariant of topological spaces: its *singular homology*. By contrast with homotopy groups which are relatively easy to define yet hard to compute the homology groups are eminently computable.

Definition

A subset A of the space $\mathbb{R}^n$ is called affine *if, for any pair of distinct points $x, x' \in A$ the line passing through x, x' is contained in A.*

Remark

Recall that $A \subset \mathbb{R}^n$ is convex if together with any pair $x, x' \in A$ the straight segment connecting x and x' lies in A. Clearly affine sets in $\mathbb{R}^n$ are convex. The intersection of any number of affine (convex) sets is affine (convex).

Thus, it makes sense to speak about the affine (convex) set in $\mathbb{R}^n$ *spanned* by a subset $X \subset \mathbb{R}^n$, namely, the intersection of all affine (convex) sets in $\mathbb{R}^n$ containing X. We will denote by $[X]$ the convex set spanned by X and by $[X]_a$ the affine set spanned by X.

Definition

An affine combination *of points $p_0, \ldots, p_m \in \mathbb{R}^n$ is a point $x := t_0 p_0 + \ldots + t_m p_m$ where $\sum_{i=0}^m t_i = 1$. A* convex combination *is an affine combination for which $t_i \geq 0$ for all i.*

A convex combination of x, x' has the form $tx + (1-t)x'$ for $0 \leq t \leq 1$.

Proposition

If $p_0, \ldots, p_m \in \mathbb{R}^n$ then $[p_0, \ldots, p_m]$ is the set of all convex combinations of $p_0, \ldots, p_m$.

Proof.

Let S denote the set of all convex combinations of $p_0, \ldots, p_m$. To show that $[p_0, \ldots, p_m] \subset S$ we need to check that S is a convex set containing each point $p_i, i = 0, \ldots, m$. If we set $t_i = 1$ and the other $t_j = 0$ then we see that $p_i \in S$ for all i. Take $\alpha = \sum a_i p_i$ and $\beta = \sum b_i p_i$ be convex combinations. Then $t\alpha + (1-t)\beta = \sum(ta_i + (1-t)b_i)p_i$ is also a convex combination (check this!) and hence lies in S.
Next we have to show that $S \subset [p_0, \ldots, p_m]$. Let X be any convex set containing $p_0, \ldots, p_m$; we will show that $S \subset X$ by induction on $m \geq 0$. The case $m = 0$ is obvious, so take $m > 0$ and consider $p = \sum_{i=0}^m t_i p_i$ with $t_i \geq 0$ and $\sum_{i=0}^m t_i = 1$. We may assume that $t_0 \neq 1$ (otherwise $p = p_0 \in X$). Let

$$q := \left(\frac{t_1}{1-t_0}\right) p_1 + \ldots + \left(\frac{t_m}{1-t_0}\right) p_m.$$

Then $q \in X$ by the inductive assumption and so $p = t_0 p_0 + (1-t_0)q \in X$, because X is convex. $\square$

Definition

An ordered set of points $p_0, \ldots, p_m \in \mathbb{R}^n$ is affine independent *if the vectors $p_1 - p_0, \ldots p_m - p_0$ are linearly independent in the vector space $\mathbb{R}^n$.*

Remark

1. *Any one-point set $\{p_0\}$ is affine independent;*
2. *a set $\{p_0, p_1\}$ is affine independent if $p_0 \neq p_1$;*
3. *a set $\{p_0, p_1, p_2\}$ is affine independent if it is not collinear;*
4. *a set $\{p_0, p_1, p_2, p_3\}$ is affine independent if it is not coplanar.*

Proposition

The following conditions on an ordered set of points $p_0, \ldots, p_m \in \mathbb{R}^n$ are equivalent:

1. *$\{p_0, \ldots, p_m\}$ is affine independent;*
2. *if $s_0, \ldots, s_m \in \mathbb{R}$ satisfy $\sum_{i=0}^m s_i p_i = 0$ and $\sum_{i=0}^m s_i = 0$, then $s_0 = s_1 = \ldots = s_m = 0$;*
3. *each $x \in [p_0, \ldots, p_m]_a$ has a unique expression as an affine combination:*

$$x = \sum_{i=0}^{m} t_i p_i \text{ with } \sum_{i=0}^{m} t_i = 1.$$

Proof. (1) $\Rightarrow$ (2). Assume that $\sum s_i = 0$ and $\sum s_i p_i = 0$. Then

$$\sum_{i=0}^{m} s_i p_i = \sum_{i=0}^{m} s_i p_i - (\sum_{i=0}^{m} s_i) p_0 = \sum_{i=0}^{m} s_i (p_i - p_0) = \sum_{i=1}^{m} s_i (p_i - p_0).$$

Affine independence of $p_0, \dots, p_m$ gives linear independence of $p_1 - p_0, \dots, p_m - p_0$ hence $s_i = 0$ for $i = 1, 2, \dots, m$. Finally $\sum s_i = 0$ implies $s_0 = 0$ as well.

(2) $\Rightarrow$ (3). Let $x \in [p_0, \dots, p_m]_a$. Then we see $x = \sum_{i=0}^m t_i p_i$ with $\sum_{i=0}^m t_i = 1$. If there is another representation of x as an affine combination of p_i's: $x = \sum_{i=0}^m t_i' p_i$ then

$$\sum_{i=0}^{m} (t_i - t_i') p_i = 0.$$

Since $\sum (t_i - t_i') = \sum t_i - \sum t_i' = 1 - 1 = 0$ it follows that $t_i = t_i'$ as desired.

(3) $\Rightarrow$ (1). Assume that each $x \in [p_0, \ldots, p_m]_a$ has a unique expression as an affine combination of $p_0, \ldots, p_m$. If the vectors $p_1 - p_0, \ldots, p_m - p_0$ were linearly dependent then there would be real numbers r_i, not all equal to zero such that

$$\sum_{i=1}^{m} r_i(p_i - p_0) = 0.$$

Let $r_j \neq 0$. Multiplying the last equation by r_j^{-1} we may assume that in fact $r_j = 1$. Now p_j has two different expressions as an affine combination of $p_0, \ldots, p_m$:

$$p_j = 1p_j;$$

$$p_j = -\sum_{i \neq j} r_i p_i + (1 + \sum_{i \neq j} r_i)p_0,$$

a contradiction. QED.

Corollary

Affine independence of the set $\{p_0, \ldots, p_m\}$ *is a property independent of the given ordering.*

Definition

Let $\{p_0, \ldots, p_m\}$ *be an affine independent subset of* $\mathbb{R}^n$*. If* $x \in [p_0, \ldots, p_m]_a$ *then there is a unique* $(m+1)$*-tuple* $(t_0, \ldots, t_m)$ *such that* $\sum t_i = 1$ *and* $x = \sum t_i p_i$*. The numbers* $t_0, \ldots, t_m$ *are called the* barycentric coordinates *of* x *(relative to the ordered set* $p_0, \ldots, p_m$*).*

Definition

Let $\{p_0, \ldots, p_m\}$ *be an affine independent subset of* $\mathbb{R}^n$*. The convex set* $[p_0, \ldots, p_m]$ *is called the m-simplex with vertices* $p_0, \ldots, p_m$*.*

Corollary

If $\{p_0, \ldots, p_m\}$ *is an affine independent set then each* x *in the m-simplex* $[p_0, \ldots, p_m]$ *has a unique expression of the form* $x = \sum t_i p_i$ *where* $\sum t_i = 1$ *and each* $t_i \geq 0$*.*

Proof.

Indeed, any $x \in [p_0, \ldots, p_m]$ is such a convex combination. If this expression had not been unique the barycentric coordinates would also have not been unique. □

Example

For $i = 0, 1, \ldots, n$ let e_i denote the point in $\mathbb{R}^{n+1}$ whose coordinates are all zeros except for 1 in the $(i+1)$st place. Clearly $\{e_0, \ldots, e_n\}$ is affine independent. The set $[e_0, \ldots, e_n]$ is called the standard n-simplex *in $\mathbb{R}^{n+1}$ and denoted by Δ^n. Thus, Δ^n consists of all convex combinations $x = \sum t_i e_i$. In this case, barycentric and cartesian coordinates of a point $x \in \Delta^n$ coincide and we see that Δ^n is a collection of points $(t_0, \ldots, t_n) \in \mathbb{R}^{n+1}$ for which $\sum t_i = 1$.*

Definition

Let $\{p_0, \ldots, p_m\} \subset \mathbb{R}^n$ be affine independent. Then an affine map *$f : [p_0, \ldots, p_m]_a \longrightarrow \mathbb{R}^k$ is a function satisfying*

$$f(\sum t_i p_i) = \sum t_i f(p_i)$$

whenever $\sum t_i = 1$. The restriction of f to $[p_0, \ldots, p_m]$ is also called an affine map.

Proposition

If $[p_0, \ldots, p_m]$ is an m-simplex, $[q_0, \ldots, q_n]$ is an n-simplex and $f : \{p_0, \ldots, p_m\} \longrightarrow \{q_0, \ldots, q_n\}$ is any function then there exists a unique affine map $\tilde{f} : [p_0, \ldots, p_m] \longrightarrow [q_0, \ldots, q_n]$ such that $\tilde{f}(p_i) = f(p_i)$ for $i = 0, 1, \ldots, m$.

Proof.

For a convex combination $\sum t_i p_i$ define $\tilde{f}(\sum t_i p_i) = \sum t_i f(p_i)$. Uniqueness is obvious. □

Definition

*Let Δ^n be the standard n-simplex. Its i*th face map *$\epsilon_i = \epsilon_i^n : \Delta^{n-1} \longrightarrow \Delta^n$ is the affine map from the standard $n-1$-simplex Δ^{n-1} to Δ^n given in the barycentric coordinates by the formula*

$$\epsilon_i^n(t_0, \ldots, t_{n-1}) = (t_0, \ldots, t_{i-1}, 0, t_i, \ldots, t_{n-1}).$$

Lemma

If $k < j$ the face maps satisfy

$$\epsilon_j^{n+1}\epsilon_k^n = \epsilon_k^{n+1}\epsilon_{j-1}^n : \Delta^{n-1} \longrightarrow \Delta^{n+1}.$$

Proof.

Just evaluate these affine maps on every vertex e_i for $0 \leq i \leq n-1$. □

Definition

Let X be a topological space. A singular n-simplex *in X is a continuous map $\sigma : \Delta^n \longrightarrow X$ where Δ^n is the standard n-simplex.*

Remark

A singular 0-simplex in X is just a point $x \in X$. A singular 1-simplex is a path $I = [0,1] \longrightarrow X$.

Definition

For a topological space X and an integer $n \geq 0$ we define the group $C_n(X)$ of singular n-chains *in X as the free abelian group generated by all singular n-simplices in X. Thus, the elements of $C_n(X)$ are linear combinations of the form $a_1\sigma_1 + \ldots + a_k\sigma_k$ (the integer k is not fixed) where σ_i are singular simplices of X. Furthermore we define the* boundary map *$d_n : C_n(X) \longrightarrow C_{n-1}(X)$ for $n > 0$ by setting*

$$d_n(\sigma) := \sum_{i=0}^{n} (-1)^i \sigma \epsilon_i^n \in C_{n-1}(X) \tag{0.1}$$

where $\sigma : \Delta^n \longrightarrow X$ is a singular n-simplex of X.

Remark

Strictly speaking we have to write d_n^X instead of d_n since these homomorphisms depend on X; this is never done, however. Furthermore we will frequently omit even the subscript n thinking of d as a collection of all d_n's.

Proposition

For all $n \geq 0$ we have $d_n d_{n+1} = 0$.

Proof. We need to prove that d applied twice is zero; applying it to an arbitrary singular n-simplex σ in X we have:

$$\begin{aligned} dd\sigma =& d\Big(\sum_j (-1)^j \sigma\epsilon_j^{n+1}\Big) \\ =& \sum_{j,k} (-1)^{j+k} \sigma\epsilon_j^{n+1}\epsilon_k^n \\ =& \sum_{j\leq k} (-1)^{j+k} \sigma\epsilon_j^{n+1}\epsilon_k^n + \sum_{j>k} (-1)^{j+k} \sigma\epsilon_j^{n+1}\epsilon_k^n \\ =& \sum_{j\leq k} (-1)^{j+k} \sigma\epsilon_j^{n+1}\epsilon_k^n + \sum_{j>k} (-1)^{j+k} \sigma\epsilon_k^{n+1}\epsilon_{j-1}^n, \end{aligned}$$

In the second sum, change variables: set $p = k$ and $q = j - 1$; it is now $\sum_{p \leq q} (-1)^{p+q+1} \sigma \epsilon_p^{n+1} \epsilon_q^n$. Each term $\sigma \epsilon_j^{n+1} \epsilon_k^n$ occurs once in the first sum and once (with the opposite sign) in the second sum. Therefore terms cancel in pairs and $dd\sigma = 0$. QED.

We see, therefore, that the sequence of abelian groups and homomorphisms

$$0 \longleftarrow C_0(X) \xleftarrow{d_1} C_1(X) \xleftarrow{d_2} \dots \xleftarrow{d_{n-1}} C_{n-1}(X) \xleftarrow{d_n} C_n(X) \xleftarrow{d_{n+1}} \dots$$

is a complex of abelian groups called the *singular complex of X*. It is denoted by $C_*(X)$ and its homology by $H_*(X)$ given by the abelian groups $H_n(X) = \operatorname{Ker} d_n / \operatorname{Im} d_{n+1}$ is called the *singular homology of the topological space X*.

Definition

Let C_* and B_* be chain complexes:

$$C_* = \{C_0 \xleftarrow{d_1} C_1 \xleftarrow{d_2} \dots \xleftarrow{d_n} C_n \xleftarrow{d_{n+1}} C_{n+1} \longleftarrow \dots\}$$

$$B_* = \{B_0 \xleftarrow{d_1} B_1 \xleftarrow{d_2} \dots \xleftarrow{d_n} B_n \xleftarrow{d_{n+1}} B_{n+1} \longleftarrow \dots\}$$

(Recall that there exists more general complexes, infinite in both directions but for our purposes it suffices to consider only those without negative components.)

Then a chain map *$f_* = \{f_n\} : C_* \longrightarrow B_*$ is a sequence of homomorphisms of abelian groups $f_n : C_n \longrightarrow B_n$ such that all squares in the diagram below are commutative:*

$$\begin{tikzcd}
C_0 \arrow[d, "f_0"'] & C_1 \arrow[l, "d_1"'] \arrow[d, "f_1"'] & \dots \arrow[l, "d_2"'] & C_n \arrow[l, "d_n"'] \arrow[d, "f_n"'] & C_{n+1} \arrow[l, "d_{n+1}"'] \arrow[d, "f_{n+1}"'] & \dots \arrow[l] \\
B_0 & B_1 \arrow[l, "d_1"'] & \dots \arrow[l, "d_2"'] & B_n \arrow[l, "d_n"'] & B_{n+1} \arrow[l, "d_{n+1}"'] & \dots \arrow[l]
\end{tikzcd}$$

Now we could form the category *Comp* whose objects are (chain) complexes of abelian groups and morphisms are chain maps between complexes. Note that a chain map $f_* : C_* \longrightarrow B_*$ is an isomorphism iff $f_n : C_n \longrightarrow B_n$ are isomorphisms of abelian groups for all n.

Definition

A complex C_* is called a subcomplex *in B_* if there exists a chain map $f_* : C_* \longrightarrow B_*$ such that $f_n : C_n \longrightarrow B_n$ is a monomorphism for each n. In that case each C_n could be identified with a subgroup $f_n(C_n)$ in B_n. Usually we will not distinguish between C_n and its image in B_n. If C_* is a subcomplex of B_* then the quotient complex B_*/C_* is the complex*

$$\ldots \longleftarrow B_{n-1}/C_{n-1} \xleftarrow{\bar{d}_n} B_n/C_n \longleftarrow \ldots$$

where $\bar{d}_n : b_n + C_n \mapsto d_n(b_n) + C_{n-1}$. If $f_* : C_* \longrightarrow B_*$ is a chain map then Ker f_* *is the subcomplex of C_**

$$\ldots \longleftarrow \operatorname{Ker} f_{n-1} \longleftarrow \operatorname{Ker} f_n \ldots$$

and Im f_* *is the subcomplex of B_*:*

$$\ldots \longleftarrow \operatorname{Im} f_{n-1} \longleftarrow \operatorname{Im} f_n \ldots$$

Exercise

Prove the analogue of the theorem on homomorphisms in the category of complexes Comp: if $f_* : C_* \longrightarrow B_*$ is a chain map then there is a chain isomorphism $C_*/\operatorname{Ker} f_* \cong \operatorname{Im} f_*$.

Recall that each complex has associated *homology groups*: $H_n(C_*) := \operatorname{Ker} d_n / \operatorname{Im} d_{n+1}$. The subgroup $\operatorname{Ker} d_n \subset C_n$ is called the subgroup of *n-cycles* whereas the subgroup $\operatorname{Im} d_{n+1} \subset C_n$ is called the subgroup of *n-boundaries*. Thus homology of a complex is the quotient of its cycles modulo the boundaries.

Exercise

Show that any chain map $C_* \longrightarrow B_*$ determines in a natural way a collection of homomorphisms $\{H_n(C_*) \longrightarrow H_n(B_*)\}$. Hint: note that under a chain map cycles in C_* map to cycles in B_* and boundaries in C_* map to boundaries in B_*.

Remark

The previous exercise shows that the correspondence $C_* \mapsto H_n(C_*)$ is a functor Comp $\longrightarrow \mathcal{A}b$. For a chain map $f_* : C_* \longrightarrow B_*$ the induced map on homology is denoted by $H_*(f) : H_*(C_*) \longrightarrow H_*(B_*)$. Sometimes we will abuse the notation and denote the induced maps on homology simply by f_*.

Let us return to the topological setting. Recall that we associated to any topological space X a chain complex $C_*(X)$, the singular complex of X. Our next aim is to show that this correspondence defines actually a functor $\mathcal{T}op \longrightarrow Comp$. Let $f : X \longrightarrow Y$ be a continuous map and $\sigma : \Delta^n \longrightarrow X$ be a singular n-simplex in X. Composing it with f we obtain a singular simplex $f \circ \sigma : \Delta^n \longrightarrow Y$. Extending by linearity gives a homomorphism $f_n : C_n(X) \longrightarrow C_n(Y)$. We denote by f_* the collection $\{f_n : C_n(X) \longrightarrow C_n(Y)\}$.

Proposition

f_* is a chain map $C_*(X) \longrightarrow C_*(Y)$. In other words, the following diagram is commutative for each n:

$$\begin{tikzcd}
C_{n-1}(X) \arrow[d, "f_{n-1}"] & C_n(X) \arrow[l, "d_n"'] \arrow[d, "f_n"] \\
C_{n-1}(Y) & C_n(Y) \arrow[l, "d_n"']
\end{tikzcd}$$

Proof.

It suffices to evaluate each composite on a generator σ of $C_*(X)$. We have:

$$f_{n-1}d_n\sigma = f_{n-1}\sum(-1)^i\sigma\epsilon_i = \sum(-1)^i f_{n-1}(\sigma\epsilon_i);$$

$$d_nf_n\sigma = d_n(f_n\sigma) = \sum(-1)^i(f_n\sigma)\epsilon_i.$$

□

Clearly to the composition of continuous maps $X \longrightarrow Y \longrightarrow Z$ there corresponds the composition of chain maps $C_*(X) \longrightarrow C_*(Y) \longrightarrow C_*(Z)$ and the identity map $X \longrightarrow X$ corresponds to the identity chain map $C_*(X) \longrightarrow C_*(X)$. That shows that the correspondence $X \mapsto C_*(X)$ is indeed a functor $\mathcal{T}op \longrightarrow Comp$. In particular we have the following

Corollary

If X and Y are homeomorphic then $H_n(X) \cong H_n(Y)$ for all $n \geq 0$.

The computation of singular homology of a topological space is not an easy task in general. However the case of H_0 is straightforward.

Proposition

For a nonempty topological space X the group $H_0(X)$ is the free abelian group whose generators are in $1-1$ correspondence with the set of path components of X.

Proof.

Let us consider first the case when X is path-connected. The group $C_0(X)$ is the free abelian group whose generators are the points in X. What is the subgroup of boundaries $B_0(X) \subset C_0(X)$? Take a 1-simplex $\sigma : \Delta^1 = I \longrightarrow X$. This is just a path in X from $x_1 = \sigma(1,0)$ to $x_2 = \sigma(0,1)$ where $(0,1)$ and $(1,0)$ are the two faces of $\Delta^1 = I$, that is its two endpoints written in barycentric coordinates. Note that $d_1(\sigma) = x_2 - x_1$ (check this!) This shows that the group $B_0(X)$ is spanned in $C_0(X)$ by all differences of the form $x_2 - x_1$ whenever x_1 and x_2 could be connected by a path. Pick a point $x \in X$. Since *all* points in X can be connected with x by a path we see that any 0-chain $c = \sum a_i x_i$ is *homologous* to (i.e. determines the same homology class as) $\sum a_i x$. Further clearly, chains of the form $ax, a \in \mathbb{Z}$ are pairwise nonhomologous. This shows that $H_0(X) \cong \mathbb{Z}$ and the generator corresponds to the 0-chain x.

Now suppose that X is not path-connected and denote by $\{X_\alpha\}$ its set of path components. Pick a point $x_\alpha \in X_\alpha$. Arguing as before we see that any 0-chain in X is homologous to the unique chain of the form $\sum a_\alpha x_\alpha$ (where the sum is, of course, finite). Moreover the chain $\sum a_\alpha x_\alpha$ is homologous to zero iff all $a_\alpha = 0$. So $H_0(X)$ is the free abelian group on the set of generators x_α. □

Suppose that $\{X_i\}$ is the collection of path components of a space X. What is the relation between the homology of X and the homology of the X_i's?
We need to discuss the notion of the *direct sum* of chain complexes. Recall that if A, B are two abelian groups then their *direct sum* $A \bigoplus B$ is the set of pairs $(a, b), a \in A, b \in B$ with componentwise addition. The direct sum of a finite number of abelian groups is defined similarly. In the case of an infinite collection $A_1, A_2, \ldots$ of abelian groups we define $\bigoplus_{i=1}^{\infty} A_i$ to be the set of sequences $(a_1, a_2, \ldots)$ where only finitely many of a_i's are nonzero. These sequences are added componentwise. (Note that if we allowed *arbitrary* sequences then the resulting object would be much bigger. It is called the *direct product* of the groups $A_1, A_2, \ldots$. The direct product of finitely many abelian groups coincides with their direct sum.) Similarly one can introduce direct sums of arbitrary (possibly uncountable) collections of abelian groups.

Definition

Let $\{C_*^i\}$ be a collection of chain complexes. Their direct sum $\bigoplus_i C_*^i$ is the complex

$$\ldots \xleftarrow{d_n} \bigoplus_i C_n^i \xleftarrow{d_{n+1}} \bigoplus_i C_{n+1}^i \xleftarrow{d_{n+2}} \ldots$$

with differentials

$$d_n(a_1, a_2, \ldots) = (d_n(a_1), d_n(a_2), \ldots).$$

Exercise

Show that $H_n(\bigoplus_i C^i_*) \cong \bigoplus_i H_n(C^i_*)$ *for all* n.

Now let $\sigma : \Delta^n \longrightarrow X$ be a singular n-simplex in X. Since the image of a connected space is connected σ is actually a singular simplex in one of the connected components of X. If $c = \sum a_i \sigma_i$ is a singular n-chain in X then grouping together the singular simplices belonging to the same connected component of X we could rewrite it as

$$c = \sum a^1_i \sigma^1_i + \sum a^2_i \sigma^2_i + \dots$$

where $c^k := \sum a^k_i \sigma^k_i$ is a singular n-chain in the kth connected component of X.

Thus we established a correspondence $c \mapsto (c^1, c^2, \dots)$. This correspondence is clearly $1-1$ and gives an isomorphism $C_*(X) \mapsto \bigoplus_i C_*(X_i)$ where X_i are the connected components of X. (Note: check that the last map is a chain map!)
We proved the following:

Proposition

The singular complex of a space X *is chain isomorphic to* $\bigoplus_i C_*(X_i)$ *where* $\{X_i\}$ *are the connected components of* X. *Therefore* $H_*(X) \cong \bigoplus_i H_*(X_i)$.

Proposition

If X is a one-point space then $H_n(X) \cong 0$ for all $n > 0$ and $H_0(X) \cong \mathbb{Z}$.

Proof.

For each $n \geq 0$ there is only one singular n-simplex $\sigma_n : \Delta^n \longrightarrow X$, namely, the constant map. Therefore $C_n(X) = \langle \sigma_n \rangle$, the infinite cyclic group generated by σ_n.

Let us compute the boundary operators

$$d_n \sigma_n = \sum_{i=0}^{n} (-1)^i \sigma_n \epsilon_i = \left[\sum_{i=0}^{n} (-1)^i \right] \sigma_{n-1},$$

(because $\sigma_n \epsilon_i$ is an $(n-1)$-simplex in X and σ_{n-1} is the only one such).

Therefore

$$d_n \sigma = \begin{cases} 0 & \text{if } n \text{ is odd} \\ \sigma_{n-1} & \text{if } n \text{ is even and positive} \end{cases}$$

It follows that the complex $C_*(X)$ has the form

$$\mathbb{Z} \xleftarrow{0} \mathbb{Z} \xleftarrow{\cong} \mathbb{Z} \xleftarrow{0} \mathbb{Z} \xleftarrow{\cong} \mathbb{Z} \xleftarrow{0} \dots$$

Clearly $H_0(X) \cong \mathbb{Z}$ (we already knew this) and $H_n(X) \cong 0$ for $n > 0$. $\square$

Our goal here is to prove that singular homology is isomorphic for homotopy equivalent spaces. We start with a preliminary result which will be used to prove the general case.

Proposition

If X is a convex subspace of a euclidean space then $H_n(X) = 0$ for all $n \geq 1$.

Proof. Choose a point $b \in X$ and for any singular simplex $\sigma : \Delta^n \longrightarrow X$ define a singular $(n+1)$-simplex $b\sigma : \Delta^{n+1} \longrightarrow X$ as follows:

$$(b\sigma)(t_0, \ldots, t_{n+1}) = \begin{cases} b & \text{if } t_0 = 1; \\ t_0 b + (1 - t_0)\sigma\left(\frac{t_1}{1-t_0}, \ldots, \frac{t_{n+1}}{1-t_0}\right) & \text{if } t_0 \neq 1 \end{cases}$$

Here $(t_0, \ldots, t_{n+1})$ are barycentric coordinates of points in Δ^{n+1}. The singular simplex $b\sigma$ is well-defined because X is convex. (Geometrically $b\sigma$ is the cone over σ with vertex b)
Now define $C_n(X) \longrightarrow C_{n+1}(X)$ by setting $c_n(\sigma) = b\sigma$ and extending by linearity. We claim that, for all $n \geq 1$ and any n-simplex σ in X

$$d_{n+1}c_n(\sigma) = \sigma - c_{n-1}d_n(\sigma). \tag{0.2}$$

The claim readily implies the desired conclusion. Indeed, if $\xi \in C_n(X)$ then we find $\xi = dc(\xi) + c(d\xi)$. If $d\xi = 0$ then $\xi = dc(\xi)$. Therefore the group of n-cycles coincides with the group of n-boundaries and $H_n(X) = 0$.

Let us now compute the faces of $c_n(\sigma) = b\sigma$. We have:

$$(b\sigma)\epsilon_0^{n+1}(t_0, \ldots, t_n) = \sigma(t_0, \ldots, t_n).$$

Now let $i > 0$. Then

$$(b\sigma)\epsilon_i^{n+1}(t_0, \ldots, t_n) = (b\sigma)(t_0, \ldots, t_{i-1}, 0, t_i, \ldots, t_n).$$

If, in addition, $t_0 = 1$ then

$$(b\sigma)\epsilon_i^{n+1}(t_0, \ldots, t_n) = b;$$

if $t_0 \neq 1$ then the right hand side above is equal to

$$t_0 b + (1 - t_0)\sigma\left(\frac{t_1}{1-t_0}, \ldots, \frac{t_{i-1}}{1-t_0}, 0, \frac{t_i}{1-t_0}, \ldots, \frac{t_n}{1-t_0}\right)$$

$$= t_0 b + (1 - t_0)\sigma\epsilon_{i-1}^n\left(\frac{t_1}{1-t_0}, \ldots, \frac{t_n}{1-t_0}\right) = c_{n-1}(\sigma\epsilon_{i-1}^n)(t_0, \ldots, t_n).$$

We conclude, after evaluating each side on $(t_0, \ldots, t_n)$ that

$$(c_n\sigma)\epsilon_0^{n+1} = \sigma \text{ and } (c_n\sigma)\epsilon_i^{n+1} = c_{n-1}\epsilon_{i-1}^n \text{ if } i > 0.$$

The rest is a routine calculation. QED.

Question

What is the condition on two chain maps $f_*, g_* : C_* \longrightarrow B_*$ ensuring that the induced maps on homology $H_*(f_*), H_*(g_*) : H_*(C_*) \longrightarrow H_*(B_*)$ coincide?

The answer is formulated in terms of *chain homotopy*.

Definition

The chain maps $f_*, g_* : C_* \longrightarrow B_*$ are chain homotopic *if there is a sequence of homomorphisms $s_n : C_n \longrightarrow B_{n+1}$ such that for all $n \in \mathbb{Z}$*

$$d_{n+1}s_n + s_{n-1}d_n = f_n - g_n. \tag{0.3}$$

The collection $s_* = \{s_n\}$ is called a chain homotopy *between f_* and g_*. We will write $f_* \sim g_*$ if there exists a chain homotopy between f_* and g_*.*

Remark

This definition is applicable to chain complexes infinite in both directions. We consider complexes C_* for which $C_n = 0$ if $n < 0$.

Proposition

The relation $\sim$ is an equivalence relation on the set of chain maps $C_* \to B_*$.

Proof.

1. Reflexivity: $f_* \sim f_*$ via $s_* := 0$.
2. Symmetry: if s_* is a chain homotopy between f_* and g_* then $-s_*$ is a chain homotopy between g_* and f_*.
3. Transitivity: if $s_* : f_* \sim g_*$ and $s'_* : g_* \sim h_*$ then $(s_* + s'_*) : f_* \sim h_*$.

□

The notion of chain homotopy is analogous to the notion of homotopy for continuous maps between topological spaces. In particular we have the following statement whose proof is left as an exercise:

Exercise

If s_* is a homotopy between $f_*, f'_* : C_* \longrightarrow B_*$ and s'_* is a homotopy between $g_*, g'_* : B_* \longrightarrow A_*$ then the chain maps $g_* \circ f_*$ and $g'_* \circ f'_*$ are homotopic through the chain homotopy $g_* \circ s_* + s'_* \circ f'_*$.

The main property of chain homotopies is that they induce identical maps on homology:

Proposition

If s_* is a homotopy between $f_*, g_* : C_* \longrightarrow B_*$ then the homomorphisms $H_n(f_*)$ and $H_n(g_*) : H_n(C_*) \longrightarrow H_n(B_*)$ coincide for any n.

Proof.

If $c \in Z_n(C_*)$ then $d_n c = 0$ and, therefore, we have $f_n(c) - g_n(c) = d(s_n c)$. In other words the cycles $f_n(c)$ and $g_n(c)$ are homologous in B_*. It follows that $f_n(c)$ and $g_n(c)$ determine the same homology class in $H_n(B_*)$ □

Just as for topological spaces we could introduce the notion of *chain homotopy equivalence* as follows:

Definition

Let C_*, B_* be two chain complexes and $f_* : C_* \longrightarrow B_*, g_* : B_* \longrightarrow C_*$ be chain maps such that $f_* \circ g_*$ is chain homotopic to id_{B_*} and $g_* \circ f_*$ is chain homotopic to id_{C_*}. Then C_* and B_* are called chain homotopy equivalent*.*

It follows that chain homotopy equivalent complexes C_* and B_* have isomorphic homology: $H_n(C_*) \cong H_n(B_*)$ (check this!).
Furthermore we say that a complex C_* is *chain contractible* if C_* is chain homotopy equivalent to the zero complex. Clearly C_* is chain contractible iff the identity map on C_* is chain homotopic to the zero map. The corresponding homotopy is called *contracting homotopy* for C_*.

We effectively constructed a contracting homotopy for the complex

$$\tilde{C}_*(X) = \{\mathbb{Z} \xleftarrow{d_0} C_0(X) \xleftarrow{d_1} C_1(X) \xleftarrow{d_2} C_2(X) \xleftarrow{d_3} \dots\} \qquad (0.4)$$

where d_0 is defined by the formula

$$d_0(\sum a_i x_i) = (\sum a_i) \cdot 1 \in \mathbb{Z}.$$

The complex is called the *augmented singular complex* of X (where X is any topological space). Thus, we showed that the augmented singular complex of a convex space is contractible (and hence, has zero homology).

Exercise

Show that the augmented sigular complex is indeed a complex, i.e that $d_0 \circ d_1 = 0$.

Definition

The homology of the augmented singular complex is called the reduced *singular homology of a space X. The nth reduced homology of X is denoted by $\tilde{H}_*(X) = H_*(\tilde{C}_*(X))$.*

Exercise

Show that $\tilde{H}_n(X) = H_n(X)$ if $n > 0$. Furthermore, show that $\tilde{H}_0(X) = 0$ if X is connected.

Theorem

Let X, Y be topological spaces. If $f, g : X \longrightarrow Y$ are homotopic then $H_n(f) = H_n(g)$ for all n.

Proof. Assume that f and g are homotopic. The following lemma allows one to replace the space Y with $X \times I$:

Lemma

Let $t, b : X \longrightarrow X \times I$ denote the maps $t(x) = (x, 1)$ and $b(x) = (x, 0)$. Then if $H_n(b) = H_n(t)$ then $H_n(f) = H_n(g)$

Proof of Lemma. Let $F : X \times I \longrightarrow Y$ be the homotopy between f and g. Consider the diagram

$$\begin{tikzcd} X \arrow[r, "t"] \arrow[d, "b"'] & X \times I \arrow[d, "F"] \\ X \times I \arrow[r, "F"] & Y \end{tikzcd}$$

Clearly $F \circ t = f$ and $F \circ b = g$. Applying to this diagram the homology functor H_n we obtain

$$\begin{tikzcd}
H_n(X) \arrow[r, "H_n(t)"] \arrow[d, "H_n(b)"'] & H_n(X \times I) \arrow[d, "H_n(F)"] \\
H_n(X \times I) \arrow[r, "H_n(F)"] & H_n(Y)
\end{tikzcd}$$

The latter diagram is commutative by assumption. Therefore

$$H_n(f) = H_n(F) \circ H_n(t) = H_n(F) \circ H_n(b) = H_n(g).$$

QED.

Returning to the proof of the theorem note according to the previous lemma that all we need to prove is that $H_n(b) = H_n(t)$. In other words the space Y has been eliminated from the picture.

Consider the induced maps $t_*, b_* : C_*(X) \longrightarrow C_*(X \times I)$. We will prove that t_*, b_* give the same maps in homology by showing that there exists a *chain homotopy* between t_* and b_*. In other words we will construct homomorphisms $s_n^X : C_n(X) \longrightarrow C_{n+1}(X \times I)$ such that

$$t_*^X - b_*^X = d_{n+1}s_n^X + s_{n-1}^X d_n. \tag{0.1}$$

We will prove this for *all* spaces X using induction on n.

In fact we will prove more: *Claim*: For all spaces X there exist homomorphisms $s_n^X : C_n(X) \longrightarrow C_{n+1}(X \times I)$ satisfying the above condition (0.1) and such that the following diagram commutes for every singular simplex $\sigma : \Delta^n \longrightarrow X$:

$$\begin{tikzcd} C_n(\Delta^n) \arrow[r, "s_n^{\Delta^n}"] \arrow[d, "\sigma_*"'] & C_{n+1}(\Delta^n \times I) \arrow[d, "(\sigma \times id)_*"] \\ C_n(X) \arrow[r, "s_n^X"] & C_{n+1}(X \times I) \end{tikzcd} \tag{0.2}$$

Let $n = 0$ and define $s_{-1}^X = 0$ (note that we don't have a choice here since $C_{-1}(X) = 0$ by definition). Now given $\sigma : \Delta^0 = pt \longrightarrow X$ we define $s_0^X : \Delta^1 = I \longrightarrow X$ by $t \longrightarrow (\sigma(pt), t)$ and then extend by linearity to the whole $C_0(X)$. Check:

$$d_1 s_0^X(\sigma) = (\sigma(pt), 1) - (\sigma(pt), 0) = t^X \circ \sigma - b^X \circ \sigma = t_*^X(\sigma) - b_*^X(\sigma),$$

and thus the condition (0.1) holds since $s_{-1}^X = 0$. Note that there is only one 0-simplex in Δ^0, the identity function $\delta : pt \longrightarrow pt$. To check commutativity evaluate each composite on δ.

We have:

$$s_0^X\sigma_*(\delta) = s_0^X(\sigma \circ \delta) = s_0^X(\sigma) : t \longrightarrow (\sigma(pt), t),$$

$$(\sigma \times id)_* s_0^{\Delta^0}(\delta) : t \longrightarrow (\sigma \times id)_*(\delta(pt), t) = (\sigma \times id)_*(pt, t) = (\sigma(pt), t).$$

Now assume that $n > 0$. If the above condition (0.1) holds then $(t_*^{\Delta^n} - b_*^{\Delta^n} - s_{n-1}^{\Delta^n} d_n)(\xi)$ would be a cycle for any $\xi \in C_n(X)$. But this is true:

$$\begin{aligned} d_n(t_*^{\Delta^n} - b_*^{\Delta^n} - s_{n-1}^{\Delta^n} d_n) =& t_*^{\Delta^n} d_n - b_*^{\Delta^n} d_n - d_n s_{n-1}^{\Delta^n} d_n \\ =& t_*^{\Delta^n} d_n - b_*^{\Delta^n} d_n - (t_*^{\Delta^n} - b_*^{\Delta^n} - s_{n-2}^{\Delta^n} d_{n-1}) d_n \end{aligned}$$

by the inductive assumption. But the last expression is clearly zero because $d_{n-1} \circ d_n = 0$.

Now let $\delta = id : \Delta^n \longrightarrow \Delta^n$ be the identity map on Δ^n considered as a singular n-simplex in Δ^n.

It follows that $d_n(t_*^{\Delta^n} - b_*^{\Delta^n} - s_{n-1}^{\Delta^n} d_n)(\delta)$ is a singular n-cycle in $\Delta^n \times I$. Since the latter is a convex set all cycles in $\Delta^n \times I$ are boundaries and therefore there exists $\beta_{n+1} \in C_{n+1}(\Delta^n \times I)$ for which

$$d_{n+1}\beta_{n+1} = d_n(t_*^{\Delta^n} - b_*^{\Delta^n} - s_{n-1}^{\Delta^n} d_n)(\delta).$$

Define $s_n^X : C_n(X) \longrightarrow C_{n+1}(X \times I)$ by

$$s_n^X(\sigma) = (\sigma \times id)_*(\beta_{n+1})$$

where σ is an n-simplex in X and extend by linearity.
Check condition (0.1); here $\sigma : \Delta^n \longrightarrow X$ is a singular n-simplex in X:

$$\begin{aligned}
d_{n+1}s_n^X(\sigma) =& d_{n+1}(\sigma \times id)_*(\beta_{n+1}) \\
=& (\sigma \times id)_* d_{n+1}(\beta_{n+1}) \\
=& (\sigma \times id)_*(t_*^{\Delta^n} - b_*^{\Delta^n} - s_{n-1}^{\Delta^n} d_n)(\delta) \\
=& (\sigma \times id)_* t_*^{\Delta^n} - (\sigma \times id)_* b_*^{\Delta^n} - (\sigma \times id)_* s_{n-1}^{\Delta^n} d_n(\delta) \\
=& (\sigma \times id)_* t_*^{\Delta^n} - (\sigma \times id)_* b_*^{\Delta^n} - s_{n-1}^X \sigma_* d_n(\delta) \\
=& t^X\sigma - b^X\sigma - s_{n-1}^X d_n \sigma_*(\delta) \\
=& (t^X - b^X - s_{n-1}^X d_n)(\sigma).
\end{aligned}$$

Here $\tau : \Delta^n \longrightarrow \Delta^n$ is a singular n-simplex in Δ^n and $\sigma : \Delta^n \longrightarrow X$ is a singular n-simplex in X:

$$\begin{gathered}
(\sigma \times id)_* s_n^{\Delta^n}(\tau) = (\sigma \times id)_*(\tau \times id)_*(\beta_{n+1}) = (\sigma\tau \times id)(\beta_{n+1}) \\
= s_n^X(\sigma\tau) = s_n^X \sigma_*(\tau).
\end{gathered}$$

QED.

Homotopy invariance of singular homology

We see, that the homology functors $H_n(?)$ could be lifted to functors $h\mathcal{T}op \mapsto \mathcal{A}b$.

Any functor respects categorical isomorphisms and we obtain immediately the following

Corollary

If X and Y are homotopy equivalent then $H_n(X) \cong H_n(Y)$ for any $n \geq 0$. In particular, a contractible space has the same homology as the one-point space.

We now come to the most important property of singular homology called *excision*. We start by defining the *relative singular complex*. Note that if A is a subspace of X then $C_*(A)$ is a subcomplex in $C_*(X)$.

Definition

The relative singular complex of a pair of topological spaces (X, A) *with* $A \subset X$ *is the complex* $C_*(X, A) := C_*(X)/C_*(A)$*. The corresponding homology is called the* relative homology *of the pair* (X, A) *and denoted by* $H_*(X, A)$*.*

Our aim is to prove the following theorem (excision); here $\bar{U}$ denotes the closure of U in X and A^o the interior of A in X:

Theorem

Assume that $U \subset A \subset X$ *are subspaces with* $\bar{U} \subset A^o$*. Then the inclusion* $i : (X \setminus U, A \setminus U) \hookrightarrow (X, A)$ *induces isomorphisms*

$$i_* : H_n(X \setminus U, A \setminus U) \longrightarrow H_n(X, A).$$

We are a rather long way away from this goal yet. Before going further we need to study the category *Comp* in some more detail.

Definition

Let C_*, D_*, A_* be chain complexes and $f_* : C_* \longrightarrow D_*, g_* : D_* \longrightarrow A_*$ be chain maps. Then the sequence

$$0 \longleftarrow A_* \xleftarrow{g_*} D_* \xleftarrow{f_*} C_* \longleftarrow 0 \tag{0.3}$$

is called a short exact sequence of complexes *if for each n the sequence of abelian groups*

$$0 \longleftarrow A_n \xleftarrow{g_n} D_n \xleftarrow{f_n} C_n \longleftarrow 0$$

is exact.

Given a short exact sequence (0.3) we will construct homomorphisms $\partial_n : H_n(A_*) \longrightarrow H_{n-1}(C_*)$ (called *connecting homomorphisms*) as follows: For $\xi \in H_n(A_*)$ choose its representative $\xi_1 \in Z_n(A_*) \subset A_n$. Since g_n is an epimorphism there exists $\xi_2 \in D_n$ such that $g_n(\xi_2) = \xi_1$. Consider the element $d_n(\xi_2) \in B_{n-1}(D_*)$. Since g_* is a chain map $g_{n-1}(d_n(\xi_2)) = 0$ (check this!). Therefore $\xi_3 := d_n(\xi_2)$ is in the kernel of g_{n-1} and it follows that ξ_3 is in the image of f_{n-1}. So there exists a unique $\xi_4 \in C_{n-1}$ such that $f_{n-1}(\xi_4) = \xi_3$. Since f_* is a chain map and $d_{n-1} \circ d_n = 0$ we conclude that $d_{n-1}(\xi_4) = 0$ (check this!). In other words $\xi_4 \in Z_{n-1}(C_*)$. Take its homology class $[\xi_4] \in H_{n-1}(C_*)$ and define $\partial_n(\xi) := [\xi_4]$.

Exercise

Show that the homomorphism ∂_n is independent of the choices involved, i.e.

- *of the choice of a ξ_1 in the homology class of ξ;*
- *of the choice of $\xi_2 \in D_n$.*

More precisely, show that different choices lead to a change in ξ_4 but not in $[\xi_4]$, that is various ξ_4's differ by an element in $B_{n-1}(C_*)$.

Proposition

Let

$$0 \longleftarrow A_* \xleftarrow{g_*} D_* \xleftarrow{f_*} C_* \longleftarrow 0$$

be a short exact sequence of complexes. Then the sequence of abelian groups and homomorphisms

$$\dots \xleftarrow{\partial_n} H_n(A_*) \xleftarrow{H_n(g_*)} H_n(D_*) \xleftarrow{H_n(f_*)} H_n(C_*) \xleftarrow{\partial_{n+1}} H_{n+1}(A_*) \longleftarrow \dots$$

is exact.

Proof.

1. *Check that* $\operatorname{Ker} H_n(g_*) = \operatorname{Im} H_n(f_*)$. *Take* $\xi \in H_n(D_*)$ *and its representative* $\xi_1 \in Z_n(D_*)$. *Suppose that* $g_*(\xi_1) \in B_n(A_*)$ *that is* $g_*(\xi_1) = d_{n+1}(\xi_2)$ *for* $\xi_2 \in A_{n+1}$. *Let* $\xi_3 \in D_{n+1}$ *be such that* $g_*\xi_3 = \xi_2$. *Then* $g_*(\xi - d_{n+1}\xi_3) = 0$ *and there exists* $\xi_4 \in C_n$ *such that* $f_*(\xi_4) = \xi - d_{n+1}\xi_3$. *That shows that* $\operatorname{Ker} H_n(g_*) \subset \operatorname{Im} H_n(f_*)$.

   The fact that $\operatorname{Im} H_n(f_*) \subset \operatorname{Ker} H_n(g_*)$ *follows from* $H_n(g_*) \circ H_n(f_*) = 0$. *The latter equality holds because* H_n *is a functor and* $g_* \circ f_* = 0$.
2. *Check that* $\operatorname{Ker} \partial_n = \operatorname{Im} H_n(g_*)$. *Let* $\xi \in H_n(A_*)$ *and choose a representative* $\xi_1 \in Z_n(A_*)$. *Recall that* $\partial_n(\xi)$ *is defined as the homology class of* $f_*^{-1}(d_n(g_*^{-1}(\xi_1)))$ *in* $H_{n-1}(A_*)$. *Suppose that* $f_*^{-1}(d_n(g_*^{-1}(\xi_1))) = d_n(\xi_2)$ *for some* $\xi_2 \in A_n$. *Consider the element* $g_*^{-1}(\xi_1) \in D_n$. *If this element is a cycle then we could stop. Otherwise replace it with the element* $\xi_3 = g_*^{-1}(\xi_1) - f_*(\xi_2)$. *Then* $\xi_3 \in Z_n(D_*)$ *and* $g_*(\xi_3) = \xi_1$. *This shows that* $\operatorname{Ker} \partial_n \subset \operatorname{Im} H_n(g_*)$. *The inclusion* $\operatorname{Im} H_n(g_*) \subset \operatorname{Ker} \partial_n$ *is easy.*
3. *Check that* $\operatorname{Ker} H_n(f_*) = \operatorname{Im} \partial_n$. *Let* $\xi_1 \in Z_n(C_*)$ *be a representative of* $\xi \in H_n(C_*)$ *and assume that* $f_*(\xi_1) = d_{n-1}(\xi_2)$ *for* $\xi_2 \in D_{n+1}$. *Set* $\xi_3 := g_*(\xi_2) \in A_n$. *Since* $d_{n+1}(\xi_3) = g_* \circ d_{n+1}(\xi_2) = 0$ *we conclude that* $\xi_3 \in Z_{n+1}(A_*)$. *Moreover,* $\partial_n(\xi_3) = \xi_1$. *This shows that* $\operatorname{Ker} H_n(f_*) \subset \operatorname{Im} \partial_n$. *The inclusion* $\operatorname{Im} \partial_n \subset \operatorname{Ker} H_n(f_*)$ *is easy. QED.*

The following result complements what we have just proved.

Proposition (Naturality of the homology long exact sequence.)

Assume that there is a commutative diagram in Comp with exact rows:

$$\begin{tikzcd}
0 & A_* \arrow[l] \arrow[d, "a"'] & D_* \arrow[l, "g_*"'] \arrow[d, "d"'] & C_* \arrow[l, "f_*"'] \arrow[d, "c"'] & 0 \arrow[l] \\
0 & A'_* \arrow[l] & D'_* \arrow[l, "g'_*"'] & C'_* \arrow[l, "f'_*"'] & 0 \arrow[l]
\end{tikzcd}$$

Then there is a commutative diagram of abelian groups with exact rows:

$$\begin{tikzcd}
\dots & H_n(A_*) \arrow[l] \arrow[d, "H_n(a)"'] & H_n(D_*) \arrow[l, "H_n(g_*)"'] \arrow[d, "H_n(d)"'] & H_n(C_*) \arrow[l, "H_n(f_*)"'] \arrow[d, "H_n(c)"'] & H_{n+1}(A_*) \arrow[l, "\partial_{n+1}"'] \arrow[d, "H_{n+1}(a)"'] & \dots \arrow[l] \\
\dots & H_n(A'_*) \arrow[l] & H_n(D'_*) \arrow[l, "H_n(g'_*)"'] & H_n(C'_*) \arrow[l, "H_n(f'_*)"'] & H_{n+1}(A'_*) \arrow[l, "\partial'_{n+1}"'] & \dots \arrow[l]
\end{tikzcd}$$

Proof.

Exactness of the rows was already proven before. The first two squares commute because H_n is a functor. The commutativity of the third square can be seen as follows. Take $\xi \in H_{n+1}(A_*)$ and $\xi' \in H_{n+1}(A'_*)$ such that $H_n(a)(\xi) = \xi'$. Choose a representative $\xi_1 \in Z_{n+1}(A_*)$, then $a(\xi_1) \in Z_{n+1}(A'_*)$ is a representative of ξ'. Take $\xi_2 \in D_{n+1}$ such that $f_*(\xi_2) = \xi_1$; then for $d(\xi_2) \in D'_{n+1}$ we have $f'_*(\xi'_2) = \xi'_1$. Set $\xi_3 := d_{n+1}(\xi_2)$, then for $\xi'_3 := d(\xi_3)$ we have $\xi'_3 = d_{n+1}(\xi'_2)$. Finally choose $\xi_4 \in C_n$ for which $f_*(\xi_4) = \xi_3$. Clearly then for $\xi'_4 := c(\xi_4)$ we have $f'_*(\xi'_4) = \xi'_3$. Therefore
$H_n(c) \circ \partial_{n+1}(\xi) = c([\xi_4]) = [\xi'_4] = \partial'_{n+1}(\xi') = \partial'_{n+1} \circ H_{n+1}(a)(\xi)$. □

Let us now go back to topology. For a space X and its subspace A we have the following short exact sequence of complexes:

$$0 \longleftarrow C_*(X,A) \longleftarrow C_*(X) \longleftarrow C_*(A) \longleftarrow 0\ .$$

Therefore we have the following

Corollary

There exists a long exact sequence (called the long exact sequence of a pair *(X,A):*

$$\ldots \longleftarrow H_n(X,A) \longleftarrow H_n(X) \longleftarrow H_n(A) \longleftarrow H_{n+1}(X,A) \longleftarrow \ldots$$

Let us now formulate a small variation on the homological long exact sequence involving a triple of complexes. It is usually called the *long exact sequence of a triple*.

Proposition

Let $B \subset A \subset X$ be inclusions of spaces. Then there is a (natural in all arguments) long exact sequence

$$\ldots \leftarrow H_n(X, A) \leftarrow H_n(A, B) \leftarrow H_n(X, B) \leftarrow H_{n+1}(X, A) \leftarrow \ldots$$

Proof.

This is just the long exact sequence associated with the short exact sequence of complexes:

$$0 \leftarrow C_*(X)/C_*(A) \leftarrow C_*(A)/C_*(B) \leftarrow C_*(X)/C_*(B) \leftarrow 0.$$

In some cases the relative homology can be reduced to the absolute one. □

Definition

Let A be a subspace in a topological space X. The pair (X, A) is called good *if A has a neighborhood V in X; $i : A \hookrightarrow V$ of which it is a* deformation retract, *i.e. there is a projection $j : V \to A$ such that $j \circ i = id_A$ and $i \circ j$ is homotopic to id_V.*

Theorem

Suppose the pair (X, A) is good. Then the quotient map $q : (X, A) \to (X/A, A/A)$ induces isomorphisms in homology $q_* : H_n(X, A) \to H_n(X/A, A/A) \cong \tilde{H}_n(X/A)$ for all n.

Proof. Here we assume excision (to be proved later!). Consider the following commutative diagram:

$$\begin{tikzcd}
H_n(X, A) \arrow[r] \arrow[d, "q_*"] & H_n(X, V) \arrow[d, "q_*"] & H_n(X \setminus A, V \setminus A) \arrow[l] \arrow[d, "q_*"] \\
H_n(X/A, A/A) \arrow[r] & H_n(X/A, V/A) & H_n(X/A \setminus A/A, V/A \setminus A/A) \arrow[l]
\end{tikzcd}$$

The upper left horizontal map is an isomorphism since in the long exact sequence of the triple (X, V, A) the groups $H_n(V, A)$ are all zero because a deformation retraction of V onto A gives a chain equivalence of complexes $C_*(V)/C_*(A)$ and $C_*(A)/C_*(A) = 0$. The same argument shows that the lower left horizontal map is an isomorphism as well. The other two horizontal maps are isomorphisms by excision and the rightmost vertical map is an isomorphism since q restricts to a homeomorphism on the complement of A. It follows from the commutativity of the diagram that the leftmost vertical map is an isomorphism as required. QED.

Corollary

For a wedge sum $\vee_\alpha X_\alpha$ *the inclusions* $X_\alpha \hookrightarrow \vee_\alpha X_\alpha$ *induce an isomorphism*

$$\bigoplus_\alpha i_{\alpha *} : \bigoplus_\alpha \tilde{H}_n(X_\alpha) \to \tilde{H}_n(\vee_\alpha X_\alpha)$$

where the wedge sum is formed at basepoints $x_\alpha \in X_\alpha$ *such that the pairs* (X_α, x_α) *are all good.*

Proof.

This follows from the above by taking $(X, A) = (\amalg_\alpha X_\alpha, \amalg_\alpha \{x_\alpha\})$. $\square$

Let $f : (X, A) \longrightarrow (Y, B)$ be a map of pairs, i.e. $A \subset X$, $B \subset Y$, and $f : X \longrightarrow Y$ is a map for which $f(A) \subset B$. Then, clearly, f induces a chain map $C_*(X, A) \longrightarrow C_*(Y, B)$ and the corresponding map on homology:

$$f_* : H_*(X, A) \longrightarrow H_*(Y, B).$$

Theorem

Let X_1 *and* X_2 *be subspaces of* X *with* $X = X_1^o \bigcup X_2^o$*. Then the inclusion of pairs* $i : (X_1, X_1 \bigcap X_2) \hookrightarrow (X, X_2)$ *induces isomorphisms in homology for all* n

$$i_* : H_n(X_1, X_1 \bigcap X_2) \cong H_n(X, X_2)$$

Proposition

The two theorems above are equivalent to each other.

Proof.

Assume the first theorem is true and let $X = X_1^o \bigcup X_2^o$. Set $A = X_2$ and $U = X \setminus X_1$. Then the pair $(X \setminus U, A \setminus U)$ is the pair $(X_1, X_1 \bigcap X_2)$ and the pair (X, A) is the pair (X, X_2) (check this!). The inclusions coincide and therefore induce the same map in homology.
Now assume the second theorem is true and let $\bar{U} \subset A^o$. Set $X_2 = A$ and $X_1 = X \setminus U$. Then

$$X_1^o \bigcup X_2^o = (X \setminus U)^o \bigcup A^o \supset (X \setminus \bar{U})^o \bigcup A^o \supset (X \setminus A^o) \bigcup A^o = X.$$

Finally we have $(X_1, X_1 \bigcap X_2) = (X \setminus U, A \setminus U)$ and $(X_1, X_2) = (X, A)$. □

We need a little bit more of homological algebra in order to deal with long exact sequences. So, here is a classic lemma from homological algebra.

Lemma

Consider the following commutative diagram with exact rows:

$$\begin{tikzcd}
\dots & A_n \arrow[l, "g_n"'] \arrow[d, "k_n"'] & D_n \arrow[l, "f_n"'] \arrow[d, "s_n"'] & C_n \arrow[l, "h_n"'] \arrow[d, "t_n"'] & A_{n+1} \arrow[l, "g_{n+1}"'] \arrow[d, "k_{n+1}"'] & \dots \arrow[l] \\
\dots & A'_n \arrow[l, "g'_n"'] & D'_n \arrow[l, "f'_n"'] & C'_n \arrow[l, "h'_n"'] & A'_{n+1} \arrow[l, "g'_{n+1}"'] & \dots \arrow[l]
\end{tikzcd}$$

in which every third map s_n is an isomorphism. Then the following sequence is exact:

$$\dots \longleftarrow A_n \xleftarrow{f_n s_n^{-1} h'_n} C'_n \xleftarrow{t_n - g'_{n+1}} C_n \oplus A'_{n+1} \xleftarrow{(g_{n+1}, k_{n+1})} A_{n+1} \longleftarrow \dots$$

Proof.

Let us check exactness at the place corresponding to $C_n \oplus A'_{n+1}$. Suppose that $(t_n - g'_{n+1})(c_n, a_{n+1}) = 0$, that is, $t_n(c_n) - g'_{n+1}(a_{n+1}) = 0$. It follows that $h_n(c_n) = 0$ and therefore there exists an element $a_{n+1} \in A_{n+1}$ such that $g_{n+1}(a_{n+1}) = c_n$. Consider $k_{n+1}(c_n) + a_{n+1}$. It is a cycle in the lower row and therefore there exists $\xi \in D'_{n+1}$ such that $f'_{n+1}(\xi) = k_{n+1}(c_n) + a_{n+1}$. Now set $\xi_1 := f_{n+1}(s_{n+1}^{-1}(\xi))$. Clearly $(g_{n+1}, k_{n+1})(\xi_1) = (c_n, a_{n+1})$. In other words $\mathrm{Ker}((t_n - g'_{n+1})) \subset \mathrm{Im}((g_{n+1}, k_{n+1}))$. The other inclusions are checked similarly. $\square$

Corollary (Mayer-Vietoris sequence)

If X_1, X_2 are subspaces of X with $X_1^o \bigcup X_2^o = X$ then the following sequence is exact:

$$\ldots \longleftarrow H_n(X_1 \bigcap X_2) \xleftarrow{\partial h_*^{-1} q_*} H_{n+1}(X) \xleftarrow{g_* - j_*} H_{n+1}(X_1) \oplus H_{n+1}(X_2) \xleftarrow{(i_{1\,*}, i_{2\,*})} \ldots$$

Here i_1, i_2 are the inclusions $X_1 \bigcap X_2 \longrightarrow X_1$ and $X_1 \bigcap X_2 \longrightarrow X_2$, g, j are the inclusions $X_1 \longrightarrow X$ and $X_2 \longrightarrow X$, q_* is induced by the projection $C_*(X) \longrightarrow C_*(X, X_2)$, h_* is the excision isomorphism $H_n(X_1, X_1 \bigcap X_2) \cong H_n(X, X_2)$ and ∂ is the connecting homomorphism of the pair $(X_1, X_1 \bigcap X_2)$.

Proof. We have the following map of topological pairs:

$$(X_1, X_1 \bigcap X_2) \longrightarrow (X, X_2).$$

This map induces a chain map between long exact sequences corresponding to the pairs $(X_1, X_1 \bigcap X_2)$ and (X, X_2).

So we get a commutative diagram whose rows are exact:

$$\begin{tikzcd}
\dots & H_n(X_1 \bigcap X_2) \arrow[l] \arrow[d] & H_{n+1}(X_1, X_1 \bigcap X_2) \arrow[l] \arrow[d, "h_*"'] & H_{n+1}(X_1) \arrow[l] \arrow[d] & \dots \arrow[l] \\
\dots & H_n(X_2) \arrow[l] & H_{n+1}(X, X_2) \arrow[l] & H_{n+1}(X) \arrow[l] & \dots \arrow[l]
\end{tikzcd}$$

By the second excision theorem each map h_* is an isomorphism and the result follows from the lemma above. QED.

Remark

Note that exactness of the Mayer-Vietoris sequence is a result concerning absolute *homology groups even though in the process relative homology were used. We will use it to compute homology groups of spheres.*

Exercise

Show that for X, X_1, X_2 as in the above corollary there exists an exact sequence

$$\dots \longleftarrow \tilde{H}_n(X_1 \bigcap X_2) \xleftarrow{\partial h_*^{-1} q_*} \tilde{H}_{n+1}(X) \xleftarrow{g_* - j_*} \tilde{H}_{n+1}(X_1) \oplus \tilde{H}_{n+1}(X_2) \xleftarrow{(i_{1\,*}, i_{2\,*})} \dots$$

The end of this sequence is (in contrast with the Mayer-Vietoris sequence for unreduced homology*):*

$$0 \longleftarrow \tilde{H}_0(X) \longleftarrow \tilde{H}_0(X_1) \oplus \tilde{H}_0(X_2) \longleftarrow \dots$$

Theorem

Let S^n be the n-sphere where $n \geq 0$. Then

1. $H_0(S^0) = \mathbb{Z} \oplus \mathbb{Z}$ *while* $H_i(S^0) = 0$ *for* $i > 0$.
2. *For* $n > 0$ *we have* $H_n(S^n) = H_0(S^n) = \mathbb{Z}$ *while* $H_i(S^n) = 0$ *if* $i \neq 0, n$.

Remark

Using reduced homology the result could be reformulated more concisely: $\tilde{H}_n(S^n) = \mathbb{Z}$ *while* $\tilde{H}_i(S^n) = 0$ *if* $i \neq n$.

Proof. We prove that the reduced homology of S^n is as claimed using induction on n. We know the result is true for $n = 0$ since S^0 is just a union of two points.
Now assume that $n > 0$. Let N and S be the north and south poles of S^n. Set $X_1 = S^n \setminus N$ and $X_2 = S^n \setminus S$. Clearly $S^n = X_1^o \bigcup X_2^o$. Furthermore $X_1 \bigcap X_2$ has the same homotopy type as the equator S^{n-1} (check this!). Applying the Mayer-Vietoris sequence for reduced homology we get an exact sequence:

$$\tilde{H}_i(X_1) \oplus \tilde{H}_i(X_2) \longleftarrow \tilde{H}_i(X_1 \bigcap X_2) \longleftarrow \tilde{H}_{i+1}(S^n) \longleftarrow \tilde{H}_{i+1}(X_1) \oplus \tilde{H}_{i+1}(X_2)$$

It follows from contractibility of X_1 and X_2 that the left and right terms in the above sequence are both zero and therefore

$$\tilde{H}_{i+1}(S^n) \cong \tilde{H}_i(X_1 \bigcap X_2) \cong \tilde{H}_i(S^{n-1}).$$

(Note that the above sequence is exact also for $i = 0$.) By induction $\tilde{H}_{i+1}(S^n) = \tilde{H}_i(S^{n-1}) = \mathbb{Z}$ if $i + 1 = n$ and 0 otherwise. QED.

Having established the necessary properties of the singular homology functor H_* we get as a corollary the Brouwer fixed point theorem discussed in the introduction:

Theorem (Brouwer fixed point theorem)

Let D^n be the n-dimensional disc. Let $f : D^n \longrightarrow D^n$ be a continuous map. Then f has at least one fixed point, i.e. there exists a point $x \in D^n$ for which $f(x) = x$.

Let's draw some other corollaries:

Corollary

If $m \neq n$ then S^n and S^m are not homotopy equivalent. In particular they are not homeomorphic. Indeed, S^n and S^m have different homology.

Proof.

The space $\mathbb{R}^n \setminus \textit{point}$ has the same homotopy type as S^{n-1} (why?). Likewise $\mathbb{R}^m \setminus \textit{point}$ is homotopically equivalent to S^{m-1}. If $\mathbb{R}^n \setminus \textit{point}$ and $\mathbb{R}^m \setminus \textit{point}$ were homeomorphic then S^{n-1} and S^{m-1} would also be homeomorphic. But, as we saw in the previous corollary, this is not true. $\square$

Now we will prove the excision property, namely the second excision theorem:

Theorem (Excision)

Let X_1 and X_2 be subspaces of X with $X = X_1^o \bigcup X_2^o$. Then the inclusion of pairs $i : (X_1, X_1 \bigcap X_2) \hookrightarrow (X, X_2)$ induces isomorphisms in homology for all n

$$i_* : H_n(X_1, X_1 \bigcap X_2) \cong H_n(X, X_2)$$

for all n.

Let X_1, X_2 be subspaces of X. Then clearly $C_*(X_1)$ and $C_*(X_2)$ are subcomplexes of $C_*(X)$. Denote by $C_*(X_1) + C_*(X_2)$ the subcomplex of $C_*(X)$ consisting of all sums $c_1 + c_2 \in C_*(X)$ where $c_1 \in C_n(X_1)$, $c_2 \in C_n(X_2)$ for some n.

Lemma

If the inclusion $C_*(X_1) + C_*(X_2) \hookrightarrow C_*(X)$ induces an isomorphism in homology the excision holds for the subspaces X_1, X_2 of X.

Proof of Excision

Proof. The short exact sequence of complexes

$$0 \longrightarrow C_*(X_1) + C_*(X_2) \longrightarrow C_*(X) \longrightarrow C_*(X)/(C_*(X_1) + C_*(X_2)) \longrightarrow 0$$

leads to the long exact sequence in homology from which it follows that the complex $C_*(X)/(C_*(X_1) + C_*(X_2))$ has zero homology (Check this!). Now consider the short exact sequence of complexes

$$0 \longrightarrow \frac{C_*(X_1) + C_*(X_2)}{C_*(X_2)} \longrightarrow \frac{C_*(X)}{C_*(X_2)} \longrightarrow \frac{C_*(X)}{C_*(X_1) + C_*(X_2)} \longrightarrow 0.$$

The associated long exact sequence in homology has every third term zero from which it follows that the map $\frac{C_*(X_1)+C_*(X_2)}{C_*(X_2)} \longrightarrow \frac{C_*(X)}{C_*(X_2)}$ induces an isomorphism in homology.

Finally consider the following commutative diagram of complexes:

$$\begin{tikzcd}
\frac{C_*(X_1)}{C_*(X_1 \bigcap X_2)} \arrow[rr] \arrow[dr] & & \frac{C_*(X)}{C_*(X_2)} \\
& \frac{C_*(X_1)+C_*(X_2)}{C_*(X_2)} \arrow[ur] &
\end{tikzcd}$$

Proof of Excision

We just showed that the northeast arrow induces an isomorphism in homology. Furthermore since $C_*(X_1 \bigcap X_2) \cong C_*(X_1) \bigcap C_*(X_2)$ the southwest arrow is actually an isomorphism of complexes, in particular it induces an isomorphism in homology. It follows that the horizontal arrow induces an isomorphism in homology which is what the excision property asserts. QED.

So it remains to prove that the inclusion $C_*(X_1) + C_*(X_2) \longrightarrow C_*(X)$ induces an isomorphism in homology whenever $X = X_1^o \bigcup X_2^o$. This is where the real difficulty lies. To overcome this difficulty we need an idea. The idea is to replace a singular simplex of X by a sum of small simplices which belong either to X_1 or X_2. To do that we need the notion of *barycentric subdivision*.

Definition

The barycenter *of an n-simplex is the point having barycentric coordinates* $(\frac{1}{n+1}, \ldots, \frac{1}{n+1})$.

In particular the barycenter of a 1-simplex, or a line segment, is its middle point, the barycenter of a 2-simplex, or a triangle, is the intersection of its medians etc.

Definition

The barycentric subdivision *of an affine simplex Σ^n is a collection* $\operatorname{Sd}\Sigma^n$ *of simplices defined inductively:*

1. $\operatorname{Sd}\Sigma^0 = \Sigma^0$;
2. *if $f_0, \ldots f_{n+1}$ are the n-dimesional faces of Σ^{n+1} then $\operatorname{Sd}\Sigma^{n+1}$ consists of all the $(n+1)$-dimensional simplices spanned by the barycenter of Σ^{n+1} and the n-simplices in $\operatorname{Sd} f_i$, $i = 0, \ldots, n+1$.*

Remark

Note that

- $\operatorname{Sd}\Sigma^n$ *consists of exactly $(n+1)!$ simplices;*
- *every n-simplex of $\operatorname{Sd}\Sigma^n$ has an ordering on the set of its vertices. Namely its first vertex is the barycenter of Σ^n. Its second vertex corresponds to the barycenter of σ^{n-1}, some $(n-1)$-dimensional face of Σ^n. Let us denote this vertex by $b_{\sigma^{n-1}}$. The third vertex $b_{\sigma^{n-2}}$ of our simplex corresponds to the barycenter of some σ^{n-2} etc. Thus, any n-simplex of $\operatorname{Sd}\Sigma^n$ has the form $[b_{\sigma^n}, b_{\sigma^{n-1}}, \ldots, b_{\sigma^0}]$ where the σ_i form a nested system: $\sigma^n = \Sigma^n \supset \sigma^{n-1} \supset \ldots \supset \sigma^0$.*

We want to define a map $\mathrm{Sd}_n : C_n(X) \longrightarrow C_n(X)$ for all $n \geq 0$. We'll do it in stages: first, assuming that X is convex and then in general.

Definition

Let X be a convex set in $\mathbb{R}^m$ and $e_0, \ldots, e_n$ are vertices of the standard n-simplex Δ^n. We say that a singular simplex $\sigma : \Delta^n \longrightarrow X$ is affine *if $\sigma(\sum t_i e_i) = \sum t_i \sigma(e_i)$ where $\sum t_i = 1$ and $t_i \geq 0$. A (finite) linear combination of singular affine simplices in X is called an* affine singular chain*. The set of all affine singular chains in X will be denoted by $C_n^{aff}(X)$*

Remark

Briefly, a singular simplex σ is affine if it is affine as a map $\Delta^n = [e_0 \ldots, e_n] \longrightarrow X \hookrightarrow \mathbb{R}^m$. The set $C_*^{aff}(X)$ of affine singular chains is a free abelian group that is a subgroup in the group of all singular chains. Moreover this subgroup is actually a subcomplex (why?)

Definition

Let X be a convex set. The barycentric subdivision is a homomorphism $\mathrm{Sd}_n : C_n^{aff}(X) \longrightarrow C_n^{aff}(X)$ defined inductively on generators $\sigma : \Delta^n \longrightarrow X$:

1. *if $n = 0$ then $\mathrm{Sd}_0(\sigma) = \sigma$*
2. *if $n > 0$ then $\mathrm{Sd}_n(\sigma) = \sigma(b_n)\,\mathrm{Sd}_{n-1}(d\sigma)$ where b_n is the barycenter of Δ^n. (Recall that $\sigma(b)_n\,\mathrm{Sd}_{n-1}(d\sigma)$ is the 'cone over $\mathrm{Sd}_{n-1}(d\sigma)$ with vertex b_n')*

We now define the barycentric subdivision of $C_*(X)$ where X is an *arbitrary* space.

Definition

If X is any space then we define the homomorphism $\mathrm{Sd}_n : C_n(X) \longrightarrow C_n(X)$ on generators $\sigma : \Delta^n \longrightarrow X$ by the formula

$$\mathrm{Sd}_n(\sigma) = \sigma_*\,\mathrm{Sd}(\delta^n),$$

where $\delta^n : \Delta^n \longrightarrow \Delta^n$ is the identity map. (Note that Δ^n is convex and δ^n is an affine simplex so $\mathrm{Sd}(\delta^n)$ has already been defined).

Remark

Note, that the operation Sd *is natural with respect to continuous maps* $X \longrightarrow Y$.
In other words, the following diagram commutes for all $n \geq 0$ *(check this!):*

$$\begin{tikzcd} C_n(X) \arrow[r, "\mathrm{Sd}_n"] \arrow[d, "f_*"] & C_n(X) \arrow[d, "f_*"] \\ C_n(Y) \arrow[r, "\mathrm{Sd}_n"] & C_n(Y) \end{tikzcd}$$

Lemma

The morphism $\mathrm{Sd} : C_*(X) \longrightarrow C_*(X)$ *is a chain map.*

Proof. Assume first that X is convex and let $\sigma : \Delta^n \longrightarrow X$ be an affine n-simplex. We will prove by induction that

$$\mathrm{Sd}_{n-1}\, d_n\sigma = d_n\, \mathrm{Sd}_n\, \sigma.$$

Since $\mathrm{Sd}_{-1} = 0$ and $d_0 = 0$ the base of induction $n = 0$ is clear. Now let $n > 0$, then

$$d_n\, \mathrm{Sd}_n\, \sigma = d_n(\sigma(b_n)\, \mathrm{Sd}_{n-1}(d_n\sigma)) = \mathrm{Sd}_{n-1}\, d_n\sigma - \sigma(b_n)((d_{n-1}\, \mathrm{Sd}_{n-1}))d_n\sigma. \quad (0.4)$$

(In the last equality we used the identity $d(b\xi) = \xi - bd\xi$ which was checked before.).

By the inductive assumption we have that:

$$\mathrm{Sd}_{n-1}\, d_n\sigma - \sigma(b_n)(\mathrm{Sd}_{n-2}\, d_{n-1}d_n\sigma) = \mathrm{Sd}_{n-1}\, d_n\sigma.$$

Now let X be a not necessarily convex space and $\sigma : \Delta^n \longrightarrow X$ be a singular n-simplex in X.

Proof of Excision

Then we have the following:

$$
\begin{aligned}
d\,\mathrm{Sd}(\sigma) =& d\sigma_*\,\mathrm{Sd}(\delta^n) \\
=& \sigma_*\, d\,\mathrm{Sd}(\delta^n) \\
=& \sigma_*\,\mathrm{Sd}\, d(\delta^n) \text{ (because } \Delta^n \text{ is convex)} \\
=& \mathrm{Sd}\,\sigma_* d(\delta^n) \\
=& \mathrm{Sd}\, d\sigma_*(\delta^n) \\
=& \mathrm{Sd}\, d\sigma.
\end{aligned}
$$

QED.

The following lemma is crucial; it shows that the barycentric subcomplex $\mathrm{Sd}\, C_*(X) \subset C_*(X)$ has the same homology as $C_*(X)$:

Lemma

For each $n \geq 0$ the induced homomorphism $H_n(\mathrm{Sd}) : H_n(X) \longrightarrow H_n(X)$ is the identity.

Proof. We show that the map $\mathrm{Sd} : C_*(X) \longrightarrow C_*(X)$ is chain homotopic to the identity map. In other words, we will construct homomorphisms $s_n : C_n(X) \longrightarrow C_{n+1}(X)$ such that $d_{n+1}s_n + s_{n-1}d_n = id - \mathrm{Sd}_n$.

Assume first that X is convex and prove the desired formula (for the affine singular complex) by induction on $n \geq 0$. Define $s_0 : C_0^{aff}(X) \longrightarrow C_1^{aff}(X)$ to be the zero map. The base of induction ($n = 0$) is obvious and we assume that $n > 0$. For any $\xi \in C_n^{aff}(X)$ we need to define s_n so that

$$ds_n\xi = \xi - \mathrm{Sd}\,\xi - s_{n-1}d\xi. \tag{0.5}$$

Note that the right-hand side above is a cycle. Indeed,

$$d(\xi - \mathrm{Sd}\,\xi - s_{n-1}d\xi) = d\xi - d\,\mathrm{Sd}\,\xi - (id - \mathrm{Sd} - s_{n-2}d)d\xi = 0.$$

(Here we used the inductive assumption and the identity $d \circ d = 0$).
Since a convex set has zero homology all cycles are boundaries and we can find an element in $C_n(X)$ whose boundary is $\xi - \mathrm{Sd}\,\xi - s_{n-1}d\xi$. We call this element $s_n(\xi)$. Specifically, set $s_n(\xi) = b(\xi - \mathrm{Sd}\,\xi - s_{n-1}d\xi)$. (Note that $s_n(\xi) \in C_n^{aff}(X)$.) Then the above equation (0.5) is satisfied (why?).

Proof of Excision

Now let X be any space and $\sigma : \Delta^n \longrightarrow X$ be a singular n-simplex. Then define

$$s_n(\sigma) = \sigma_* s_n(\delta^n) \in C_{n+1}(X),$$

where, as usual, we denoted by δ^n the identity singular simplex on Δ^n.

What remains is to prove that formula (0.5) holds for so defined s_n. To see this first notice that the following diagram is commutative for any continuous map $X \longrightarrow Y$:

$$\begin{tikzcd} C_n(X) \arrow[r, "f_*"] \arrow[d, "s_n"] & C_n(Y) \arrow[d, "s_n"] \\ C_{n+1}(X) \arrow[r, "f_*"] & C_{n+1}(Y) \end{tikzcd}$$

Using this naturality property and the fact that formula (0.5) is proved for the simplex Δ^n we see that it holds in the general case (Check it!). This finishes the proof. QED.

The next result we are going to discuss makes precise the intuitively obvious picture we have in mind: *the simplices of the barycentric subdivision are smaller then the original simplex.*
Moreover, by iterating the operation Sd one gets *arbitrarily small* simplices.

Definition

The *diameter* of a simplex in $\mathbb{R}^m$ is the maximal distance between any two points in it.

Proposition

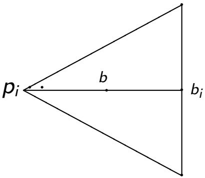
Let $\sigma = [p_0, \ldots, p_n]$ be an n-simplex in $\mathbb{R}^m$. Then the diameter of any simplex in $\operatorname{Sd}\sigma$ *is at most $\frac{n}{n+1}$ times the diameter of σ.*

Proof. Note that the diameter of σ equals the maximal distance between any of its vertices. This fact is geometrically obvious and could be proved rigorously using the triangle inequality (Check it!).
So we have to check that the distance between any two q_i, q_j of the barycentric subdivision of σ is at most $\frac{n}{n+1}$ times the diameter of σ. If neither q_i nor q_j is the barycenter of σ then these two points lie in a proper face of σ and obvious induction on n gives the result (Check it!).

Proof of Excision

So suppose that q_i is the barycenter b. We could also suppose q_j to be one of the vertices p_i of σ, again by the triangle inequality (Check it!).

Let b_i be the barycenter of the face $[p_0, \ldots, \hat{p}_i, \ldots, p_n]$. Then $b = \frac{1}{n+1} p_i + \frac{n}{n+1} b_i$. The sum of two coefficents is 1 so b lies on the line segment $[p_i, b_i]$ from p_i to b_i:



Furthermore the distance from b to p_i is $\frac{n}{n+1}$ times the length of $[p_i, b_i]$. Therefore $|b, p_i|$ is bounded by $\frac{n}{n+1}$ times the diameter of σ. QED.

Proof of Excision

Now let $\mathrm{Sd}^d : C_*(X) \longrightarrow C_*(X)$ be the dth iteration of the operation Sd. The previous result implies that the diameter of any simplex in $\mathrm{Sd}^d(\sigma)$ is at most $\left(\frac{n}{n+1}\right)^d$ the diameter of σ. In particular the simplices in $\mathrm{Sd}^d(\sigma)$ become arbitrarily small as d gets bigger. We have the following

Corollary

If X_1, X_2 are subspaces in X with $X = X_1^o \bigcup X_2^o$ and σ is a singular n-simplex in X then for a large enough d we will have $\mathrm{Sd}^d \sigma \in C_n(X_1) \bigcup C_n(X_2)$.

Proof.

Consider the covering of Δ^n by two open sets $X_1' = \sigma^{-1}(X_1)$ and $X_2' = \sigma^{-1}(X_2)$. Standard considerations using compactness of Δ^n shows that any set $U \subset \Delta^n$ whose diameter is small enough must be contained in X_1' or X_2' (check the details!).
Therefore there exists an integer d for which every simplex in $\mathrm{Sd}^d(\Delta^n)$ is contained in X_1' or X_2'. It follows that the image of every singular simplex entering in $\mathrm{Sd}^d \sigma$ is contained in $C_n(X_1)$ or $C_n(X_2)$. □

Proof of Excision

We can now complete the proof of the excision property. Recall that we only need to prove that $i : C_*(X_1) + C_*(X_2) \hookrightarrow C_*(X)$ induces an isomorphism in homology.

1. The map $i_* : H_n(C_*(X_1) + C_*(X_2)) \longrightarrow H_n(X)$ is surjective. Let $\xi \in H_n(X)$ and ξ_1 be the cycle representing ξ. Since $\mathrm{Sd} : C_*(X) \longrightarrow C_*(X)$ is chain homotopic to the identity map the cochain $\mathrm{Sd}(\xi_1)$ is a cycle which is homologous to ξ_1. Iterating we see that $\mathrm{Sd}^d(\xi_1)$ is a cycle homologous to ξ_1 for any integer d. But we just saw that $\mathrm{Sd}^d(\xi_1) \in C_n(X_1) + C_n(X_2)$. So we found a cycle which lies in $C_n(X_1) + C_n(X_2)$ and is homologous to ξ_1, hence i is surjective in homology.
2. The map $i_* : H_n(C_*(X_1) + C_*(X_2)) \longrightarrow H_n(X)$ is injective. Let $\xi_1 + \xi_2 \in \mathrm{Ker}\, i_*$ and take a representative cycle $\xi_1' + \xi_2'$ of $\xi_1 + \xi_2$. Then $i(\xi_1' + \xi_2') \in C_n(X)$ is a boundary: $i(\xi_1' + \xi_2') = d(\eta)$ for $\eta \in C_{n+1}(X)$. Since

$$\eta - \mathrm{Sd}\,\eta = (sd + ds)\eta$$

we have, after taking d of both sides.

$$d\eta - d(\mathrm{Sd}\,\eta) = dsd(\eta).$$

We conclude that

$$\xi_1' + \xi_2' = d\eta = d(\mathrm{Sd}\,\eta + s(d\eta)) = d(\mathrm{Sd}\,\eta + s(\xi_1') + s(\xi_2')).$$

Proof of Excision

So we proved that $\xi_1' + \xi_2'$ is a boundary of an element in $C_*(X_1) + C_*(X_2)$ and we are done (assuming that $\operatorname{Sd} \eta \in C_{n+1}(X_1) + C_{n+1}(X_2)$).
If this is not the case note that there exists an integer d for which $\operatorname{Sd}^d \eta \in C_{n+1}(X_1) + C_{n+1}(X_2)$ and the map $\operatorname{Sd}^d : C_*(X) \longrightarrow C_*(X)$ is still homotopic to the identity map. So we could argue as before replacing Sd with Sd^d. This completes the proof. QED.

The relationship between homology and the fundamental group

Question

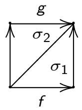
What is the relationship between the fundamental group of a topological space and it first homology group?

A map $f : I \to X$ can be viewed either as a path or as a 1-simplex in X. If $f(0) = f(1)$ then this singular simplex is a 1-cycle. This idea gives rise to a homomorphism between $\pi_1(X)$ and $H_1(X)$.

Theorem

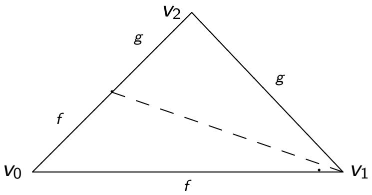
1. *The above construction determines a homomorphism $h : \pi_1(X, x_0) \to H_1(X)$.*
2. *If X is path-connected then h is surjective and its kernel is the commutator subgroup of $\pi_1(X)$, i.e. the normal subgroup generated by all commutators $[a, b] = aba^{-1}b^{-1}$ where $a, b \in \pi_1(X)$.*

Proof. Let us first prove well-definedness. Note that the constant path viewed as a 1-simplex is equal to the boundary of the constant 2-simplex with the same image and thus, is homologous to zero.
Next, let two paths f and g be homotopic. Consider a homotopy $F : I \times I \to X$ from f to g and subdivide the square $I \times I$ into two triangles as shown on the following picture:



When one computes $\partial(\sigma_1 - \sigma_2)$ the two restrictions of F onto the diagonal cancel, leaving $f - g$ together with two constant singular 1-simplices from the left and right edges of the square. Since constant singular 1-simplices are boundaries it follows that $f - g$ is a boundary also.

To show that h is a homomorphism consider the singular 2-simplex $\sigma : \Delta^2 \to X$ given as the composition of the orthogonal projection of $\Delta^2 = [v_0, v_1, v_2]$ onto the edge $[v_0, v_2]$ followed by $fg : [v_0, v_2] \to X$ then $\partial\sigma = g - fg + f$.



Further we have $f + f^{-1}$ is homologous to ff^{-1} which is homologous to zero and it follows that f^{-1} is homologous to $-f$.

Now show that h is surjective (if X is path-connected). Let $\sum n_i\sigma_i$ be a 1-cycle representing a given homology class. After relabeling we can assume that in fact all n_i are ± 1 and since inverse paths correspond to negative of the corresponding chains we can assume that all n_i are 1. If some of the σ_i is not a loop then since $\partial(\sum \sigma_i) = 0$ there must be another σ_j in the sum such that its combined path $\sigma_i\sigma_j$ is defined and we can, therefore, decrease the number of summands until all of them will be loops. Since X is path-connected we can replace all these loops by the homologous ones and based at the same point x_0. Then we can take the composition of all these loops obtaining a single loop representing our original homology class.

The final part is to prove that the kernel of h is the commutator subgroup of $\pi_1(X)$. Since $H_1(X)$ is an abelian group we conclude that the commutator is inside the kernel. It remains to show that any element $[f] \in \pi_1(X)$ that is in the kernel of h must be homotopic to products of commutators.

If an element $[f] \in \pi_1(X)$ is in the kernel of h then it is, as a 1-chain, a boundary of a 2-chain $\sum n_i\sigma_i$. As before, we can assume that $n_i = \pm 1$. We will now construct a certain topological space (a 2-dimensional surface in fact) by taking 2-simplices – triangles – one for each σ_i and glueing them together. To do that write $\partial\sigma_i = \tau_{i0} - \tau_{i1} + \tau_{i2}$ for the corresponding singular simplices τ_{ij}. It follows that in the formula for $\partial\sigma_i$ all singular simplices, except for one that is equal to f, could be divided into pairs so that each pair consists of a singular 1-simplex τ_{ij} plus itself taken with coefficient -1 (resulting in cancelation, of course).

This gives a scheme for glueing faces of our two-dimensional simplices: we identify the corresponding edges of our triangles preserving their orientation. There results a space K; the maps σ_i fit together to get a map $K \to X$.

It is clear that K is a two-dimensional surface with boundary corresponding to f since glueing triangles along their edges will always give rise to a surface. We claim that K is an oriented surface. Indeed, we can glue a disc along f, triangulate this disc and also assume that the partition of the obtained closed surface $\tilde{K}$ is in fact a triangulation by taking barycentric subdivisions if needed. Then $\tilde{K}$ has the property (ensured by the equation $f = \partial(\sum n_i\sigma_i)$) that the sum of all triangles in the triangulation together with appropriate signs (viewed as a singular 2-chain) is a 2-cycle.

This property will clearly be preserved under any refinement of the triangulation. It will also hold for an orientable surface as its representation as a $4g$-gon makes clear and it does not hold for an unorientable surface by the same reason.
So we proved that the loop $f : S^1 \to X$ extends to a map from an orientable surface K whose boundary is S^1. Let g be the genus of K, then its fundamental group is the free group on $2g$ generators $a_1 \ldots, a_g, b_1, \ldots, b_g$ and the class of the boundary circle is represented by the product of commutators $a_1 b_1 a_1^{-1} b_1^{-1} \ldots a_g b_g a_g^{-1} b_g^{-1}$. Therefore the class of f inside $\pi_1(X)$ also belongs to the commutator subgroup as required. QED.

Remark

Note the following useful observation used in the proof above: a (based) map $f : S^1 \to X$ lies in the commutator subgroups of $\pi_1(X)$ if and only if it extends to a map from an orientable surface whose boundary is S^1. More precisely, if such a map can be represented as a product of n commutators then this surface could be taken to have have genus n. The genus 0 surface corresponds to maps homotopic to zero.

Question

What can we say about a map that can be extended to a map from an unorientable surface?

The relationship between homology and the fundamental group

Developing this line of thinking further one could ask for a similar interpretation of an element in the fundamental group G of a space which lies not in the commutator subgroup $G' := [G, G]$ of G, but in the smaller subgroup $[G', G]$ or in $[G', G']$. Moreover, one could go still further and consider an iteration of this procedure; e.g. when does a given element in the fundamental group lie in the nth member of the lower central series of G?

These question lead to the notion of a *grope* which are certain two-dimensional spaces (not surfaces) obtained by certain simple glueings of surfaces. These spaces play an important role in knot theory and low-dimensional topology.

Corollary

Let S_g be a two-dimensional surface of genus g. Recall that $\pi_1(S_g)$ is a group with generators $a_i, b_i, i = 1, 2, \ldots g$ subject to the relation $a_1 b_1 a_1^{-1} b_1^{-1} \ldots a_g b_g a_g^{-1} b_g^{-1} = 1$. The singular homology $H_1(S_g)$ of S_g is the quotient of $\pi_1(S_g)$ by the commutator subgroup and a free abelian group on $2g$ generators a_i, b_i.