

Algebraic Topology

Lecture Notes

Anton Mellit

2026–2027

Contents

1	Topological spaces and homotopy	3
1.1	The study of shapes	3
1.2	Topologies and metric spaces	3
1.3	Continuous maps	4
1.4	Subspaces and products	5
1.5	Homeomorphisms	5
1.6	Connectedness and invariants	6
1.7	Homotopy: a movie of maps	6
1.8	Homotopy equivalence	7
1.9	Contractible spaces	8
1.10	The punctured plane and the circle	8
1.11	Retractions and deformation retractions	9
2	The homotopy category and the first invariants	10
2.1	Categories: objects and morphisms	10
2.2	The homotopy category	11
2.3	Isomorphisms in a category	11
2.4	Homotopy equivalence as isomorphism	12
2.5	Paths and path components	13
2.6	The invariant π_0	13
2.7	Connectedness and local path-connectedness	14
2.8	Free loops and simple connectedness	17
2.9	Based loops and $\pi_1(X, x_0)$ as a set	18
3	The fundamental group, coverings, and degree	19
3.1	Based loops and their multiplication	19
3.2	The group axioms as square homotopies	20
3.3	Why free homotopy does not suffice	21
3.4	Covering maps: local copies of the base	22
3.5	Lifting paths and homotopies	23
3.6	Degree of a loop in S^1	23
3.7	Homotopy invariance of degree	24
4	The circle, functoriality, and change of basepoint	26
4.1	The fundamental group of the circle	26
4.2	Application: the Brouwer fixed-point theorem	26
4.3	Functors	27
4.4	The fundamental group as a functor	27
4.5	Changing the basepoint	28
4.6	Dependence on the connecting path	29
4.7	Conjugation and words in two generators	30

5	Proof of the homotopy lifting theorem	31
5.1	The lifting problem	31
5.2	Good sets and two lifting principles	31
5.3	A lift near $\{z\} \times I$	32
5.4	Uniqueness and global gluing	33

1 Topological spaces and homotopy

Lecture 1, 1 October 2026

Throughout these notes, $I = [0, 1]$ denotes the unit interval.

1.1 The study of shapes

Topology studies shapes and continuous maps between them. Familiar examples include Euclidean space $\mathbb{R}^n$, the spheres

$$S^n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\},$$

the torus (the surface of a doughnut), and a surface of genus two (a surface with two handles). We would like to say when two spaces have the same shape, and to find ways of proving that their shapes are different.

There are several possible meanings of “the same shape.” A homeomorphism is a continuous change of coordinates with a continuous inverse. Homotopy equivalence, introduced later in this lecture, allows more flexibility: for example, a solid ball will have the same homotopy type as a point. Algebraic topology studies spaces by associating algebraic objects to them. Such objects can provide *invariants*: quantities or structures that remain unchanged under the chosen notion of equivalence. Our first task is to make the topological language precise.

1.2 Topologies and metric spaces

In analysis, continuity is expressed using distances and ε - δ conditions. Topology extracts the part of that language which depends only on open sets.

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

1. $\emptyset$ and X belong to $\mathcal{T}$;
2. an arbitrary union of members of $\mathcal{T}$ belongs to $\mathcal{T}$;
3. a finite intersection of members of $\mathcal{T}$ belongs to $\mathcal{T}$.

The pair $(X, \mathcal{T})$ is a *topological space*, and the members of $\mathcal{T}$ are its *open sets*. We usually write simply X when the topology is understood.

Arbitrary unions include infinite unions. Intersections, by contrast, are required to be open only when finitely many open sets are involved. For example, in the usual topology on $\mathbb{R}$,

$$\bigcap_{m=1}^{\infty} (-1/m, 1/m) = \{0\}$$

is not open.

Definition 1.2. A *metric* on a set X is a function $d : X \times X \rightarrow [0, \infty)$ such that, for all $x, y, z \in X$,

$$\begin{aligned} d(x, y) = 0 &\iff x = y, \\ d(x, y) &= d(y, x), \\ d(x, z) &\leq d(x, y) + d(y, z). \end{aligned}$$

A set equipped with a metric is a *metric space*. Its open ball with centre x and radius $r > 0$ is

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

A subset $U \subseteq X$ is *open in the metric topology* if, for every $x \in U$, there is an $r > 0$ such that $B_d(x, r) \subseteq U$.

Proposition 1.3. *The open subsets of a metric space form a topology.*

Proof. The empty set satisfies the condition vacuously, and X contains every ball. If $x \in \bigcup_{\alpha} U_{\alpha}$, choose α with $x \in U_{\alpha}$. A ball about x contained in U_{α} is also contained in the union. If $x \in U_1 \cap \cdots \cap U_k$, choose $r_i > 0$ with $B_d(x, r_i) \subseteq U_i$. The ball of radius $\min(r_1, \dots, r_k) > 0$ lies in the intersection. The empty intersection is X . $\square$

The balls themselves are open. If $y \in B_d(x, r)$, then $r - d(x, y) > 0$, and the triangle inequality gives

$$B_d(y, r - d(x, y)) \subseteq B_d(x, r).$$

Thus a subset is open exactly when it is a union of open balls.

Example 1.4. Euclidean space has the metric

$$d(x, y) = \|x - y\| = \left(\sum_{i=1}^n (x_i - y_i)^2 \right)^{1/2}.$$

Unless otherwise specified, subsets of $\mathbb{R}^n$ carry the topology inherited from this metric.

Definition 1.5. A subset F of a topological space X is *closed* if $X \setminus F$ is open.

De Morgan's laws give the corresponding closed-set properties: $\emptyset$ and X are closed, arbitrary intersections of closed sets are closed, and finite unions of closed sets are closed. A subset can be both open and closed, or neither.

1.3 Continuous maps

Definition 1.6. A map $f : X \rightarrow Y$ between topological spaces is *continuous* if $f^{-1}(V)$ is open in X for every open subset $V \subseteq Y$.

Here $f^{-1}(V) = \{x \in X : f(x) \in V\}$ denotes an inverse image; the map f need not be invertible. The definition concerns inverse images and does not require images of open sets to be open.

Proposition 1.7. *For metric spaces (X, d_X) and (Y, d_Y) , a map $f : X \rightarrow Y$ is continuous if and only if, for every $x \in X$ and every $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$d_X(x, x') < \delta \implies d_Y(f(x), f(x')) < \varepsilon.$$

Proof. Suppose first that f is continuous in the open-set sense. The inverse image of $B_{d_Y}(f(x), \varepsilon)$ is an open set containing x . It therefore contains some $B_{d_X}(x, \delta)$, which is precisely the required implication.

Conversely, assume the ε - δ condition. Let $V \subseteq Y$ be open and let $x \in f^{-1}(V)$. Choose $\varepsilon > 0$ with $B_{d_Y}(f(x), \varepsilon) \subseteq V$, and choose $\delta > 0$ from the assumed condition. Then $B_{d_X}(x, \delta) \subseteq f^{-1}(V)$. Thus $f^{-1}(V)$ is open. $\square$

Identity maps and constant maps are continuous. If $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ are continuous, then so is $g \circ f$, since

$$(g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W)).$$

A map is continuous if and only if inverse images of closed sets are closed: use $f^{-1}(Y \setminus F) = X \setminus f^{-1}(F)$.

1.4 Subspaces and products

We need topologies on subsets and products to formulate the examples and the definition of homotopy.

Definition 1.8. For a subset $A \subseteq X$, the *subspace topology* on A consists of all sets $A \cap U$ with U open in X .

The topology axioms follow by taking intersections with A in the axioms for X . Openness here is relative to A : for example, $[0, 1/2)$ is open in I , because it equals $I \cap (-1, 1/2)$. A subset of A is closed in A exactly when it has the form $A \cap F$ with F closed in X .

The inclusion $i : A \hookrightarrow X$ is continuous. A map $h : Z \rightarrow A$ is continuous if and only if $i \circ h : Z \rightarrow X$ is continuous, because

$$h^{-1}(A \cap U) = (i \circ h)^{-1}(U).$$

This lets us check continuity into a subspace in the ambient space, provided we first check that the values lie in the subspace.

Definition 1.9. The *product topology* on $X \times Y$ consists of arbitrary unions of sets $U \times V$, where $U \subseteq X$ and $V \subseteq Y$ are open.

Finite intersections of open rectangles are open rectangles, so this is a topology. Both coordinate projections are continuous. A map $h : Z \rightarrow X \times Y$, written $h(z) = (h_1(z), h_2(z))$, is continuous exactly when h_1 and h_2 are continuous. One implication follows by composing with projections; for the other, use

$$h^{-1}(U \times V) = h_1^{-1}(U) \cap h_2^{-1}(V)$$

and take unions. In particular, this applies to $X \times I$, where I has its usual subspace topology in $\mathbb{R}$.

1.5 Homeomorphisms

Definition 1.10. A *homeomorphism* from X to Y is a bijection $f : X \rightarrow Y$ such that both f and f^{-1} are continuous. If such a map exists, the spaces are *homeomorphic*, written $X \cong Y$.

A homeomorphism identifies the topologies: a subset $U \subseteq X$ is open if and only if $f(U) \subseteq Y$ is open. Continuity of the inverse is part of the definition, not a consequence of continuity and bijectivity alone.

Example 1.11 (Euclidean space and the open ball). Let $B^n = \{x \in \mathbb{R}^n : \|x\| < 1\}$. Define

$$f : B^n \longrightarrow \mathbb{R}^n, \quad f(x) = \frac{x}{1 - \|x\|}, \quad g : \mathbb{R}^n \longrightarrow B^n, \quad g(y) = \frac{y}{1 + \|y\|}.$$

The denominator of f is positive, and

$$\|g(y)\| = \frac{\|y\|}{1 + \|y\|} < 1,$$

so both maps are well defined and continuous. If $r = \|x\| < 1$, then

$$\|f(x)\| = \frac{r}{1 - r}, \quad g(f(x)) = \frac{x/(1 - r)}{1 + r/(1 - r)} = x.$$

Similarly, writing $\rho = \|y\|$ gives

$$f(g(y)) = \frac{y/(1 + \rho)}{1 - \rho/(1 + \rho)} = y.$$

Thus f and g are inverse homeomorphisms. Geometrically, f moves points along rays from the origin, replacing the radius r by $r/(1 - r)$. Radii approaching 1 are sent to arbitrarily large radii.

1.6 Connectedness and invariants

How can we show that two spaces are *not* homeomorphic? Trying and failing to construct a homeomorphism proves nothing. Instead, we look for a property preserved by every homeomorphism which differs between the spaces.

One such idea is *connectedness*: informally, whether a space consists of one piece or several separate pieces. A homeomorphism cannot join separate pieces or split a piece into two. For now we use this geometric intuition; precise definitions and the relevant proofs will come in Section 2.7 of Lecture 2.

Example 1.12 (The circle is not an interval). The circle S^1 and the interval I are each one piece, so this alone does not distinguish them. Instead, remove a point.

Removing any point p from the circle leaves an arc, still in one piece. Removing an interior point $a \in (0, 1)$ from the interval leaves two separate pieces:

$$I \setminus \{a\} = [0, a) \sqcup (a, 1].$$

If a homeomorphism from the circle to the interval existed, removing an interior point of the interval and its corresponding point on the circle would still leave homeomorphic spaces. But one would have two pieces and the other only one. The same argument applies to any real interval containing more than one point, by removing a point between two of its points.

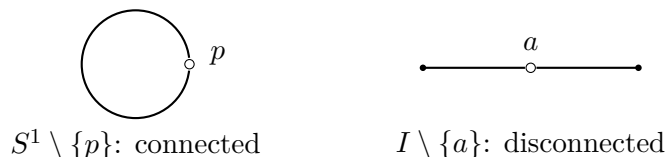


Figure 1: Connectedness after deleting a point distinguishes the circle from an interval. An open circle marks the deleted point.

This illustrates the method of invariants. Sometimes a basic invariant, such as the number of connected components, suffices. Sometimes we must apply it to related spaces, such as spaces obtained by removing points. Later, algebraic invariants will give more powerful ways to distinguish spaces.

1.7 Homotopy: a movie of maps

Fix two spaces X and Y . We want to express the idea that one continuous map $X \rightarrow Y$ can be continuously deformed into another.

Definition 1.13. Two continuous maps $f, g : X \rightarrow Y$ are *homotopic* if there exists a continuous map

$$h : X \times I \longrightarrow Y$$

such that $h(x, 0) = f(x)$ and $h(x, 1) = g(x)$ for every $x \in X$. We call h a *homotopy from f to g* , and write $f \simeq g$.

Think of $s \in I$ as *time*. At time s the map is

$$h_s : X \rightarrow Y, \quad h_s(x) = h(x, s).$$

Thus a homotopy is a *movie of maps*: the first frame is f , the last frame is g , and the intervening frames describe the deformation. For a fixed $x \in X$, the curve $s \mapsto h(x, s)$ traces the motion of that point's image.

The requirement is joint continuity in (x, s) . Each frame is continuous, since $x \mapsto (x, s)$ is continuous, and each point traces a continuous path. These separate observations do not replace joint continuity of h .

Proposition 1.14. *Homotopy is an equivalence relation on the set of continuous maps from X to Y .*

Proof. Reflexivity. The stationary movie $h(x, s) = f(x)$ is a homotopy from f to itself.

Symmetry. If h is a homotopy from f to g , the reversed movie

$$\bar{h}(x, s) = h(x, 1 - s)$$

is a homotopy from g to f . It is continuous as a composition of continuous maps.

Transitivity. Suppose h_1 is a homotopy from f_0 to f_1 and h_2 is a homotopy from f_1 to f_2 . Run h_1 during the first half of the time interval and h_2 during the second half:

$$h(x, s) = \begin{cases} h_1(x, 2s), & 0 \leq s \leq \frac{1}{2}, \\ h_2(x, 2s - 1), & \frac{1}{2} \leq s \leq 1. \end{cases}$$

At the joining time the expressions agree: $h_1(x, 1) = f_1(x) = h_2(x, 0)$. Each formula is continuous on its own half, and the two halves are closed subsets covering $X \times I$. The pasting lemma below proves continuity of h . Its initial and final maps are f_0 and f_2 . $\square$

Lemma 1.15 (Pasting lemma for a finite closed cover). *Let $Z = A_1 \cup \dots \cup A_k$, where each A_i is closed in Z . If continuous maps $u_i : A_i \rightarrow Y$ agree on every overlap, the map $u : Z \rightarrow Y$ obtained by using u_i on A_i is continuous.*

Proof. Agreement on overlaps makes u well defined. For a closed subset $F \subseteq Y$, each $u_i^{-1}(F)$ is closed in A_i , and hence in Z because A_i is closed in Z . Thus

$$u^{-1}(F) = \bigcup_{i=1}^k u_i^{-1}(F)$$

is a finite union of closed subsets of Z , so is closed. $\square$

In the transitivity argument the closed subsets are $X \times [0, \frac{1}{2}]$ and $X \times [\frac{1}{2}, 1]$. They are inverse images of closed intervals under the projection to I . No separation assumption on X is needed.

1.8 Homotopy equivalence

Definition 1.16. Spaces X and Y are *homotopy equivalent* if there are continuous maps $f : X \rightarrow Y$ and $g : Y \rightarrow X$ with

$$g \circ f \simeq \text{id}_X, \quad f \circ g \simeq \text{id}_Y.$$

The map f is a *homotopy equivalence*, and g a *homotopy inverse*. We write $X \simeq Y$ and say that X and Y have the same *homotopy type*.

For a homeomorphism the two compositions equal the identity maps. For a homotopy equivalence they need only be homotopic to the identity maps. Every homeomorphism is a homotopy equivalence, but a homotopy equivalence need not be a bijection.

The notation $\simeq$ has two related uses: between maps it means that the maps are homotopic; between spaces it means that they are homotopy equivalent.

Remark 1.17 (Compatibility with composition). If $f_0 \simeq f_1 : X \rightarrow Y$ and $a : W \rightarrow X$, $b : Y \rightarrow Z$ are continuous, then $b \circ f_0 \circ a \simeq b \circ f_1 \circ a$. A homotopy h between f_0 and f_1 gives the homotopy $(w, s) \mapsto b(h(a(w), s))$ between the composites.

Consequently, homotopy equivalence of spaces is itself an equivalence relation. Reflexivity and symmetry follow directly from the definition. For transitivity, suppose $f : X \rightarrow Y$ and $u : Y \rightarrow Z$ have homotopy inverses g and v . Then $g \circ v$ is a homotopy inverse to $u \circ f$:

$$(gv)(uf) = g(vu)f \simeq gf \simeq \text{id}_X, \quad (uf)(gv) = u(fg)v \simeq uv \simeq \text{id}_Z.$$

Composition symbols have been omitted in these chains.

1.9 Contractible spaces

Definition 1.18. A space X is *contractible* if it is homotopy equivalent to a one-point space $\{*\}$.

Proposition 1.19. A space X is contractible if and only if there exist $x_0 \in X$ and a homotopy $h : X \times I \rightarrow X$ with

$$h(x, 0) = x, \quad h(x, 1) = x_0 \quad \text{for every } x \in X.$$

Such a homotopy is called a *contraction* of X .

Proof. There is a unique map $p : X \rightarrow \{*\}$. A map $j : \{*\} \rightarrow X$ amounts to a choice of $x_0 = j(*)$. Both maps are continuous, and $p \circ j = \text{id}_{\{*\}}$, whereas $j \circ p$ is the constant map c_{x_0} . Thus p and j are homotopy inverses exactly when $\text{id}_X \simeq c_{x_0}$. $\square$

Contractibility requires X to be nonempty. A contraction moves every point to x_0 through points of X . The definition does not require x_0 to remain fixed at intermediate times; in our examples below it does.

Example 1.20 (Convex subsets of Euclidean space). A subset $C \subseteq \mathbb{R}^n$ is *convex* if $(1-s)x + sy \in C$ for all $x, y \in C$ and $s \in I$. Every nonempty convex subset is contractible. Choose $x_0 \in C$ and set

$$h : C \times I \longrightarrow C, \quad h(x, s) = (1-s)x + sx_0.$$

Convexity ensures that the values lie in C . The formula is continuous as a map into $\mathbb{R}^n$, hence as a map into the subspace C . Its endpoints are $h(x, 0) = x$ and $h(x, 1) = x_0$.

This contracts C along straight line segments. In particular, $\mathbb{R}^n$, open and closed balls, and nonempty intervals are contractible. For $n \geq 1$, $\mathbb{R}^n$ is not homeomorphic to a point, since no bijection exists. Thus homotopy equivalence is strictly weaker than homeomorphism.

1.10 The punctured plane and the circle

Let $X = \mathbb{R}^2 \setminus \{0\}$ and let $i : S^1 \hookrightarrow X$ be inclusion. Radial projection gives a continuous map

$$r : X \longrightarrow S^1, \quad r(x) = \frac{x}{\|x\|}.$$

It fixes points of the circle, so $r \circ i = \text{id}_{S^1}$. The other composite sends each nonzero vector to the unit vector in the same direction. It is homotopic to the identity by radial movement:

$$h : X \times I \longrightarrow X, \quad h(x, s) = (1-s)x + s \frac{x}{\|x\|} = \left(1-s + \frac{s}{\|x\|}\right)x. \quad (1.1)$$

The scalar in parentheses is strictly positive for $x \neq 0$ and $s \in I$, so the homotopy never passes through the omitted origin. The formula is continuous and satisfies

$$h(x, 0) = x, \quad h(x, 1) = i(r(x)).$$

Thus $i \circ r \simeq \text{id}_X$, and i and r are homotopy inverses. We have proved

$$\mathbb{R}^2 \setminus \{0\} \simeq S^1.$$

At time s , a point of radius ρ has radius $(1-s)\rho + s$ and keeps its direction. Points outside the circle move inward, points inside move outward, and points on the circle remain fixed throughout.

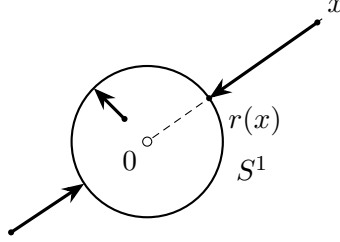


Figure 2: Radial deformation onto the unit circle. Every point travels along its ray, and no trajectory reaches the origin.

1.11 Retractions and deformation retractions

The preceding example has more structure than an arbitrary homotopy equivalence. One space is a subspace of the other, the map back fixes that subspace, and the deformation leaves it fixed at every time.

Definition 1.21. Let $A \subseteq X$ be a subspace, with inclusion $i : A \hookrightarrow X$. A *retraction* onto A is a continuous map $r : X \rightarrow A$ with $r(a) = a$ for all $a \in A$, or equivalently $r \circ i = \text{id}_A$. If such a map exists, A is called a *retract* of X .

Thus *retraction* names the map and *retract* names the subspace. A retraction alone does not provide a homotopy from id_X to $i \circ r$.

Definition 1.22. A *deformation retraction* of X onto a subspace A consists of a retraction $r : X \rightarrow A$ and a continuous map $h : X \times I \rightarrow X$ such that

$$h(x, 0) = x, \quad h(x, 1) = i(r(x)), \quad h(a, s) = a \quad (a \in A, s \in I).$$

When these data exist, A is a *deformation retract* of X .

Remark 1.23 (Terminology). We require A to remain fixed throughout the deformation. This is also called a *strong deformation retraction*. Some authors use “deformation retraction” without this requirement and reserve “strong” for the definition above. The convention here is always the one in Definition 1.22.

Proposition 1.24. If A is a deformation retract of X , then inclusion $i : A \rightarrow X$ is a homotopy equivalence, with homotopy inverse $r : X \rightarrow A$.

Proof. The retraction identity gives $r \circ i = \text{id}_A$. The deformation supplies $i \circ r \simeq \text{id}_X$. These are exactly the two conditions for a homotopy equivalence. $\square$

Both main examples have this form. In Example 1.20, $h(x, s) = (1-s)x + sx_0$ fixes x_0 at every time, so $\{x_0\}$ is a deformation retract of the nonempty convex set C . In (1.1), if $a \in S^1$ then $\|a\| = 1$ and $h(a, s) = a$. Hence S^1 is a deformation retract of the punctured plane.

Deformation retractions provide a concrete way to prove homotopy equivalences: exhibit a subspace and a continuous motion of the whole space onto it, keeping that subspace fixed.

2 The homotopy category and the first invariants

Lecture 2, 7 October 2026

2.1 Categories: objects and morphisms

Our next goal is to interpret homotopy equivalence categorically. In a suitable category, saying that two spaces are homotopy equivalent will mean exactly that they are isomorphic. We first introduce the language needed to make this statement precise.

Definition 2.1. A *category* $\mathcal{C}$ consists of:

1. a collection of *objects*;
2. for each pair of objects X, Y , a set $\text{Hom}_{\mathcal{C}}(X, Y)$ of *morphisms* from X to Y ;
3. for each triple X, Y, Z , a composition operation

$$\text{Hom}_{\mathcal{C}}(Y, Z) \times \text{Hom}_{\mathcal{C}}(X, Y) \longrightarrow \text{Hom}_{\mathcal{C}}(X, Z), \quad (g, f) \longmapsto g \circ f;$$

4. for every object X , an identity morphism $\text{id}_X \in \text{Hom}_{\mathcal{C}}(X, X)$.

These data satisfy *associativity*,

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

for every composable triple, and the *identity laws*,

$$\text{id}_Y \circ f = f = f \circ \text{id}_X \quad (f : X \rightarrow Y).$$

We write $f : X \rightarrow Y$ for a morphism with source X and target Y , and sometimes write gf for $g \circ f$: first apply f , then g . Objects may initially be thought of as “things,” with morphisms specifying the permitted ways to relate them. A morphism need not be a function between underlying sets, even though it is one in many familiar examples.

Remark 2.2 (Sets and classes). The objects of a category need not form a set. There is no set of all sets in ordinary set theory: unrestricted formation of such collections leads to Russell’s paradox. Indeed, a set R of all sets which do not contain themselves would satisfy $R \in R$ if and only if $R \notin R$. We therefore use the word *class* or *collection* when discussing all sets, all topological spaces, or all groups.

The definition above adopts the convention that each individual $\text{Hom}_{\mathcal{C}}(X, Y)$ is a set; this is called *local smallness*. All categories used here have this property. A category is *small* if its objects also form a set.

Example 2.3. The following are categories, with ordinary composition of maps and ordinary identity maps. In the vector-space example a field k is fixed.

Category	Objects	Morphisms
Top	Topological spaces	Continuous maps
Groups	Groups	Group homomorphisms
Vec_k	Vector spaces over k	k -linear maps
Set	Sets	Functions

For each row, compositions stay within the stated class of maps, identities belong to that class, and the category axioms follow from the corresponding laws for functions.

2.2 The homotopy category

In the first lecture we proved that homotopy is an equivalence relation on the continuous maps from X to Y . Write $[f]$ for the homotopy class of a continuous map $f : X \rightarrow Y$.

Definition 2.4. The *homotopy category* $\mathbf{HTop}$ has topological spaces as objects and

$$\mathrm{Hom}_{\mathbf{HTop}}(X, Y) = \{\text{continuous maps } X \rightarrow Y\} / \simeq$$

as its morphisms. Composition and identities are defined by

$$[g] \circ [f] = [g \circ f], \quad \mathrm{id}_X^{\mathbf{HTop}} = [\mathrm{id}_X].$$

We must verify that composition does not depend on the chosen representatives. We do this in two separate steps.

1. *Replacing f .* If $f \simeq f' : X \rightarrow Y$ and $g : Y \rightarrow Z$ is fixed, then $g \circ f \simeq g \circ f'$. Indeed, if $h : X \times I \rightarrow Y$ is a homotopy from f to f' , then

$$(x, s) \mapsto g(h(x, s))$$

is a homotopy from $g \circ f$ to $g \circ f'$. It is continuous as the composite $g \circ h$.

2. *Replacing g .* If $g \simeq g' : Y \rightarrow Z$ and $f : X \rightarrow Y$ is fixed, then $g \circ f \simeq g' \circ f$. Indeed, if $h : Y \times I \rightarrow Z$ is a homotopy from g to g' , then

$$(x, s) \mapsto h(f(x), s)$$

is a homotopy from $g \circ f$ to $g' \circ f$. It is continuous as the composite $h \circ (f \times \mathrm{id}_I)$.

Consequently, if both representatives are replaced, transitivity gives

$$g \circ f \simeq g \circ f' \simeq g' \circ f'.$$

Thus $[g \circ f] = [g' \circ f']$, so composition is well defined.

The remaining axioms descend directly from composition of continuous maps:

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

and

$$[\mathrm{id}_Y] \circ [f] = [f] = [f] \circ [\mathrm{id}_X].$$

This completes the verification that $\mathbf{HTop}$ is a category. Its objects are the same as those of $\mathbf{Top}$; it is the morphisms which have changed. Two homotopic maps become the same morphism.

2.3 Isomorphisms in a category

Definition 2.5. A morphism $f : X \rightarrow Y$ in a category is an *isomorphism* if there exists a morphism $g : Y \rightarrow X$ such that

$$gf = \mathrm{id}_X, \quad fg = \mathrm{id}_Y.$$

Then g is an *inverse* of f , denoted f^{-1} , and the objects X, Y are *isomorphic*.

For example, isomorphisms in $\mathbf{Top}$ are homeomorphisms, isomorphisms in $\mathbf{Groups}$ are group isomorphisms, and isomorphisms in $\mathbf{Vec}_k$ are invertible linear maps.

Lemma 2.6 (Left and right inverses). *If $f : X \rightarrow Y$ has a left inverse $\ell : Y \rightarrow X$ and a right inverse $r : Y \rightarrow X$, then $\ell = r$ and f is an isomorphism. In particular, the inverse of an isomorphism is unique.*

Proof. The assumptions are $\ell f = \text{id}_X$ and $fr = \text{id}_Y$. Associativity gives

$$\ell = \ell \text{id}_Y = \ell(fr) = (\ell f)r = \text{id}_X r = r.$$

Thus the common morphism is a two-sided inverse. Applying the same argument to any two inverses proves uniqueness. $\square$

Lemma 2.7 (Two out of three). *Let $X \xrightarrow{f} Y \xrightarrow{g} Z$ be composable morphisms. If any two of f , g , and gf are isomorphisms, so is the third.*

Proof. If f and g are isomorphisms, then $(gf)^{-1} = f^{-1}g^{-1}$: composing in either order gives the appropriate identity.

If f and gf are isomorphisms, write $g = (gf)f^{-1}$. This is a composite of isomorphisms, so is an isomorphism; its inverse is $f(gf)^{-1}$.

If g and gf are isomorphisms, write $f = g^{-1}(gf)$. Its inverse is $(gf)^{-1}g$. $\square$

2.4 Homotopy equivalence as isomorphism

We can now return to the motivating question.

Proposition 2.8. *A continuous map $f : X \rightarrow Y$ is a homotopy equivalence if and only if $[f] : X \rightarrow Y$ is an isomorphism in **HTop**.*

Proof. An inverse to $[f]$ is represented by a continuous map $g : Y \rightarrow X$. The inverse identities say

$$[g][f] = [\text{id}_X], \quad [f][g] = [\text{id}_Y].$$

By definition of equality of homotopy classes, these say precisely

$$gf \simeq \text{id}_X, \quad fg \simeq \text{id}_Y.$$

Thus g represents an inverse exactly when it is a homotopy inverse of f . $\square$

This is the promised categorical interpretation: homotopy equivalence is simply isomorphism in **HTop**. The isomorphism lemmas therefore apply to homotopy equivalences. In particular, homotopy equivalences satisfy the two-out-of-three property.

Isomorphic objects are indistinguishable from the point of view of their category. The qualification “in their category” matters: a nonempty convex subset of $\mathbb{R}^n$ and a point are isomorphic in **HTop**, even though they usually are not homeomorphic. Likewise, the punctured plane and the circle are isomorphic in **HTop** by the radial deformation constructed in Lecture 1.

Here is a precise way to use this principle. If $u : X \rightarrow Y$ is an isomorphism in any category and A is an object, then composition with u gives a bijection

$$\text{Hom}_C(A, X) \longrightarrow \text{Hom}_C(A, Y), \quad a \longmapsto u \circ a. \quad (2.1)$$

Its inverse is composition with u^{-1} . We will apply this first with A a point, and then with A a circle, in **HTop**.

2.5 Paths and path components

Definition 2.9. A *path* from x to y in a space X is a continuous map $\gamma : I \rightarrow X$ with $\gamma(0) = x$ and $\gamma(1) = y$. A space is *path-connected* if any two of its points can be joined by a path.

In particular, the empty space is path-connected: the condition is vacuous. Paths can be reversed and joined together. If γ runs from x to y , $\gamma^{-1}(t) = \gamma(1 - t)$ runs from y to x . If δ runs from y to z , their concatenation is

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

The formulas agree at $t = 1/2$, so the pasting lemma (Lemma 1.15) proves continuity. Here $\gamma\delta$ means first traverse γ , then δ .

Lemma 2.10. *The relation “can be joined by a path” is an equivalence relation on X .*

Proof. Constant paths give reflexivity, reversal gives symmetry, and concatenation gives transitivity. $\square$

Definition 2.11. The *path component* of $x \in X$ is

$$C_x = \{y \in X : \text{there is a path from } x \text{ to } y\}.$$

The set of all path components is denoted

$$\pi_0(X) = \{C_x : x \in X\}.$$

As equivalence classes, any two sets C_x, C_y are either equal or disjoint, and together they partition X . Each C_x is itself path-connected: paths from x to points of C_x stay inside C_x , since every intermediate point can be reached by restricting and reparametrizing the path. Reversal and concatenation then join any two points of C_x within C_x .

The notation $\pi_0(X)$ means a *set*, not just the number of its elements. A space is path-connected exactly when this set has at most one element: zero if the space is empty, and one otherwise.

2.6 The invariant π_0

Let $*$ denote a one-point space. A continuous map $* \rightarrow X$ is exactly a choice of a point $x \in X$; write c_x for this map. A homotopy $* \times I \rightarrow X$ between c_x and c_y is exactly a path from x to y . We therefore have a canonical bijection

$$\text{Hom}_{\mathbf{HTop}}(*, X) \longrightarrow \pi_0(X), \quad [c_x] \longmapsto C_x. \quad (2.2)$$

It is well defined and injective because $c_x \simeq c_y$ exactly when $C_x = C_y$, and it is surjective because every path component contains a point.

A continuous map $f : X \rightarrow Y$ sends paths to paths by composition, hence induces a map

$$\pi_0(f) : \pi_0(X) \longrightarrow \pi_0(Y), \quad C_x \longmapsto C_{f(x)}.$$

This does not assert that $f(C_x)$ is the whole component of $f(x)$; it says that $f(C_x)$ is contained in that component. Under (2.2), this construction is composition with $[f]$. It follows that

$$\pi_0(\text{id}_X) = \text{id}_{\pi_0(X)}, \quad \pi_0(gf) = \pi_0(g) \circ \pi_0(f).$$

Moreover, homotopic maps induce the same map on components. Directly, if h is a homotopy from f to g , then $s \mapsto h(x, s)$ is a path from $f(x)$ to $g(x)$, so $C_{f(x)} = C_{g(x)}$.

Proposition 2.12. *A homotopy equivalence $f : X \rightarrow Y$ induces a bijection $\pi_0(f) : \pi_0(X) \rightarrow \pi_0(Y)$.*

Proof. If g is a homotopy inverse, the preceding identities give

$$\pi_0(g)\pi_0(f) = \pi_0(gf) = \pi_0(\text{id}_X), \quad \pi_0(f)\pi_0(g) = \pi_0(fg) = \pi_0(\text{id}_Y).$$

Thus $\pi_0(g)$ is an inverse. Equivalently, apply (2.1) in **HTop** with $A = *$. $\square$

Consequently, homotopy equivalent spaces have the same number of path components, whether finite or infinite. Path-connectedness is a homotopy invariant. In particular, a contractible space is path-connected. The action on maps is part of the invariant: π_0 does more than assign an unrelated set to each space. Assignments preserving identities and composition in this way are called *functors*; thus π_0 can be viewed as a functor from **HTop** to **Set**.

2.7 Connectedness and local path-connectedness

We now make precise the connectedness discussion from Lecture 1. Connectedness concerns separation by open sets; path-connectedness instead asks for actual continuous paths.

Definition 2.13. A space X is *connected* if it cannot be expressed as a union $X = U \cup V$ of two disjoint, nonempty open subsets. Such a decomposition is called a *separation*. A subset of X is connected if it is connected in its subspace topology.

Definition 2.14 (Connected component). A *connected component* of a space X is a nonempty subset $C \subseteq X$ which is *open in X* , *closed in X* , and *connected* in its subspace topology.

Terminology. This is the convention of this course. Standard terminology calls any maximal nonempty connected subset a connected component; such subsets need not be open. Our components are exactly the open standard components. In an arbitrary space, they need not cover all of X .

Proposition 2.15. *Any two connected components of a space either are disjoint or coincide.*

Proof. Let C_1, C_2 be connected components with $C_1 \cap C_2 \neq \emptyset$. Since C_2 is open and closed in X , the sets $C_1 \cap C_2$ and $C_1 \setminus C_2$ are disjoint open subsets of C_1 whose union is C_1 . Connectedness of C_1 forces the second set to be empty, so $C_1 \subseteq C_2$. Interchanging C_1 and C_2 gives the reverse inclusion. Hence $C_1 = C_2$. $\square$

Lemma 2.16. *The continuous image of a connected space is connected.*

Proof. Let $f : X \rightarrow Y$ be continuous, with X connected, and regard f as a map onto the subspace $f(X)$. A separation of $f(X)$ would pull back to two disjoint, nonempty open subsets covering X , a contradiction. $\square$

A homeomorphism therefore preserves connectedness. It also preserves openness, closedness, and nonemptiness, so it sends connected components in our convention bijectively to connected components. Hence their number is a topological invariant.

Intervals in $\mathbb{R}$ are connected. Indeed, if an interval J had a separation $U \cup V$, the function equal to 0 on U and 1 on V would be continuous, since it is constant on an open neighbourhood of each point. Between a point of U and a point of V , this contradicts the intermediate value theorem.

For example, the continuous surjection

$$(0, 2\pi) \longrightarrow S^1 \setminus \{(1, 0)\}, \quad t \longmapsto (\cos t, \sin t)$$

shows that the punctured circle is connected. Rotating the circle gives the same conclusion for any deleted point. This justifies the geometric observation in Example 1.12.

Proposition 2.17. *Every path-connected space is connected.*

Proof. Suppose a path-connected space X had a separation $X = U \cup V$. Choose $x \in U$, $y \in V$, and a path γ from x to y . Since I is connected, Lemma 2.16 implies that $\gamma(I)$ is connected. But its intersections with U and V are disjoint nonempty relatively open subsets covering $\gamma(I)$, a contradiction. $\square$

The converse is false: a connected space need not be path-connected.

Example 2.18 (The topologist's sine curve). Let

$$G = \{(x, \sin(1/x)) : 0 < x \leq 1\}, \quad V = \{0\} \times [-1, 1], \quad T = G \cup V.$$

The graph oscillates increasingly rapidly as x approaches zero, accumulating on the whole vertical interval V . The space T is connected but not path-connected.

Connectedness. The graph G is the continuous image of the connected interval $(0, 1]$, so G is connected. Its closure in $\mathbb{R}^2$ is exactly T . Indeed, for each $y \in [-1, 1]$, choose θ with $\sin \theta = y$. For all sufficiently large n ,

$$x_n = \frac{1}{\theta + 2\pi n} \in (0, 1], \quad (x_n, \sin(1/x_n)) = (x_n, y) \longrightarrow (0, y).$$

Thus every point of V is a limit point of G . There are no other new limit points: for a positive limiting first coordinate, continuity of $\sin(1/x)$ keeps the limit on G , and all second coordinates lie in the closed interval $[-1, 1]$.

The closure of a connected set is connected. To see this, a separation of the closure would have two nonempty relatively open sides, both of which meet the dense original set; their intersections would separate that set. Hence $T = \overline{G}$ is connected.

Failure of path-connectedness. Suppose a path $\gamma : I \rightarrow T$, written $\gamma(t) = (x(t), y(t))$, starts in V and ends in G . Let

$$a = \max\{t \in I : x(t) = 0\}.$$

This maximum exists because the set is nonempty and closed in I . We have $a < 1$ and $x(t) > 0$ for every $t > a$.

By continuity of y at a , choose b with $a < b \leq 1$ so close to a that

$$|y(t) - y(a)| < \frac{1}{2} \quad (a \leq t \leq b).$$

But $x(a) = 0$ and $x(b) > 0$. The intermediate value theorem implies that x takes every value between 0 and $x(b)$ on $[a, b]$. Among these positive values are numbers u, v as small as desired with $\sin(1/u) = 1$ and $\sin(1/v) = -1$. At times when $x(t)$ equals u or v , the path lies in G , so $y(t)$ equals 1 or -1 . Both values cannot be within $1/2$ of $y(a)$, a contradiction.

Each of G and V is path-connected: G is homeomorphic to $(0, 1]$ by projection onto the first coordinate, and V is an interval. They are therefore precisely the two path components of T . On the other hand, T itself is its single connected component. This illustrates the difference between the two notions.

Definition 2.19. A space X is *locally path-connected* if for every $x \in X$ and every open neighbourhood U of x , there is a path-connected open subset V of X with $x \in V \subseteq U$.

Thus each point must have arbitrarily small path-connected open neighbourhoods. Having just one such neighbourhood is insufficient: a path-connected space always provides a large one, namely itself. Euclidean spaces are locally path-connected because open balls form a neighbourhood basis and are convex.

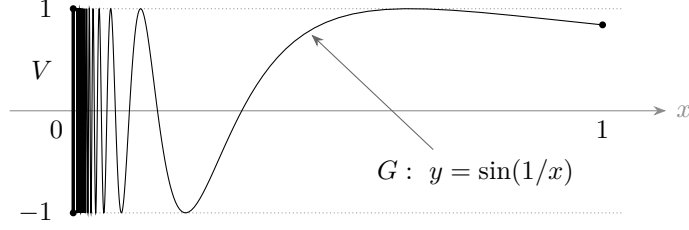


Figure 3: The topologist's sine curve includes the limiting vertical interval. The oscillations continue arbitrarily close to $x = 0$; only a finite portion of the graph is drawn.

Example 2.20 (The comb). Consider the subspace

$$K = (I \times \{0\}) \cup (\{0\} \times I) \cup \bigcup_{n \geq 1} (\{1/n\} \times I) \subset \mathbb{R}^2.$$

This is a horizontal base with vertical teeth at 0 and at each $1/n$. It is path-connected: travel vertically down to the base, along the base, and vertically up to the desired point. Each segment stays inside K , and concatenating the segments gives a continuous path.

It is not locally path-connected at any $p = (0, a)$ with $0 < a \leq 1$. Take the open neighbourhood

$$U = K \cap \{(u, v) \in \mathbb{R}^2 : v > a/2\}.$$

A path lying in U must have constant first coordinate. Indeed, that coordinate takes values only in $\{0\} \cup \{1/n : n \geq 1\}$, which contains no interval of positive length. The intermediate value theorem therefore forces it to be constant. So no path in U joins p to a point of a tooth at $1/n$. But every open neighbourhood V of p in K contains $(1/n, a)$ for sufficiently large n , since these points converge to p . Consequently no such $V \subseteq U$ can be path-connected.

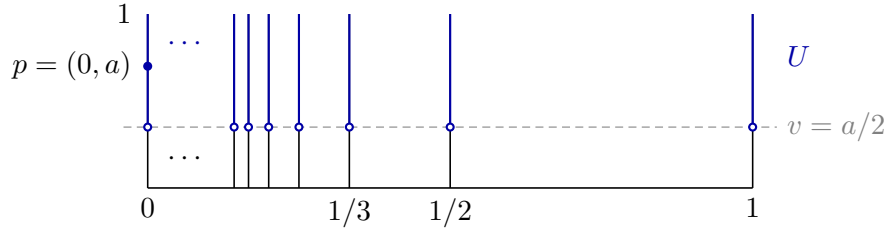


Figure 4: The comb K : the dots indicate the remaining teeth at $1/n$, accumulating at the included vertical segment over 0. Blue shows U ; the hollow markers belong to K but not to U . The dashed horizontal line is only a guide, not part of the comb.

Proposition 2.21. *In a locally path-connected space, path components are open and closed, and they coincide with connected components.*

Proof. Let C be a path component. For each $x \in C$, local path-connectedness gives a path-connected open neighbourhood V_x of x . Every point of V_x can be joined to x , so $V_x \subseteq C$. Thus C is open. All path components are open, and they partition X , so $X \setminus C$ is a union of open components. Hence C is also closed.

The set C is nonempty and connected, so it is a connected component by our definition.

Conversely, let K be a connected component and choose $x \in K$. Let C be the path component of x . Since K is open and closed, connectedness of C forces $C \subseteq K$. Since C is open and closed, connectedness of K forces $K \subseteq C$. Hence $K = C$. The path components are therefore exactly the connected components. $\square$

In particular, a connected, locally path-connected space is path-connected. The local hypothesis is essential to this argument. For locally path-connected spaces, the connected components in our convention also agree with the standard connected components: the path components partition the space into connected open-and-closed subsets, each of which is maximal connected.

2.8 Free loops and simple connectedness

The next object we would like to understand is π_1 . Following the construction of π_0 , it is natural to look at

$$\mathrm{Hom}_{\mathbf{HTop}}(S^1, X).$$

Its elements are *free homotopy classes of loops*: a loop here is a continuous map $S^1 \rightarrow X$, and homotopies impose no condition on a chosen point of the circle. This is a useful invariant, but this construction alone gives a set, without the specified basepoint and multiplication needed for the fundamental group.

Before imposing a basepoint, we can formulate simple connectedness.

Definition 2.22. A space X is *simply connected* if it is path-connected and every continuous map $f : S^1 \rightarrow X$ is homotopic to a constant map. A map homotopic to a constant map is called *null-homotopic*.

The homotopies in this definition are free homotopies. The path-connectedness requirement is part of the definition. Connectedness alone does not substitute for it in arbitrary spaces. The empty space is simply connected: it is path-connected and has no loops, so the condition on loops is vacuous.

Proposition 2.23. *A space X is simply connected if and only if $\mathrm{Hom}_{\mathbf{HTop}}(S^1, X)$ has at most one element.*

Proof. If $X = \emptyset$, both conditions hold. Assume now that $X \neq \emptyset$.

Suppose X is simply connected. Every loop is homotopic to a constant. Any two constant maps $c_x, c_y : S^1 \rightarrow X$ are homotopic: if γ is a path from x to y , use $h(z, s) = \gamma(s)$. Thus all loops represent the same class, and at least one loop exists because X is nonempty.

Conversely, suppose the Hom-set has at most one element. Since X is nonempty, a constant map exists, so it has exactly one element. For any $x, y \in X$, the constant maps c_x and c_y are homotopic. Evaluating such a homotopy at a fixed point of S^1 gives a path from x to y , so X is path-connected. Every loop is homotopic to a constant because there is just one homotopy class. $\square$

Remark 2.24. For every nonempty X a constant map $S^1 \rightarrow X$ exists, so this Hom-set cannot have zero elements. If $X = \emptyset$, the Hom-set is empty. Thus a simply connected space has zero free homotopy classes of loops if it is empty, and exactly one otherwise.

Corollary 2.25. *Simple connectedness is a homotopy invariant.*

Proof. A homotopy equivalence $X \simeq Y$ induces a bijection

$$\mathrm{Hom}_{\mathbf{HTop}}(S^1, X) \longrightarrow \mathrm{Hom}_{\mathbf{HTop}}(S^1, Y)$$

by composition, as in (2.1). One of these sets has at most one element if and only if the other does. $\square$

Example 2.26. A point is simply connected. Therefore every contractible space is simply connected. In particular, every nonempty convex subset $C \subseteq \mathbb{R}^n$ is simply connected. Directly, line segments join its points, and a loop $f : S^1 \rightarrow C$ contracts to a chosen $x_0 \in C$ by

$$h(z, s) = (1 - s)f(z) + sx_0.$$

Remark 2.27 (Results announced for later proof). The circle S^1 is not simply connected, and consequently neither is $\mathbb{R}^2 \setminus \{0\}$, since it is homotopy equivalent to S^1 . The higher-dimensional spheres S^n , $n \geq 2$, are simply connected. These statements are announced here; their proofs require further tools and are not supplied in this lecture.

2.9 Based loops and $\pi_1(X, x_0)$ as a set

To prepare for the fundamental group, we choose a point $x_0 \in X$. We use the complex model

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

with distinguished point $1 \in S^1$.

Definition 2.28. A *based loop* in (X, x_0) is a continuous map $f : S^1 \rightarrow X$ with $f(1) = x_0$. Two based loops f, g are *based homotopic* if there is a continuous map $h : S^1 \times I \rightarrow X$ with

$$h(z, 0) = f(z), \quad h(z, 1) = g(z), \quad h(1, s) = x_0 \quad \text{for all } s \in I.$$

In the movie interpretation, every frame must remain a based loop: the image of 1 stays fixed at x_0 throughout. This is stronger than merely requiring the first and last frames to be based. During a free homotopy, that distinguished point is allowed to move.

Based homotopy is an equivalence relation. The stationary, reversed, and concatenated homotopies used in Proposition 1.14 all preserve the condition $h(1, s) = x_0$ when the original homotopies do. We may therefore form the quotient set

$$\pi_1(X, x_0) = \{f : S^1 \rightarrow X \text{ continuous} : f(1) = x_0\} / \text{based homotopy}. \quad (2.3)$$

The constant based loop $c_{x_0}(z) = x_0$ supplies a distinguished element. A based loop f also gives a path

$$\gamma_f(t) = f(e^{2\pi it}), \quad 0 \leq t \leq 1,$$

which starts and ends at x_0 .

Forgetting the basepoint condition gives a well-defined map

$$\pi_1(X, x_0) \longrightarrow \mathrm{Hom}_{\mathbf{HTop}}(S^1, X), \quad [f]_{\text{based}} \longmapsto [f]_{\text{free}},$$

because every based homotopy is a free homotopy. The two quotients use different equivalence relations and should not be identified without further argument.

At this stage, $\pi_1(X, x_0)$ has been defined as a set of based homotopy classes. A multiplication and the proof of the group axioms remain to be constructed.

3 The fundamental group, coverings, and degree

Lecture 3, 8 October 2026

3.1 Based loops and their multiplication

We now give the set $\pi_1(X, x_0)$ from Lecture 2 a group structure. Recall that a based loop was defined as a continuous map

$$f : S^1 \longrightarrow X, \quad f(1) = x_0,$$

We identify the circle with the interval whose endpoints are glued, using the homeomorphism

$$I/(0 \sim 1) \longrightarrow S^1, \quad [t] \longmapsto e^{2\pi it}.$$

The identified endpoints $[0] = [1]$ correspond to the point $1 \in S^1$.

From now on, we write a based loop as a continuous map

$$\gamma : I \longrightarrow X, \quad \gamma(0) = \gamma(1) = x_0.$$

A *based homotopy* from γ to δ is a continuous map $h : I \times I \rightarrow X$ satisfying

$$h(t, 0) = \gamma(t), \quad h(t, 1) = \delta(t), \quad h(0, s) = h(1, s) = x_0.$$

We write $\gamma \simeq_* \delta$. The horizontal coordinate t runs along the loop; the vertical coordinate s is homotopy time. Both vertical sides of the parameter square must stay at x_0 .

For two based loops, define their *concatenation* by

$$(\gamma\delta)(t) = \begin{cases} \gamma(2t), & 0 \leq t \leq \frac{1}{2}, \\ \delta(2t - 1), & \frac{1}{2} \leq t \leq 1. \end{cases} \quad (3.1)$$

Thus $\gamma\delta$ travels along γ first and δ second. The formulas agree at $t = \frac{1}{2}$, since both give x_0 . The pasting lemma therefore proves continuity.

Proposition 3.1. *The rule $[\gamma][\delta] = [\gamma\delta]$ is a well-defined operation on $\pi_1(X, x_0)$.*

Proof. We check independence of the two representatives separately.

1. *Replacing γ .* Suppose $\gamma \simeq_* \gamma'$ and fix δ . If h is a based homotopy from γ to γ' , then

$$(t, s) \longmapsto \begin{cases} h(2t, s), & 0 \leq t \leq \frac{1}{2}, \\ \delta(2t - 1), & \frac{1}{2} \leq t \leq 1 \end{cases}$$

is a based homotopy from $\gamma\delta$ to $\gamma'\delta$.

2. *Replacing δ .* Suppose $\delta \simeq_* \delta'$ and fix γ . If h is a based homotopy from δ to δ' , then

$$(t, s) \longmapsto \begin{cases} \gamma(2t), & 0 \leq t \leq \frac{1}{2}, \\ h(2t - 1, s), & \frac{1}{2} \leq t \leq 1 \end{cases}$$

is a based homotopy from $\gamma\delta$ to $\gamma\delta'$.

In each case the formulas agree along $t = \frac{1}{2}$, where both equal x_0 . Continuity follows by pasting on the two closed rectangles, and both vertical sides stay at x_0 . If both representatives change, transitivity gives

$$\gamma\delta \simeq_* \gamma'\delta \simeq_* \gamma'\delta'.$$

□

3.2 The group axioms as square homotopies

Let $e(t) = x_0$ be the constant loop and let $\gamma^{-1}(t) = \gamma(1-t)$ be the reversed loop. Concatenation changes the speed of travel and introduces pauses. The group laws assert that these changes disappear on based homotopy classes.

The identity. The loop $e\gamma$ waits at x_0 for half the time and then traverses γ . Shrink the waiting interval from length $\frac{1}{2}$ to zero. Put $a(s) = (1-s)/2$ and define

$$h(t, s) = \begin{cases} x_0, & 0 \leq t \leq a(s), \\ \gamma\left(\frac{t-a(s)}{1-a(s)}\right), & a(s) \leq t \leq 1. \end{cases} \quad (3.2)$$

The denominator is always at least $1/2$. The regions are closed and the formulas agree on their common boundary, so h is continuous. Its bottom edge is $e\gamma$, its top edge is γ , and its two vertical sides have value x_0 . Thus $e\gamma \simeq_* \gamma$.

For the right identity, put $b(s) = (1+s)/2$ and use

$$h(t, s) = \begin{cases} \gamma(t/b(s)), & 0 \leq t \leq b(s), \\ x_0, & b(s) \leq t \leq 1. \end{cases}$$

This removes the pause at the end and gives $\gamma e \simeq_* \gamma$.

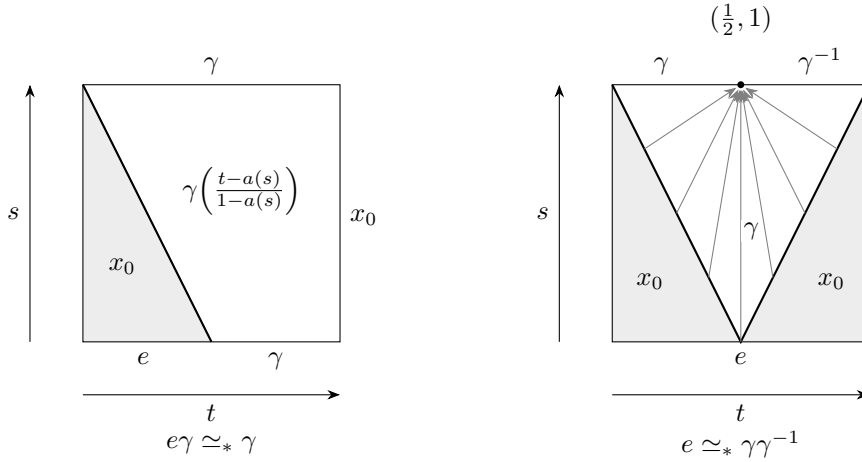


Figure 5: Parameter squares for the identity and inverse laws, redrawn from the blackboard constructions. Read each horizontal slice as a loop. In the inverse square, each arrow carries γ towards $(\frac{1}{2}, 1)$; the two outer triangles are constant at x_0 .

Inverses. We construct a based homotopy from e to $\gamma\gamma^{-1}$. In the parameter square, draw the triangle with vertices $(0, 1)$, $(\frac{1}{2}, 0)$, and $(1, 1)$. Put x_0 on the two outer right triangles. Fill the middle triangle by placing γ on every segment from its sloping boundary to $(\frac{1}{2}, 1)$, with parameter 0 at the boundary and 1 at $(\frac{1}{2}, 1)$.

A formula for this construction is

$$h(t, s) = \gamma(\max\{0, s - |2t - 1|\}). \quad (3.3)$$

Inside the middle triangle, $s - |2t - 1|$ runs linearly from 0 to 1 along each such segment; outside it the maximum is 0. The formula is continuous and satisfies

$$h(t, 0) = x_0, \quad h(t, 1) = \gamma(1 - |2t - 1|) = (\gamma\gamma^{-1})(t),$$

as well as $h(0, s) = h(1, s) = x_0$. Hence $e \simeq_* \gamma\gamma^{-1}$. Applying the same construction to γ^{-1} , with $(\gamma^{-1})^{-1} = \gamma$, gives $e \simeq_* \gamma^{-1}\gamma$. Also, reversing a based homotopy h by replacing t with $1 - t$ shows that $[\gamma^{-1}]$ depends only on $[\gamma]$.

Associativity. The loops $(\gamma_1\gamma_2)\gamma_3$ and $\gamma_1(\gamma_2\gamma_3)$ traverse the same three loops in the same order. Their time allocations differ:

	γ_1	γ_2	γ_3
$(\gamma_1\gamma_2)\gamma_3$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
$\gamma_1(\gamma_2\gamma_3)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$

Move the two breakpoints linearly, setting $a(s) = (1 + s)/4$ and $b(s) = (2 + s)/4$. Define

$$h(t, s) = \begin{cases} \gamma_1(t/a(s)), & 0 \leq t \leq a(s), \\ \gamma_2((t - a(s))/(b(s) - a(s))), & a(s) \leq t \leq b(s), \\ \gamma_3((t - b(s))/(1 - b(s))), & b(s) \leq t \leq 1. \end{cases} \quad (3.4)$$

All denominators are positive. On each seam both formulas give x_0 , so the finite closed-cover pasting lemma proves continuity. The bottom and top edges give the two parenthesizations, and the vertical sides stay at x_0 . This is the required based homotopy.

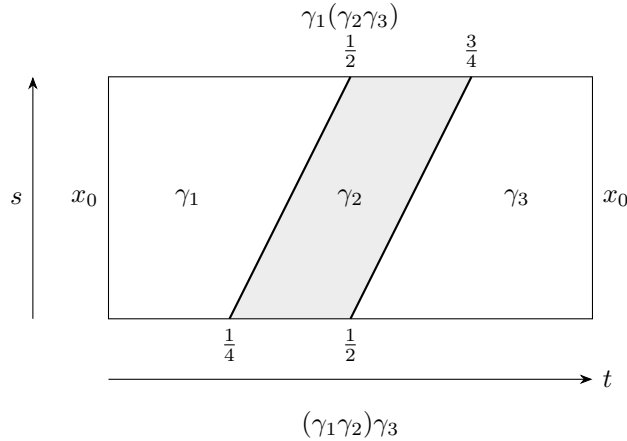


Figure 6: Associativity changes the time assigned to each loop. Within each region, traverse the indicated loop once as t crosses that region.

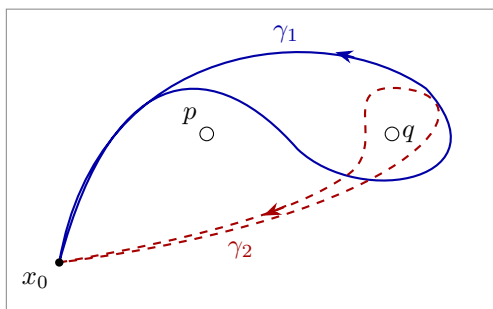
Theorem 3.2. *For every pointed space (X, x_0) , concatenation makes $\pi_1(X, x_0)$ into a group, called the fundamental group. Its identity is the class of the constant loop, and $[\gamma]^{-1} = [\gamma^{-1}]$.*

Proof. The operation is well defined by Proposition 3.1. The square homotopies above prove both identity laws, both inverse laws, and associativity. $\square$

3.3 Why free homotopy does not suffice

Consider the loops γ_1, γ_2 in the rectangle with punctures p, q shown in Figure 7. They are freely homotopic: if the basepoint may move, each can be deformed to the same loop around q .

Convince yourself from the picture that $\gamma_1\gamma_2^{-1}$ is not homotopic to a constant map. In contrast, $\gamma_2\gamma_2^{-1} \simeq_* e$. Thus replacing γ_1 by the freely homotopic loop γ_2 changes the homotopy class of the product. This is why we use based homotopies to define multiplication.



$\gamma_1 \simeq \gamma_2$ freely, but
 $\gamma_1 \gamma_2^{-1} \neq e$.

Figure 7: Two freely homotopic loops based at x_0 in a twice-punctured rectangle. Follow the arrows for their orientations.

The picture-hanging problem. How can a picture be hung by a string over two nails so that it stays up with both nails in place, but falls if either one is removed? Use $\gamma_1 \gamma_2^{-1}$ as a guide: check that this loop becomes homotopic to a constant when either puncture is filled in.

3.4 Covering maps: local copies of the base

Our next goal is the computation

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

We begin with the map

$$f : \mathbb{R} \longrightarrow S^1, \quad f(x) = (\cos(2\pi x), \sin(2\pi x)) = e^{2\pi i x}. \quad (3.5)$$

One complete positive turn corresponds to an increase of 1 in the real coordinate. Although f is not globally one-to-one, a small arc of the circle has disjoint copies in $\mathbb{R}$, one in each turn. This is the local picture of a covering map.

Definition 3.3. A continuous map $f : X \rightarrow Y$ is a *covering map* if every $y \in Y$ has an open neighborhood U such that

$$f^{-1}(U) = \coprod_{\alpha \in A} V_\alpha,$$

where the V_α are pairwise disjoint open subsets of X and each restriction $f|_{V_\alpha} : V_\alpha \rightarrow U$ is a homeomorphism. The index set A may be empty; surjectivity is not required. The space Y is the *base*, X is the *cover* or *covering space*, and U is an *evenly covered neighborhood*. The subsets V_α are its *sheets*.

To read the definition, first choose one small open set U downstairs. Then examine its *entire* inverse image upstairs. This inverse image must split into disjoint open pieces, each mapped bijectively and homeomorphically onto the whole of U . After identifying each piece with U by this homeomorphism, the restriction of f becomes the identity. The sheets are local: they need not extend to disjoint copies of all of Y inside X .

Example 3.4 (The real line covering the circle). Choose $a \in \mathbb{R}$ and $0 < \varepsilon < 1/2$. The image $U = f((a - \varepsilon, a + \varepsilon))$ is an open arc containing $f(a)$, and

$$f^{-1}(U) = \coprod_{n \in \mathbb{Z}} (a - \varepsilon + n, a + \varepsilon + n).$$

On each interval, f is a homeomorphism onto U : its inverse chooses the continuous angle along that arc, in the indicated interval. These intervals are disjoint because their length is less than 1. Thus (3.5) is a covering map. Notice that $f(x) = f(x')$ if and only if $x - x' \in \mathbb{Z}$.

Example 3.5 (Two disks over one disk). Let $D = \{u \in \mathbb{R}^2 : \|u\| < 1\}$ and give $\{1, 2\}$ the discrete topology. Then

$$f : D \times \{1, 2\} \longrightarrow D, \quad f(u, j) = u$$

is a covering. We may take $U = D$ itself: its inverse image has the two open sheets $D \times \{1\}$ and $D \times \{2\}$. This example consists of two separate copies of the whole base.

Example 3.6 (The double covering of the circle). The map $f : S^1 \rightarrow S^1$, $f(z) = z^2$, is a covering with two sheets over each small arc. To see this, write that arc as

$$U = \{e^{2\pi it} : a - \varepsilon < t < a + \varepsilon\}, \quad 0 < \varepsilon < 1/2.$$

Its inverse image is the disjoint union of the arcs

$$V_0 = \{e^{\pi it} : a - \varepsilon < t < a + \varepsilon\}, \quad V_1 = -V_0.$$

The two inverse branches are $e^{2\pi it} \mapsto e^{\pi it}$ and $e^{2\pi it} \mapsto -e^{\pi it}$, using the unique t in the chosen interval. Each branch is continuous. In contrast to the two-disk example, this covering space is connected: the local sheets belong to a single circle.

Example 3.7 (A projection which is not a covering). The projection $p : \mathbb{R}^2 \rightarrow \mathbb{R}$, $p(u, v) = u$, is not a covering map. For a nonempty open interval $U \subset \mathbb{R}$, the inverse image $p^{-1}(U) = U \times \mathbb{R}$ is connected but is not homeomorphic to U .

3.5 Lifting paths and homotopies

A *lift* of a map $g : Z \rightarrow Y$ through $f : X \rightarrow Y$ is a continuous map $\tilde{g} : Z \rightarrow X$ such that $f \circ \tilde{g} = g$. It makes the same motion upstairs that g describes downstairs. For the exponential covering, lifting means choosing a continuous real-valued angle, measured in turns.

Theorem 3.8 (Homotopy lifting). *Let $f : X \rightarrow Y$ be a covering map, let $h : Z \times I \rightarrow Y$ be a continuous homotopy, and let $\alpha : Z \rightarrow X$ be a continuous initial lift, so that $f(\alpha(z)) = h(z, 0)$. There is a unique continuous map $\tilde{h} : Z \times I \rightarrow X$ such that*

$$f \circ \tilde{h} = h, \quad \tilde{h}(z, 0) = \alpha(z).$$

The proof will be given in a later lecture.

The data and the conclusion have different roles. The homotopy h is prescribed downstairs. At time $s = 0$, α tells us which point above $h(z, 0)$ to use for every z . The theorem says that these choices determine the entire motion upstairs, continuously and uniquely. It does not assert that every map $Z \rightarrow Y$ admits a lift: the initial lift is part of the hypothesis.

Corollary 3.9 (Path lifting). *If $\gamma : I \rightarrow Y$ is a path and $x \in X$ satisfies $f(x) = \gamma(0)$, then there is a unique path $\tilde{\gamma} : I \rightarrow X$ such that*

$$f \circ \tilde{\gamma} = \gamma, \quad \tilde{\gamma}(0) = x.$$

Proof. Take Z to be a one-point space in the homotopy lifting theorem. □

3.6 Degree of a loop in S^1

Return to the covering $f : \mathbb{R} \rightarrow S^1$, $f(x) = e^{2\pi ix}$. Let $\gamma : I \rightarrow S^1$ be a based loop at 1. Path lifting gives a unique $\tilde{\gamma} : I \rightarrow \mathbb{R}$ with

$$\tilde{\gamma}(0) = 0, \quad e^{2\pi i \tilde{\gamma}(t)} = \gamma(t).$$

Since $\gamma(1) = 1$, its endpoint lies in $f^{-1}(1) = \mathbb{Z}$.

Definition 3.10. The *degree*, or *winding number*, of γ is

$$\deg(\gamma) = \tilde{\gamma}(1) \in \mathbb{Z},$$

where the lift starts at 0.

The real coordinate records the total angle without reducing it modulo one turn. The constant loop has degree 0. The positive loop $t \mapsto e^{2\pi it}$ has lift $t \mapsto t$ and degree 1; the negative loop $t \mapsto e^{-2\pi it}$ has degree -1 .

For a loop with any starting point on the circle, choose a lift starting at any real preimage of that point and set

$$\deg(\gamma) = \tilde{\gamma}(1) - \tilde{\gamma}(0). \quad (3.6)$$

The difference is an integer, because the two endpoints have the same image under f . It is independent of the chosen starting lift: two possible starting points differ by an integer n , and path uniqueness says that the corresponding lifts differ by n everywhere. Subtracting endpoints removes that common integer. For a based loop at 1, this agrees with Definition 3.10.

3.7 Homotopy invariance of degree

Proposition 3.11. *Based homotopic loops in $(S^1, 1)$ have the same degree.*

Proof. Let $h : I \times I \rightarrow S^1$ be a based homotopy from γ_1 to γ_2 , so

$$h(t, 0) = \gamma_1(t), \quad h(t, 1) = \gamma_2(t), \quad h(0, s) = h(1, s) = 1.$$

Now view h as a homotopy in the t -direction, with s as the parameter. Apply homotopy lifting to $(s, t) \mapsto h(t, s)$, with $Z = I$ and the constant initial lift $\alpha(s) = 0$. Writing the lifted map in the original coordinate order gives a continuous map $\tilde{h} : I \times I \rightarrow \mathbb{R}$ satisfying

$$e^{2\pi i \tilde{h}(t, s)} = h(t, s), \quad \tilde{h}(0, s) = 0.$$

Since $h(1, s) = 1$, the function

$$s \mapsto \tilde{h}(1, s)$$

is continuous and takes values in $\mathbb{Z}$, hence is constant. For each s , the path $t \mapsto \tilde{h}(t, s)$ is the unique lift starting at 0. Therefore

$$\deg(\gamma_1) = \tilde{h}(1, 0) = \tilde{h}(1, 1) = \deg(\gamma_2). \quad \square$$

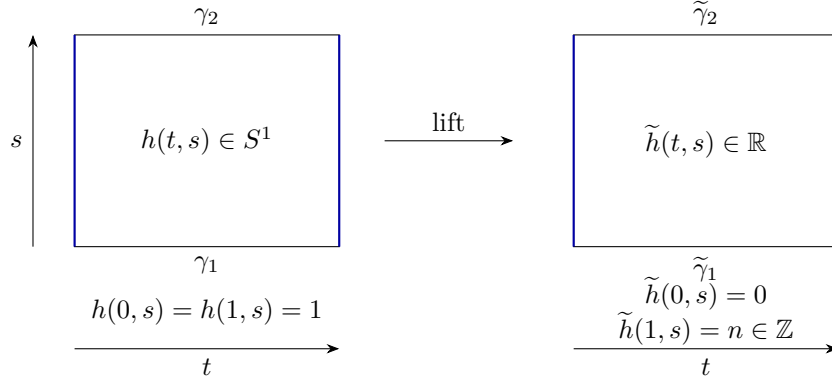


Figure 8: Lift in the t -direction from the constant initial lift on the left edge. The right edge is integer-valued and therefore constant.

The distinction in Figure 8 is essential: a loop has equal endpoint values, but its lift need not. The lifted endpoints can differ by a nonzero integer. For example, the loop $\gamma(t) = e^{2\pi it}$ lifts to $\tilde{\gamma}(t) = t$, whose endpoint values are 0 and 1.

We have therefore constructed a well-defined map of sets

$$\deg : \pi_1(S^1, 1) \longrightarrow \mathbb{Z}, \quad [\gamma] \longmapsto \deg(\gamma).$$

This is the endpoint of the lecture. To complete the announced isomorphism $\pi_1(S^1, 1) \cong \mathbb{Z}$, it remains to show that this map respects the group operation and is bijective.

4 The circle, functoriality, and change of basepoint

Lecture 4, 14 October 2026

4.1 The fundamental group of the circle

In Lecture 3 we constructed the degree map

$$\deg : \pi_1(S^1, 1) \longrightarrow \mathbb{Z}.$$

For a loop $\gamma : I \rightarrow S^1$ based at 1, its degree is the endpoint of the unique lift $\tilde{\gamma} : I \rightarrow \mathbb{R}$ starting at 0, under the covering $f(x) = e^{2\pi i x}$. Proposition 3.11 shows that this depends only on the based homotopy class of γ .

Theorem 4.1. *The degree map is an isomorphism of groups*

$$\deg : \pi_1(S^1, 1) \xrightarrow{\cong} \mathbb{Z}.$$

Proof. Surjectivity. For $n \in \mathbb{Z}$, consider

$$\gamma_n(t) = e^{2\pi i n t}, \quad t \in I.$$

The path $t \mapsto nt$ starts at 0 and satisfies $f(nt) = \gamma_n(t)$, so it is the required lift. Its endpoint is n ; hence $\deg(\gamma_n) = n$. The constant loop has degree 0.

Compatibility with multiplication. Let γ_1, γ_2 be based loops, and let $\tilde{\gamma}_1, \tilde{\gamma}_2$ be their lifts starting at 0. Recall that

$$(\gamma_1\gamma_2)(t) = \begin{cases} \gamma_1(2t), & 0 \leq t \leq \frac{1}{2}, \\ \gamma_2(2t-1), & \frac{1}{2} \leq t \leq 1. \end{cases}$$

Its lift starting at 0 is

$$\widetilde{\gamma_1\gamma_2}(t) = \begin{cases} \tilde{\gamma}_1(2t), & 0 \leq t \leq \frac{1}{2}, \\ \deg(\gamma_1) + \tilde{\gamma}_2(2t-1), & \frac{1}{2} \leq t \leq 1. \end{cases}$$

At $t = \frac{1}{2}$ both expressions equal $\deg(\gamma_1)$, so the pieces paste continuously. Adding the integer $\deg(\gamma_1)$ does not change the image under f , which verifies the lifting identity on the second half. Taking the endpoint gives

$$\deg(\gamma_1\gamma_2) = \deg(\gamma_1) + \deg(\gamma_2).$$

Thus $\deg$ is a group homomorphism.

Injectivity. It remains to show that its kernel is trivial. Suppose $\deg(\gamma) = 0$. Then the lift satisfies $\tilde{\gamma}(0) = \tilde{\gamma}(1) = 0$, so it is itself a based loop in $\mathbb{R}$. Contract it to 0 inside $\mathbb{R}$, keeping both endpoints at 0. Composing this contraction with f gives a based homotopy from γ to the constant loop at 1. Hence $[\gamma]$ is the identity, proving injectivity. $\square$

4.2 Application: the Brouwer fixed-point theorem

Write $D^2 = \{x \in \mathbb{R}^2 : \|x\| \leq 1\}$ for the closed unit disk. The following application was also the subject of Sheet 1, Problem 4.

Theorem 4.2 (Brouwer fixed-point theorem in dimension two). *Every continuous map $g : D^2 \rightarrow D^2$ has a fixed point.*

Proof sketch. First, there can be no retraction $r : D^2 \rightarrow S^1$. Indeed, the loop $\gamma(t) = e^{2\pi it}$ contracts to 1 inside the disk by

$$h(t, s) = (1 - s)\gamma(t) + s.$$

This fixes its endpoints. If r were a retraction, $r \circ h$ would contract the same loop inside S^1 , contradicting its degree 1.

Now suppose that g has no fixed point. For each $x \in D^2$, draw the ray from $g(x)$ through x , and let $r(x)$ be its intersection with S^1 beyond or at x . Since $g(x) \neq x$, this construction is well defined and varies continuously with x . If $x \in S^1$, then $r(x) = x$. Thus r would be a retraction, a contradiction. $\square$

4.3 Functors

In the proof of Theorem 4.1, composing with the covering map carried a contraction in $\mathbb{R}$ to a contraction in S^1 . More generally, maps between spaces induce maps between their algebraic invariants. The notion of a functor expresses the compatibility of these constructions with composition.

Definition 4.3. Let $\mathcal{C}, \mathcal{D}$ be categories. A *functor* $F : \mathcal{C} \rightarrow \mathcal{D}$ assigns an object $F(X)$ to every object X , and a morphism $F(f) : F(X) \rightarrow F(Y)$ to every morphism $f : X \rightarrow Y$, such that

$$F(\text{id}_X) = \text{id}_{F(X)}, \quad F(g \circ f) = F(g) \circ F(f)$$

whenever f, g are composable.

Lemma 4.4. A functor takes isomorphic objects to isomorphic objects.

Proof. If $f : X \rightarrow Y$ and $g : Y \rightarrow X$ are mutually inverse, then

$$F(g) \circ F(f) = F(g \circ f) = \text{id}_{F(X)}, \quad F(f) \circ F(g) = F(f \circ g) = \text{id}_{F(Y)}.$$

Thus $F(f)$ and $F(g)$ are mutually inverse. $\square$

4.4 The fundamental group as a functor

Let $f : X \rightarrow Y$ be continuous with $f(x_0) = y_0$. A loop γ based at x_0 gives the loop $f \circ \gamma$ based at y_0 .

Proposition 4.5. Composition with f induces a group homomorphism

$$f_* : \pi_1(X, x_0) \longrightarrow \pi_1(Y, y_0), \quad [\gamma] \longmapsto [f \circ \gamma].$$

Proof. If h is a based homotopy from γ_1 to γ_2 , then $f \circ h$ is a based homotopy from $f \circ \gamma_1$ to $f \circ \gamma_2$. Thus f_* is well defined. The identity

$$f \circ (\gamma_1 \gamma_2) = (f \circ \gamma_1)(f \circ \gamma_2)$$

shows that it preserves multiplication, and the constant loop maps to the constant loop. $\square$

Definition 4.6. The category **PTop** of *pointed topological spaces* has objects (X, x_0) , where X is a topological space and $x_0 \in X$. A morphism $(X, x_0) \rightarrow (Y, y_0)$ is a continuous map $f : X \rightarrow Y$ with $f(x_0) = y_0$. Identities and composition are those of continuous maps.

For composable pointed maps f, g , and for every loop γ ,

$$(g \circ f)_*([\gamma]) = [g \circ f \circ \gamma] = g_*(f_*([\gamma])).$$

Also $(\text{id}_X)_* = \text{id}_{\pi_1(X, x_0)}$. Consequently, the assignments $(X, x_0) \mapsto \pi_1(X, x_0)$ and $f \mapsto f_*$ define a functor

$$\pi_1 : \mathbf{PTop} \longrightarrow \mathbf{Groups}.$$

4.5 Changing the basepoint

We now compare fundamental groups at different basepoints.

Theorem 4.7. *If $x_1, x_2 \in X$ are joined by a path, then*

$$\pi_1(X, x_1) \cong \pi_1(X, x_2).$$

Warning. *Which isomorphism?* The theorem asserts that the groups are isomorphic, but does not single out an isomorphism. We will construct one by choosing a path, and different paths can give different isomorphisms. In Figure 10, the small loop γ around the rightmost puncture is evident at x_1 . To view it from x_2 , we must choose a route past the left puncture. The two choices δ, δ' give different images of $[\gamma]$, as we will see below.

Choose a path $\delta : I \rightarrow X$ from x_1 to x_2 . For a loop γ based at x_1 , define

$$\gamma_\delta(t) = \begin{cases} \delta(1 - 3t), & 0 \leq t \leq \frac{1}{3}, \\ \gamma(3t - 1), & \frac{1}{3} \leq t \leq \frac{2}{3}, \\ \delta(3t - 2), & \frac{2}{3} \leq t \leq 1. \end{cases}$$

The formulas agree at $t = \frac{1}{3}, \frac{2}{3}$, where their value is x_1 , and both endpoints have value x_2 . Thus γ_δ is a continuous loop based at x_2 : it follows δ backwards, traverses γ , and returns along δ . We set

$$\beta_\delta : \pi_1(X, x_1) \longrightarrow \pi_1(X, x_2), \quad \beta_\delta([\gamma]) = [\gamma_\delta]. \quad (4.1)$$

Proof of Theorem 4.7. Well-definedness. If h is a based homotopy from γ_1 to γ_2 at x_1 , replace the middle expression $\gamma(3t - 1)$ by $h(3t - 1, s)$ and leave the two outer thirds fixed. This gives a based homotopy from $(\gamma_1)_\delta$ to $(\gamma_2)_\delta$ at x_2 (Figure 9). The values agree at the seams, so continuity follows by pasting.

The homomorphism property. Put the loop $(\gamma_1)_\delta(\gamma_2)_\delta$ along the bottom of a square. Its six successive pieces occupy equal subintervals. The middle two follow δ forwards and then backwards. Contract this excursion using the triangular construction from Section 3.2, keeping the first two and last two pieces fixed (Figure 11, left). The vertical sides stay at x_2 . Removing the constant pause at the top and adjusting the speed gives $(\gamma_1\gamma_2)_\delta$. Consequently,

$$\beta_\delta([\gamma_1])\beta_\delta([\gamma_2]) = [(\gamma_1)_\delta(\gamma_2)_\delta] = [(\gamma_1\gamma_2)_\delta] = \beta_\delta([\gamma_1][\gamma_2]).$$

An inverse. Apply the same three-piece formula to the reversed path $\delta^{-1}(t) = \delta(1 - t)$. The loop $(\gamma_\delta)_{\delta^{-1}}$ first follows δ forwards and backwards, then traverses γ , and finally follows δ forwards and backwards once more. After adjusting the time intervals, put the triangular contraction in each outer strip and keep γ fixed in the middle (Figure 11, right). Both vertical sides stay at x_1 , and the top edge is γ with constant pauses at its ends. Hence

$$(\beta_{\delta^{-1}} \circ \beta_\delta)([\gamma]) = [(\gamma_\delta)_{\delta^{-1}}] = [\gamma].$$

The same construction with δ^{-1} in place of δ gives $\beta_\delta \circ \beta_{\delta^{-1}} = \text{id}$. Thus β_δ is an isomorphism, with inverse $\beta_{\delta^{-1}}$. $\square$

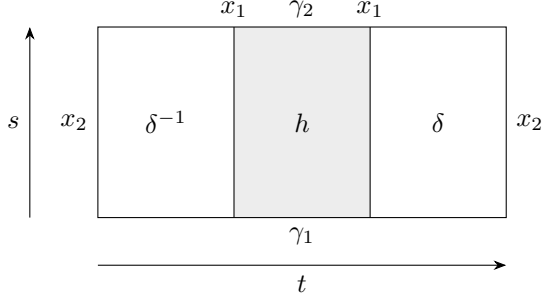


Figure 9: A based homotopy after changing the basepoint. The outer paths stay fixed, and the middle strip carries h . Each strip has width $\frac{1}{3}$, as in the defining formula.

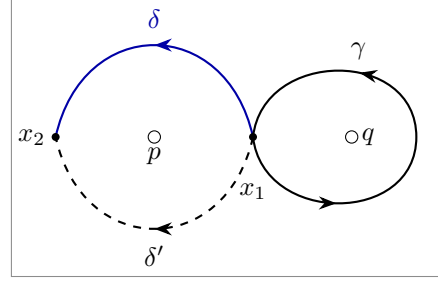


Figure 10: The loop γ is based at x_1 , beside the rightmost puncture. The two paths from x_1 to x_2 give different images of $[\gamma]$, and hence different isomorphisms.

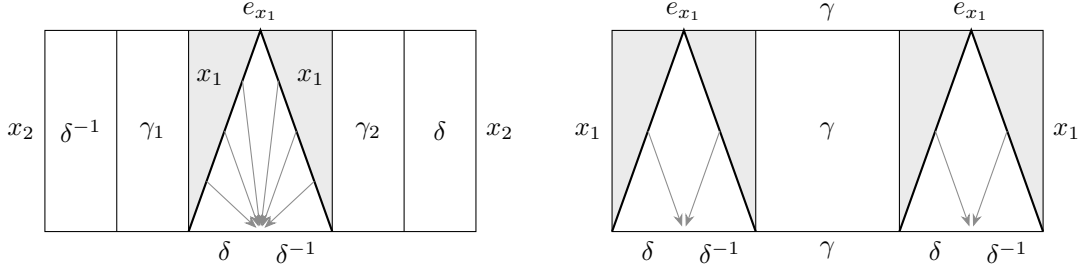


Figure 11: Cancellation squares, read from bottom to top. Left: the homomorphism property. Right: $\beta_{\delta^{-1}} \circ \beta_{\delta}$ is the identity, after rescaling the time intervals. The shaded triangles are constant at x_1 ; each fan carries δ towards its lower vertex.

4.6 Dependence on the connecting path

Let δ, δ' be two paths from x_1 to x_2 . Applying the three-piece formula twice gives the loop $(\gamma_{\delta})_{(\delta')^{-1}}$, representing $(\beta_{\delta'}^{-1} \circ \beta_{\delta})([\gamma])$. Its successive pieces follow δ' , then δ backwards, traverse γ , return along δ , and follow δ' backwards.

Define a loop u based at x_1 by

$$u(t) = \begin{cases} \delta(2t), & 0 \leq t \leq \frac{1}{2}, \\ \delta'(2-2t), & \frac{1}{2} \leq t \leq 1. \end{cases}$$

Draw the five pieces along the top of a rectangle, put γ on the bottom, and put u on both vertical sides, directed upwards (Figure 12).

To fill the rectangle at height s , start at $u(s)$, follow the initial segment of u backwards to x_1 , traverse γ , and return along the same initial segment to $u(s)$. The two outer pieces shrink to length zero at the bottom. The middle region carries γ , with a change of speed. Thus the basepoint of each horizontal loop is $u(s)$.

At the top, the first two pieces traverse u^{-1} , and the last two traverse u . All three of u^{-1}, γ, u are loops at x_1 , so their loop product is defined. Reading the top edge, up to a change of speed, gives

$$(\beta_{\delta'}^{-1} \circ \beta_{\delta})([\gamma]) = [u^{-1}\gamma u] = [u]^{-1}[\gamma][u].$$

The resulting automorphism of $\pi_1(X, x_1)$ is conjugation.

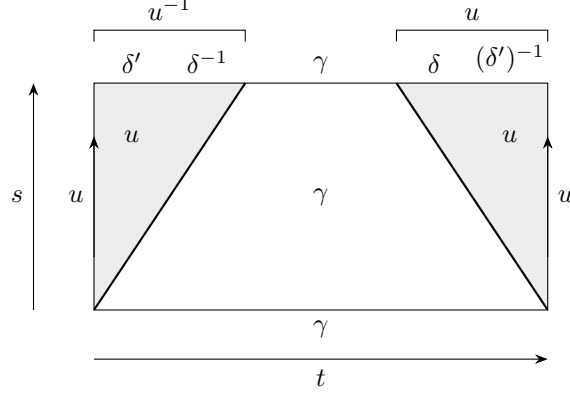


Figure 12: Complete the composite loop to a rectangle and fill it using γ and initial segments of u . The basepoint follows u along the vertical sides. The top edge reads $u^{-1}\gamma u$.

4.7 Conjugation and words in two generators

Let X be a rectangle with two punctures p, q , and choose a basepoint x_0 to their left. Let $a, b \in \pi_1(X, x_0)$ be represented by the loops drawn in Figure 13. The loop labelled a goes counterclockwise around p . The loop labelled b passes above p , goes counterclockwise around q , and returns along the same route. Outgoing and returning portions of a shared route are drawn slightly apart.

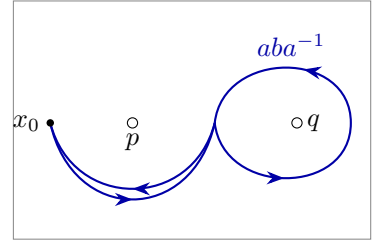
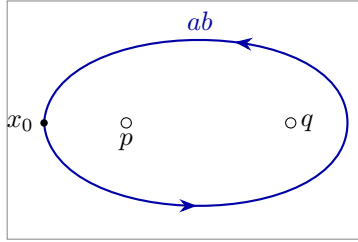
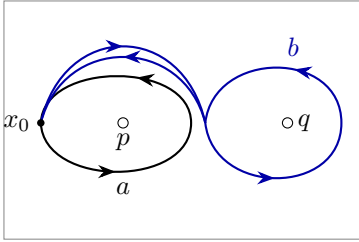


Figure 13: The generators a and b .

Figure 14: The product ab .

Figure 15: The conjugate aba^{-1} .

We will prove later that $\pi_1(X, x_0)$ is the *free group* generated by a and b . Its elements are reduced words in a, a^{-1}, b, b^{-1} : no letter is immediately followed by its inverse. Multiplication is concatenation followed by cancellation.

For example, ab is represented by following the loop labelled a , then the one labelled b . A based homotopy turns this into the single counterclockwise loop around both punctures in Figure 14. For aba^{-1} , follow these by the loop labelled a backwards. A based homotopy gives the representative in Figure 15, which reaches q below p . The reduced words b and aba^{-1} are distinct.

In any group, conjugation by a fixed element is an automorphism, called an *inner automorphism*. Here the inner automorphism $g \mapsto aga^{-1}$ takes b to aba^{-1} .

Returning to Figure 10, take $x_0 = x_2$. The upper path δ gives b , whereas the lower path δ' gives aba^{-1} :

$$\beta_\delta([\gamma]) = b, \quad \beta_{\delta'}([\gamma]) = aba^{-1}.$$

Thus the two connecting paths give different isomorphisms. Their difference is described by conjugation by a , giving a geometric interpretation of this inner automorphism.

5 Proof of the homotopy lifting theorem

Lecture 5, 15 October 2026

5.1 The lifting problem

We now prove Theorem 3.8. Let $f : X \rightarrow Y$ be a covering map and let

$$h : Z \times I \longrightarrow Y$$

be continuous. A *lift* of h is a continuous map $\tilde{h} : Z \times I \rightarrow X$ such that $f \circ \tilde{h} = h$. More generally, for a subset $A \subseteq Z \times I$, a *lift* of $h|_A$ is a continuous map $\tilde{h} : A \rightarrow X$ such that $f \circ \tilde{h} = h|_A$.

We are given an initial lift $\alpha : Z \rightarrow X$, satisfying

$$f(\alpha(z)) = h(z, 0) \quad (z \in Z).$$

We must construct a unique lift $\tilde{h}$ with $\tilde{h}(z, 0) = \alpha(z)$.

The proof has three parts: construct lifts on $U \times I$ for sufficiently small open neighborhoods U of each point of Z , prove that they agree on overlaps, and glue them. Uniqueness also enters the local construction, so we begin with two principles that will be used throughout.

5.2 Good sets and two lifting principles

We first express the covering condition in product form. Every point of Y has an open neighborhood V for which there are a discrete space Λ and a homeomorphism

$$\varphi : V \times \Lambda \xrightarrow{\cong} f^{-1}(V), \quad f(\varphi(y, \lambda)) = y.$$

Thus, in these coordinates, f is the projection onto V .

Definition 5.1. An open subset $V \subseteq Y$ with such a description is called *good*. A subset $A \subseteq Z \times I$ is called *good* if $h(A)$ is contained in a good subset of Y .

The inverse images under h of good open subsets of Y form an open cover of $Z \times I$ by good subsets.

Lemma 5.2 (Extension over a good set). *Let $A \subseteq Z \times I$ be good, let $p \in A$, and choose $x \in X$ with $f(x) = h(p)$. There is a lift of $h|_A$ whose value at p is x .*

Proof. Choose a good set V containing $h(A)$ and a homeomorphism $\varphi : V \times \Lambda \rightarrow f^{-1}(V)$ as above. There is a unique $\lambda \in \Lambda$ such that $x = \varphi(h(p), \lambda)$. The formula

$$\tilde{h}(q) = \varphi(h(q), \lambda), \quad q \in A,$$

defines the required continuous lift. □

Lemma 5.3 (Agreement of lifts). *Let $\tilde{h}_1, \tilde{h}_2 : A \rightarrow X$ be two lifts of $h|_A$, where $A \subseteq Z \times I$. Their agreement set*

$$E = \{p \in A : \tilde{h}_1(p) = \tilde{h}_2(p)\}$$

is both open and closed in A . In particular, if A is connected and the two lifts agree at one point, then they agree everywhere.

Proof. Fix $p \in A$ and choose a good neighborhood V of $h(p)$. On the relative open neighborhood $A \cap h^{-1}(V)$, write

$$\varphi^{-1}(\tilde{h}_j(q)) = (h(q), \lambda_j(q)), \quad j = 1, 2.$$

Each λ_j is continuous with values in the discrete space Λ , so both are constant on a sufficiently small neighborhood of p . If the two lifts agree at p , their λ -values agree on this neighborhood; if they differ at p , their λ -values remain different there. Thus both E and its complement are open. If A is connected and E is nonempty, it follows that $E = A$. □

5.3 A lift near $\{z\} \times I$

Lemma 5.4 (Tube lemma). *Let U be a topological space, let K be compact, and let $z \in U$. If an open subset $O \subseteq U \times K$ contains $\{z\} \times K$, then there is an open neighborhood W of z in U such that*

$$W \times K \subseteq O.$$

Proof. For each $k \in K$, choose open neighborhoods $W_k \subseteq U$ of z and $V_k \subseteq K$ of k such that $W_k \times V_k \subseteq O$. Compactness gives finitely many $k_1, \dots, k_m$ with $K = V_{k_1} \cup \dots \cup V_{k_m}$. Then

$$W = W_{k_1} \cap \dots \cap W_{k_m}$$

is an open neighborhood of z , and $W \times K \subseteq O$. $\square$

Fix $z \in Z$. We will find an open neighborhood U of z and a lift on $U \times I$ with initial value $\alpha|_U$.

Choosing finitely many good rectangles. For each $t \in I$, choose an open neighborhood U_t of z and an open subinterval J_t of I containing t such that

$$U_t \times \overline{J_t}$$

is good. This is possible by first choosing a product neighborhood inside a good open set and then taking a smaller interval whose closure stays inside that neighborhood. All closures here are taken in I . Open subintervals at the endpoints may have the forms $[0, b)$ and $(a, 1]$; the whole interval I is allowed as well.

By compactness of I , the open intervals J_t admit a finite subcover $J_1, \dots, J_n$, so $I = J_1 \cup \dots \cup J_n$. Write the corresponding good rectangles as

$$U_i \times \overline{J_i}, \quad i = 1, \dots, n.$$

Discard intervals contained in others and order the remaining ones by their left endpoints. Their right endpoints are then in increasing order as well. The first interval contains 0, the last contains 1, and consecutive intervals overlap.

Extending successive lifts. By Lemma 5.2, the prescribed value $\alpha(z)$ at $(z, 0)$ extends to a lift

$$\tilde{h}_1 : U_1 \times \overline{J_1} \longrightarrow X.$$

Having constructed $\tilde{h}_i$, choose $t_i \in J_i \cap J_{i+1}$ and apply the same lemma to $U_{i+1} \times \overline{J_{i+1}}$, prescribing the value $\tilde{h}_i(z, t_i)$ at (z, t_i) . This produces lifts $\tilde{h}_i$ on all the rectangles.

Making the lifts agree on overlaps. Set

$$K_i = \overline{J_i} \cap \overline{J_{i+1}}, \quad i = 1, \dots, n-1.$$

The interval K_i is connected, and $\tilde{h}_i$ and $\tilde{h}_{i+1}$ agree at (z, t_i) . Lemma 5.3 therefore shows that they agree on $\{z\} \times K_i$.

Their agreement set is also open in $(U_i \cap U_{i+1}) \times K_i$. Since K_i is compact, Lemma 5.4 gives an open neighborhood $W_i \subseteq U_i \cap U_{i+1}$ of z such that they agree on

$$W_i \times K_i.$$

This is the reason for using the closures of the intervals: the overlap used in this shrinking step is compact.

We must also retain the initial condition. On $U_1 \times \{0\}$, the two lifts $(u, 0) \mapsto \tilde{h}_1(u, 0)$ and $(u, 0) \mapsto \alpha(u)$ agree at $(z, 0)$. Their agreement set is open, so there is an open neighborhood $U_0 \subseteq U_1$ of z on which

$$\tilde{h}_1(u, 0) = \alpha(u).$$

Take the intersection of all the finitely many neighborhoods:

$$U = U_0 \cap \bigcap_{i=1}^n U_i \cap \bigcap_{i=1}^{n-1} W_i.$$

On $U \times J_i$, the restricted lifts agree for neighboring indices. They agree on every pairwise overlap as well: if $i < j < k$, the ordering of the interval endpoints gives

$$J_i \cap J_k \subseteq J_j.$$

At a point belonging to $J_i \cap J_k$, agreement therefore passes through all the intermediate lifts.

The sets $U \times J_i$ form an open cover of $U \times I$. By Lemma 5.5, the lifts on them glue to a continuous lift

$$\tilde{h}_U : U \times I \longrightarrow X, \quad \tilde{h}_U(u, 0) = \alpha(u).$$

This completes the local construction.

5.4 Uniqueness and global gluing

For each $z \in Z$, the interval $\{z\} \times I$ is connected. Consequently, two lifts on this interval with the same value at $(z, 0)$ agree everywhere, by Lemma 5.3.

The local construction gives an open cover $\{U_\sigma\}$ of Z and lifts

$$\tilde{h}_\sigma : U_\sigma \times I \longrightarrow X$$

with the prescribed initial values. If $z \in U_\sigma \cap U_\tau$, both lifts restrict to lifts on $\{z\} \times I$ starting at $\alpha(z)$, so they agree there. Thus the local lifts agree on every overlap $(U_\sigma \cap U_\tau) \times I$.

Recall the gluing principle used already in the local construction.

Lemma 5.5 (Gluing over an open cover). *Let $\{A_\sigma\}$ be an open cover of a space A . If continuous maps $g_\sigma : A_\sigma \rightarrow B$ agree on every overlap, then there is a unique continuous map $g : A \rightarrow B$ whose restriction to each A_σ is g_σ .*

Proof. For $a \in A$, choose σ with $a \in A_\sigma$ and set $g(a) = g_\sigma(a)$. Agreement on overlaps makes this well-defined and also gives uniqueness. For an open set $O \subseteq B$,

$$g^{-1}(O) = \bigcup_{\sigma} g_\sigma^{-1}(O).$$

Each set in this union is open in A_σ , hence in A . Therefore g is continuous. □

Applying Lemma 5.5 to the open cover $\{U_\sigma \times I\}$ gives a continuous map

$$\tilde{h} : Z \times I \longrightarrow X, \quad f \circ \tilde{h} = h, \quad \tilde{h}(z, 0) = \alpha(z).$$

Finally, any two such global lifts agree on every $\{z\} \times I$ by the same connectedness argument. Hence the lift is unique, proving Theorem 3.8.