Lecture 3
Fundamental group

In this lecture we will define the fundamental group of a space. This group is invariant under homotopy equivalence and so allows us to distinguish between spaces which are not homotopy equivalent. As a rule of thumb, most computational problems are easy for spaces with a trivial fundamental group and are hard otherwise.

1 Loops and the group operation

Start with a space X and fix a **basepoint** $x_0 \in X$. A **loop at** x_0 is a path $\alpha : I \to X$ with $\alpha(0) = \alpha(1) = x_0$. Write $[\alpha]$ for its homotopy class relative to $\{0, 1\}$. Thus $[\alpha]$ is the set of all loops at x_0 which can be continuously deformed into α while keeping both ends fixed at the basepoint.

Given two loops at x_0, say α and β we can first follow α and then β. This defines their **concatenation** denoted $\alpha \cdot \beta$, which is also a loop at x_0. We can also follow α backwards, getting again a loop at x_0 denoted by $\overline{\alpha}$. Formally

$$(\alpha \cdot \beta)(s) = \begin{cases} \alpha(2s), & 0 \le s \le \frac{1}{2}, \\ \beta(2s-1), & \frac{1}{2} \le s \le 1, \end{cases} \qquad \overline{\alpha}(s) = \alpha(1-s).$$

It is easy to check that if $\alpha' \simeq \alpha$ then $\alpha' \cdot \beta \simeq \alpha \cdot \beta$ and $\overline{\alpha'} \simeq \overline{\alpha}$. Thus, concatenation and reversing are well-defined for homotopy classes of loops as well:

$$[\alpha] \cdot [\beta] := [\alpha \cdot \beta] \qquad \overline{[\alpha]} := [\overline{\alpha}].$$

Definition 3.1. The **fundamental group** $\pi_1(X, x_0)$ of X with the basepoint x_0 is the set of the homotopy classes of loops at x_0 with the concatenation as the multiplication operation and reversing as the inverse operation. The unit of the group is the constant loop $I \to X$ given by $s \mapsto x_0$.

Proposition 3.2. *With this definition $\pi_1(X, x_0)$ is indeed a group.*

Proof. We will only check that the product of a loop with its inverse is equal to the unit of the group. I.e., that the concatenation of any loop with its inverse is homotopic to the constant loop.

Let α be a loop at x_0. Then $\alpha \cdot \overline{\alpha}$ is a loop which first goes forward along α and then goes back the same way. Consider a family of loops c_t, $t \in I$, which only follow α up to the time t and then go back. We see that c_1 is $\alpha \cdot \overline{\alpha}$ and c_0 is the constant loop which spends all the time at x_0. Hence, $\alpha \cdot \overline{\alpha}$ is homotopic to the constant loop, see Figure 1 (right). □

Here is an important criterion of triviality of a loop: *a loop in X is trivial if it bounds a 2-disk D^2 in X.*

Let us be more precise. Parametrize the boundary of D^2 with the parameter $s \in I$, where $s = 0$ and $s = 1$ correspond to the same point. This way the boundary ∂D^2 is identified with I with both ends glued together. We say that a loop $\alpha : I \to X$ **bounds a disk** in X if

$$\alpha : I = \partial D^2 \to X \quad \text{extends to a map} \quad D^2 \to X.$$

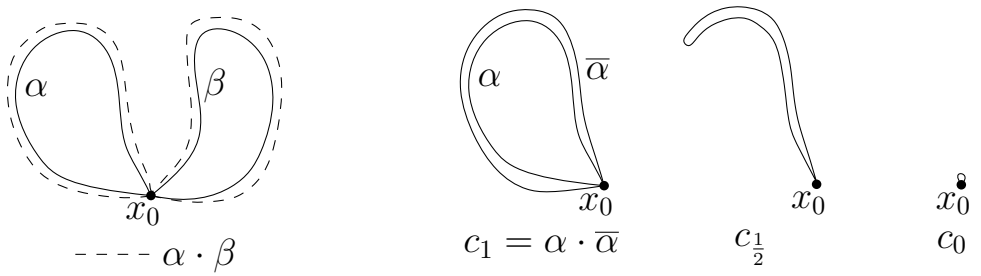


Figure 1: Concatenation of two loops (left), the product of a loop with its inverse is homotopic to the constant loop (right).

Proposition 3.3. *A loop $\alpha : I \to X$ is homotopic to a constant loop if and only if it bounds a disk in X.*

Two paths $\alpha, \beta : I \to X$ with the same ends $\alpha(0) = \beta(0)$ and $\alpha(1) = \beta(1)$ are homotopic if and only if the loop $\overline{\alpha}\beta$ bounds a disk.

Proof. See Figure 2. □

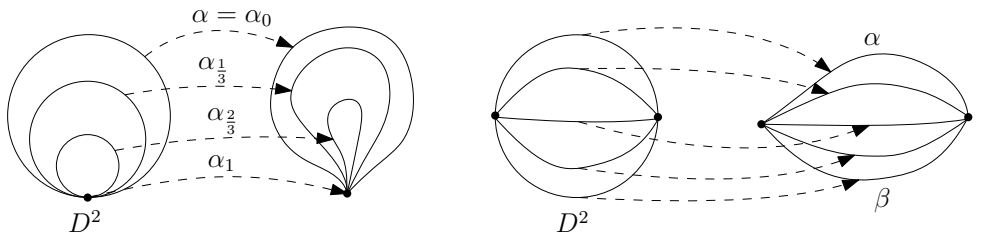


Figure 2: Proof of Proposition 3.3. If loop α bounds a disk then there is a family of paths α_t with $\alpha_0 = \alpha$ and α_1 the constant path (left). If $\overline{\alpha}\beta$ bounds a disk then there is homotopy between α and β (right).

Example 3.4 (Convex sets)**.** In a nonempty convex subset $C \subseteq \mathbb{R}^n$, the straight-line homotopy from a loop to the constant loop fixes its endpoints. Hence $\pi_1(C, x_0) = 1$, where 1 denotes the trivial group. In particular, $\pi_1(\mathbb{R}^n, x_0) = \pi_1(D^n, x_0) = 1$.

2 Maps, basepoints, and homotopies

A based map $f : (X, x_0) \to (Y, y_0)$ between spaces (*based* means $f(x_0) = y_0$), induces a map between their fundamental groups:

$$f_* : \pi_1(X, x_0) \to \pi_1(Y, y_0).$$

The induced map is given by the formula:

$$f_*[\alpha] = [f\alpha],$$

where $f\alpha : I \to Y$ is a loop in Y at y_0.

Homotopic loops have homotopic images: just compose their homotopy with f. Thus f_* is well defined. It is a **group homomorphism** because $f(\alpha \cdot \beta) = (f\alpha) \cdot (f\beta)$. We also have $(gf)_* = g_* f_*$ and $(\mathrm{id}_X)_* = \mathrm{id}$.

Our definition of the fundamental group singles out the basepoint, so it is important to understand what happens when the basepoint changes. A connecting path lets us turn a loop at the new basepoint into a loop at the old one.

Proposition 3.5 (Change of basepoint)**.** *Suppose $x_0, x_1 \in X$ are two different basepoints and suppose there is a connecting path h from x_0 to x_1.*

Then

$$\beta_h : \pi_1(X, x_1) \to \pi_1(X, x_0), \qquad \beta_h[\alpha] := [h \cdot \alpha \cdot \overline{h}]$$

is a group isomorphism. Its inverse is $\beta_{\overline{h}} : \pi_1(X, x_0) \to \pi_1(X, x_1)$.

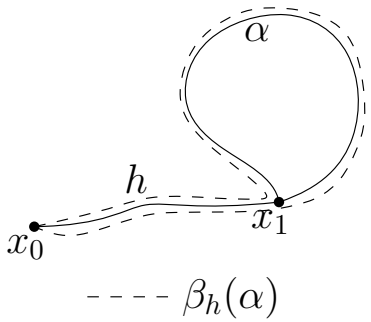


Figure 3: Change of the basepoint isomorphism $\beta_h : \pi_1(X, x_1) \to \pi_1(X, x_0)$.

Proof. Adding h and $\overline{h}$ to homotopic loops keeps them homotopic, so β_h is well defined. In a product of two such loops, we go back along h and immediately follow it again:

$$[h\alpha\overline{h}]\,[h\gamma\overline{h}] = [h\alpha\overline{h}h\gamma\overline{h}] = [h\alpha\gamma\overline{h}].$$

Here the backtracking pair $\overline{h}\cdot h$ cancels, just as in Proposition 3.2. This proves the homomorphism law.

Applying β_h and then $\beta_{\overline{h}}$ to a loop $[\alpha]$ we get $[\overline{h} \cdot h \cdot \alpha \cdot \overline{h} \cdot h]$. The $\overline{h} \cdot h$ part on both ends cancel again, so we get back to $[\alpha]$. This means that $\beta_{\overline{h}}\beta_h$ is the identity homomorphism of $\pi_1(X, x_1)$. □

For a path-connected space, the groups at different basepoints are therefore isomorphic, and we often write just $\pi_1(X)$. There need not be a preferred isomorphism. If k is another path from x_0 to x_1, cancellation gives

$$\beta_k[\alpha] := [k \cdot \alpha \cdot \overline{k}] = [k \cdot \overline{h} \cdot h \cdot \alpha \cdot \overline{h} \cdot h \cdot \overline{k}] = [k \cdot \overline{h}]\,\beta_h(a)\,[k \cdot \overline{h}]^{-1}.$$

So, the isomorphisms $\beta_k[\alpha]$ and $\beta_h[\alpha]$ differ by conjugation with $[k \cdot \overline{h}]$.

Definition 3.6. A nonempty space is **simply connected** if it is path connected and its fundamental group is trivial.

Simple connectivity also tells us about paths that are not loops: their homotopy class is determined by their endpoints.

Proposition 3.7. *A nonempty path-connected space is simply connected if and only if any two paths with the same endpoints are homotopic fixing endpoints.*

Proof. Suppose any two paths are homotopic fixing their endpoints. In particular it means that any two loops at the basepoint x_0 are homotopic and so the fundamental group is trivial.

Suppose now that the fundamental group with the basepoint x_0 is trivial. Consider two paths α, β from y to z. The fundamental group with the basepoint y is isomorphic to the one with the basepoint x_0 and so is trivial as well. This means that the loop $\overline{\alpha} \cdot \beta$ at y is trivial and hence bounds a disk by Proposition 3.3. Again, by Proposition 3.3 the paths α and β are homotopic along this disk. □

To make the fundamental group an invariant of homotopy type, we need to understand how a homotopy affects the induced homomorphism.

Proposition 3.8 (Homotopies)**.** *Suppose $H : X \times I \to Y$ is a homotopy between $f_0, f_1 : X \to Y$. Consider the path $h : I \to Y$ with $h(t) := H_t(x_0)$ which is traversed by the image of the basepoint $x_0 \in X$. We have*

$$(f_0)_* = \beta_h \circ (f_1)_*.$$

In particular, a homotopy fixing the basepoint gives $(f_0)_* = (f_1)_*$.

Proof. Consider a loop α at x_0 in X. At time t, its image $H_t\alpha$ is based at $h(t)$, which may move. Let h_t be the part of the path h up to time t. Consider now the loop

$$\alpha_t := h_t \cdot H_t\alpha \cdot \overline{h_t}.$$

For every t, we have that α_t is a loop based at $f_0(x_0)$. Also, $\alpha_0 = \alpha$ and $\alpha_1 = h \cdot f_1\alpha \cdot \overline{h} = \beta_h((f_1)_*(\alpha))$, so $[\alpha] = \beta_h \circ (f_1)_*[\alpha]$. □

Corollary 3.9. *A homotopy equivalence $f : X \to Y$ induces an isomorphism $\pi_1(X, x_0) \cong \pi_1(Y, f(x_0))$. In particular, contractible spaces are simply connected.*

3 The fundamental group of the circle S^1

View S^1 as the unit circle in $\mathbb{C}$, with basepoint 1. The map $p : \mathbb{R} \to S^1$, $p(u) = e^{2\pi i u}$, wraps the real line around the circle.

For any point $x \in S^1$ its preimage $p^{-1}(x)$ is a discrete set of points in $\mathbb{R}$ with any two adjacent points differing by 1. In particular, $p^{-1}(1) = \mathbb{Z}$.

Proposition 3.10. *Let $\alpha : I \to S^1$ be a path starting at $x := \alpha(0)$. Suppose α is **short** , meaning that its image $U := \alpha(I)$ is not the full circle. Then for any choice of $\widetilde{x} \in p^{-1}(x)$ there is a unique **lifting** $\widetilde{\alpha} : I \to \mathbb{R}$ of α such that*

$$\widetilde{\alpha}(0) = \widetilde{x} \quad \text{and} \quad p\widetilde{\alpha} = \alpha.$$

Moreover, the length of $\widetilde{\alpha}(I)$ is the same as the length of $\alpha(I)$.

Proof. The set $p^{-1}(U)$ consists of a countable many segments in $\mathbb{R}$. The restriction of p to each of these segments is a homeomorphism to U. Pick the segment S containing $\widetilde{x}$. Then $p|_S : S \to U$ is a homeomorphism. The only way to lift α is to set $\widetilde{\alpha} := (p|_S)^{-1}\alpha$. □

Proposition 3.11. *Let $\alpha : I \to S^1$ be a path starting at $x := \alpha(0)$. Then for any choice of $\widetilde{x} \in p^{-1}(x)$ there is a unique lifting $\widetilde{\alpha} : I \to \mathbb{R}$ of α such that*

$$\widetilde{\alpha}(0) = \widetilde{x} \quad \textit{and} \quad p\widetilde{\alpha} = \alpha.$$

Proof. Split α into a concatenation of short paths $\alpha_1 \cdot \alpha_2 \cdot \ldots \cdot \alpha_k$. We can assume that these paths are short by choosing k large enough. Use the previous proposition to first lift α_1, then α_2, and so on.

To see that the lifting is a unique, i.e., does not depend on our splitting of α, pick any point $s \in I$. The set of possible values of $\widetilde{\alpha}(s)$ is the discrete set of points $p^{-1}(\alpha(s))$. Varying our splitting of α continuously, the value of $\widetilde{\alpha}(s)$ must also change continuously. But since it lies in a discrete set it must stay the same. □

Let α be a loop in S^1 at the basepoint 1. Consider its unique lifting $\widetilde{\alpha} : I \to \mathbb{R}$ starting at $\widetilde{\alpha}(0) = 0$. Define the **winding number** of α by $w(\alpha) = \widetilde{\alpha}(1)$. Because $p\widetilde{\alpha}(1) = \alpha(1) = \alpha(0) = 1$, we have that $w(\alpha)$ is an integer.

The winding number depends only on the homotopy type of the loop:

Proposition 3.12. *Suppose α_t is a continuous family of loops in S^1 at the basepoint 1 parametrized by $t \in I$. Then $w(\alpha_0) = w(\alpha_1)$.*

Proof. Consider a dense enough grid in the unit square, see Figure 4. The bottom side of the square is mapped to S^1 by α_0, the top by α_1 and for every t the horizontal segment at height t is mapped by α_t.

We already have a unique lifting $\widetilde{\alpha_0}$ on the bottom. The left side of the square is mapped to $1 \in S^1$, so we can simply lift the whole left side to $0 \in \mathbb{R}$. Let us now construct the lifting on the rest of the grid.

Do it square by square, starting from the bottom left one and moving from left to right and from the bottom to the top. Once we arrive at a square, we already have a lifting at its bottom and left sides, Figure 5.

The square is small, so it is mapped to a very short arc in S^1. The top side of the square is a short path with its start (the top left vertex) already lifted, so we can lift this path using Proposition 3.10. Likewise, the right side is a short path with its start (the bottom right vertex) already lifted. Lift this side again using Proposition 3.10.

We now have two different liftings of the top right vertex. These liftings differ by an integer in $\mathbb{R}$. However, since the square is very small, the lifting of each of its sides is very short. In particular, both liftings of the top right vertex are very close to the single lifting of the bottom left vertex. And if two numbers are very close but also differ by an integer, then they are the same. Hence, both liftings of the top right vertex coincide.

We have now constructed the lifting on the whole grid. Because $\alpha_t(1) = 1$ for each t, we have that the lifting of the whole right side of the square is a single point, namely, $w(\alpha_0)$. In particular, the top right vertex of the square is lifted to $w(\alpha_0)$. But the lifting of the top side of the square is the lifting of α_1, so by the definition of w the top right vertex of the square is lifted to $w(\alpha_1)$. So, $w(\alpha_0) = w(\alpha_1)$. □

Lemma 3.13. *The winding number w gives an isomorphism $\pi_1(S^1, 1) \cong \mathbb{Z}$.*

Proof. First, lets us prove that $w : \pi_1(S^1, 1) \to \mathbb{Z}$ is injective. Suppose two loops α, β have liftings $\widetilde{\alpha}$ and $\widetilde{\beta}$ with the same winding numbers, i.e., $\widetilde{\alpha}(1) = \widetilde{\beta}(1)$ is the same integer. For injectivity we need to prove that $\alpha \simeq \beta$.

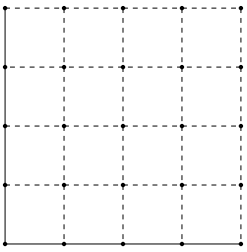


Figure 4: Dense grid. Initially the lifting is defined on the bottom and the left sides.

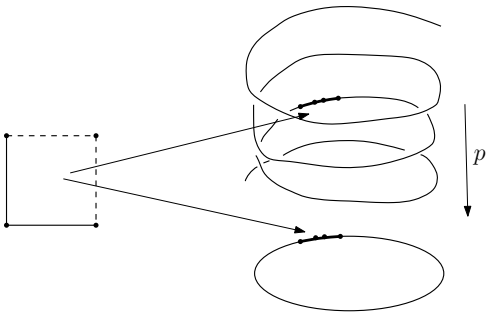


Figure 5: The square is small, so it is mapped to a very short arc in S^1.

Indeed, $\overline{\widetilde{\alpha}} \cdot \widetilde{\beta}$ is a closed loop in $\mathbb{R}$. Using linear homotopy we see that $\overline{\widetilde{\alpha}} \cdot \widetilde{\beta}$ is null-homotopic and hence bounds a disk in $\mathbb{R}$. This means that $\overline{\widetilde{\alpha}} \cdot \widetilde{\beta}$ composed with p, which is just $\overline{\alpha} \cdot \beta$, bounds a disk in S^1. So $\overline{\alpha} \cdot \beta$ bounds a disk in S^1 meaning $\alpha \simeq \beta$.

For surjectivity notice that the loop $s \mapsto e^{2\pi i n s}$ lifts to $s \mapsto ns$ and has winding number n, for every $n \in \mathbb{Z}$.

Finally, to prove that w is a homomorphism, notice that the lifting of the concatenations of two loops, $\alpha \cdot \beta$, can be constructed as follows. First we lift α lifting its start to $0 \in \mathbb{R}$. Its end is lifted to $w(\alpha) \in \mathbb{Z}$. Then we lift β by lifting its start to $w(\alpha)$. The end of β is thus lifted to $w(\alpha) + w(\beta)$. □

Brouwer's fixed-point theorem

Lemma 3.14. *Suppose that $r : X \to A$ is a retraction and the basepoint x_0 lies in A. Then the map $i_* : \pi_1(A, x_0) \to \pi_1(X, x_0)$ induced by the inclusion $i : A \to X$ is injective.*

Proof. Let $\alpha, \beta : I \to A$ be loops in A and suppose α and β are homotopic as loops in X. Then $\overline{\alpha}\beta$ bounds a disk in X, meaning there is a map $f : D^2$ whose restriction to ∂D^2 is $\overline{\alpha}\beta$. Composing f with r we get a map $rf : D^2 \to A$ whose restriction to ∂D^2 is again $\overline{\alpha}\beta$. Hence, $\overline{\alpha}\beta$ bound a disk in A and α is homotopic to β in A. □

Corollary 3.15 (Brouwer fixed point theorem)**.** *Every continuous map $f : D^2 \to D^2$ has a fixed point: there is an $x \in D^2$ such that $f(x) = x$.*

Proof. Suppose that $f(x) \neq x$ for every $x \in D^2$. Start a ray at $f(x)$ and pass through x. Let $r(x)$ be the point of ∂D^2 where this ray leaves the disk. It is easy to check that $r : D^2 \to \partial D^2 \cong S^1$ is continuous and is identity on ∂D^2, hence a retraction.

By Lemma 3.14, we get an injective map from $\pi_1(S^1) \cong \mathbb{Z}$ to $\pi_1(D^2) = 1$, contradiction. □

Reference. Hatcher, *Algebraic Topology*, Section 1.1, pp. 25–32 and 36–38.