MA3427—Algebraic Topology I
School of Mathematics, Trinity College
Michaelmas Term 2018
Section 3: The Fundamental Group of a Topological Spaces

David R. Wilkins

3. The Fundamental Group of a Topological Space

3.1. Homotopies between Continuous Maps

Definition

Let $f\colon X \to Y$ and $g\colon X \to Y$ be continuous maps between topological spaces X and Y. The maps f and g are said to be *homotopic* if there exists a continuous map $H\colon X \times [0,1] \to Y$ such that $H(x,0) = f(x)$ and $H(x,1) = g(x)$ for all $x \in X$. If the maps f and g are homotopic then we denote this fact by writing $f \simeq g$. The map H with the properties stated above is referred to as a *homotopy* between f and g.

Continuous maps f and g from X to Y are homotopic if and only if it is possible to 'continuously deform' the map f into the map g.

Definition

Let X and Y be topological spaces, and let A be a subset of X. Let $f\colon X \to Y$ and $g\colon X \to Y$ be continuous maps from X to some topological space Y, where $f|A = g|A$ (i.e., $f(a) = g(a)$ for all $a \in A$). We say that f and g are *homotopic relative to A* (denoted by $f \simeq g \ \mathrm{rel}\ A$) if and only if there exists a (continuous) homotopy $H\colon X \times [0,1] \to Y$ such that $H(x,0) = f(x)$ and $H(x,1) = g(x)$ for all $x \in X$ and $H(a,t) = f(a) = g(a)$ for all $a \in A$.

Proposition 3.1

Let X and Y be topological spaces, and let A be a subset of X. The relation of being homotopic relative to the subset A is then an equivalence relation on the set of all continuous maps from X to Y.

Proof

Given $f\colon X \to Y$, let $H_0\colon X \times [0,1] \to Y$ be defined so that $H_0(x,t) = f(x)$ for all $x \in X$ and $t \in [0,1]$. Then $H_0(x,0) = H_0(x,1) = f(x)$ for all $x \in X$ and $H_0(a,t) = f(a)$ for all $a \in A$ and $t \in [0,1]$, and therefore $f \simeq f \text{ rel } A$. Thus the relation of homotopy relative to A is reflexive.

Let f and g be continuous maps from X to Y that satisfy $f(a) = g(a)$ for all $a \in A$. Suppose that $f \simeq g \ \mathrm{rel}\ A$. Then there exists a homotopy $H\colon X \times [0,1] \to Y$ with the properties that $H(x,0) = f(x)$ and $H(x,1) = g(x)$ for all $x \in X$ and $H(a,t) = f(a) = g(a)$ for all $a \in A$ and $t \in [0,1]$. Let $K\colon X \times [0,1] \to Y$ be defined so that $K(x,t) = H(x,1-t)$ for all $t \in [0,1]$. Then K is a homotopy between g and f, and $K(a,t) = g(a) = f(a)$ for all $a \in A$ and $t \in [0,1]$. It follows that $g \simeq f \ \mathrm{rel}\ A$. Thus the relation of homotopy relative to A is symmetric.

Finally let f, g and h be continuous maps from X to Y with the property that $f(a) = g(a) = h(a)$ for all $a \in A$. Suppose that $f \simeq g \ \mathrm{rel}\ A$ and $g \simeq h \ \mathrm{rel}\ A$. Then there exist homotopies $H_1 \colon X \times [0,1] \to Y$ and $H_2 \colon X \times [0,1] \to Y$ satisfying the following properties:

$$\begin{aligned} H_1(x,0) &= f(x), \\ H_1(x,1) &= g(x) = H_2(x,0), \\ H_2(x,1) &= h(x) \end{aligned}$$

for all $x \in X$;

$$H_1(a,t) = H_2(a,t) = f(a) = g(a) = h(a)$$

for all $a \in A$ and $t \in [0,1]$.

Define $H\colon X \times [0,1] \to Y$ by

$$H(x,t) = \begin{cases} H_1(x,2t) & \text{if } 0 \leq t \leq \frac{1}{2}; \\ H_2(x,2t-1) & \text{if } \frac{1}{2} \leq t \leq 1. \end{cases}$$

Now $H|X \times [0,\frac{1}{2}]$ and $H|X \times [\frac{1}{2},1]$ are continuous. It follows from the Pasting Lemma (Lemma 1.24) that H is continuous on $X \times [0,1]$. Moreover $H(x,0) = f(x)$ and $H(x,1) = h(x)$ for all $x \in X$. Thus $f \simeq h \,\mathrm{rel}\, A$. Thus the relation of homotopy relative to the subset A of X is transitive. This relation has now been shown to be reflexive, symmetric and transitive. It is therefore an equivalence relation. $\blacksquare$

Remark

Let X and Y be topological spaces, and let $H\colon X \times [0,1] \to Y$ be a function whose restriction to the sets $X \times [0, \frac{1}{2}]$ and $X \times [\frac{1}{2}, 1]$ is continuous. Then the function H is continuous on $X \times [0,1]$. The Pasting Lemma (Lemma 1.24) was applied in the proof of Proposition 3.1 to justify this assertion. We consider in more detail how the Pasting Lemma guarantees the continuity of this function. Let $x \in X$. If $t \in [0,1]$ and $t \neq \frac{1}{2}$ then the point (x,t) is contained in an open subset of $X \times [0,1]$ over which the function H is continuous, and therefore the function H is continuous at (x,t). In order to complete the proof that the function H is continuous everywhere on $X \times [0,1]$ it suffices to verify continuity of H at $(x, \frac{1}{2})$, where $x \in X$.

Let V be an open set in Y for which $f(x, \frac{1}{2}) \in V$. Then the continuity of the restrictions of H to $X \times [0, \frac{1}{2}]$ and $X \times [\frac{1}{2}, 1]$ ensures the existence of open sets W_1 and W_2 in $X \times [0, 1]$ such that $(x, \frac{1}{2}) \in W_1 \cap W_2$, $H(W_1 \cap (X \times [0, \frac{1}{2}])) \subset V$ and $H(W_2 \cap (X \times [\frac{1}{2}, 1])) \subset V$. Let $W = W_1 \cap W_2$. Then $H(W) \subset V$. This completes the verification that the function H is continuous at $(x, \frac{1}{2})$. The Pasting Lemma is a basic tool for establishing the continuity of functions occurring in algebraic topology that are similar in nature to the function $H\colon X \times [0, 1] \to Y$ considered in this discussion. The continuity of such functions can typically be established directly using arguments analogous to that employed here.

Corollary 3.2

Let X and Y be topological spaces. The homotopy relation $\simeq$ is an equivalence relation on the set of all continuous maps from X to Y.

Proof

This result follows on applying Proposition 3.1 in the case where homotopies are relative to the empty set. ∎

3.2. The Fundamental Group of a Topological Space

Definition

Let X be a topological space, and let x_0 and x_1 be points of X. A *path* in X from x_0 to x_1 is defined to be a continuous map $\gamma\colon [0,1] \to X$ for which $\gamma(0) = x_0$ and $\gamma(1) = x_1$. A *loop* in X based at x_0 is defined to be a continuous map $\gamma\colon [0,1] \to X$ for which $\gamma(0) = \gamma(1) = x_0$.

We can concatenate paths. Let $\gamma_1\colon [0,1] \to X$ and $\gamma_2\colon [0,1] \to X$ be paths in some topological space X. Suppose that $\gamma_1(1) = \gamma_2(0)$. We define the *product path* $\gamma_1 \,.\, \gamma_2\colon [0,1] \to X$ by

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

If $\gamma\colon [0,1] \to X$ is a path in X then we define the *inverse path* $\gamma^{-1}\colon [0,1] \to X$ by $\gamma^{-1}(t) = \gamma(1-t)$. (Thus if γ is a path from the point x_0 to the point x_1 then γ^{-1} is the path from x_1 to x_0 obtained by traversing γ in the reverse direction.)

Let X be a topological space, and let $x_0 \in X$ be some chosen point of X. We define an equivalence relation on the set of all (continuous) loops based at the basepoint x_0 of X, where two such loops γ_0 and γ_1 are equivalent if and only if $\gamma_0 \simeq \gamma_1 \operatorname{rel} \{0, 1\}$. We denote the equivalence class of a loop $\gamma\colon [0, 1] \to X$ based at x_0 by $[\gamma]$. This equivalence class is referred to as the *based homotopy class* of the loop γ. The set of equivalence classes of loops based at x_0 is denoted by $\pi_1(X, x_0)$. Thus two loops γ_0 and γ_1 represent the same element of $\pi_1(X, x_0)$ if and only if $\gamma_0 \simeq \gamma_1 \operatorname{rel} \{0, 1\}$ (i.e., there exists a homotopy $F\colon [0, 1] \times [0, 1] \to X$ between γ_0 and γ_1 which maps $(0, \tau)$ and $(1, \tau)$ to x_0 for all $\tau \in [0, 1]$).

Theorem 3.3

Let X be a topological space, let x_0 be some chosen point of X, and let $\pi_1(X, x_0)$ be the set of all based homotopy classes of loops based at the point x_0. Then $\pi_1(X, x_0)$ is a group, the group multiplication on $\pi_1(X, x_0)$ being defined according to the rule $[\gamma_1][\gamma_2] = [\gamma_1 . \gamma_2]$ for all loops γ_1 and γ_2 based at x_0.

Proof

First we show that the group operation on $\pi_1(X, x_0)$ is well-defined. Let γ_1, γ_1', γ_2 and γ_2' be loops in X based at the point x_0. Suppose that $[\gamma_1] = [\gamma_1']$ and $[\gamma_2] = [\gamma_2']$. Let the map $F\colon [0,1] \times [0,1] \to X$ be defined by

$$F(t,\tau) = \begin{cases} F_1(2t,\tau) & \text{if } 0 \leq t \leq \frac{1}{2}, \\ F_2(2t-1,\tau) & \text{if } \frac{1}{2} \leq t \leq 1, \end{cases}$$

where $F_1\colon [0,1] \times [0,1] \to X$ is a homotopy between γ_1 and γ_1', $F_2\colon [0,1] \times [0,1] \to X$ is a homotopy between γ_2 and γ_2', and where the homotopies F_1 and F_2 map $(0,\tau)$ and $(1,\tau)$ to x_0 for all $\tau \in [0,1]$. Then F is itself a homotopy from $\gamma_1 \,.\, \gamma_2$ to $\gamma_1' \,.\, \gamma_2'$, and maps $(0,\tau)$ and $(1,\tau)$ to x_0 for all $\tau \in [0,1]$. Thus $[\gamma_1 \,.\, \gamma_2] = [\gamma_1' \,.\, \gamma_2']$, showing that the group operation on $\pi_1(X, x_0)$ is well-defined.

Next we show that the group operation on $\pi_1(X, x_0)$ is associative. Let γ_1, γ_2 and γ_3 be loops based at x_0, and let $\alpha = (\gamma_1.\gamma_2).\gamma_3$. Then $\gamma_1.(\gamma_2.\gamma_3) = \alpha \circ \theta$, where

$$\theta(t) = \begin{cases} \frac{1}{2}t & \text{if } 0 \leq t \leq \frac{1}{2}; \\ t - \frac{1}{4} & \text{if } \frac{1}{2} \leq t \leq \frac{3}{4}; \\ 2t - 1 & \text{if } \frac{3}{4} \leq t \leq 1. \end{cases}$$

Thus the map $G\colon [0,1] \times [0,1] \to X$ defined by $G(t,\tau) = \alpha((1-\tau)t + \tau\theta(t))$ is a homotopy between $(\gamma_1.\gamma_2).\gamma_3$ and $\gamma_1.(\gamma_2.\gamma_3)$, and moreover this homotopy maps $(0,\tau)$ and $(1,\tau)$ to x_0 for all $\tau \in [0,1]$. It follows that $(\gamma_1 \,.\, \gamma_2) \,.\, \gamma_3 \simeq \gamma_1 \,.\, (\gamma_2 \,.\, \gamma_3) \text{ rel } \{0,1\}$ and hence $([\gamma_1][\gamma_2])[\gamma_3] = [\gamma_1]([\gamma_2][\gamma_3])$. This shows that the group operation on $\pi_1(X, x_0)$ is associative.

Let $\varepsilon\colon [0,1] \to X$ denote the constant loop at x_0, defined by $\varepsilon(t) = x_0$ for all $t \in [0,1]$. Then $\varepsilon \,.\, \gamma = \gamma \circ \theta_0$ and $\gamma \,.\, \varepsilon = \gamma \circ \theta_1$ for any loop γ based at x_0, where

$$\theta_0(t) = \begin{cases} 0 & \text{if } 0 \leq t \leq \frac{1}{2}, \\ 2t - 1 & \text{if } \frac{1}{2} \leq t \leq 1, \end{cases}$$

$$\theta_1(t) = \begin{cases} 2t & \text{if } 0 \leq t \leq \frac{1}{2}, \\ 1 & \text{if } \frac{1}{2} \leq t \leq 1, \end{cases}$$

for all $t \in [0,1]$. But the continuous map $(t,\tau) \mapsto \gamma((1-\tau)t + \tau\theta_j(t))$ is a homotopy between γ and $\gamma \circ \theta_j$ for $j = 0, 1$ which sends $(0,\tau)$ and $(1,\tau)$ to x_0 for all $\tau \in [0,1]$. Therefore $\varepsilon \,.\, \gamma \simeq \gamma \simeq \gamma \,.\, \varepsilon \text{ rel } \{0,1\}$, and hence $[\varepsilon][\gamma] = [\gamma] = [\gamma][\varepsilon]$. We conclude that $[\varepsilon]$ represents the identity element of $\pi_1(X, x_0)$.

It only remains to verify the existence of inverses. Now the map $K\colon [0,1] \times [0,1] \to X$ defined by

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

is a homotopy between the loops ε and $\gamma\,.\,\gamma^{-1}$, and moreover this homotopy sends $(0,\tau)$ and $(1,\tau)$ to x_0 for all $\tau \in [0,1]$. Therefore $\varepsilon \simeq \gamma\,.\,\gamma^{-1} \,\mathrm{rel}\{0,1\}$, and thus $[\gamma][\gamma^{-1}] = [\gamma\,.\,\gamma^{-1}] = [\varepsilon]$. On replacing γ by γ^{-1}, we see also that $[\gamma^{-1}][\gamma] = [\varepsilon]$, and thus $[\gamma^{-1}] = [\gamma]^{-1}$, as required. ∎

Let x_0 be a point of some topological space X. The group $\pi_1(X, x_0)$ is referred to as the *fundamental group* of X based at the point x_0.

Let $f\colon X \to Y$ be a continuous map between topological spaces X and Y, and let x_0 be a point of X. Then f induces a homomorphism $f_\#\colon \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$, where $f_\#([\gamma]) = [f \circ \gamma]$ for all loops $\gamma\colon [0,1] \to X$ based at x_0. If x_0, y_0 and z_0 are points belonging to topological spaces X, Y and Z, and if $f\colon X \to Y$ and $g\colon Y \to Z$ are continuous maps satisfying $f(x_0) = y_0$ and $g(y_0) = z_0$, then the induced homomorphisms $f_\#\colon \pi_1(X, x_0) \to \pi_1(Y, y_0)$ and $g_\#\colon \pi_1(Y, y_0) \to \pi_1(Z, z_0)$ satisfy $g_\# \circ f_\# = (g \circ f)_\#$. It follows easily from this that any homeomorphism of topological spaces induces a corresponding isomorphism of fundamental groups, and thus the fundamental group is a topological invariant.

3.3. Simply-Connected Topological Spaces

Definition

A topological space X is said to be *simply-connected* if it is path-connected, and any continuous map $f: \partial D \to X$ mapping the boundary circle ∂D of a closed disc D into X can be extended continuously over the whole of the disk.

Example

$\mathbb{R}^n$ is simply-connected for all n. Indeed any continuous map $f: \partial D \to \mathbb{R}^n$ defined over the boundary ∂D of the closed unit disk D can be extended to a continuous map $F: D \to \mathbb{R}^n$ over the whole disk by setting $F(r\mathbf{x}) = rf(\mathbf{x})$ for all $\mathbf{x} \in \partial D$ and $r \in [0, 1]$.

Let E be a topological space that is homeomorphic to the closed disk D, and let $\partial E = h(\partial D)$, where ∂D is the boundary circle of the disk D and $h\colon D \to E$ is a homeomorphism from D to E. Then any continuous map $g\colon \partial E \to X$ mapping ∂E into a simply-connected space X extends continuously to the whole of E. Indeed there exists a continuous map $F\colon D \to X$ which extends $g \circ h\colon \partial D \to X$, and the map $F \circ h^{-1}\colon E \to X$ then extends the map g.

Theorem 3.4

A path-connected topological space X is simply-connected if and only if $\pi_1(X, x)$ is trivial for all $x \in X$.

Proof

Suppose that the space X is simply-connected. Let $\gamma\colon [0,1] \to X$ be a loop based at some point x of X. Now the unit square is homeomorphic to the unit disk, and therefore any continuous map defined over the boundary of the square can be continuously extended over the whole of the square. It follows that there exists a continuous map $F\colon [0,1] \times [0,1] \to X$ such that $F(t,0) = \gamma(t)$ and $F(t,1) = x$ for all $t \in [0,1]$, and $F(0,\tau) = F(1,\tau) = x$ for all $\tau \in [0,1]$. Thus $\gamma \simeq \varepsilon_x \operatorname{rel}\{0,1\}$, where ε_x is the constant loop at x, and hence $[\gamma] = [\varepsilon_x]$ in $\pi_1(X,x)$. This shows that $\pi_1(X,x)$ is trivial.

Conversely suppose that X is path-connected and $\pi_1(X, x)$ is trivial for all $x \in X$. Let $f\colon \partial D \to X$ be a continuous function defined on the boundary circle ∂D of the closed unit disk D in $\mathbb{R}^2$. We must show that f can be extended continuously over the whole of D. Let $x = f(1,0)$. There exists a continuous map $G\colon [0,1] \times [0,1] \to X$ such that $G(t,0) = f(\cos(2\pi t), \sin(2\pi t))$ and $G(t,1) = x$ for all $t \in [0,1]$ and $G(0,\tau) = G(1,\tau) = x$ for all $\tau \in [0,1]$, since $\pi_1(X,x)$ is trivial. Moreover $G(t_1,\tau_1) = G(t_2,\tau_2)$ whenever $q(t_1,\tau_1) = q(t_2,\tau_2)$, where

$$q(t,\tau) = \bigl((1-\tau)\cos(2\pi t) + \tau, (1-\tau)\sin(2\pi t)\bigr)$$

for all $t, \tau \in [0,1]$. It follows that there is a well-defined function $F\colon D \to X$ such that $F \circ q = G$.

However $q\colon [0,1] \times [0,1] \to D$ is a continuous surjection from a compact space to a Hausdorff space and is therefore an identification map. It follows that $F\colon D \to X$ is continuous (since a basic property of identification maps ensures that a function $F\colon D \to X$ is continuous if and only if $F \circ q\colon [0,1] \times [0,1] \to X$ is continuous). Moreover $F\colon D \to X$ extends the map f. We conclude that the space X is simply-connected, as required. ∎

One can show that, if two points x_1 and x_2 in a topological space X can be joined by a path in X then $\pi_1(X, x_1)$ and $\pi_1(X, x_2)$ are isomorphic. On combining this result with Theorem 3.4, we see that a path-connected topological space X is simply-connected if and only if $\pi_1(X, x)$ is trivial for some $x \in X$.

Theorem 3.5

Let X be a topological space, and let U and V be open subsets of X, with $U \cup V = X$. Suppose that U and V are simply-connected, and that $U \cap V$ is non-empty and path-connected. Then X is itself simply-connected.

Proof

We must show that any continuous function $f: \partial D \to X$ defined on the unit circle ∂D can be extended continuously over the closed unit disk D. Now the preimages $f^{-1}(U)$ and $f^{-1}(V)$ of U and V are open in ∂D (since f is continuous), and $\partial D = f^{-1}(U) \cup f^{-1}(V)$. It follows from the Lebesgue Lemma that there exists some $\delta > 0$ such that any arc in ∂D whose length is less than δ is entirely contained in one or other of the sets $f^{-1}(U)$ and $f^{-1}(V)$.

Choose points $z_1, z_2, \ldots, z_n$ around ∂D such that the distance from z_i to z_{i+1} is less than δ for $i = 1, 2, \ldots, n - 1$ and the distance from z_n to z_1 is also less than δ. Then, for each i, the short arc joining z_{i-1} to z_i is mapped by f into one or other of the open sets U and V.

Let x_0 be some point of $U \cap V$. Now the sets U, V and $U \cap V$ are all path-connected. Therefore we can choose paths $\alpha_i \colon [0, 1] \to X$ for $i = 1, 2, \ldots, n$ such that $\alpha_i(0) = x_0$, $\alpha_i(1) = f(z_i)$, $\alpha_i([0, 1]) \subset U$ whenever $f(z_i) \in U$, and $\alpha_i([0, 1]) \subset V$ whenever $f(z_i) \in V$. For convenience let $\alpha_0 = \alpha_n$.

Now, for each i, consider the sector T_i of the closed unit disk bounded by the line segments joining the centre of the disk to the points z_{i-1} and z_i and by the short arc joining z_{i-1} to z_i. Now this sector is homeomorphic to the closed unit disk, and therefore any continuous function mapping the boundary ∂T_i of T_i into a simply-connected space can be extended continuously over the whole of T_i. In particular, let F_i be the function on ∂T_i defined by

$$F_i(z) = \begin{cases} f(z) & \text{if } z \in T_i \cap \partial D, \\ \alpha_{i-1}(t) & \text{if } z = tz_{i-1} \text{ for some } t \in [0,1], \\ \alpha_i(t) & \text{if } z = tz_i \text{ for some } t \in [0,1], \end{cases}$$

Note that $F_i(\partial T_i) \subset U$ whenever the short arc joining z_{i-1} to z_i is mapped by f into U, and $F_i(\partial T_i) \subset V$ whenever this short arc is mapped into V.

Now U and V are both simply-connected. It follows that each of the functions F_i can be extended continuously over the whole of the sector T_i. Moreover the functions defined in this fashion on each of the sectors T_i agree with one another wherever the sectors intersect, and can therefore be pieced together to yield a continuous map defined over the the whole of the closed disk D which extends the map f, as required. ∎

Example

The n-dimensional sphere S^n is simply-connected for all $n > 1$, where $S^n = \{\mathbf{x} \in \mathbb{R}^{n+1} : |\mathbf{x}| = 1\}$. Indeed let $U = \{\mathbf{x} \in S^n : x_{n+1} > -\frac{1}{2}\}$ and $V = \{\mathbf{x} \in S^n : x_{n+1} < \frac{1}{2}\}$. Then U and V are homeomorphic to an n-dimensional ball, and are therefore simply-connected. Moreover $U \cap V$ is path-connected, provided that $n > 1$. It follows that S^n is simply-connected for all $n > 1$.

3.4. The Fundamental Group of the Circle

Proposition 3.6

Let S^1 be the unit circle in the Euclidean plane, defined so that

$$S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\},$$

and let $\gamma\colon [a, b] \to S^1$ be a continuous map into S^1 defined on a closed bounded interval $[a, b]$. Then there exists a continuous real-valued function $\tilde{\gamma}\colon [a, b] \to \mathbb{R}$ on the interval $[a, b]$ with the property that

$$(\cos 2\pi\tilde{\gamma}(t), \sin 2\pi\tilde{\gamma}(t)) = \gamma(t)$$

for all $t \in [a, b]$.

Proof

Let $\gamma(t) = \big(\gamma_1(t), \gamma_2(t)\big)$ for all $t \in [a, b]$ and let $\eta\colon [a, b] \to \mathbb{C}$ be the continuous map into the complex plane defined such that $\eta(t) = \gamma_1(t) + i\gamma_2(t)$ for all $t \in [a, b]$, where $i^2 = -1$. Now $|\eta(t)| = 1$ for all $t \in [a, b]$. It follows from the path-lifting property of the exponential map (Theorem 2.5) that there exists a continuous map $\tilde{\eta}\colon [a, b] \to \mathbb{C}$ with the property that $\exp(\tilde{\eta}(t)) = \eta(t)$ for all $t \in [a, b]$. Moreover $\mathrm{Re}[\tilde{\eta}(t)] = 0$ for all $t \in [a, b]$ (where $\mathrm{Re}[\tilde{\eta}(t)]$ denotes the real part of $\tilde{\eta}(t)$), because $|\eta(t)| = 1$ for all $t \in [a, b]$. Therefore there exists a continuous map $\tilde{\gamma}\colon [a, b] \to \mathbb{R}$ such that $\tilde{\eta}(t) = 2\pi i\,\tilde{\gamma}(t)$ for all $t \in [a, b]$. Then

$$
\begin{aligned}
\cos 2\pi\tilde{\gamma}(t) + i\,\sin 2\pi\tilde{\gamma}(t) &= \exp(2\pi i\,\tilde{\gamma}(t)) = \exp(\tilde{\eta}(t)) \\
&= \eta(t) = \gamma_1(t) + i\,\gamma_2(t)
\end{aligned}
$$

for all $t \in [a, b]$. The result follows. ∎

Let

$$S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

and let $p\colon \mathbb{R} \to S^1$ be defined so that $p(t) = (\cos 2\pi t, \sin 2\pi t)$ for all $t \in \mathbb{R}$. This function p has the following periodicity property:

> *real numbers s and t satisfy $p(s) = p(t)$ if and only if $s - t$ is an integer.*

It follows from Proposition 3.6 that, given any loop $\gamma\colon [0, 1] \to S^1$ in the circle S^1, there exists a continuous real-valued function $\tilde{\gamma}\colon [0, 1] \to \mathbb{R}$ with the property that $p \circ \tilde{\gamma} = \gamma$. Then $p(\tilde{\gamma}(1)) = p(\tilde{\gamma}(0))$. It follows from the periodicity property of the function p that $\tilde{\gamma}(1) - \tilde{\gamma}(0)$ is an integer. We now that the value of this integer is determined by the loop γ, and does not depend on the choice of function $\tilde{\gamma}$, provided that $p \circ \tilde{\gamma} = \gamma$.

If $\eta\colon [0,1] \to \mathbb{R}$ is a continuous function with the property that $p \circ \eta = \gamma$ then $p \circ \eta = p \circ \tilde{\gamma}$ and therefore

$$\eta(t) - \tilde{\gamma}(t) \in \mathbb{Z}$$

for all $t \in [0,1]$. But $\eta(t) - \tilde{\gamma}(t)$ is a continuous function of t on $[0,1]$, and the connectedness of $[0,1]$ ensures that every continuous integer-valued function on $[0,1]$ is constant (Corollary 1.58). It follows that there exists some integer m with the property that $\eta(t) = \tilde{\gamma}(t) + m$ for all $t \in [0,1]$, where the value of m is independent of t. But then $\eta(1) - \eta(0) = \tilde{\gamma}(1) - \tilde{\gamma}(0)$. It follows that the loop γ determines a well-defined integer $n(\gamma)$ characterized by the property that $n(\gamma) = \tilde{\gamma}(1) - \tilde{\gamma}(0)$ for all continuous real-valued functions $\tilde{\gamma}\colon [0,1] \to \mathbb{R}$ on $[0,1]$ that satisfy $p \circ \tilde{\gamma} = \gamma$.

Definition

Let $\gamma\colon [0,1] \to S^1$ be a loop in the circle S^1, where

$$S^1 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

The *winding number* $n(\gamma)$ of γ is defined to be unique integer characterized by the property that

$$n(\gamma) = \tilde{\gamma}(1) - \tilde{\gamma}(0)$$

for all continuous functions $\tilde{\gamma}\colon [0,1] \to \mathbb{R}$ that satisfy

$$(\cos 2\pi\tilde{\gamma}(t), \sin 2\pi\tilde{\gamma}(t)) = \gamma(t)$$

for all $t \in [0,1]$.

Proposition 3.7

Let

$$S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\},$$

let $H\colon [0,1] \times [0,1] \to S^1$ *be a continuous map that satisfies* $H(0,\tau) = H(1,\tau)$ *for all* $\tau \in [0,1]$*. Also, for each* $\tau \in [0,1]$*, let* $n(\gamma_\tau)$ *be the winding number of the loop* γ_τ *in* S^1 *defined such that* $\gamma_\tau(t) = H(t,\tau)$ *for all* $t \in [0,1]$*. Then* $n(\gamma_0) = n(\gamma_1)$*.*

Proof

Let $G = T \circ H$, where $T\colon \mathbb{R}^2 \to \mathbb{C}$ is defined so that $T(x, y) = x + iy$ for all real numbers x and y. Then $G(t, \tau) = T \circ \gamma_\tau(t)$ for all $t \in [0, 1]$ and $\tau \in [0, 1]$. Moreover $n(\gamma_\tau) = n(T \circ \gamma_\tau, 0)$ for all $\tau \in [0, 1]$, where $n(T \circ \gamma_\tau, 0)$ denotes the winding number of the closed curve $T \circ \gamma_\tau$ around zero. It therefore follows from Proposition 2.9 that

$$n(\gamma_0) = n(T \circ \gamma_0, 0) = n(T \circ \gamma_1, 0) = n(\gamma_1),$$

as required. ∎

Corollary 3.8

Let S^1 be the unit circle in the Euclidean plane, defined so that

$$S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\},$$

and let $\mathbf{b}$ *be a point of* S^1*. Let* α *and* β *be loops in* S^1 *based at* $\mathbf{b}$*. Suppose that* $\alpha \simeq \beta \ \mathrm{rel}\ \{0, 1\}$*. Then* $n(\alpha) = n(\beta)$*, where* $n(\alpha)$ *and* $n(\beta)$ *denote the winding numbers of the loops* α *and* β *respectively.*

Proof

The loops α and β satisfy $\alpha \simeq \beta \ \mathrm{rel}\ \{0, 1\}$ if and only if there exists a homotopy $H\colon [0, 1] \times [0, 1] \to S^1$ with the following properties: $H(t, 0) = \alpha(t)$ and $H(t, 1) = \beta(t)$ for all $t \in [0, 1]$; $H(0, \tau) = H(1, \tau) = \mathbf{b}$ for all $\tau \in [0, 1]$. The result therefore follows directly from Proposition 3.7. ∎

Theorem 3.9

Let S^1 be the unit circle in the Euclidean plane, defined so that

$$S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\},$$

and let $\mathbf{b}$ be a point of S^1. Then the function sending each loop γ in S^1 based at $\mathbf{b}$ to its winding number $n(\gamma)$ induces an isomorphism from the fundamental group $\pi_1(S^1, \mathbf{b})$ of the circle S^1 to the group $\mathbb{Z}$ of integers.

Proof

Let $p\colon \mathbb{R} \to S^1$ denote the function from $\mathbb{R}$ to S^1 defined so that

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

for all real numbers t. Also, for each loop $\gamma\colon [0,1] \to S^1$ in S^1 based at $\mathbf{b}$ let $[\gamma]$ denote the element of the fundamental group $\pi_1(S^1, \mathbf{b})$ determined by γ, and let $n(\gamma)$ denote the winding number of γ. Every element of $\pi_1(S^1, \mathbf{b})$ is the based homotopy class $[\gamma]$ of some loop γ in S^1 based at $\mathbf{b}$. If $\tilde{\gamma}\colon [0,1] \to \mathbb{R}$ is a real-valued function for which $p \circ \tilde{\gamma} = \gamma$ then $n(\gamma) = \tilde{\gamma}(1) - \tilde{\gamma}(0)$.

Let α and β be loops in S^1 based at $\mathbf{b}$. Suppose that $[\alpha] = [\beta]$. Then $\alpha \simeq \beta \ \mathrm{rel}\ \{0,1\}$. It then follows from Corollary 3.8 that $n(\alpha) = n(\beta)$. It follows from this that there is a well-defined function $\lambda\colon \pi_1(S^1, \mathbf{b}) \to \mathbb{Z}$ characterized by the property that $\lambda([\gamma]) = n(\gamma)$ for all loops γ in S^1 based at $\mathbf{b}$.

Next we show that the function $\lambda\colon \pi_1(S^1, \mathbf{b}) \to \mathbb{Z}$ is a homomorphism. Let $\alpha\colon [0,1] \to S^1$ and $\beta\colon [0,1] \to S^1$ be loops in S^1 based at $\mathbf{b}$. Then there exists a continuous real-valued function $\eta\colon [0,1] \to \mathbb{R}$ with the property that

$$p(\eta(t)) = \begin{cases} \alpha(2t) & \text{if } 0 \le t \le \frac{1}{2}, \\ \beta(2t-1) & \text{if } \frac{1}{2} \le t \le 1, \end{cases}$$

where $p(t) = (\cos 2\pi t, \sin 2\pi t)$ for all $t \in \mathbb{R}$ (see Proposition 3.6). Then $\alpha(t) = p(\eta(\frac{1}{2}t))$ for all $t \in [0,1]$. It follows from the definition of winding numbers that $n(\alpha) = \eta(\frac{1}{2}) - \eta(0)$. Also $\beta(t) = p(\eta(\frac{1}{2}(t+1)))$ for all $t \in [0,1]$, and therefore $n(\beta) = \eta(1) - \eta(\frac{1}{2})$. It follows that

$$n(\alpha) + n(\beta) = \eta(1) - \eta(0) = n(p \circ \eta) = n(\alpha\,.\,\beta),$$

where $\alpha\,.\,\beta$ is the concatenation of the loops α and β. It follows that

$$\lambda([\alpha]) + \lambda([\beta]) = n(\alpha) + n(\beta) = n(\alpha \,.\, \beta) = \lambda([\alpha \,.\, \beta]) = \lambda([\alpha][\beta]).$$

We conclude that $\lambda\colon \pi_1(S^1, \mathbf{b}) \to \mathbb{Z}$ is a homomorphism.

Next we show that $\lambda\colon \pi_1(S^1, \mathbf{b}) \to \mathbb{Z}$ is injective. Let α and β be loops in S^1 for which $n(\alpha) = n(\beta)$. Then there exist real-valued functions $\tilde{\alpha}\colon [0,1] \to \mathbb{R}$ and $\tilde{\beta}\colon [0,1] \to \mathbb{R}$ for which $\alpha = p \circ \tilde{\alpha}$ and $\beta = p \circ \tilde{\beta}$ (Proposition 3.6). Moreover

$$\tilde{\alpha}(1) - \tilde{\alpha}(0) = n(\alpha) = n(\beta) = \tilde{\beta}(1) - \tilde{\beta}(0).$$

Also $p(\tilde{\alpha}(0)) = \mathbf{b} = p(\tilde{\beta}(0))$, and therefore there exists some integer m for which $\tilde{\beta}(0) = \tilde{\alpha}(0) + m$. Then

$$\tilde{\beta}(1) = \tilde{\beta}(1) - \tilde{\beta}(0) + \tilde{\alpha}(0) + m = \tilde{\alpha}(1) + m.$$

Let
$$F(t,\tau) = (1-\tau)\tilde{\alpha}(t) + \tau(\tilde{\beta}(t) - m).$$
Then $F(t,0) = \tilde{\alpha}(t)$ and $F(t,1) = \tilde{\beta}(t) - m$ for all $t \in [0,1]$. Also $F(0,\tau) = \tilde{\alpha}(0)$ and $F(1,\tau) = \tilde{\alpha}(1)$ for all $\tau \in [0,1]$. Let $H\colon [0,1] \times [0,1] \to S^1$ be defined so that $H(t,\tau) = p(F(t,\tau))$ for all $t \in [0,1]$ and $\tau \in [0,1]$. Then $H(t,0) = \alpha(t)$ and $H(t,1) = \beta(t)$ for all $t \in [0,1]$. Also $H(0,\tau) = H(1,\tau) = \mathbf{b}$ for all $\tau \in [0,1]$. It follows that $\alpha \simeq \beta \ \mathrm{rel}\ \{0,1\}$ and therefore $[\alpha] = [\beta]$ in $\pi_1(X,\mathbf{b})$. We conclude therefore that $\lambda\colon \pi_1(S^1,\mathbf{b}) \to \mathbb{Z}$ is injective.

Let m be an integer, let t_0 be a real number for which $p(t_0) = \mathbf{b}$, and let $\gamma(t) = p(t_0 + mt)$ for all $t \in [0, 1]$. Then $\gamma\colon [0, 1] \to S^1$ is a loop in S^1 based at $\mathbf{b}$, and $\lambda([\gamma]) = n(\gamma) = m$. We conclude that $\lambda\colon \pi_1(S^1, \mathbf{b}) \to \mathbb{Z}$ is surjective. We have now shown that the function λ is a homomorphism that is both injective and surjective. It follows that $\lambda\colon \pi_1(S^1, \mathbf{b}) \to \mathbb{Z}$ is an isomorphism. This completes the proof. $\blacksquare$

Proposition 3.10

Let $X = \mathbb{R}^2 \setminus \{(0,0)\}$. *Then* $\pi_1(X, (1,0)) \cong \mathbb{Z}$.

Proof

Let

$$S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1,$$

let $i\colon S^1 \to X$ be the inclusion map, and let $r\colon X \to S^1$ be the radial projection map, defined such that

$$r(x, y) = \left(\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}} \right)$$

for all $(x, y) \in X$. Now the composition map $r \circ i$ is the identity map of S^1. Let

$$u(x, y, \tau) = \frac{1 - \tau}{\sqrt{x^2 + y^2}} + \tau$$

for all $(x, y) \in X$ and $\tau \in [0, 1]$. Then the function $F\colon X \times [0, 1] \to X$ that sends $((x, y), \tau) \in X \times [0, 1]$ to $(u(x, y, \tau)x, u(x, y, \tau)y)$ is a homotopy between the composition map $i \circ r$ and the identity map of the punctured plane X. Moreover $F((x, y), \tau) = (x, y)$ for all $(x, y) \in S^1$ and $\tau \in [0, 1]$.

Let $\gamma\colon [0, 1] \to X$ be a loop in X based at $(1, 0)$ and let $H\colon [0, 1] \times [0, 1] \to X$ be defined so that $H(t, \tau) = F(\gamma(t), \tau)$ for all $t \in [0, 1]$ and $\tau \in [0, 1]$. Then $H(t, 0) = r(\gamma(t))$ and $H(t, 1) = \gamma(t)$ for all $t \in [0, 1]$, and $H(0, \tau) = H(1, \tau) = (1, 0)$ for all $\tau \in [0, 1]$, and therefore $i \circ r \circ \gamma \simeq \gamma \ \mathrm{rel}\ \{0, 1\}$.

Now the continuous maps $i\colon S^1 \to X$ and $r\colon X \to S^1$ induce well-defined homomorphisms $i_\#\colon \pi_1(S^1,(1,0)) \to \pi_1(X,(1,0))$ and $r_\#\colon \pi_1(X,(1,0)) \to \pi_1(S^1,(1,0))$, where $i_\#[\eta] = [i \circ \eta]$ for all loops η in S^1 based at $(1,0)$ and $r_\#[\gamma] = [r \circ \gamma]$ for all loops γ in X based at $(1,0)$. Moreover

$$i_\#(r_\#([\gamma]) = i_\#([r \circ \gamma]) = [i \circ r \circ \gamma] = [\gamma]$$

for all loops γ in X based at $(1,0)$, and

$$r_\#(i_\#([\eta])r_\#[i \circ \eta] = [r \circ i \circ \eta] = [\eta]$$

for all loops η in S^1 based at $(1,0)$. It follows that the homomorphism $i_\#\colon \pi_1(S^1,(1,0)) \to \pi_1(X,(1,0))$ is an isomorphism whose inverse is the homomorphism $r_\#\colon \pi_1(X,(1,0)) \to \pi_1(S^1,(1,0))$, and therefore

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

as required. $\blacksquare$

Example

Let D be the closed unit disk in $\mathbb{R}^2$ and let ∂D be its boundary circle, where

$$\begin{aligned} D^2 &= \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}, \\ \partial D^2 = S^1 &= \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}, \end{aligned}$$

let $i\colon \partial D \to D$ be the inclusion map, and let $\mathbf{b} = (1,0)$. Suppose there were to exist a continuous map $r\colon D \to \partial D$ with the property that $r(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \partial D$. Then $r \circ i\colon \partial D \to \partial D$ would be the identity map of the unit circle ∂D. It would then follow that $r_\# \circ i_\#$ would be the identity isomorphism of $\pi_1(\partial D, \mathbf{b})$, where $i_\#\colon \pi_1(\partial D, \mathbf{b}) \to \pi_1(D,)$ and $r_\#\colon \pi_1(D, \mathbf{b}) \to \pi_1(\partial D,)$ denote the homomorphisms of fundamental groups induced by $i\colon \partial D \to D$ and $r\colon D \to \partial D$ respectively.

But $\pi_1(D, \mathbf{b})$ is the trivial group, because D is a convex set in $\mathbb{R}^2$, and $\pi_1(\partial D, \mathbf{b}) \cong \mathbb{Z}$ (Theorem 3.9). It follows that the identity homomorphism of $\pi_1(D, \mathbf{b})$ cannot be expressed as a composition of two homomorphisms $\theta \circ \varphi$ where θ is a homomorphism from $\pi_1(\partial D, \mathbf{b})$ to $\pi_1(D, \mathbf{b})$ and φ is a homomorphism from $\pi_1(D, \mathbf{b})$ to $\pi_1(\partial D, \mathbf{b})$. Therefore there cannot exist any continous map $r \colon D \to \partial D$ with the property that $r(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \partial D$. This result has already been established (see Corollary 2.15). Moreover the result is used to establish the Brouwer Fixed Point Theorem in the two-dimensional case (Theorem 2.16) which ensures that every continuous map from the two-dimensional closed disk D^2 to itself has a fixed point.