Course 421: Algebraic Topology
Section 2: Homotopies and the Fundamental Group

David R. Wilkins

Contents

2 Homotopies and the Fundamental Group

2.1 Homotopies

Definition Let $f: X \to Y$ and $g: X \to Y$ 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 $H: X \times [0,1] \to Y$ 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.

Lemma 2.1 *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 Clearly $f \simeq f$, since $(x,t) \mapsto f(x)$ is a homotopy between f and itself. Thus the relation is reflexive. If $f \simeq g$ then there exists a homotopy $H: X \times [0,1] \to Y$ between f and g (so that $H(x,0) = f(x)$ and $H(x,1) = g(x)$ for all $x \in X$). But then $(x,t) \mapsto H(x, 1-t)$ is a homotopy between g and f. Therefore $f \simeq g$ if and only if $g \simeq f$. Thus the relation is symmetric. Finally, suppose that $f \simeq g$ and $g \simeq h$. Then there exist homotopies $H_1: X \times [0,1] \to Y$ and $H_2: X \times [0,1] \to Y$ such that $H_1(x,0) = f(x)$, $H_1(x,1) = g(x) = H_2(x,0)$ and $H_2(x,1) = h(x)$ for all $x \in X$. Define $H: X \times [0,1] \to Y$ 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 elementary point set topology 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 $f \simeq h$. Thus the relation is transitive. The relation $\simeq$ is therefore an equivalence relation. ∎

Definition Let X and Y be topological spaces, and let A be a subset of X. Let $f: X \to Y$ and $g: X \to Y$ 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 $f \simeq g$ rel A) if and only if there exists a (continuous) homotopy $H: X \times [0,1] \to Y$ 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 $a \in A$.

Homotopy relative to a chosen subset of X is also an equivalence relation on the set of all continuous maps between topological spaces X and Y.

2.2 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

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

(The continuity of $\gamma_1.\gamma_2$ may be deduced from Lemma 2.1.)

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 $\gamma_0 \simeq \gamma_1$ 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 $\gamma_0 \simeq \gamma_1$ 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 2.2 *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

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

where $F_1: [0,1] \times [0,1] \to X$ is a homotopy between γ_1 and γ_1', $F_2: [0,1] \times [0,1] \to X$ 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

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

Thus the map $G: [0,1] \times [0,1] \to X$ 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 $(\gamma_1.\gamma_2).\gamma_3 \simeq \gamma_1.(\gamma_2.\gamma_3)$ 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 $\varepsilon: [0,1] \to X$ 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

$$\theta_0(t) = \begin{cases} 0 & \text{if } 0 \le t \le \frac{1}{2}, \\ 2t - 1 & \text{if } \frac{1}{2} \le t \le 1, \end{cases} \qquad \theta_1(t) = \begin{cases} 2t & \text{if } 0 \le t \le \frac{1}{2}, \\ 1 & \text{if } \frac{1}{2} \le t \le 1, \end{cases}$$

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 $\varepsilon.\gamma \simeq \gamma \simeq \gamma.\varepsilon$ 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 $K: [0,1] \times [0,1] \to X$ defined by

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

is a homotopy between the loops $\gamma.\gamma^{-1}$ and ε, and moreover this homotopy sends $(0,\tau)$ and $(1,\tau)$ to x_0 for all $\tau \in [0,1]$. Therefore $\gamma.\gamma^{-1} \simeq \varepsilon \,\mathrm{rel}\{0,1\}$, 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 $f: X \to Y$ be a continuous map between topological spaces X and Y, and let x_0 be a point of X. Then f induces a homomorphism $f_\#: \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$, where $f_\#([\gamma]) = [f \circ \gamma]$ for all loops $\gamma: [0,1] \to X$ based at x_0. If x_0, y_0 and z_0 are points belonging to topological spaces X, Y and Z, and if $f: X \to Y$ and $g: Y \to Z$ are continuous maps satisfying $f(x_0) = y_0$ and $g(y_0) = z_0$, then the induced homomorphisms $f_\#: \pi_1(X, x_0) \to \pi_1(Y, y_0)$ and $g_\#: \pi_1(Y, x_0) \to \pi_1(Z, z_0)$ 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.

2.3 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 $h: D \to E$ is a homeomorphism from D to E. Then any continuous map $g: \partial E \to X$ mapping ∂E into a simply-connected space X extends continuously to the whole of E. Indeed there exists a continuous map $F: D \to X$ which extends $g \circ h: \partial D \to X$, and the map $F \circ h^{-1}: E \to X$ then extends the map g.

Theorem 2.3 *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 $\gamma: [0,1] \to X$ 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 $F: [0,1] \times [0,1] \to X$ 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 $f: \partial D \to X$ 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 $G: [0, 1] \times [0, 1] \to X$ 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

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

for all $t, \tau \in [0, 1]$. It follows that there is a well-defined function $F: D \to X$ such that $F \circ q = G$. However $q: [0, 1] \times [0, 1] \to D$ is a continuous surjection from a compact space to a Hausdorff space and is therefore an identification map. It follows that $F: D \to X$ is continuous (since a basic property of identification maps ensures that a function $F: D \to X$ is continuous if and only if $F \circ q: [0, 1] \times [0, 1] \to X$ is continuous). Moreover $F: D \to X$ 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 2.3, 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 2.4 *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 $\alpha_i: [0, 1] \to X$ 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

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

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. But 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 $n > 1$.