Course 421: Academic Year 1998-9
Part II: Covering Maps and the Fundamental Group

D. R. Wilkins

Contents

6 Homotopies and Covering Maps

6.1 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 6.1 *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 $f \simeq g \text{ rel } A$) 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.

6.2 Covering Maps

Definition Let X and $\tilde{X}$ be topological spaces and let $p: \tilde{X} \to X$ be a continuous map. An open subset U of X is said to be *evenly covered* by the map p if and only if $p^{-1}(U)$ is a disjoint union of open sets of $\tilde{X}$ each of which is mapped homeomorphically onto U by p. The map $p: \tilde{X} \to X$ is said to be a *covering map* if $p: \tilde{X} \to X$ is surjective and in addition every point of X is contained in some open set that is evenly covered by the map p.

If $p: \tilde{X} \to X$ is a covering map, then we say that $\tilde{X}$ is a *covering space* of X.

Example Let S^1 be the unit circle in $\mathbb{R}^2$. Then the map $p: \mathbb{R} \to S^1$ defined by

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

is a covering map. Indeed let $\mathbf{n}$ be a point of S^1. Consider the open set U in S^1 containing $\mathbf{n}$ defined by $U = S^1 \setminus \{-\mathbf{n}\}$. Now $\mathbf{n} = (\cos 2\pi t_0, \sin 2\pi t_0)$ for some $t_0 \in \mathbb{R}$. Then $p^{-1}(U)$ is the union of the disjoint open sets J_n for all integers n, where

$$J_n = \{t \in \mathbb{R} : t_0 + n - \tfrac{1}{2} < t < t_0 + n + \tfrac{1}{2}\}.$$

Each of the open sets J_n is mapped homeomorphically onto U by the map p. This shows that $p: \mathbb{R} \to S^1$ is a covering map.

Example The map $p: \mathbb{C} \to \mathbb{C} \setminus \{0\}$ defined by $p(z) = \exp(z)$ is a covering map. Indeed, given any $\theta \in [-\pi, \pi]$ let us define

$$U_\theta = \{z \in \mathbb{C} \setminus \{0\} : \arg(-z) \neq \theta\}.$$

Then $p^{-1}(U_\theta)$ is the disjoint union of the open sets

$$\{z \in \mathbb{C} : |\mathrm{Im}\, z - \theta - 2\pi n| < \pi\},$$

for all integers n, and p maps each of these open sets homeomorphically onto U_θ. Thus U_θ is evenly covered by the map p.

Example Consider the map $\alpha: (-2, 2) \to S^1$, where $\alpha(t) = (\cos 2\pi t, \sin 2\pi t)$ for all $t \in (-2, 2)$. It can easily be shown that there is no open set U containing the point $(1, 0)$ that is evenly covered by the map α. Indeed suppose that there were to exist such an open set U. Then there would exist some δ satisfying $0 < \delta < \frac{1}{2}$ such that $U_\delta \subset U$, where

$$U_\delta = \{(\cos 2\pi t, \sin 2\pi t) : -\delta < t < \delta\}.$$

The open set U_δ would then be evenly covered by the map α. However the connected components of $\alpha^{-1}(U_\delta)$ are $(-2, -2+\delta)$, $(-1-\delta, -1+\delta)$, $(-\delta, \delta)$, $(1-\delta, 1+\delta)$ and $(2-\delta, 2)$, and neither $(-2, -2+\delta)$ nor $(2-\delta, 2)$ is mapped homeomorphically onto U_δ by α.

Lemma 6.2 *Let $p: \tilde{X} \to X$ be a covering map. Then $p(V)$ is open in X for every open set V in $\tilde{X}$. In particular, a covering map $p: \tilde{X} \to X$ is a homeomorphism if and only if it is a bijection.*

Proof Let V be open in $\tilde{X}$, and let $x \in p(V)$. Then $x = p(v)$ for some $v \in V$. Now there exists an open set U containing the point x which is evenly covered by the covering map p. Then $p^{-1}(U)$ is a disjoint union of open sets, each of which is mapped homeomorphically onto U by the covering map p. One of these open sets contains v; let $\tilde{U}$ be this open set, and let $N_x = p(V \cap \tilde{U})$. Now N_x is open in X, since $V \cap \tilde{U}$ is open in $\tilde{U}$ and $p|\tilde{U}$ is a homeomorphism from $\tilde{U}$ to U. Also $x \in N_x$ and $N_x \subset p(V)$. It follows that $p(V)$ is the union of the open sets N_x as x ranges over all points of $p(V)$, and thus $p(V)$ is itself an open set, as required. The result that a bijective covering map is a homeomorphism then follows directly from the fact that a continuous bijection is a homeomorphism if and only if it maps open sets to open sets. ∎

6.3 Path Lifting and the Monodromy Theorem

Let $p: \tilde{X} \to X$ be a covering map over a topological space X. Let Z be a topological space, and let $f: Z \to X$ be a continuous map from Z to X. A continuous map $\tilde{f}: Z \to \tilde{X}$ is said to be a *lift* of the map $f: Z \to X$ if and only if $p \circ \tilde{f} = f$. We shall prove various results concerning the existence and uniqueness of such lifts.

Proposition 6.3 *Let $p: \tilde{X} \to X$ be a covering map, let Z be a connected topological space, and let $g: Z \to \tilde{X}$ and $h: Z \to \tilde{X}$ be continuous maps. Suppose that $p \circ g = p \circ h$ and that $g(z) = h(z)$ for some $z \in Z$. Then $g = h$.*

Proof Let $Z_0 = \{z \in Z : g(z) = h(z)\}$. Note that Z_0 is non-empty, by hypothesis. We show that Z_0 is both open and closed in Z.

Let z be a point of Z. There exists an open set U in X containing the point $p(g(z))$ which is evenly covered by the covering map p. Then $p^{-1}(U)$ is a disjoint union of open sets, each of which is mapped homeomorphically onto U by the covering map p. One of these open sets contains $g(z)$; let this set be denoted by $\tilde{U}$. Also one of these open sets contains $h(z)$; let this open set be denoted by $\tilde{V}$. Let $N_z = g^{-1}(\tilde{U}) \cap h^{-1}(\tilde{V})$. Then N_z is an open set in Z containing z.

Consider the case when $z \in Z_0$. Then $g(z) = h(z)$, and therefore $\tilde{V} = \tilde{U}$. It follows from this that both g and h map the open set N_z into $\tilde{U}$. But $p \circ g = p \circ h$, and $p|\tilde{U}: \tilde{U} \to U$ is a homeomorphism. Therefore $g|N_z = h|N_z$, and thus $N_z \subset Z_0$. We have thus shown that, for each $z \in Z_0$, there exists an open set N_z such that $z \in N_z$ and $N_z \subset Z_0$. We conclude that Z_0 is open.

Next consider the case when $z \in Z \setminus Z_0$. In this case $\tilde{U} \cap \tilde{V} = \emptyset$, since $g(z) \neq h(z)$. But $g(N_z) \subset \tilde{U}$ and $h(N_z) \subset \tilde{V}$. Therefore $g(z') \neq h(z')$ for all $z' \in N_z$, and thus $N_z \subset Z \setminus Z_0$. We have thus shown that, for each $z \in Z \setminus Z_0$, there exists an open set N_z such that $z \in N_z$ and $N_z \subset Z \setminus Z_0$. We conclude that $Z \setminus Z_0$ is open.

The subset Z_0 of Z is therefore both open and closed. Also Z_0 is non-empty by hypothesis. We deduce that $Z_0 = Z$, since Z is connected. Thus $g = h$, as required. ∎

Lemma 6.4 *Let $p\colon \tilde{X} \to X$ be a covering map, let Z be a topological space, let A be a connected subset of Z, and let $f\colon Z \to X$ and $g\colon A \to \tilde{X}$ be continuous maps with the property that $p \circ g = f|A$. Suppose that $f(Z) \subset U$, where U is an open subset of X that is evenly covered by the covering map p. Then there exists a continuous map $\tilde{f}\colon Z \to \tilde{X}$ such that $\tilde{f}|A = g$ and $p \circ \tilde{f} = f$.*

Proof The open set U is evenly covered by the covering map p, and therefore $p^{-1}(U)$ is a disjoint union of open sets, each of which is mapped homeomorphically onto U by the covering map p. One of these open sets contains $g(a)$ for some $a \in A$; let this set be denoted by $\tilde{U}$. Let $\sigma\colon U \to \tilde{U}$ be the inverse of the homeomorphism $p|\tilde{U}\colon \tilde{U} \to U$, and let $\tilde{f} = \sigma \circ f$. Then $p \circ \tilde{f} = f$. Also $p \circ \tilde{f}|A = p \circ g$ and $\tilde{f}(a) = g(a)$. It follows from Proposition 6.3 that $\tilde{f}|A = g$, since A is connected. Thus $\tilde{f}\colon Z \to \tilde{X}$ is the required map. ∎

Theorem 6.5 (Path Lifting Theorem) *Let $p\colon \tilde{X} \to X$ be a covering map, let $\gamma\colon [0,1] \to X$ be a continuous path in X, and let w be a point of $\tilde{X}$ satisfying $p(w) = \gamma(0)$. Then there exists a unique continuous path $\tilde{\gamma}\colon [0,1] \to \tilde{X}$ such that $\tilde{\gamma}(0) = w$ and $p \circ \tilde{\gamma} = \gamma$.*

Proof The map $p\colon \tilde{X} \to X$ is a covering map; therefore there exists an open cover $\mathcal{U}$ of X such that each open set U belonging to X is evenly covered by the map p. Now the collection consisting of the preimages $\gamma^{-1}(U)$ of the open sets U belonging to $\mathcal{U}$ is an open cover of the interval $[0,1]$. But $[0,1]$ is compact, by the Heine-Borel Theorem. It follows from the Lebesgue Lemma that there exists some $\delta > 0$ such that every subinterval of length less than δ is mapped by γ into one of the open sets belonging to $\mathcal{U}$. Partition the interval $[0,1]$ into subintervals $[t_{i-1}, t_i]$, where $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = 1$, and where the length of each subinterval is less than δ. Then each subinterval $[t_{i-1}, t_i]$ is mapped by γ into some open set in X that is evenly covered by the map p. It follows from Lemma 6.4 that once $\tilde{\gamma}(t_{i-1})$ has been determined, we can extend $\tilde{\gamma}$ continuously over the ith subinterval $[t_{i-1}, t_i]$. Thus by extending $\tilde{\gamma}$ successively over $[t_0, t_1]$, $[t_1, t_2]$,..., $[t_{n-1}, t_n]$, we can lift the path $\gamma\colon [0,1] \to X$ to a path $\tilde{\gamma}\colon [0,1] \to \tilde{X}$ starting at w. The uniqueness of $\tilde{\gamma}$ follows from Proposition 6.3. ∎

Theorem 6.6 (The Monodromy Theorem) *Let $p\colon \tilde{X} \to X$ be a covering map, let $H\colon [0,1] \times [0,1] \to X$ be a continuous map, and let w be a point of $\tilde{X}$ satisfying $p(w) = H(0,0)$. Then there exists a unique continuous map $\tilde{H}\colon [0,1] \times [0,1] \to \tilde{X}$ such that $\tilde{H}(0,0) = w$ and $p \circ \tilde{H} = H$.*

Proof The unit square $[0,1] \times [0,1]$ is compact. By applying the Lebesgue Lemma to an open cover of the square by preimages of evenly covered open sets in X (as in the proof of Theorem 6.5), we see that there exists some $\delta > 0$ with the property that any square contained in $[0,1] \times [0,1]$ whose sides have length less than δ is mapped by H into some open set in X which is evenly covered by the covering map p. It follows from Lemma 6.4 that if the lift $\tilde{H}$ of H has already been determined over a corner, or along one side, or along two adjacent sides of a square whose sides have length less than δ, then $\tilde{H}$ can be extended over the whole of that square. Thus if we subdivide $[0,1] \times [0,1]$ into squares $S_{j,k}$, where

$$S_{j,k} = \left\{ (s,t) \in [0,1] \times [0,1] : \frac{j-1}{n} \leq s \leq \frac{j}{n} \text{ and } \frac{k-1}{n} \leq t \leq \frac{k}{n} \right\},$$

and $1/n < \delta$, then we can extend the map g to a lift $\tilde{H}$ of H by successively extending $\tilde{H}$ in turn over each of these smaller squares. (Indeed the map $\tilde{H}$ can be extended successively over the squares

$$S_{1,1}, S_{1,2}, \ldots, S_{1,n}, S_{2,1}, S_{2,2}, \ldots, S_{2,n}, S_{3,1}, \ldots, S_{n-1,n}, \ldots, S_{n,1}, S_{n,2}, \ldots, S_{n,n}.)$$

The uniqueness of $\tilde{H}$ follows from Proposition 6.3. ∎

7 Winding Numbers

7.1 Winding Numbers of Closed Curves

Let $\gamma\colon [0,1] \to \mathbb{C}$ be a continuous closed curve in the complex plane which is defined on some closed interval $[0,1]$ (so that $\gamma(0) = \gamma(1)$), and let w be a complex number which does not belong to the image of the closed curve γ. It then follows from the Path Lifting Theorem (Theorem 6.5) that there exists a continuous path $\tilde{\gamma}\colon [0,1] \to \mathbb{C}$ in $\mathbb{C}$ such that $\gamma(t) - w = \exp(\tilde{\gamma}(t))$ for all $t \in [0,1]$. Let us define

$$n(\gamma, w) = \frac{\tilde{\gamma}(1) - \tilde{\gamma}(0)}{2\pi i}.$$

Now $\exp(\tilde{\gamma}(1)) = \gamma(1) - w = \gamma(0) - w = \exp(\tilde{\gamma}(0))$ (since γ is a closed curve). It follows from this that $n(\gamma, w)$ is an integer. This integer is known as the *winding number* of the closed curve γ about w.

Lemma 7.1 *The value of the winding number $n(\gamma, w)$ does not depend on the choice of the lift $\tilde{\gamma}$ of the curve γ.*

Proof Let $\sigma\colon [0,1] \to \mathbb{C}$ be a continuous curve in $\mathbb{C}$ with the property that $\exp(\sigma(t)) = \gamma(t) - w = \exp(\tilde{\gamma}(t))$ for all $t \in [0,1]$. Then

$$\frac{\sigma(t) - \tilde{\gamma}(t)}{2\pi i}$$

is an integer for all $t \in [0,1]$. But the map sending $t \in [0,1]$ to $\sigma(t) - \tilde{\gamma}(t)$ is continuous on $[0,1]$. This map must therefore be a constant map, since the interval $[0,1]$ is connected. Thus there exists some integer m with the property that $\sigma(t) = \tilde{\gamma}(t) + 2\pi i m$ for all $t \in [0,1]$. But then

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

This proves that the value of the winding number $n(\gamma, w)$ of the closed curve γ about w is indeed independent of the choice of the lift $\tilde{\gamma}$ of γ. ∎

A continuous curve is said to be *piecewise* C^1 if it is made up of a finite number of continuously differentiable segments. We now show how the winding number of a piecewise C^1 closed curve in the complex plane can be expressed as a contour integral.

Proposition 7.2 *Let $\gamma\colon [0,1] \to \mathbb{C}$ be a piecewise C^1 closed curve in the complex plane, and let w be a point of $\mathbb{C}$ that does not lie on the curve γ. Then the winding number $n(\gamma, w)$ of γ about w is given by*

$$n(\gamma, w) = \frac{1}{2\pi i} \int_\gamma \frac{dz}{z - w}.$$

Proof By definition

$$n(\gamma, w) = \frac{\sigma(1) - \sigma(0)}{2\pi i},$$

where $\sigma\colon [0,1] \to \mathbb{C}$ is a path in $\mathbb{C}$ such that $\gamma(t) - w = \exp(\sigma(t))$ for all $t \in [0,1]$. Taking derivatives, we see that

$$\gamma'(t) = \exp(\sigma(t))\sigma'(t) = (\gamma(t) - w)\sigma'(t).$$

Thus

$$\begin{aligned} n(\gamma, w) &= \frac{\sigma(1) - \sigma(0)}{2\pi i} = \frac{1}{2\pi i} \int_0^1 \sigma'(t)\, dt \\ &= \frac{1}{2\pi i} \int_0^1 \frac{\gamma'(t)\, dt}{\gamma(t) - w} = \frac{1}{2\pi i} \int_\gamma \frac{dz}{z - w}. \quad \blacksquare \end{aligned}$$

One of the most important properties of winding numbers of closed curves in the complex plane is their invariance under continuous deformations of the closed curve.

Proposition 7.3 *Let w be a complex number and, for each $\tau \in [0,1]$, let $\gamma_\tau\colon [0,1] \to \mathbb{C}$ be a closed curve in $\mathbb{C}$ which does not pass through w. Suppose that the map sending $(t, \tau) \in [0,1] \times [0,1]$ to $\gamma_\tau(t)$ is a continuous map from $[0,1] \times [0,1]$ to $\mathbb{C}$. Then $n(\gamma_\tau, w) = n(\gamma_0, w)$ for all $\tau \in [0,1]$. In particular, $n(\gamma_1, w) = n(\gamma_0, w)$.*

Proof Let $H\colon [0,1] \times [0,1] \to \mathbb{C} \setminus \{0\}$ be defined by $H(t, \tau) = \gamma_\tau(t) - w$. It follows from the Monodromy Theorem (Theorem 6.6) that there exists a continuous map $\tilde{H}\colon [0,1] \times [0,1] \to \mathbb{C}$ such that $H = \exp \circ \tilde{H}$. But then

$$\tilde{H}(1, \tau) - \tilde{H}(0, \tau) = 2\pi i n(\gamma_\tau, w)$$

for all $\tau \in [0,1]$, and therefore the function $\tau \mapsto n(\gamma_\tau, w)$ is a continuous function on the interval $[0,1]$ taking values in the set $\mathbb{Z}$ of integers. But such a function must be constant on $[0,1]$, since the interval $[0,1]$ is connected. Thus $n(\gamma_0, w) = n(\gamma_1, w)$, as required. $\blacksquare$

Corollary 7.4 *Let $\gamma_0\colon [0,1] \to \mathbb{C}$ and $\gamma_1\colon [0,1] \to \mathbb{C}$ be continuous closed curves in $\mathbb{C}$, and let w be a complex number which does not lie on the images of the closed curves γ_0 and γ_1. Suppose that $|\gamma_1(t) - \gamma_0(t)| < |w - \gamma_0(t)|$ for all $t \in [0,1]$. Then $n(\gamma_0, w) = n(\gamma_1, w)$.*

Proof Let $\gamma_\tau(t) = (1-\tau)\gamma_0(t) + \tau\gamma_1(t)$ for all $t \in [0,1]$ and $\tau \in [0,1]$. Then

$$|\gamma_\tau(t) - \gamma_0(t)| = \tau|\gamma_1(t) - \gamma_0(t)| < |w - \gamma_0(t)|,$$

for all $t \in [0,1]$ and $\tau \in [0,1]$, and thus the closed curve γ_τ does not pass through w. The result therefore follows from Proposition 7.3. ∎

Corollary 7.5 *Let $\gamma\colon [0,1] \to \mathbb{C}$ be a continuous closed curve in $\mathbb{C}$, and let $\sigma\colon [0,1] \to \mathbb{C}$ be a continuous path in $\mathbb{C}$ whose image does not intersect the image of γ. Then $n(\gamma, \sigma(0)) = n(\gamma, \sigma(1))$. Thus the function $w \mapsto n(\gamma, w)$ is constant over each path-component of the set $\mathbb{C} \setminus \gamma([0,1])$.*

Proof For each $\tau \in [0,1]$, let $\gamma_\tau\colon [0,1] \to \mathbb{C}$ be the closed curve given by $\gamma_\tau(t) = \gamma(t) - \sigma(\tau)$. Then the closed curves γ_τ do not pass through 0 (since the curves γ and σ do not intersect), and the map from $[0,1] \times [0,1]$ to $\mathbb{C}$ sending (t, τ) to $\gamma_\tau(t)$ is continuous. It follows from Proposition 7.3 that

$$n(\gamma, \sigma(0)) = n(\gamma_0, 0) = n(\gamma_1, 0) = n(\gamma, \sigma(1)),$$

as required. ∎

7.2 The Fundamental Theorem of Algebra

Theorem 7.6 (The Fundamental Theorem of Algebra) *Let $P\colon \mathbb{C} \to \mathbb{C}$ be a non-constant polynomial with complex coefficients. Then there exists some complex number z_0 such that $P(z_0) = 0$.*

Proof The result is trivial if $P(0) = 0$. Thus it suffices to prove the result when $P(0) \neq 0$.

For any $r \geq 0$, let the closed curve σ_r denote the circle about zero of radius r, traversed once in the anticlockwise direction, given by $\sigma_r(t) = r\exp(2\pi i t)$ for all $t \in [0,1]$. Consider the winding number $n(P \circ \sigma_r, 0)$ of $P \circ \sigma_r$ about zero. We claim that this winding number is equal to m for large values of r, where m is the degree of the polynomial P.

Let $P(z) = a_0 + a_1 z + \cdots + a_m z^m$, where $a_1, a_2, \ldots, a_n$ are complex numbers, and where $a_m \neq 0$. We write $P(z) = P_m(z) + Q(z)$, where $P_m(z) = a_m z^m$ and $Q(z) = a_0 + a_1 z + \cdots + a_{m-1} z^{m-1}$. Let

$$R = \frac{|a_0| + |a_1| + \cdots + |a_m|}{|a_m|}.$$

If $|z| > R$ then

$$\left|\frac{Q(z)}{P_m(z)}\right| = \frac{1}{|a_m z|}\left|\frac{a_0}{z^{m-1}} + \frac{a_1}{z^{m-2}} + \cdots + a_{m-1}\right| < 1,$$

since $R \geq 1$, and thus $|P(z) - P_m(z)| < |P_m(z)|$. It follows from Corollary 7.4 that $n(P \circ \sigma_r, 0) = n(P_m \circ \sigma_r, 0) = m$ for all $r > R$.

Given $r \geq 0$, let $\gamma_\tau = P \circ \sigma_{\tau r}$ for all $\tau \in [0,1]$. Then $n(\gamma_0, 0) = 0$, since γ_0 is a constant curve with value $P(0)$. Thus if the polynomial P were everywhere non-zero, then it would follow from Proposition 7.3 that $n(\gamma_1, 0) = n(\gamma_0, 0) = 0$. But $n(\gamma_1, 0) = n(P \circ \sigma_r, 0) = m$ for all $r > R$, and $m > 0$. Therefore the polynomial P must have at least one zero in the complex plane. ∎

The proof of the Fundamental Theorem of Algebra given above depends on the continuity of the polynomial P, together with the fact that the winding number $n(P \circ \sigma_r, 0)$ is non-zero for sufficiently large r, where σ_r denotes the circle of radius r about zero, described once in the anticlockwise direction. We can therefore generalize the proof of the Fundamental Theorem of Algebra in order to obtain the following result (sometimes referred to as the *Kronecker Principle*).

Proposition 7.7 *Let $f\colon D \to \mathbb{C}$ be a continuous map defined on the closed unit disk D in $\mathbb{C}$, and let $w \in \mathbb{C} \setminus f(D)$. Then $n(f \circ \sigma, w) = 0$, where $\sigma\colon [0,1] \to \mathbb{C}$ is the parameterization of unit circle defined by $\sigma(t) = \exp(2\pi i t)$, and $n(f \circ \sigma, w)$ is the winding number of $f \circ \sigma$ about w.*

Proof Define $\gamma_\tau(t) = f(\tau \exp(2\pi i t))$ for all $t \in [0,1]$ and $\tau \in [0,1]$. Then none of the closed curves γ_τ passes through w, and γ_0 is the constant curve with value $f(0)$. It follows from Proposition 7.3 that

$$n(f \circ \sigma, w) = n(\gamma_1, w) = n(\gamma_0, w) = 0,$$

as required. ∎

7.3 The Brouwer Fixed Point Theorem

We now use Proposition 7.7 to show that there is no continuous 'retraction' mapping the closed unit disk onto its boundary circle.

Corollary 7.8 *There does not exist a continuous map $r\colon D \to \partial D$ with the property that $r(z) = z$ for all $z \in \partial D$, where ∂D denotes the boundary circle of the closed unit disk D.*

Proof Let $\sigma\colon [0,1] \to \mathbb{C}$ be defined by $\sigma(t) = \exp(2\pi i t)$. If a continuous map $r\colon D \to \partial D$ with the required property were to exist, then $r(z) \neq 0$ for all $z \in D$ (since $r(D) \subset \partial D$), and therefore $n(\sigma, 0) = n(r \circ \sigma, 0) = 0$, by Proposition 7.7. But $\sigma = \exp \circ \tilde{\sigma}$, where $\tilde{\sigma}(t) = 2\pi i t$ for all $t \in [0,1]$, and thus

$$n(\sigma, 0) = \frac{\tilde{\sigma}(1) - \tilde{\sigma}(0)}{2\pi i} = 1.$$

This shows that there cannot exist any continuous map r with the required property. ∎

Theorem 7.9 (The Brouwer Fixed Point Theorem in Two Dimensions) *Let $f\colon D \to D$ be a continuous map which maps the closed unit disk D into itself. Then there exists some $z_0 \in D$ such that $f(z_0) = z_0$.*

Proof Suppose that there did not exist any fixed point z_0 of $f\colon D \to D$. Then one could define a continuous map $r\colon D \to \partial D$ as follows: for each $z \in D$, let $r(z)$ be the point on the boundary ∂D of D obtained by continuing the line segment joining $f(z)$ to z beyond z until it intersects ∂D at the point $r(z)$. Then $r\colon D \to \partial D$ would be a continuous map, and moreover $r(z) = z$ for all $z \in \partial D$. But Corollary 7.8 shows that there does not exist any continuous map $r\colon D \to \partial D$ with this property. We conclude that $f\colon D \to D$ must have at least one fixed point. ∎

Remark The Brouwer Fixed Point Theorem is also valid in higher dimensions. This theorem states that any continuous map from the closed n-dimensional ball into itself must have at least one fixed point. The proof of the theorem for $n > 2$ is analogous to the proof for $n = 2$, once one has shown that there is no continuous map from the closed n-dimensional ball to its boundary which is the identity map on the boundary. However winding numbers cannot be used to prove this result, and thus more powerful topological techniques need to be employed.

7.4 The Borsuk-Ulam Theorem

Lemma 7.10 *Let $f\colon S^1 \to \mathbb{C} \setminus \{0\}$ be a continuous function defined on S^1, where $S^1 = \{z \in \mathbb{C} : |z| = 1\}$. Suppose that $f(-z) = -f(z)$ for all $z \in S^1$. Then the winding number $n(f \circ \sigma, 0)$ of $f \circ \sigma$ about 0 is odd, where $\sigma\colon [0,1] \to S^1$ is given by $\sigma(t) = \exp(2\pi i t)$.*

Proof It follows from the Path Lifting Theorem (Theorem 6.5) that there exists a continuous path $\tilde{\gamma}\colon [0,1] \to \mathbb{C}$ in $\mathbb{C}$ such that $\exp(\tilde{\gamma}(t)) = f(\sigma(t))$ for all $t \in [0,1]$. Now $f(\sigma(t + \frac{1}{2})) = -f(\sigma(t))$ for all $t \in [0, \frac{1}{2}]$, since $\sigma(t + \frac{1}{2}) =$

$-\sigma(t)$ and $f(-z) = -f(z)$ for all $z \in \mathbb{C}$. Thus $\exp(\tilde{\gamma}(t+\frac{1}{2})) = \exp(\tilde{\gamma}(t)+\pi i)$ for all $t \in [0, \frac{1}{2}]$. It follows that $\tilde{\gamma}(t+\frac{1}{2}) = \tilde{\gamma}(t) + (2m+1)\pi i$ for some integer m. (The value of m for which this identity is valid does not depend on t, since every continuous function from $[0, \frac{1}{2}]$ to the set of integers is necessarily constant.) Hence

$$n(f \circ \sigma, 0) = \frac{\tilde{\gamma}(1) - \tilde{\gamma}(0)}{2\pi i} = \frac{\tilde{\gamma}(1) - \tilde{\gamma}(\frac{1}{2})}{2\pi i} + \frac{\tilde{\gamma}(\frac{1}{2}) - \tilde{\gamma}(0)}{2\pi i} = 2m + 1.$$

Thus $n(f \circ \sigma, 0)$ is an odd integer, as required. ∎

We shall identify the space $\mathbb{R}^2$ with $\mathbb{C}$, identifying $(x, y) \in \mathbb{R}^2$ with the complex number $x + iy \in \mathbb{C}$ for all $x, y \in \mathbb{R}$. This is permissible, since we are interested in purely topological results concerning continuous functions defined on appropriate subsets of these spaces. Under this identification the closed unit disk D is given by

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

As usual, we define

$$S^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}.$$

Lemma 7.11 *Let $f: S^2 \to \mathbb{R}^2$ be a continuous map with the property that $f(-\mathbf{n}) = -f(\mathbf{n})$ for all $\mathbf{n} \in S^2$. Then there exists some point $\mathbf{n}_0$ of S^2 with the property that $f(\mathbf{n}_0) = 0$.*

Proof Let $\varphi: D \to S^2$ be the map defined by

$$\varphi(x, y) = (x, y, +\sqrt{1 - x^2 - y^2}).$$

(Thus the map φ maps the closed disk D homeomorphically onto the upper hemisphere in $\mathbb{R}^3$.) Let $\sigma: [0, 1] \to S^2$ be the parameterization of the equator in S^2 defined by

$$\sigma(t) = (\cos 2\pi t, \sin 2\pi t, 0)$$

for all $t \in [0, 1]$. Let $f: S^2 \to \mathbb{R}^2$ be a continuous map with the property that $f(-\mathbf{n}) = -f(\mathbf{n})$ for all $\mathbf{n} \in S^2$. If $f(\sigma(t_0)) = 0$ for some $t_0 \in [0, 1]$ then the function f has a zero at $\sigma(t_0)$. It remains to consider the case in which $f(\sigma(t)) \neq 0$ for all $t \in [0, 1]$. In that case the winding number $n(f \circ \sigma, 0)$ is an odd integer, by Lemma 7.10, and is thus non-zero. It follows from Proposition 7.7, applied to $f \circ \varphi: D \to \mathbb{R}^2$, that $0 \in f(\varphi(D))$, (since otherwise the winding number $n(f \circ \sigma, 0)$ would be zero). Thus $f(\mathbf{n}_0) = 0$ for some $\mathbf{n}_0 = \varphi(D)$, as required. ∎

Theorem 7.12 (Borsuk-Ulam) *Let $f\colon S^2 \to \mathbb{R}^2$ be a continuous map. Then there exists some point $\mathbf{n}$ of S^2 with the property that $f(-\mathbf{n}) = f(\mathbf{n})$.*

Proof This result follows immediately on applying Lemma 7.11 to the continuous function $g\colon S^2 \to \mathbb{R}^2$ defined by $g(\mathbf{n}) = f(\mathbf{n}) - f(-\mathbf{n})$. ∎

Remark It is possible to generalize the Borsuk-Ulam Theorem to n dimensions. Let S^n be the unit n-sphere centered on the origin in $\mathbb{R}^n$. The Borsuk-Ulam Theorem in n-dimensions states that if $f\colon S^n \to \mathbb{R}^n$ is a continuous map then there exists some point $\mathbf{x}$ of S^n with the property that $f(\mathbf{x}) = f(-\mathbf{x})$.

8 The Fundamental Group

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 6.1.)

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 8.1 *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 $F_1\colon [0,1] \times [0,1] \to X$ 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 $(\gamma_1.\gamma_2).\gamma_3 \simeq \gamma_1.(\gamma_2.\gamma_3)$ 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

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

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 $\varepsilon.\gamma \simeq \gamma \simeq \gamma.\varepsilon$ 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 $\gamma.\gamma^{-1}$ and ε, and moreover this homotopy sends $(0,\tau)$ and $(1,\tau)$ to x_0 for all $\tau \in [0,1]$. Therefore $\gamma.\gamma^{-1} \simeq \varepsilon \operatorname{rel}\{0,1\}$, 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.

8.1 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 8.2 *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 8.2, 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 8.3 *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 $n > 1$.

8.2 The Fundamental Group of the Circle

Theorem 8.4 *$\pi_1(S^1, b) \cong \mathbb{Z}$ for any $b \in S^1$.*

Proof We regard S^1 as the unit circle in $\mathbb{R}^2$. Without loss of generality, we can take $b = (1, 0)$. Now the map $p\colon \mathbb{R} \to S^1$ which sends $t \in \mathbb{R}$ to $(\cos 2\pi t, \sin 2\pi t)$ is a covering map, and $b = p(0)$. Moreover $p(t_1) = p(t_2)$ if and only if $t_1 - t_2$ is an integer; in particular $p(t) = b$ if and only if t is an integer.

Let α and β be loops in S^1 based at b, and let $\tilde\alpha$ and $\tilde\beta$ be paths in $\mathbb{R}$ that satisfy $p \circ \tilde\alpha = \alpha$ and $p \circ \tilde\beta = \beta$. Suppose that α and β represent the same element of $\pi_1(S^1, b)$. Then there exists a homotopy $F\colon [0,1] \times [0,1] \to S^1$ such that $F(t, 0) = \alpha(t)$ and $F(t, 1) = \beta(t)$ for all $t \in [0,1]$, and $F(0, \tau) =$

$F(1,\tau) = b$ for all $\tau \in [0,1]$. It follows from the Monodromy Theorem (Theorem 6.6) that this homotopy lifts to a continuous map $G\colon [0,1]\times[0,1] \to \mathbb{R}$ satisfying $p \circ G = F$. Moreover $G(0,\tau)$ and $G(1,\tau)$ are integers for all $\tau \in [0,1]$, since $p(G(0,\tau)) = b = p(G(1,\tau))$. Also $G(t,0)-\tilde{\alpha}(t)$ and $G(t,1)-\tilde{\beta}(t)$ are integers for all $t \in [0,1]$, since $p(G(t,0)) = \alpha(t) = p(\tilde{\alpha}(t))$ and $p(G(t,1)) = \beta(t) = p(\tilde{\beta}(t))$. Now any continuous integer-valued function on $[0,1]$ is constant, by the Intermediate Value Theorem. In particular the functions sending $\tau \in [0,1]$ to $G(0,\tau)$ and $G(1,\tau)$ are constant, as are the functions sending $t \in [0,1]$ to $G(t,0)-\tilde{\alpha}(t)$ and $G(t,1)-\tilde{\beta}(t)$. Thus

$$G(0,0) = G(0,1), \qquad G(1,0) = G(1,1),$$

$$G(1,0)-\tilde{\alpha}(1) = G(0,0)-\tilde{\alpha}(0), \qquad G(1,1)-\tilde{\beta}(1) = G(0,1)-\tilde{\beta}(0).$$

On combining these results, we see that

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

We conclude from this that there exists a well-defined function $\lambda\colon \pi_1(S^1,b) \to \mathbb{Z}$ characterized by the property that $\lambda([\alpha]) = \tilde{\alpha}(1)-\tilde{\alpha}(0)$ for all loops α based at b, where $\tilde{\alpha}\colon [0,1] \to \mathbb{R}$ is any path in $\mathbb{R}$ satisfying $p \circ \tilde{\alpha} = \alpha$.

Next we show that λ is a homomorphism. Let α and β be any loops based at b, and let $\tilde{\alpha}$ and $\tilde{\beta}$ be lifts of α and β. The element $[\alpha][\beta]$ of $\pi_1(S^1,b)$ is represented by the product path $\alpha.\beta$, where

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

Define a continuous path $\sigma\colon [0,1] \to \mathbb{R}$ by

$$\sigma(t) = \begin{cases} \tilde{\alpha}(2t) & \text{if } 0 \leq t \leq \frac{1}{2}; \\ \tilde{\beta}(2t-1)+\tilde{\alpha}(1)-\tilde{\beta}(0) & \text{if } \frac{1}{2} \leq t \leq 1. \end{cases}$$

(Note that $\sigma(t)$ is well-defined when $t = \frac{1}{2}$.) Then $p \circ \sigma = \alpha.\beta$ and thus

$$\begin{aligned} \lambda([\alpha][\beta]) &= \lambda([\alpha.\beta]) = \sigma(1)-\sigma(0) = \tilde{\alpha}(1)-\tilde{\alpha}(0)+\tilde{\beta}(1)-\tilde{\beta}(0) \\ &= \lambda([\alpha])+\lambda([\beta]). \end{aligned}$$

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

Now suppose that $\lambda([\alpha]) = \lambda([\beta])$. Let $F\colon [0,1]\times[0,1] \to S^1$ be the homotopy between α and β defined by

$$F(t,\tau) = p\left((1-\tau)\tilde{\alpha}(t)+\tau\tilde{\beta}(t)\right),$$

where $\tilde{\alpha}$ and $\tilde{\beta}$ are the lifts of α and β respectively starting at 0. Now $\tilde{\beta}(1) = \lambda([\beta]) = \lambda([\alpha]) = \tilde{\alpha}(1)$, and $\tilde{\beta}(0) = \tilde{\alpha}(0) = 0$. Therefore $F(0, \tau) = b = p(\tilde{\alpha}(1)) = F(1, \tau)$ for all $\tau \in [0, 1]$. Thus $\alpha \simeq \beta$ rel $\{0, 1\}$, and therefore $[\alpha] = [\beta]$. This shows that $\lambda: \pi_1(S^1, b) \to \mathbb{Z}$ is injective.

The homomorphism λ is surjective, since $n = \lambda([\gamma_n])$ for all $n \in \mathbb{Z}$, where the loop $\gamma_n: [0, 1] \to S^1$ is given by $\gamma_n(t) = p(nt) = (\cos 2\pi nt, \sin 2\pi nt)$ for all $t \in [0, 1]$. We conclude that $\lambda: \pi_1(S^1, b) \to \mathbb{Z}$ is an isomorphism. ∎

We now show that every continuous map from the closed disk D to itself has at least one fixed point. This is the two-dimensional version of the *Brouwer Fixed Point Theorem.*

Theorem 8.5 *Let $f: D \to D$ be a continuous map which maps the closed disk D into itself. Then $f(\mathbf{x}_0) = \mathbf{x}_0$ for some $\mathbf{x}_0 \in D$.*

Proof Let ∂D denote the boundary circle of D. The inclusion map $i: \partial D \hookrightarrow D$ induces a corresponding homomorphism $i_\#: \pi_1(\partial D, \mathbf{b}) \to \pi_1(D, \mathbf{b})$ of fundamental groups for any $\mathbf{b} \in \partial D$.

Suppose that it were the case that the map f has no fixed point in D. Then one could define a continuous map $r: D \to \partial D$ as follows: for each $\mathbf{x} \in D$, let $r(\mathbf{x})$ be the point on the boundary ∂D of D obtained by continuing the line segment joining $f(\mathbf{x})$ to $\mathbf{x}$ beyond $\mathbf{x}$ until it intersects ∂D at the point $r(\mathbf{x})$. Note that $r|\partial D$ is the identity map of ∂D.

Let $r_\#: \pi_1(D, \mathbf{b}) \to \pi_1(\partial D, \mathbf{b})$ be the homomorphism of fundamental groups induced by $r: D \to \partial D$. Now $(r \circ i)_\#: \pi_1(\partial D, \mathbf{b}) \to \pi_1(\partial D, \mathbf{b})$ is the identity isomorphism of $\pi_1(\partial D, \mathbf{b})$, since $r \circ i: \partial D \to \partial D$ is the identity map. But it follows directly from the definition of induced homomorphisms that $(r \circ i)_\# = r_\# \circ i_\#$. Therefore $i_\#: \pi_1(\partial D, \mathbf{b}) \to \pi_1(D, \mathbf{b})$ is injective, and $r_\#: \pi_1(D, \mathbf{b}) \to \pi_1(\partial D, \mathbf{b})$ is surjective. But this is impossible, since $\pi_1(\partial D, \mathbf{b}) \cong \mathbb{Z}$ (Theorem 8.4) and $\pi_1(D, \mathbf{b})$ is the trivial group. This contradiction shows that the continuous map $f: D \to D$ must have at least one fixed point. ∎