Module MA342R: Covering Spaces and Fundamental Groups
Hilary Term 2017
Part II (Sections 2 to 4)

D. R. Wilkins

Contents

2 Winding Numbers of Closed Paths in the Complex Plane

2.1 Paths in the Complex Plane

Let D be a subset of the complex plane $\mathbb{C}$. We define a *path* in D to be a continuous complex-valued function $\gamma\colon [a,b] \to D$ defined over some closed interval $[a,b]$. We shall denote the range $\gamma([a,b])$ of the function γ defining the path by $[\gamma]$. Now it follows from the Heine-Borel Theorem (Theorem 1.37) that the closed bounded interval $[a,b]$ is compact. Moreover continuous functions map compact sets to compact sets (see Lemma 1.39). It follows that $[\gamma]$ is a closed bounded subset of the complex plane.

Lemma 2.1 *Let $\gamma\colon [a,b] \to \mathbb{C}$ be a path in the complex plane, and let w be a complex number that does not lie on the path γ. Then there exists some positive real number ε_0 such that $|\gamma(t) - w| \geq \varepsilon_0 > 0$ for all $t \in [a,b]$.*

Proof The closed unit interval $[a,b]$ is a closed bounded subset of $\mathbb{R}$. Now any continuous real-valued function on a compact set is bounded above and below on that set (Lemma 1.40). Therefore there exists some positive real number M such that $|\gamma(t) - w|^{-1} \leq M$ for all $t \in [a,b]$. Let $\varepsilon_0 = M^{-1}$. Then the positive real number ε_0 has the required property. ∎

Definition A path $\gamma\colon [a,b] \to \mathbb{C}$ in the complex plane is said to be *closed* if $\gamma(a) = \gamma(b)$.

Remark The use of the technical term *closed* as in the above definition has no relation to the notions of open and closed sets.) Thus a *closed path* is a path that returns to its starting point.

Let $\gamma\colon [a,b] \to \mathbb{C}$ be a path in the complex plane. We say that a complex number w *lies on* the path γ if $w \in [\gamma]$, where $[\gamma] = \gamma([a,b])$.

2.2 The Exponential Map

The exponential map $\exp\colon \mathbb{C} \to \mathbb{C}$ is defined on the complex plane so that

$$\exp(x + iy) = e^x \cos y + i\, e^x \sin y$$

for all real numbers x and y, where $i^2 = -1$. Then

$$\exp(x + iy) = u(x,y) + i\, v(x,y)$$

where

$$u(x,y) = e^x \cos y, \quad v(x,y) = e^x \sin y$$

for all real numbers x and y. The functions $u\colon \mathbb{R}^2 \to \mathbb{R}$ and $v\colon \mathbb{R}^2 \to \mathbb{R}$ satisfy the partial differential equations

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x} \end{aligned}$$

These partial differential equations are the *Cauchy-Riemann equations* that are satisfied by the real and imaginary parts of a function of a complex variable if and only if that function is holomorphic.

Lemma 2.2 *The exponential map* $\exp\colon \mathbb{C} \to \mathbb{C}$ *satisfies the identities* $\exp(z+w) = \exp(z)\,\exp(w)$ *and* $\exp(-z) = \exp(z)^{-1}$ *for all complex numbers* z *and* w.

Proof Let $z = x + iy$ and $w = u + iv$, where x, y, u and v are real numbers and $i^2 = -1$. Then

$$\begin{aligned} \exp(z+w) &= e^{x+u}\,(\cos(y+v) + i\sin(y+v)) \\ &= e^x e^u (\cos y\,\cos v - \sin y\,\sin v + i\sin y\,\cos v + i\cos y\,\sin v) \\ &= e^x e^u (\cos y + i\sin y)(\cos v + i\sin v) \\ &= \exp(z)\,\exp(w). \end{aligned}$$

Applying this result with $w = -z$, we see that $\exp(z)\exp(-z) = \exp(0) = 1$, and therefore $\exp(-z) = \exp(z)^{-1}$, as required. ∎

Lemma 2.3 *Let* z *and* w *be complex numbers. Then* $\exp(z) = \exp(w)$ *if and only if* $w = z + 2\pi i n$ *for some integer* n.

Proof Suppose that $w = z + 2\pi i n$ for some integer n. Then

$$\begin{aligned} \exp(w) &= \exp(z)\exp(2\pi i n) = \exp(z)(\cos 2\pi n + i\sin 2\pi n) \\ &= \exp(z). \end{aligned}$$

Conversely suppose that $\exp(w) = \exp(z)$. Let $w - z = u + iv$, where u and v are real numbers. Then

$$e^u(\cos v + i\sin v) = \exp(w-z) = \exp(w)\exp(z)^{-1} = 1.$$

Taking the modulus of both sides, we see that $e^u = 1$, and thus $u = 0$. Also $\cos v = 1$ and $\sin v = 0$, and therefore $v = 2\pi n$ for some integer n. The result follows. ∎

Remark The infinite series $\displaystyle\sum_{n=0}^{+\infty} \frac{z^n}{n!}$ converges absolutely for all complex numbers z. Standard theorems concerning power series then ensure that the infinite series converges uniformly in z over any closed disk of positive radius about zero in the complex plane. A standard theorem of analysis regarding Cauchy products of absolutely convergent infinite series then ensures that

$$\left(\sum_{n=0}^{+\infty} \frac{z^n}{n!}\right)\left(\sum_{n=0}^{+\infty} \frac{w^n}{n!}\right) = \left(\sum_{n=0}^{+\infty} \frac{(z+w)^n}{n!}\right)$$

for all complex numbers z. It follows that if $z = x + iy$, where x and y are real numbers and $i^2 = -1$, then

$$\begin{aligned} \sum_{n=0}^{+\infty} \frac{z^n}{n!} &= \left(\sum_{n=0}^{+\infty} \frac{x^n}{n!}\right)\left(\sum_{k=0}^{+\infty} \frac{(-1)^k y^{2k}}{(2k)!} + i \sum_{k=0}^{+\infty} \frac{(-1)^k y^{2k+1}}{(2k+1)!}\right) \\ &= e^x(\cos y + i \sin y) \end{aligned}$$

for all real numbers x and y. Thus

$$\exp z = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$$

for all complex numbers z.

Lemma 2.4 *Let w be a non-zero complex number, and let*

$$D_{w,|w|} = \{z \in \mathbb{C} : |z - w| < |w|\}.$$

Then there exists a continuous function $F_w \colon D_{w,|w|} \to \mathbb{C}$ with the property that $\exp(F_w(z)) = z$ *for all $z \in D_{w,|w|}$.*

Proof Let $U = \mathbb{C} \setminus \{x \in \mathbb{R} : x \leq 0\}$, and let $\log \colon U \to \mathbb{C}$ be the "principal branch" of the logarithm function, defined so that $\log(re^{i\theta}) = \log r + i\theta$ for all real numbers r and θ satisfying $r > 0$ and $\pi < \theta < \pi$. Then the function $\log \colon U \to \mathbb{C}$ is continuous, and $\exp(\log z) = z$ for all $z \in U$. Let ζ be a complex number satisfying $\exp \zeta = w$. Then $z/w \in U$ for all $z \in D_{w,|w|}$. Let $F_w \colon D_{w,|w|} \to \mathbb{C}$ be defined so that $F_w(z) = \zeta + \log(z/w)$ for all $z \in D_{w,|w|}$. Then

$$\exp(F_w(z)) = \exp(\zeta)\exp(\log(z/w)) = w(z/w) = z$$

for all $z \in D(w, |w|)$, as required. ∎

2.3 Path-Lifting with respect to the Exponential Map

Theorem 2.5 *Let $\gamma\colon [a,b] \to \mathbb{C} \setminus \{0\}$ be a path in the set $\mathbb{C} \setminus \{0\}$ of non-zero complex numbers. Then there exists a path $\tilde{\gamma}\colon [a,b] \to \mathbb{C}$ in the complex plane which satisfies $\exp(\tilde{\gamma}(t)) = \gamma(t)$ for all $t \in [a,b]$.*

Proof The complex number $\gamma(t)$ is non-zero for all $t \in [a,b]$, and therefore there exists some positive number ε_0 such that $|\gamma(t)| \geq \varepsilon_0$ for all $t \in [a,b]$. (Lemma 2.1). Now any continuous complex-valued function on a closed bounded interval is uniformly continuous. (This follows, for example, from Theorem 1.48.) Therefore there exists some positive real number δ such that $|\gamma(t) - \gamma(s)| < \varepsilon_0$ for all $s, t \in [a,b]$ satisfying $|t - s| < \delta$.

Let m be a positive integer satisfying $m > |b - a|/\delta$, and let $t_j = a + j(b-a)/m$ for $j = 0, 1, 2, \ldots, m$. Then $|t_j - t_{j-1}| < \delta$ for $j = 1, 2, \ldots, m$. It follows from this that

$$|\gamma(t) - \gamma(t_j)| < \varepsilon_0 \leq |\gamma(t_j)|$$

for all $t \in [t_{j-1}, t_j]$, and thus

$$\gamma([t_{j-1}, t_j]) \subset D_{\gamma(t_j), |\gamma(t_j)|}$$

for $j = 1, 2, \ldots, n$, where

$$D_{w,|w|} = \{z \in \mathbb{C} : |z - w| < |w|\}$$

for all $w \in \mathbb{C}$.

Now there exist continuous functions $F_j\colon D_{\gamma(t_j),|\gamma(t_j)|} \to \mathbb{C}$ with the property that $\exp(F_j(z)) = z$ for all $z \in D_{\gamma(t_j),|\gamma(t_j)|}$ (see Lemma 2.4). Let $\tilde{\gamma}_j(t) = F_j(\gamma(t))$ for all $t \in [t_{j-1}, t_j]$. Then, for each integer j between 1 and m, the function $\tilde{\gamma}_j\colon [t_{j-1}, t_j] \to \mathbb{C}$ is continuous, and is thus a path in the complex plane with the property that $\exp(\tilde{\gamma}_j(t)) = \gamma(t)$ for all $t \in [t_{j-1}, t_j]$. Now

$$\exp(\tilde{\gamma}_j(t_j)) = \gamma(t_j) = \exp(\tilde{\gamma}_{j+1}(t_j))$$

for each integer j between 1 and $m - 1$. The periodicity properties of the exponential function (Lemma 2.3) therefore ensure that there exist integers $k_1, k_2, \ldots, k_{m-1}$ such that $\tilde{\gamma}_{j+1}(t_j) = \tilde{\gamma}_j(t_j) + 2\pi i k_j$ for $j = 1, 2, \ldots, m - 1$. Then

$$\tilde{\gamma}_{j+1}(t_j) - 2\pi i \sum_{r=1}^{j} k_r = \tilde{\gamma}_j(t_j) - 2\pi i \sum_{r=1}^{j-1} k_r$$

for $j = 1, 2, \ldots, m-1$. The Pasting Lemma (Lemma 1.24) then ensures the existence of a continuous function $\tilde{\gamma}: [a,b] \to \mathbb{C}$ defined so that $\tilde{\gamma}(t) = \tilde{\gamma}_1(t)$ whenever $t \in [a, t_1]$, and

$$\tilde{\gamma}(t) = \tilde{\gamma}_j(t) - 2\pi i \sum_{r=1}^{j-1} k_r$$

whenever $t \in [t_{j-1}, t_j]$ for some integer j between 2 and m. Moreover $\exp(\tilde{\gamma}(t)) = \gamma(t)$ for all $t \in [a,b]$. We have thus proved the existence of a path $\tilde{\gamma}$ in the complex plane with the required properties. ∎

2.4 Winding Numbers

Let $\gamma: [a,b] \to \mathbb{C}$ be a closed path in the complex plane, and let w be a complex number that does not lie on γ. Then there exists a path $\tilde{\gamma}_w: [a,b] \to \mathbb{C}$ in the complex plane such that $\exp(\tilde{\gamma}_w(t)) = \gamma(t) - w$ for all $t \in [a,b]$ (Theorem 2.5). Now the definition of closed paths ensures that $\gamma(b) = \gamma(a)$. Also two complex numbers z_1 and z_2 satisfy $\exp z_1 = \exp z_2$ if and only if $(2\pi i)^{-1}(z_2 - z_1)$ is an integer (Lemma 2.3). It follows that there exists some integer $n(\gamma, w)$ such that $\tilde{\gamma}_w(b) = \tilde{\gamma}_w(a) + 2\pi i n(\gamma, w)$.

Now let $\varphi: [a,b] \to \mathbb{C}$ be any path with the property that $\exp(\varphi(t)) = \gamma(t) - w$ for all $t \in [a,b]$. Then the function sending $t \in [a,b]$ to $(2\pi i)^{-1}(\varphi(t) - \tilde{\gamma}_w(t))$ is a continuous integer-valued function on the interval $[a,b]$, and is therefore constant on this interval (Corollary 1.58). It follows that

$$\varphi(b) - \varphi(a) = \tilde{\gamma}_w(b) - \tilde{\gamma}_w(a) = 2\pi i n(\gamma, w).$$

It follows from this that the value of the integer $n(\gamma, w)$ depends only on the choice of γ and w, and is independent of the choice of path $\tilde{\gamma}_w$ satisfying $\exp(\tilde{\gamma}_w(t)) = \gamma(t) - w$ for all $t \in [a,b]$.

Definition Let $\gamma: [a,b] \to \mathbb{C}$ be a closed path in the complex plane, and let w be a complex number that does not lie on γ. The *winding number* of γ about w is defined to be the unique integer $n(\gamma, w)$ with the property that $\varphi(b) - \varphi(a) = 2\pi i n(\gamma, w)$ for all paths $\varphi: [a,b] \to \mathbb{C}$ in the complex plane that satisfy $\exp(\varphi(t)) = \gamma(t) - w$ for all $t \in [a,b]$.

Example Let n be an integer, and let $\gamma_n: [0,1] \to \mathbb{C}$ be the closed path in the complex plane defined by $\gamma_n(t) = \exp(2\pi i n t)$. Then $\gamma_n(t) = \exp(\varphi_n(t))$ for all $t \in [0,1]$ where $\varphi_n: [0,1] \to \mathbb{C}$ is the path in the complex plane defined such that $\varphi_n(t) = 2\pi i n t$ for all $t \in [0,1]$. It follows that $n(\gamma_n, 0) = (2\pi i)^{-1}(\varphi_n(1) - \varphi_n(0)) = n$.

Given a closed path γ, and given a complex number w that does not lie on γ, the winding number $n(\gamma, w)$ measures the number of times that the path γ winds around the point w of the complex plane in the anticlockwise direction.

Lemma 2.6 (Dog-Walking Lemma) *Let $\gamma_1\colon [a,b] \to \mathbb{C}$ and $\gamma_2\colon [a,b] \to \mathbb{C}$ be closed paths in the complex plane, and let w be a complex number that does not lie on γ_1. Suppose that $|\gamma_2(t) - \gamma_1(t)| < |\gamma_1(t) - w|$ for all $t \in [a,b]$. Then $n(\gamma_2, w) = n(\gamma_1, w)$.*

Proof Note that the inequality satisfied by the functions γ_1 and γ_2 ensures that w does not lie on the path γ_2. Let $\tilde{\gamma}_1\colon [0,1] \to \mathbb{C}$ be a path in the complex plane such that $\exp(\tilde{\gamma}_1(t)) = \gamma_1(t) - w$ for all $t \in [a,b]$, and let

$$\rho(t) = \frac{\gamma_2(t) - w}{\gamma_1(t) - w}$$

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

$$|\rho(t) - 1| = \left| \frac{\gamma_2(t) - \gamma_1(t)}{\gamma_1(t) - w} \right| < 1$$

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

Now it follows from Lemma 2.4 that there exists a continuous function $F\colon \{z \in \mathbb{C} : |z - 1| < 1\} \to \mathbb{C}$ with the property that $\exp(F(z)) = z$ for all complex numbers z satisfying $|z - 1| < 1$. Let $\tilde{\gamma}_2\colon [0,1] \to \mathbb{C}$ be the path in the complex plane defined such that $\tilde{\gamma}_2(t) = F(\rho(t)) + \tilde{\gamma}_1(t)$ for all $t \in [a,b]$. Then

$$\begin{aligned} \exp(\tilde{\gamma}_2(t)) &= \exp(F(\rho(t)))\exp(\tilde{\gamma}_1(t)) = \rho(t)(\gamma_1(t) - w) \\ &= \gamma_2(t) - w. \end{aligned}$$

Now $\rho(b) = \rho(a)$. It follows that

$$\begin{aligned} 2\pi i n(\gamma_2, w) &= \tilde{\gamma}_2(b) - \tilde{\gamma}_2(a) \\ &= F(\rho(b)) + \tilde{\gamma}_1(b) - F(\rho(a)) - \tilde{\gamma}_1(a) \\ &= \tilde{\gamma}_1(b) - \tilde{\gamma}_1(a) \\ &= 2\pi i n(\gamma_1, w), \end{aligned}$$

as required. ∎

Remark Imagine that you are exercising a dog in a park. You walk along a path close to the perimeter of the park that remains at all times at at least 200 metres from an oak tree in the centre of the park. Your dog runs around in your vicinity, but remains at all times within 100 metres of you. In order to leave the park you and your dog return to the point at which you entered the park. The Dog-Walking Lemma then ensures that the number of times that your dog went around the oak tree in the centre of the park is equal to the number of times that you yourself went around that tree.

Example Let $\gamma\colon [0,1] \to \mathbb{C}$ be the closed curve in the complex plane defined such that

$$\gamma(t) = 3\cos 6\pi t + 4i \sin 6\pi t + (\sin 16\pi t)(\sin 8\pi t) - 2ie^{\cos 8\pi t - 1} \cos 8\pi t$$

for all $t \in [0,1]$, where $i^2 = -1$. Let

$$\gamma_1(t) = 3\cos 6\pi t + 4i \sin 6\pi t$$

for all $t \in [0,1]$. Then $|\gamma_1(t)| \geq 3$ for all $t \in [0,1]$. Also $|\sin 16\pi t| \leq 1$ and $0 \leq e^{\cos 8\pi t - 1} \leq 1$ for all $t \in [0,1]$, and therefore

$$\left|(\sin 16\pi t)(\sin 8\pi t) - 2ie^{\cos 8\pi t - 1} \cos 8\pi t\right|^2 \leq \sin^2 8\pi t + 4\cos^2 8\pi t \leq 4$$

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

$$\begin{aligned} |\gamma(t) - \gamma_1(t)| &= \left|(\sin 16\pi t)(\sin 8\pi t) - 2ie^{\cos 8\pi t - 1} \cos 8\pi t\right| \\ &\leq 2 < |\gamma_1(t)| \end{aligned}$$

for all $t \in [0,1]$. The Dog-Walking Lemma (Lemma 2.6) then ensures that $n(\gamma, 0) = n(\gamma_1, 0)$. Another application of the Dog-Walking Lemma then ensures that $n(\gamma_1, 0) = n(\gamma_2, 0)$, where

$$\gamma_2(t) = 3(\cos 6\pi t + i \sin 6\pi t)$$

for all $t \in [0,1]$. Moreover $\gamma_2 = \exp \circ \tilde{\gamma}_2$ where $\tilde{\gamma}_2\colon [0,1] \to \mathbb{C}$ is the path in $\mathbb{C}$ defined so that $\tilde{\gamma}_2(t) = \log 3 + 6\pi t$ for all $t \in [0,1]$. The definition of winding number ensures that

$$n(\gamma_2, 0) = (2\pi i)^{-1}(\tilde{\gamma}_2(1) - \tilde{\gamma}_2(0)) = 3.$$

Therefore $n(\gamma, 0) = 3$.

Lemma 2.7 *Let $\gamma\colon [a,b] \to \mathbb{C}$ be a closed path in the complex plane and let W be the set $\mathbb{C} \setminus [\gamma]$ of all points of the complex plane that do not lie on the curve γ. Then the function that sends $w \in W$ to the winding number $n(\gamma, w)$ of γ about w is a continuous function on W.*

Proof Let $w \in W$. It then follows from Lemma 2.1 that there exists some positive real number ε_0 such that $|\gamma(t) - w| \geq \varepsilon_0 > 0$ for all $t \in [a, b]$. Let w_1 be a complex number satisfying $|w_1 - w| < \varepsilon_0$, and let $\gamma_1 \colon [a, b] \to \mathbb{C}$ be the closed path in the complex plane defined such that $\gamma_1(t) = \gamma(t) + w - w_1$ for all $t \in [a, b]$. Then $\gamma(t) - w_1 = \gamma_1(t) - w$ for all $t \in [a, b]$, and therefore $n(\gamma, w_1) = n(\gamma_1, w)$. Also $|\gamma_1(t) - \gamma(t)| < |\gamma(t) - w|$ for all $t \in [a, b]$. It follows from the Dog-Walking Lemma (Lemma 2.6) that $n(\gamma, w_1) = n(\gamma_1, w) = n(\gamma, w)$. This shows that the function sending $w \in W$ to $n(\gamma, w)$ is continuous on W, as required. ∎

Lemma 2.8 *Let $\gamma \colon [a, b] \to \mathbb{C}$ be a closed path in the complex plane, and let R be a positive real number with the property that $|\gamma(t)| < R$ for all $t \in [a, b]$. Then $n(\gamma, w) = 0$ for all complex numbers w satisfying $|w| \geq R$.*

Proof Let $\gamma_0 \colon [a, b] \to \mathbb{C}$ be the constant path defined by $\gamma_0(t) = 0$ for all $[a, b]$. If $|w| > R$ then $|\gamma(t) - \gamma_0(t)| = |\gamma(t)| < |w| = |\gamma_0(t) - w|$. It follows from the Dog-Walking Lemma (Lemma 2.6) that $n(\gamma, w) = n(\gamma_0, w) = 0$, as required. ∎

Proposition 2.9 *Let $[a, b]$ and $[c, d]$ be closed bounded intervals, and, for each $s \in [c, d]$, let $\gamma_s \colon [a, b] \to \mathbb{C}$ be a closed path in the complex plane. Let w be a complex number that does not lie on any of the paths γ_s. Suppose that the function $H \colon [a, b] \times [c, d] \to \mathbb{C}$ is continuous, where $H(t, s) = \gamma_s(t)$ for all $t \in [a, b]$ and $s \in [c, d]$. Then $n(\gamma_c, w) = n(\gamma_d, w)$.*

Proof The rectangle $[a, b] \times [c, d]$ is a closed bounded subset of $\mathbb{R}^2$, and is therefore compact. It follows that the continuous function on the closed rectangle $[a, b] \times [c, d]$ that sends a point (t, s) of the rectangle to $|H(t, s) - w|^{-1}$ is a bounded function on $[a, b] \times [c, d]$ (see, for example, Lemma 1.40). Therefore there exists some positive number ε_0 such that $|H(t, s) - w| \geq \varepsilon_0 > 0$ for all $t \in [a, b]$ and $s \in [c, d]$.

Now any continuous complex-valued function on a closed bounded subset of a Euclidean space is uniformly continuous. (This follows, for example, on combining the results of Theorem 1.50 and Theorem 1.48.) Therefore there exists some positive real number δ such that $|H(t, s) - H(t, u)| < \varepsilon_0$ for all $t \in [a, b]$ and for all $s, u \in [c, d]$ satisfying $|s - u| < \delta$. Let $s_0, s_1, \ldots, s_m$ be real numbers chosen such that $c = s_0 < s_1 < \ldots < s_m = d$ and $|s_j - s_{j-1}| < \delta$ for $j = 1, 2, \ldots, m$. Then

$$
\begin{aligned}
|\gamma_{s_j}(t) - \gamma_{s_{j-1}}(t)| &= |H(t, s_j) - H(t, s_{j-1})| \\
&< \varepsilon_0 \leq |H(t, s_{j-1}) - w| = |\gamma_{s_{j-1}}(t) - w|
\end{aligned}
$$

for all $t \in [a,b]$, and for each integer j between 1 and m. It therefore follows from the Dog-Walking Lemma (Lemma 2.6) that $n(\gamma_{s_{j-1}}, w) = n(\gamma_{s_j}, w)$ for each integer j between 1 and m. But then $n(\gamma_c, w) = n(\gamma_d, w)$, as required. ∎

Definition Let D be a subset of the complex plane, and let $\gamma\colon [a,b] \to D$ be a closed path in D. The closed path γ is said to be *contractible* in D if and only if there exists a continuous function $H\colon [a,b] \times [0,1] \to D$ and an element z_0 of D such that $H(t,0) = \gamma(t)$ and $H(t,1) = z_0$ for all $t \in [a,b]$, and $H(a,s) = H(b,s)$ for all $s \in [0,1]$.

Corollary 2.10 *Let D be a subset of the complex plane, and let $\gamma\colon [a,b] \to D$ be a closed path in D. Suppose that γ is contractible in D. Then $n(\gamma, w) = 0$ for all $w \in \mathbb{C} \setminus D$, where $n(\gamma, w)$ denotes the winding number of γ about w.*

Proof The closed curve γ is contractible, and therefore there exists an element z_0 of D and a continuous function $H\colon [a,b] \times [0,1] \to D$ such that $H(t,0) = \gamma(t)$ and $H(t,1) = z_0$ for all $t \in [a,b]$, and $H(a,s) = H(b,s)$ for all $s \in [0,1]$. For each $s \in [0,1]$ let $\gamma_s\colon [a,b] \to D$ be the closed path in D defined such that $\gamma_s(t) = H(t,s)$ for all $t \in [a,b]$. Then γ_1 is a constant path, and therefore $n(\gamma_1, w) = 0$ for all points w that do not lie on γ_1. Let w be an element of $w \in \mathbb{C} \setminus D$. Then w does not lie on any of the paths γ_s. It follows from Proposition 2.9 that

$$n(\gamma, w) = n(\gamma_0, w) = n(\gamma_1, w) = 0,$$

as required. ∎

2.5 Simply-Connected Subsets of the Complex Plane

Definition A subset D of the complex plane is said to be *path-connected* if, given any elements z_1 and z_2, there exists a path in D from z_1 and z_2.

Definition A path-connected subset D of the complex plane is said to be *simply-connected* if every closed loop in D is contractible.

Definition An subset D of the complex plane is said to be a *star-shaped* if there exists some complex number z_0 in D with the property that

$$\{(1-t)z + tz_0 : t \in [0,1]\} \subset D$$

for all $z \in D$. (Thus an open set in the complex plane is a star-shaped if and only if the line segment joining any point of D to z_0 is contained in D.)

Lemma 2.11 *Star-shaped subsets of the complex plane are simply-connected.*

Proof Let D be a star-shaped subset of the complex plane. Then there exists some element z_0 of D such that the line segment joining z_0 to z is contained in D for all $z \in D$. The star-shaped set D is obviously path-connected. Let $\gamma\colon [a,b] \to D$ be a closed path in D, and let $H(t,s) = (1-s)\gamma(t) + sz_0$ for all $t \in [a,b]$ and $s \in [0,1]$. Then $H(t,s) \in D$ for all $t \in [a,b]$ and $s \in [0,1]$, $H(t,0) = \gamma(t)$ and $H(t,1) = z_0$ for all $t \in [a,b]$. Also $\gamma(a) = \gamma(b)$, and therefore $H(a,s) = H(b,s)$ for all $s \in [0,1]$. It follows that the closed path γ is contractible. Thus D is simply-connected. ∎

The following result is an immediate consequence of Corollary 2.10

Proposition 2.12 *Let D be a simply-connected subset of the complex plane, and let γ be a closed path in D. Then $n(\gamma, w) = 0$ for all $w \in \mathbb{C} \setminus D$.*

2.6 The Fundamental Theorem of Algebra

Theorem 2.13 (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 We shall prove that any polynomial that is everywhere non-zero must be a constant polynomial.

Let $P(z) = a_0 + a_1 z + \cdots + a_m z^m$, where $a_1, a_2, \ldots, a_m$ are complex numbers and $a_m \neq 0$. 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 $|z| \geq 1$, and therefore

$$\begin{aligned} \left| \frac{Q(z)}{P_m(z)} \right| &= \frac{1}{|a_m z|} \left| \frac{a_0}{z^{m-1}} + \frac{a_1}{z^{m-2}} + \cdots + a_{m-1} \right| \\ &\leq \frac{1}{|a_m|\,|z|} \left(\left| \frac{a_0}{z^{m-1}} \right| + \left| \frac{a_1}{z^{m-2}} \right| + \cdots + |a_{m-1}| \right) \\ &\leq \frac{1}{|a_m|\,|z|} (|a_0| + |a_1| + \cdots + |a_{m-1}|) \leq \frac{R}{|z|} < 1. \end{aligned}$$

It follows that $|P(z) - P_m(z)| < |P_m(z)|$ for all complex numbers z satisfying $|z| > R$.

For each non-zero real number r, let $\gamma_r: [0, 1] \to \mathbb{C}$ and $\varphi_r: [0, 1] \to \mathbb{C}$ be the closed paths defined such that $\gamma_r(t) = P(r \exp(2\pi i t))$ and $\varphi_r(t) = P_m(r \exp(2\pi i t)) = a_m r^m \exp(2\pi i m t)$ for all $t \in [0, 1]$. If $r > R$ then $|\gamma_r(t) - \varphi_r(t)| < |\varphi_r(t)|$ for all $t \in [0, 1]$. It then follows from the Dog-Walking Lemma (Lemma 2.6) that $n(\gamma_r, 0) = n(\varphi_r, 0) = m$ whenever $r > R$.

Now if the polynomial P is everywhere non-zero then it follows on applying Proposition 2.9 that the function sending each non-negative real number r to the winding number $n(\gamma_r, 0)$ of the closed path γ_r about zero is a continuous function on the set of non-negative real numbers. But any continuous integer-valued function on an interval is necessarily constant (see Corollary 1.58). It follows that $n(\gamma_r, 0) = n(\gamma_0, 0)$ for all positive real-numbers r. But γ_0 is the constant path defined by $\gamma_0(t) = P(0)$ for all $t \in [0, 1]$, and therefore $n(\gamma_0, 0) = 0$. It follows that is the polynomial P is everywhere non-zero then $n(\gamma_r, 0) = 0$ for all non-negative real numbers r. But we have shown that $n(\gamma_r, 0) = m$ for sufficiently large values of r, where m is the degree of the polynomial P. It follows that if the polynomial P is everywhere non-zero, then it must be a constant polynomial. The result follows. ∎

2.7 The Kronecker Principle

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 2.14 *Let $f: D \to \mathbb{C}$ 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 $\sigma: [0, 1] \to \mathbb{C}$ 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 $\gamma_s(t) = f(s \exp(2\pi i t))$ for all $t \in [0, 1]$ and $s \in [0, 1]$. Then none of the closed curves γ_s passes through w, and γ_0 is the constant curve with value $f(0)$. It follows from Proposition 2.9 that

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

as required. ∎

2.8 The Brouwer Fixed Point Theorem

We now use Proposition 2.14 to show that there is no continuous 'retraction' mapping the closed unit disk onto its boundary circle.

Corollary 2.15 *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 2.14. 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 2.16 (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 2.15 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.

2.9 The Borsuk-Ulam Theorem

Lemma 2.17 *Let $f: S^1 \to \mathbb{C} \setminus \{0\}$ 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 $\sigma: [0,1] \to S^1$ is given by $\sigma(t) = \exp(2\pi i t)$.*

Proof It follows from the Path Lifting Theorem (Theorem 2.5) that there exists a continuous path $\tilde{\gamma}: [0,1] \to \mathbb{C}$ 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 2.18 *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 2.17, and is thus non-zero. It follows from Proposition 2.14, 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 2.19 (Borsuk-Ulam) *Let $f: S^2 \to \mathbb{R}^2$ 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 2.18 to the continuous function $g: S^2 \to \mathbb{R}^2$ 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 $f: S^n \to \mathbb{R}^n$ is a continuous map then there exists some point $\mathbf{x}$ of S^n with the property that $f(\mathbf{x}) = f(-\mathbf{x})$.

3 The Fundamental Group of a Topological Space

3.1 Homotopies between Continuous Maps

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.

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 $a \in A$.

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 $f: X \to Y$, let $H_0: X \times [0,1] \to Y$ 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 $f \simeq f$ 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 $f \simeq g$ rel A. Then there exists a homotopy $H: X \times [0,1] \to Y$ 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 $K: X \times [0,1] \to Y$ 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 $g \simeq f$ 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 $f \simeq g$ rel A and $g \simeq h$ 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

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

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 $f \simeq h$ 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. ∎

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 $\gamma_0 \simeq \gamma_1$ 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 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 $F: [0,1] \times [0,1] \to X$ 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 $F_1: [0,1] \times [0,1] \to X$ 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

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

Thus the map $G: [0,1] \times [0,1] \to X$ 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

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

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

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 $\varepsilon \,.\, \gamma \simeq \gamma \simeq \gamma \,.\, \varepsilon$ 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

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

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 $\varepsilon \simeq \gamma \,.\, \gamma^{-1} \operatorname{rel}\{0,1\}$, 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 $f\colon X \to Y$ 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 $f\colon \partial D \to X$ 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 $f\colon \partial D \to \mathbb{R}^n$ defined over the boundary ∂D of the closed unit disk D can be extended to a continuous map $F\colon D \to \mathbb{R}^n$ 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 $\gamma: [0,1] \to X$ 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 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 $\gamma: [a,b] \to S^1$ be a continuous map into S^1 defined on a closed bounded interval $[a,b]$. Then there exists a continuous real-valued function $\tilde{\gamma}: [a,b] \to \mathbb{R}$ 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 $\eta: [a,b] \to \mathbb{C}$ 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 $\tilde{\eta}: [a,b] \to \mathbb{C}$ 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 $\tilde{\gamma}: [a,b] \to \mathbb{R}$ 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 $p: \mathbb{R} \to S^1$ 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 $\gamma: [0,1] \to S^1$ in the circle S^1, there exists a continuous real-valued function $\tilde{\gamma}: [0,1] \to \mathbb{R}$ 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 $\eta: [0,1] \to \mathbb{R}$ 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 $\alpha \simeq \beta$* 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 $\alpha \simeq \beta$ rel $\{0, 1\}$ if and only if there exists a homotopy $H: [0, 1] \times [0, 1] \to S^1$ 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 $p: \mathbb{R} \to S^1$ 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 $\gamma: [0, 1] \to S^1$ 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 $\tilde{\gamma}: [0, 1] \to \mathbb{R}$ 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 $\alpha \simeq \beta$ 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 $\lambda: \pi_1(S^1, \mathbf{b}) \to \mathbb{Z}$ 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 $\lambda: \pi_1(S^1, \mathbf{b}) \to \mathbb{Z}$ is a homomorphism. Let $\alpha: [0, 1] \to S^1$ and $\beta: [0, 1] \to S^1$ be loops in S^1 based at $\mathbf{b}$. Then there exists a continuous real-valued function $\eta: [0, 1] \to \mathbb{R}$ 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 $\alpha \simeq \beta$ 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. ∎

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 $F: X \times [0, 1] \to X$ 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 $\gamma: [0, 1] \to X$ be a loop in X based at $(1, 0)$ and let $H: [0, 1] \times [0, 1] \to X$ 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 $i \circ r \circ \gamma \simeq \gamma$ rel $\{0, 1\}$.

Now the continuous maps $i: S^1 \to X$ and $r: X \to S^1$ induce well-defined homomorphisms $i_\#: \pi_1(S^1, (1, 0)) \to \pi_1(X, (1, 0))$ and $r_\#: \pi_1(X, (1, 0)) \to \pi_1(S^1, (1, 0))$, 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 $i_\#: \pi_1(S^1, (1, 0)) \to \pi_1(X, (1, 0))$ is an isomorphism whose inverse is the homomorphism $r_\#: \pi_1(X, (1, 0)) \to \pi_1(S^1, (1, 0))$, and therefore

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

as required. ∎

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 $i: \partial D \to D$ be the inclusion map, and let $\mathbf{b} = (1, 0)$. Suppose there were to exist a continuous map $r: D \to \partial D$ with the property that $r(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \partial D$. Then $r \circ i: \partial D \to \partial D$ 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.

4 Covering Maps

4.1 Evenly-Covered Open Sets and Covering Maps

Definition Let X and $\tilde{X}$ be topological spaces and let $p\colon \tilde{X} \to X$ 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 $p\colon \tilde{X} \to X$ is said to be a *covering map* if $p\colon \tilde{X} \to X$ is surjective and in addition every point of X is contained in some open set that is evenly covered by the map p.

If $p\colon \tilde{X} \to X$ 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 $p\colon \mathbb{R} \to S^1$ 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 $p\colon \mathbb{R} \to S^1$ is a covering map.

Example Let $p_{\exp}\colon \mathbb{C} \to \mathbb{C}\setminus\{0\}$ be the map from the complex plane $\mathbb{C}$ to the open subset $\mathbb{C} \setminus \{0\}$ of $\mathbb{C}$ defined such that $p_{\exp}(z) = \exp(z)$ for all complex numbers z. We show that $p_{\exp}(z)$ is a covering map.

Given any real number s, let

$$L_s = \{-re^{is} : r \in \mathbb{R} \text{ and } r \geq 0\}.$$

Then L_s is a ray in the complex plane starting at zero and passing through $-\cos s - i \sin s$. Moreover every complex number belonging to the complement $\mathbb{C} \setminus L_s$ of the ray L_s in $\mathbb{C}$ can be expressed uniquely in the form re^{it}, where r and t are real numbers satisfying $r > 0$ and $s - \pi < t < s + \pi$.

Let

$$W_s = \{w \in \mathbb{C} : s - \pi < \mathrm{Im}[w] < s + \pi\},$$

where $\mathrm{Im}[w]$ denotes the imaginary part of w for all complex numbers w, and let $F_s\colon \mathbb{C} \setminus L_s \to W_s$ be the complex-valued function on the open subset $\mathbb{C} \setminus L_s$ of the complex plane defined such that

$$F_s(re^{it}) = \log r + it$$

for all real numbers r and t satisfying $r > 0$ and $s - \pi < t < s + \pi$. Then $F_s\colon \mathbb{C} \setminus L_s \to W_s$ is a continuous map, $\exp(F_s(z)) = z$ for all $z \in \mathbb{C} \setminus L_s$ and $F_s(\exp(w)) = w$ for all $w \in W_s$. It follows that $F_s\colon \mathbb{C} \setminus L_s \to W_s$ is a homeomorphism between $\mathbb{C} \setminus L_s$ and W_s.

Let w be a complex number for which $\exp(w) \in \mathbb{C} \setminus L_s$. Then there exists a unique integer m such that $s + 2\pi m - \pi < \mathrm{Im}[w] < s + 2\pi m + \pi$. Then $w \in F_{s+m}(\exp w)$. It follows from this that, for each real number s, the preimage $p_{\exp}^{-1}(\mathbb{C}\setminus L_s)$ is the disjoint union of the sets $W_{s+2\pi m}$ as m ranges over the set $\mathbb{Z}$ of integers. Also $W_{s+2\pi m} \cap W_{s+2\pi n} = \emptyset$ when m and n are integers and $m \neq n$, and $p_{\exp}\colon \mathbb{C}\setminus\mathbb{C}\setminus\{0\}$ maps the open set $W_{s+2\pi m}$ homeomorphically onto $\mathbb{C} \setminus L_s$ for all integers m, where $p_{\exp}(w) = \exp(w)$ for all $w \in \mathbb{C}$. Thus $p_{\exp}\colon \mathbb{C} \to \mathbb{C} \setminus \{0\}$ is a covering map.

Example Let

$$\begin{aligned} X &= \{(x,y) \in \mathbb{R}^2 : (x,y) \neq (0,0)\}, \\ \tilde{X} &= \{(x,y,z) \in \mathbb{R}^3 : (x,y) \neq (0,0),\ \ x = \sqrt{x^2+y^2}\cos 2\pi z \text{ and } y = \sqrt{x^2+y^2}\sin 2\pi z\}, \end{aligned}$$

and let $p\colon \tilde{X} \to X$ be defined so that $p(x,y,z) = (x,y)$ for all $(x,y,z) \in \tilde{X}$. Now $\exp(w) = T(p(h(w)))$ for all $w \in \mathbb{C}$, where

$$h(u+iv) = \left(e^u \cos v, e^u \sin v, \frac{v}{2\pi}\right)$$

for all real numbers u and v and $T(x,y) = x + iy$ for all $(x,y) \in X$.

Moreover $h\colon \mathbb{C} \to \tilde{X}$ is a homeomorphism whose inverse h^{-1} satisfies

$$h^{-1}(z) = \tfrac{1}{2}\log(x^2+y^2) + 2\pi i z$$

for all $(x,y,z) \in \tilde{X}$.

The map $p\colon \tilde{X} \to X$ is a covering map. Indeed let

$$W_{s,m} = \{(x,y,z) \in \tilde{X} : s + m - \tfrac{1}{2} < z < s + m + \tfrac{1}{2}\}$$

and let $V_{s,m} = p(W_{s,0})$ for all real numbers s and integers m. Then $V_{s,0}$ is an open set in X, $p^{-1}(V_{s,0}) = \bigcup_{m\in\mathbb{Z}} W_{s,m}$ and p maps $W_{s,m}$ homeomorphically onto $V_{s,0}$ for all $s \in \mathbb{R}$ and $m \in \mathbb{Z}$.

The surface $\tilde{X}$ is a *helicoid* in $\mathbb{R}^3$.

Example Consider the map $\alpha\colon (-2,2) \to S^1$, 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.

Example Let $Z = \mathbb{C} \setminus \{1, -1\}$, let

$$\tilde{Z} = \{(z, w) \in \mathbb{C}^2 : w \neq 0 \text{ and } w^2 = z^2 - 1\},$$

and let $p\colon \tilde{Z} \to Z$ be defined such that $p(z, w) = z$ for all $(z, w) \in \tilde{Z}$. Let $(z_0, w_0) \in \tilde{Z}$, and let $z = z_0 + \zeta$. Then

$$\begin{aligned} z^2 - 1 &= z_0^2 - 1 + 2z_0\zeta + \zeta^2 = w_0^2 + 2z_0\zeta + \zeta^2 \\ &= w_0^2 \left(1 + \frac{2z_0\zeta + \zeta^2}{w_0^2}\right). \end{aligned}$$

Now the continuity at zero of the function sending each complex number ζ to $(2z_0\zeta + \zeta^2)/w_0^2$ ensures that there exists some positive real number δ such that

$$\left|\frac{2z_0\zeta + \zeta^2}{w_0^2}\right| < 1$$

whenever $|\zeta| < \delta$. Let $D(z_0, \delta)$ be the open disk of radius δ about z_0 in the complex plane, and let

$$F(z) = \frac{1}{2} \log \left(1 + \frac{2z_0(z - z_0) + (z - z_0)^2}{w_0^2}\right)$$

for all $z \in D(z_0, \delta)$, where $\log(re^{i\theta}) = \log r + i\theta$ for all real numbers r and θ satisfying $r > 0$ and $-\pi < \theta < \pi$. Then $F(z)$ is a continuous function of z on $D(z_0, \delta)$, and

$$\exp(F(z))^2 = 1 + \frac{2z_0(z - z_0) + (z - z_0)^2}{w_0^2} = \frac{z^2 - 1}{w_0^2}$$

for all $z \in D(z_0, \delta)$.

Let $(z, w) \in p^{-1}(D(z_0, \delta))$. Then $z \in D(z_0, \delta)$ and

$$w^2 = z^2 - 1 = \Bigl(w_0 \exp(F(z))\Bigr)^2,$$

and therefore $w = \pm w_0 \exp(F(z))$. It follows that $p^{-1}(D(z_0, \delta)) = W_+ \cup W_-$ where

$$\begin{aligned} W_+ &= \{(z,w) \in \mathbb{C}^2 : z \in D(z_0,\delta) \text{ and } w = w_0 \exp(F(z))\}, \\ W_- &= \{(z,w) \in \mathbb{C}^2 : z \in D(z_0,\delta) \text{ and } w = -w_0 \exp(F(z))\}, \end{aligned}$$

Now

$$\mathrm{Re}\left[1 + \frac{2z_0(z - z_0) + (z - z_0)^2}{w_0^2}\right] > 0$$

for all $z \in D(z_0, \delta)$. It follows from the definition of $F(z)$ that

$$-\tfrac{1}{4}\pi < \mathrm{Im}[F(z)] < \tfrac{1}{4}\pi$$

for all $z \in D(z_0, \delta)$, and therefore

$$\mathrm{Re}[\exp(F(z))] = \exp(\mathrm{Re}[F(z)]) \cos(\mathrm{Im}[F(z)]) > 0$$

for all $z \in D(z_0, \delta)$. It follows that

$$\begin{aligned} W_+ &= \left\{(z,w) \in \tilde{Z} : z \in D(z_0,\delta) \text{ and } \mathrm{Re}\left[\frac{w}{w_0}\right] > 0\right\}, \\ &= \left\{(z,w) \in p^{-1}\left(D(z_0,\delta)\right) : \mathrm{Re}\left[\frac{w}{w_0}\right] > 0\right\}, \\ W_- &= \left\{(z,w) \in \tilde{Z} : z \in D(z_0,\delta) \text{ and } \mathrm{Re}\left[\frac{w}{w_0}\right] < 0\right\}, \\ &= \left\{(z,w) \in p^{-1}\left(D(z_0,\delta)\right) : \mathrm{Re}\left[\frac{w}{w_0}\right] < 0\right\}. \end{aligned}$$

Now $p^{-1}(D(z_0, \delta)$ is open in $\tilde{Z}$, because the it is the preimage of the open subset $D(z_0, \delta)$ of Z under the continuous map $p\colon \tilde{Z} \to Z$. Moreover the function mapping (z, w) to the real part of w/w_0 is continuous on $p^{-1}(D(z_0, \delta))$. It follows that W_+ and W_- are open in $\tilde{Z}$. Also $W_+ \cap W_- = \emptyset$, and the map $p\colon \tilde{Z} \to Z$ maps each of the sets W_+ and W_- homeomorphically onto Z, where $Z = \mathbb{C} \setminus \{1, -1\}$. It follows that the open disk $D(z_0, \delta)$ is evenly covered by the map $p\colon \tilde{Z} \to Z$. We have therefore shown that this map is a covering map.

Let $\tilde{f}(z, w) = w$ for all $(z, w) \in \tilde{Z}$. Then

$$\tilde{f}(\tilde{z})^2 = p(\tilde{z})^2 - 1$$

for all $\tilde{z} \in \tilde{Z}$. It follows that the function $\tilde{f}\colon \tilde{Z} \to \mathbb{C}$ represents in some sense the many-valued 'function' $\sqrt{z^2 - 1}$. However this function $\tilde{z}$ is not defined on the open subset Z of the complex plane, but is instead defined over a covering space $\tilde{Z}$ of this open set. This covering space is the *Riemann surface* for the 'function' $\sqrt{z^2 - 1}$. This method of representing many-valued 'functions' of a complex variable using single-valued functions defined over a covering space was initiated and extensively developed by Bernhard Riemann (1826–1866) in his doctoral thesis.

Proposition 4.1 *Let $p: \tilde{X} \to X$ be a covering map. Then $p(V)$ is open in X for every open set V in $\tilde{X}$.*

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. ∎

Corollary 4.2 *A bijective covering map is a homeomophism.*

Proof This result follows directly from Proposition 4.1 the fact that a continuous bijection is a homeomorphism if and only if it maps open sets to open sets. ∎

4.2 Uniqueness of Lifts into Covering Spaces

Definition Let $p: \tilde{X} \to X$ be a covering map, 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 $f: Z \to X$ to the covering space $\tilde{X}$ if $p \circ \tilde{f} = f$.

Much of the general theory of covering maps is concerned with the development of necessary and sufficient conditions to determine whether or not maps into the base space of a covering map can be lifted to the covering space.

We prove that any lift of a given map from a connected topological topological space into the base space of a covering map is determined by its value at a single point of its domain.

Proposition 4.3 *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 at least one point z of 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. ∎

Corollary 4.4 *Let $p: \tilde{X} \to X$ be a covering map over a topological space X. Let Z be a connected topological space, and let $f: Z \to \tilde{X}$ be a continuous map. Suppose that $p(f(z)) = x_0$ for all $z \in Z$, where x_0 is some point of X. Then $f(z) = \tilde{x}_0$ for all $z \in Z$, where $\tilde{x}_0$ is some point of $\tilde{X}$ which satisfies $p(\tilde{x}_0) = x_0$.*

Proof Let z_0 be some point of Z. Let $\tilde{x}_0 = f(z_0)$, and let $c: Z \to \tilde{X}$ be the constant map defined by $c(z) = \tilde{x}_0$ for all $z \in Z$. Then $c(z_0) = f(z_0)$ and $p \circ c = p \circ f$. It follows from Theorem 4.3 that $f = c$, as required. ∎

4.3 The Path-Lifting Theorem

Theorem 4.5 (Path-Lifting Theorem) *Let $p: \tilde{X} \to X$ be a covering map over a topological space X. Let $\gamma: [a, b] \to X$ be a continuous map from the closed interval $[a, b]$ to X, and let w be a point of $\tilde{X}$ for which $p(w) = \gamma(a)$. Then there exists a unique continuous map $\tilde{\gamma}: [a, b] \to \tilde{X}$ for which $\tilde{\gamma}(a) = w$ and $p \circ \tilde{\gamma} = \gamma$.*

Proof Let S be the subset of $[a, b]$ defined as follows: an element c of $[a, b]$ belongs to S if and only if there exists a continuous map $\eta_c: [a, c] \to \tilde{X}$ such that $\eta_c(a) = w$ and $p(\eta_c(t)) = \gamma(t)$ for all $t \in [a, c]$. Note that S is non-empty, since a belongs to S. Let $s = \sup S$.

There exists an open neighbourhood U of $\gamma(s)$ which is evenly covered by the map p, since $p: \tilde{X} \to X$ is a covering map. It then follows from the continuity of the path γ that there exists some $\delta > 0$ such that $\gamma(J(s, \delta)) \subset U$, where

$$J(s, \delta) = \{t \in [a, b] : |t - s| < \delta\}.$$

Now $S \cap J(s, \delta)$ is non-empty, because s is the supremum of the set S. Choose some element c of $S \cap J(s, \delta)$. Then there exists a continuous map $\eta_c: [a, c] \to \tilde{X}$ such that $\eta_c(a) = w$ and $p(\eta_c(t)) = \gamma(t)$ for all $t \in [a, c]$. Now the open set U is evenly covered by the map p. Therefore $p^{-1}(U)$ is a disjoint union of open sets in $\tilde{X}$, each of which is mapped homeomorphically onto U by the covering map p. One of these open sets contains the point $\eta_c(c)$; let this open set be denoted by $\tilde{U}$. There then exists a unique continuous map $\sigma: U \to \tilde{U}$ defined such that, for all $x \in U$, $\sigma(x)$ is the unique element of $\tilde{U}$ for which $p(\sigma(x)) = x$. Then $\sigma(\gamma(c)) = \eta_c(c)$.

Then, given any $d \in J(s, \delta)$, let $\eta_d: [a, d] \to \tilde{X}$ be the function from $[a, d]$ to $\tilde{X}$ defined so that

$$\eta_d(t) = \begin{cases} \eta_c(t) & \text{if } a \leq t \leq c; \\ \sigma(\gamma(t)) & \text{if } c \leq t \leq d. \end{cases}$$

Then $\eta_d(a) = w$ and $p(\eta_d(t)) = \gamma(t)$ for all $t \in [a, d]$. The restrictions of the function $\eta_d: [a, d] \to \tilde{X}$ to the intervals $[a, c]$ and $[c, d]$ are continuous. It follows from the Pasting Lemma (Lemma 1.24) that η_d is continuous on $[a, d]$. Thus $d \in S$. We conclude from this that $J(s, \delta) \subset S$. However s is defined to be the supremum of the set S. Therefore $s = b$, and b belongs to S. It follows that that there exists a continuous map $\tilde{\gamma}: [a, b] \to \tilde{X}$ for which $\tilde{\gamma}(a) = w$ and $p \circ \tilde{\gamma} = \gamma$, as required. ∎

4.4 The Homotopy-Lifting Theorem

Theorem 4.6 (Homotopy-Lifting Theorem) *Let $p: \tilde{X} \to X$ be a covering map over a topological space X. Let Z be a topological space, and let $F: Z \times [0, 1] \to X$ and $g: Z \to \tilde{X}$ be continuous maps with the property that $p(g(z)) = F(z, 0)$ for all $z \in Z$. Then there exists a unique continuous map $G: Z \times [0, 1] \to \tilde{X}$ such that $G(z, 0) = g(z)$ for all $z \in Z$ and $p \circ G = F$.*

Proof For each $z \in Z$, consider the path $\gamma_z: [0, 1] \to Z$ defined by $\gamma_z(t) = F(z, t)$ for all $t \in [0, 1]$. Note that $p(g(z)) = \gamma_z(0)$. It follows from the Path-Lifting Theorem (Theorem 4.5) that there exists a unique continuous path

$\tilde{\gamma}_z: [0,1] \to \tilde{X}$ such that $\tilde{\gamma}_z(0) = g(z)$ for all $z \in Z$ and $p \circ \tilde{\gamma}_z = \gamma_z$. Let the function $G: Z \times [0,1] \to \tilde{X}$ be defined by $G(z,t) = \tilde{\gamma}_z(t)$ for all $z \in Z$ and $t \in [0,1]$. Then $G(z,0) = g(z)$ for all $z \in Z$ and

$$p(G(z,t)) = p(\tilde{\gamma}_z(t)) = \gamma_z(t) = F(z,t)$$

for all $z \in Z$ and $t \in [0,1]$. It remains to show that the function $G: Z \times [0,1] \to \tilde{X}$ is continuous and that it is unique.

Given any $z \in Z$, let S_z denote the set of all real numbers c belonging to the closed interval $[0,1]$ which have the following property:

> there exists an open set N in Z such that $z \in N$ and the function G is continuous on $N \times [0,c]$.

Let s_z be the supremum $\sup S_z$ (i.e., the least upper bound) of the set S_z. We prove that s_z belongs to the set S_z and that $s_z = 1$.

Choose some $z \in Z$, and let $w \in \tilde{X}$ be given by $w = G(z, s_z)$. There exists an open neighbourhood U of $p(w)$ in X which is evenly covered by the map p. Thus $p^{-1}(U)$ 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 the point w; let this open set be denoted by $\tilde{U}$. Then there exists a unique continuous map $\sigma: U \to \tilde{U}$ defined such that, for all $x \in U$, $\sigma(x)$ is the unique element of $\tilde{U}$ for which $p(\sigma(x)) = x$. Then $\sigma(F(z, s_z)) = w$. Now $F(z, s_z) = p(w)$. It follows from the continuity of the map F that there exists some positive real number δ and some open set M_1 in Z such that $z \in M_1$ and $F(M_1 \times J(s_z, \delta)) \subset U$, where

$$J(s_z, \delta) = \{t \in [0,1] : s_z - \delta < t < s_z + \delta\}.$$

Now we can choose some c belonging to S_z which satisfies $s_z - \delta < c \leq s_z$, because s_z is the least upper bound of the set S_z. It then follows from the definition of the set S_z that there exists an open set M_2 in Z such that $z \in M_2$ and the function G is continuous on $M_2 \times [0,c]$. Let

$$N = \{z' \in M_1 \cap M_2 : G(z', c) \in \tilde{U}\}.$$

Then $z \in N$, and the continuity of the function G on $M_2 \times [0,c]$ ensures that N is open in Z. Moreover the function G is continuous on $N \times [0,c]$ and $F(N \times J(s_z, \delta)) \subset U$.

Let $z' \in N$. Then $G(z', c) \in \tilde{U}$ and $p(G(z', c)) = F(z', c)$. It follows from the definition of the map $\sigma: U \to \tilde{X}$ that $G(z', c) = \sigma(F(z', c))$. Also the interval $J(s_z, \delta)$ is connected, and

$$p(G(z', t)) = F(z', t) = p(\sigma(F(z', t))$$

for all $t \in J(s_z, \delta)$. It follows from Theorem 4.3 that $G(z', t) = \sigma(F(z', t)$ for all $t \in J(s_z, \delta)$. We have thus shown that the function G is equal to the continuous function $\sigma \circ F$ on $N \times J(s_z, \delta)$. The function G is therefore continuous on both $N \times [0, c]$ and $N \times [c, t]$ for all $t \in J(s_z, \delta)$ satisfying $t \geq c$. It then follows from the Pasting Lemma (Lemma 1.24) that the function G is continuous on $N \times [0, t]$ for all $t \in J(s_z, \delta)$, and thus $J(s_z, \delta) \subset S_z$. This however contradicts the definition of S_z unless $s_z \in S_z$ and $s_z = 1$. We conclude therefore that $1 \in S_z$, and thus there exists an open set N in Z such that $z \in N$ and $G|N \times [0, 1]$ is continuous.

We conclude from this that every point of $Z \times [0, 1]$ is contained in some open subset of $Z \times [0, 1]$ on which that function G is continuous. It follows that $G: Z \times [0, 1] \to \tilde{X}$ is continuous (see Proposition 1.23).

The uniqueness of the map $G: Z \times [0, 1] \to \tilde{X}$ follows directly from the fact that for any $z \in Z$ there is a unique continuous path $\tilde{\gamma}_z: [0, 1] \to \tilde{X}$ such that $\tilde{\gamma}_z(0) = g(z)$ and $p(\tilde{\gamma}_z(t)) = F(z, t)$ for all $t \in [0, 1]$. ∎

4.5 Path-Lifting and the Fundamental Group

Let $p: \tilde{X} \to X$ be a covering map and let $\alpha: [0, 1] \to X$ and $\beta: [0, 1] \to X$ be paths in the base space X which both start at some point x_0 of X and finish at some point x_1 of X, so that

$$\alpha(0) = \beta(0) = x_0 \quad \text{and} \quad \alpha(1) = \beta(1) = x_1.$$

Let $\tilde{x}_0$ be some point of the covering space $\tilde{X}$ that projects down to x_0, so that $p(\tilde{x}_0) = x_0$. It follows from the Path-Lifting Theorem (Theorem 4.5) that there exist paths $\tilde{\alpha}: [0, 1] \to \tilde{X}$ and $\tilde{\beta}: [0, 1] \to \tilde{X}$ in the covering space $\tilde{X}$ that both start at $\tilde{x}_0$ and that are lifts of the paths α and β respectively. Thus

$$\tilde{\alpha}(0) = \tilde{\beta}(0) = \tilde{x}_0,$$

$$p(\tilde{\alpha}(t)) = \alpha(t) \quad \text{and} \quad p(\tilde{\beta}(t)) = \beta(t) \quad \text{for all } t \in [0, 1].$$

These lifts $\tilde{\alpha}$ and $\tilde{\beta}$ of the paths α and β are uniquely determined by their starting point $\tilde{x}_0$ (see Proposition 4.3).

Now, though the lifts $\tilde{\alpha}$ and $\tilde{\beta}$ of the paths α and β have been chosen such that they start at the same point $\tilde{x}_0$ of the covering space $\tilde{X}$, they need not in general end at the same point of $\tilde{X}$. However we shall prove that if $\alpha \simeq \beta$ rel $\{0, 1\}$, then the lifts $\tilde{\alpha}$ and $\tilde{\beta}$ of α and β respectively that both start at some point $\tilde{x}_0$ of $\tilde{X}$ will both finish at some point $\tilde{x}_1$ of $\tilde{x}$, so that $\tilde{\alpha}(1) = \tilde{\beta}(1) = \tilde{x}_1$. This result is established in Proposition 4.7 below.

Proposition 4.7 *Let $p\colon \tilde{X} \to X$ be a covering map over a topological space X, let $\alpha\colon [0,1] \to X$ and $\beta\colon [0,1] \to X$ be paths in X, where $\alpha(0) = \beta(0)$ and $\alpha(1) = \beta(1)$, and let $\tilde{\alpha}\colon [0,1] \to \tilde{X}$ and $\tilde{\beta}\colon [0,1] \to \tilde{X}$ be paths in $\tilde{X}$ such that $p \circ \tilde{\alpha} = \alpha$ and $p \circ \tilde{\beta} = \beta$. Suppose that $\tilde{\alpha}(0) = \tilde{\beta}(0)$ and that $\alpha \simeq \beta \text{ rel } \{0,1\}$. Then $\tilde{\alpha}(1) = \tilde{\beta}(1)$ and $\tilde{\alpha} \simeq \tilde{\beta} \text{ rel } \{0,1\}$.*

Proof Let x_0 and x_1 be the points of X given by

$$x_0 = \alpha(0) = \beta(0), \qquad x_1 = \alpha(1) = \beta(1).$$

Now $\alpha \simeq \beta \text{ rel } \{0,1\}$, and therefore there exists a homotopy $F\colon [0,1]\times[0,1] \to X$ such that

$$F(t,0) = \alpha(t) \quad \text{and} \quad F(t,1) = \beta(t) \quad \text{for all } t \in [0,1],$$

and

$$F(0,\tau) = x_0 \quad \text{and} \quad F(1,\tau) = x_1 \quad \text{for all } \tau \in [0,1].$$

It then follows from the Homotopy-Lifting Theorem (Theorem 4.6) that there exists a continuous map $G\colon [0,1] \times [0,1] \to \tilde{X}$ such that $p \circ G = F$ and $G(0,0) = \tilde{\alpha}(0)$. Then $p(G(0,\tau)) = x_0$ and $p(G(1,\tau)) = x_1$ for all $\tau \in [0,1]$. A straightforward application of Proposition 4.3 shows that any continuous lift of a constant path must itself be a constant path. Therefore $G(0,\tau) = \tilde{x}_0$ and $G(1,\tau) = \tilde{x}_1$ for all $\tau \in [0,1]$, where

$$\tilde{x}_0 = G(0,0) = \tilde{\alpha}(0), \qquad \tilde{x}_1 = G(1,0).$$

However

$$G(0,0) = G(0,1) = \tilde{x}_0 = \tilde{\alpha}(0) = \tilde{\beta}(0),$$

$$p(G(t,0)) = F(t,0) = \alpha(t) = p(\tilde{\alpha}(t))$$

and

$$p(G(t,1)) = F(t,1) = \beta(t) = p(\tilde{\beta}(t))$$

for all $t \in [0,1]$. It follows that the map that sends $t \in [0,1]$ to $G(t,0)$ is a lift of the path α that starts at $\tilde{x}_0$, and the map that sends $t \in [0,1]$ to $G(t,1)$ is a lift of the path β that also starts at $\tilde{x}_0$.

However Proposition 4.3 ensures that the lifts $\tilde{\alpha}$ and $\tilde{\beta}$ of the paths α and β are uniquely determined by their starting points. It follows that $G(t,0) = \tilde{\alpha}(t)$ and $G(t,1) = \tilde{\beta}(t)$ for all $t \in [0,1]$. In particular,

$$\tilde{\alpha}(1) = G(1,0) = \tilde{x}_1 = G(1,1) = \tilde{\beta}(1).$$

Moreover the map $G\colon [0,1] \times [0,1] \to \tilde{X}$ is a homotopy between the paths $\tilde{\alpha}$ and $\tilde{\beta}$ which satisfies $G(0,\tau) = \tilde{x}_0$ and $G(1,\tau) = \tilde{x}_1$ for all $\tau \in [0,1]$. It follows that $\tilde{\alpha} \simeq \tilde{\beta} \text{ rel } \{0,1\}$, as required. ∎

Proposition 4.8 *Let $p: \tilde{X} \to X$ be a covering map, and let $\tilde{x}_0$ be a point of the covering space $\tilde{X}$. Then the homomorphism*

$$p_{\#}: \pi_1(\tilde{X}, \tilde{x}_0) \to \pi_1(X, p(\tilde{x}_0))$$

of fundamental groups induced by the covering map p is injective.

Proof Let σ_0 and σ_1 be loops in $\tilde{X}$ based at the point $\tilde{x}_0$, representing elements $[\sigma_0]$ and $[\sigma_1]$ of $\pi_1(\tilde{X}, \tilde{x}_0)$. Suppose that $p_{\#}[\sigma_0] = p_{\#}[\sigma_1]$. Then $p \circ \sigma_0 \simeq p \circ \sigma_1$ rel $\{0, 1\}$. Also $\sigma_0(0) = \tilde{x}_0 = \sigma_1(0)$. Therefore $\sigma_0 \simeq \sigma_1$ rel $\{0, 1\}$, by Proposition 4.7, and thus $[\sigma_0] = [\sigma_1]$. We conclude that the homomorphism $p_{\#}: \pi_1(\tilde{X}, \tilde{x}_0) \to \pi_1(X, p(\tilde{x}_0))$ is injective. ∎

Proposition 4.9 *Let $p: \tilde{X} \to X$ be a covering map, let $\tilde{x}_0$ be a point of the covering space $\tilde{X}$, and let γ be a loop in X based at $p(\tilde{x}_0)$. Then $[\gamma] \in p_{\#}(\pi_1(\tilde{X}, \tilde{x}_0))$ if and only if there exists a loop $\tilde{\gamma}$ in $\tilde{X}$, based at the point $\tilde{x}_0$, such that $p \circ \tilde{\gamma} = \gamma$.*

Proof If $\gamma = p \circ \tilde{\gamma}$ for some loop $\tilde{\gamma}$ in $\tilde{X}$ based at $\tilde{x}_0$ then $[\gamma] = p_{\#}[\tilde{\gamma}]$, and therefore $[\gamma] \in p_{\#}(\pi_1(\tilde{X}, \tilde{x}_0))$.

Conversely suppose that $[\gamma] \in p_{\#}(\pi_1(\tilde{X}, \tilde{x}_0))$. We must show that there exists some loop $\tilde{\gamma}$ in $\tilde{X}$ based at $\tilde{x}_0$ such that $\gamma = p \circ \tilde{\gamma}$. Now there exists a loop σ in $\tilde{X}$ based at the point $\tilde{x}_0$ such that $[\gamma] = p_{\#}([\sigma])$ in $\pi_1(X, p(\tilde{x}_0))$. Then $\gamma \simeq p \circ \sigma$ rel $\{0, 1\}$. It follows from the Path-Lifting Theorem for covering maps (Theorem 4.5) that there exists a unique path $\tilde{\gamma}: [0, 1] \to \tilde{X}$ in $\tilde{X}$ for which $\tilde{\gamma}(0) = \tilde{x}_0$ and $p \circ \tilde{\gamma} = \gamma$. It then follows from Proposition 4.7 that $\tilde{\gamma}(1) = \sigma(1)$ and $\tilde{\gamma} \simeq \sigma$ rel $\{0, 1\}$. But $\sigma(1) = \tilde{x}_0$. Therefore the path $\tilde{\gamma}$ is the required loop in $\tilde{X}$ based the point $\tilde{x}_0$ which satisfies $p \circ \tilde{\gamma} = \gamma$. ∎

Corollary 4.10 *Let $p: \tilde{X} \to X$ be a covering map over a topological space X, let w_0 and w_1 be points of $\tilde{X}$ satisfying $p(w_0) = p(w_1)$, and let $\alpha: [0, 1] \to \tilde{X}$ be a path in $\tilde{X}$ from w_0 to w_1. Suppose that $[p \circ \alpha] \in p_{\#}(\pi_1(\tilde{X}, w_0))$. Then the path α is a loop in $\tilde{X}$, and thus $w_0 = w_1$.*

Proof It follows from Proposition 4.9 that there exists a loop β based at w_0 satisfying $p \circ \beta = p \circ \alpha$. Then $\alpha(0) = \beta(0)$. Now Proposition 4.3 ensures that the lift to $\tilde{X}$ of any path in X is uniquely determined by its starting point. It follows that $\alpha = \beta$. But then the path α must be a loop in $\tilde{X}$, and therefore $w_0 = w_1$, as required. ∎

Theorem 4.11 *Let $p: \tilde{X} \to X$ be a covering map over a topological space X. Suppose that $\tilde{X}$ is path-connected and that X is simply-connected. Then the covering map $p: \tilde{X} \to X$ is a homeomorphism.*

Proof We show that the map $p: \tilde{X} \to X$ is a bijection. This map is surjective (since covering maps are by definition surjective). We must show that it is injective. Let w_0 and w_1 be points of $\tilde{X}$ with the property that $p(w_0) = p(w_1)$. Then there exists a path $\alpha: [0,1] \to \tilde{X}$ with $\alpha(0) = w_0$ and $\alpha(1) = w_1$, since $\tilde{X}$ is path-connected. Then $p \circ \alpha$ is a loop in X based at the point x_0, where $x_0 = p(w_0)$. However $\pi_1(X, p(w_0))$ is the trivial group, since X is simply-connected. It follows from Corollary 4.10 that the path α is a loop in $\tilde{X}$ based at w_0, and therefore $w_0 = w_1$. This shows that the the covering map $p: \tilde{X} \to X$ is injective. Thus the map $p: \tilde{X} \to X$ is a bijection, and thus has a well-defined inverse $p^{-1}: X \to \tilde{X}$. But any bijective covering map is a homeomorphism (Corollary 4.2). The result follows. ∎