Course 421: Algebraic Topology
Michaelmas Term 1988

D. R. Wilkins

Contents

1 Continuous Maps and Homotopies

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

If f and g are continuous maps from a topological space X to a topological space Y then the maps f and g are homotopic if and only if it is possible to 'continuously deform' the map f into the map g.

Lemma 1.1 *Let X and Y be topological spaces. Then the relation $\simeq$ is an equivalence relation on the set of all continuous maps from X to Y (where $f \simeq g$ if and only if the maps f and g are homotopic).*

Proof We must prove that the relation $\simeq$ is reflexive, symmetric and transitive. Let f, g and h be continuous maps from X to Y.

We observe that $f \simeq f$. (This result follows directly on considering the *constant homotopy* $F_0: X \times [0,1] \to Y$ defined by $F_0(x,t) = f(x)$ for all $x \in X$ and $t \in [0,1]$.) Thus the relation $\simeq$ is *reflexive.*

Suppose that $f \simeq g$. Let $H: X \times [0,1] \to Y$ be a homotopy such that $H(x,0) = f(x)$ and $H(x,1) = g(x)$ Define $K: X \times [0,1] \to Y$ by $K(x,t) = H(x, 1-t)$ for all $x \in X$ and $t \in [0,1]$. Then the map K is continuous, and $K(x,0) = g(x)$ and $K(x,1) = h(x)$ for all $x \in X$. Hence $g \simeq f$. Thus the relation $\simeq$ is *symmetric.*

Now suppose that $f \simeq g$ and $g \simeq h$. We must show that $f \simeq h$. There exist homotopies $H_1: X \times [0,1] \to Y$ and $H_2: X \times [0,1] \to Y$ such that $H_1(x,0) = f(x)$, $H_1(x,1) = g(x) = H_2(x,0)$ and $H_2(x,1) = h(x)$. Define a map $G: X \times [0,1] \to Y$ by

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

It follows easily from the continuity of the maps H_1 and H_2 that the map G is continuous. (Indeed the required result follows from a straightforward application of Lemma A.7 in Appendix A.) Also $G(x,0) = f(x)$ and $G(x,1) = h(x)$ for all $x \in X$. Therefore $f \simeq h$. Thus the relation $\simeq$ is *transitive.*

The relation $\simeq$ is reflexive, symmetric and transitive. Therefore it is an equivalence relation on the set of all continuous maps from the topological space X to the topological space Y. ∎

Lemma 1.2 *Let W, X, Y and Z be topological spaces, and let $q: W \to X$, $f: X \to Y$, $g: X \to Y$ and $r: Y \to Z$ be continuous maps. Suppose that $f \simeq g$. Then $f \circ q \simeq g \circ q$ and $r \circ f \simeq r \circ g$.*

Proof Let $H: X \times [0, 1] \to Y$ be a homotopy between the maps f and g. Thus $H(x, 0) = f(x)$ and $H(x, 1) = g(x)$ for all $x \in X$. Let $G: W \times [0, 1] \to Y$ be the map defined by $G(w, t) = H(q(w), t)$ for all $w \in W$. Then G is a homotopy between the maps $f \circ q$ and $g \circ q$ Similarly the composition map $r \circ H: X \times [0, 1] \to Z$ is a homotopy between $r \circ f$ and $r \circ g$. ∎

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 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 $a \in A$. If the maps f and g are homotopic relative to the subset A of X then we denote this fact by writing $f \simeq g$ rel A.

Let X and Y be topological spaces. Let us choose points x_0 and y_0 of X and Y respectively. We refer to these chosen points as *basepoints* for the spaces X and Y. A continuous map $f: X \to Y$ is said to be *basepoint-preserving* if and only if $f(x_0) = y_0$. A homotopy $H: X \times [0, 1] \to Y$ between basepoint-preserving continuous maps is said to be *basepoint-preserving* if and only if $H(x_0, t) = y_0$ for all $t \in [0, 1]$. One can define an equivalence relation on the set of all basepoint-preserving continuous maps from X to Y, where two such maps are equivalent if and only if they are homotopic through a basepoint-preserving homotopy. Thus if $f: X \to Y$ and $g: X \to Y$ are basepoint-preserving continuous maps then f and g are equivalent if and only if $f \simeq g$ rel $\{x_0\}$. The set of equivalence classes of basepoint-preserving continuous maps from X to Y is often denoted by $[X, Y]_0$.

Definition A topological space X is said to be *contractible* if and only if there exists a point p of X and a continuous map $F: X \times [0, 1] \to X$ such that $F(x, 0) = p$ and $F(x, 1) = x$ for all $x \in X$.

We see that a topological space X is contractible if and only if the identity map of the space X is homotopic to the constant map which sends the whole of X to some point p of X.

1.1 Identification Maps and Homotopies

Definition Let X and Y be topological spaces and let $q\colon X \to Y$ be a map from X to Y. The map q is said to be an *identification map* if and only if the following conditions are satisfied:

(i) the map $q\colon X \to Y$ is surjective,

(ii) a subset U of Y is open in Y if and only if $q^{-1}(U)$ is open in X.

Let $q\colon X \to Y$ be an identification map, and let U be an open set in Y. Then $q^{-1}(U)$ is an open set in X. Thus $q\colon X \to Y$ is continuous. We conclude therefore that every identification map is necessarily continuous.

Lemma 1.3 *Let X be a topological space, let Y be a set, and let $q\colon X \to Y$ be a surjection. Then there is a unique topology on Y for which the map $q\colon X \to Y$ is an identification map.*

Proof Let τ be the collection consisting of all subsets U of Y for which $q^{-1}(U)$ is open in X. Now $q^{-1}(\emptyset) = \emptyset$, and $q^{-1}(Y) = X$, so that $\emptyset \in \tau$ and $Y \in \tau$. It is readily verified that any union of sets belonging to τ is itself a set belonging to τ, and that any finite intersection of sets belonging to τ is itself a set belonging to τ. Thus τ is a topology on Y, and the map $q\colon X \to Y$ is an identification map with respect to the topology τ on Y. Clearly the topology τ is the unique topology on Y for which the map $q\colon X \to Y$ is a identification map. ∎

Let X be a topological space, let Y be a set, and let $q\colon X \to Y$ be a surjection. The unique topology on Y for which the map q is an identifcation map is referred to as the *quotient topology* (or *identification topology*) on Y (with respect to the map q).

The following useful theorem is frequently used to provide examples of identification maps.

Theorem 1.4 *Let X be a compact topological space, let Y be a Hausdorff space, and let $q\colon X \to Y$ be a surjection. Suppose that the map q is continuous. Then $q\colon X \to Y$ is an identification map.*

Proof Let U be a subset of Y. We must show that U is open in Y if and only if $q^{-1}(U)$ is open in X. Now if U is open in Y, then $q^{-1}(U)$ is open in X, since the map q is continuous. Conversely suppose that $q^{-1}(U)$ is open in X. Let F be the complement $Y \setminus U$ of U in Y, and let $\tilde{F} = q^{-1}(F)$.

Then $\tilde{F} = X \setminus q^{-1}(U)$, hence $\tilde{F}$ is closed in X. But then $\tilde{F}$ is compact by Lemma A.16 of Appendix A (since X is a compact topological space). It then follows from Lemma A.17 that $q(\tilde{F})$ is a compact subset of Y. But $q(\tilde{F}) = F$, since the map $q: X \to Y$ is surjective. Thus F is compact. But every compact subset of the Hausdorff space Y is closed, by Lemma A.18. Therefore F is closed, so that U is open, as required. Thus $q: X \to Y$ is an identification map. ∎

Lemma 1.5 *Let X and Y be topological spaces and let $q: X \to Y$ be an identification map. Let Z be a topological space, and let $f: Y \to Z$ be a map from Y to Z. Then the map f is continuous if and only if the composition map $f \circ q: X \to Z$ is continuous.*

Proof Suppose that f is continuous. Then the composition map $f \circ q$ is a composition of continuous maps and hence is itself continuous.

Conversely suppose that $f \circ q$ is continuous. Let U be an open set in Z. Then $q^{-1}(f^{-1}(U))$ is open in X (since $f \circ q$ is continuous), and hence $f^{-1}(U)$ is open in Y (since the map q is an identification map). Therefore the map f is continuous, as required. ∎

Example Let S^1 be the unit circle in $\mathbb{R}^2$, defined by

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

Let $q: [0, 1] \to S^1$ be the map defined by

$$q(t) = \{(\cos 2\pi t, \sin 2\pi t) \qquad (t \in [0, 1]).$$

The map q is surjective. Moreover the closed interval $[0, 1]$ is compact, and the circle S^1 is Hausdorff. Therefore the map $q: [0, 1] \to S^1$ is an identification map. Thus a map $f: S^1 \to Z$ from the circle S^1 to some topological space Z is continuous if and only if the composition map $f \circ q: [0, 1] \to Z$ is continuous.

Let X and Y be topological spaces, and let $q: X \to Y$ be an identification map. Let Z be a topological space and let $f_0: Y \to Z$ and $f_1: Y \to Z$ be continuous maps from Y to Z. We claim that the maps f_0 and f_1 are homotopic if and only if there exists a (continuous) homotopy $\tilde{H}: X \times [0, 1] \to Y$ between the maps $f_0 \circ q$ and $f_1 \circ q$ which is consistent with the identifications represented by the map q, (so that $\tilde{H}(x_1, t) = \tilde{H}(x_2, t)$ whenever $q(x_1) = q(x_2)$). In order to prove this result, we first show that if $q: X \to Y$ is an identification map then so is the continuous map $\hat{q}: X \times [0, 1] \to Y \times [0, 1]$, where $\hat{q}(x, t) = (q(x), t)$ for all $x \in X$ and $t \in [0, 1]$. (The continuity of the map $\hat{q}$ may be verified using the definition of the product topology on $X \times [0, 1]$ and on $Y \times [0, 1]$.)

Let X and Z be topological spaces, let K be a compact subset of Z, and U be an open set in $X \times Z$. Let V be the subset of X defined by

$$V = \{x \in X : \{x\} \times K \subset U\}.$$

The Tube Lemma (Lemma A.25) shows that V is then an open set in X. We use this result in the proof of the following theorem.

Theorem 1.6 *Let X and Y be topological spaces and let $q: X \to Y$ be an identification map. Then the map $\hat{q}: X \times [0, 1] \to Y \times [0, 1]$ defined by $\hat{q}(x, t) = (q(x), t)$ is also an identification map.*

Proof The map $\hat{q}: X \times [0, 1] \to Y \times [0, 1]$ is surjective, since $q: X \to Y$ is surjective. Let U be a subset of $Y \times [0, 1]$, and let $\tilde{U} = \hat{q}^{-1}(U)$. In order to show that the map $\hat{q}$ is an identification map we must prove that U is open in $Y \times [0, 1]$ if and only if $\tilde{U}$ is open in $X \times [0, 1]$.

If U is open in $Y \times [0, 1]$ then $\tilde{U}$ is open in $X \times [0, 1]$, since the map $\hat{q}$ is continuous. Conversely suppose that $\tilde{U}$ is open in $X \times [0, 1]$. Let u_0 be a point of U. We show that there exists an open set N in $Y \times [0, 1]$ such that $u_0 \in N$ and $N \subset U$. Now $u_0 = (y_0, s)$ for some $y_0 \in Y$ and $s \in [0, 1]$. There exists some $x_0 \in X$ such that $q(x_0) = y_0$, since the identification map $q: X \to Y$ is (by definition) surjective. Then $(x_0, s) \in \tilde{U}$. Now

$$\{t \in [0, 1] : (x_0, t) \in \tilde{U}\}$$

is an open set in the interval $[0, 1]$, since $\tilde{U}$ is open in $X \times [0, 1]$. Therefore there exists some $\delta > 0$ such that $\{x_0\} \times I_{s,\delta} \subset \tilde{U}$, where

$$I_{s,\delta} = \{t \in [0, 1] : |t - s| \leq \delta\}.$$

Define

$$V = \{y \in Y : \{y\} \times I_{s,\delta} \subset U\}.$$

Then

$$q^{-1}(V) = \{x \in X : \{x\} \times I_{s,\delta} \subset \tilde{U}\}.$$

Using the Tube Lemma (Lemma A.25), we see that $q^{-1}(V)$ is open in X (since $I_{s,\delta}$ is compact and $\tilde{U}$ is open). Thus V is open in Y (since the map $q: X \to Y$ is an identification map). Moreover y_0 belongs to V (since x_0 belongs to $\tilde{V}$). Thus if we define

$$N = \{(y, t) \in Y \times [0, 1] : y \in V \text{ and } |t - s| < \delta\}$$

then N is an open neighbourhood of u_0 in $Y \times [0,1]$, where $u_0 = (y_0, s)$, and moreover $N \subset U$, as required. We have thus shown that if U is a subset of $Y \times [0,1]$ with the property that $\hat{q}^{-1}(U)$ is open in $X \times [0,1]$ then, given any point u_0 of U, there exists an open neighbourhood N of u_0 in $Y \times [0,1]$ for which $N \subset U$. Thus if $\hat{q}^{-1}(U)$ is open in $X \times [0,1]$ then U is open in $Y \times [0,1]$. Thus the map $\hat{q}: X \times [0,1] \to Y \times [0,1]$ is an identification map, as required. ∎

Let X, Y and Z be topological spaces, and let $q: X \to Y$ be an identification map. Let $F: Y \times [0,1] \to Z$ be a map from $Y \times [0,1]$ to Z. It follows from Theorem 1.6 and Lemma 1.5 that the map F is continuous if and only if the composition map $F \circ \hat{q}: X \times [0,1] \to Z$ is continuous (where $\hat{q}(x,t) = (q(x), t)$ for all $x \in X$ and $t \in [0,1]$). We deduce immediately the following result.

Corollary 1.7 *Let X and Y be topological spaces and let $q: X \to Y$ be an identification map. Let $f_0: Y \to Z$ and and $f_1: Y \to Z$ be continuous maps from Y to Z. Let $H: Y \times [0,1] \to Z$ be a function from $Y \times [0,1]$ to Z such that $H(y,0) = f_0(y)$ and $H(y,1) = f_1(y)$ for all $y \in Y$. Then the function H is continuous (and so defines a homotopy between the maps f_0 and f_1) if and only if the corresponding function $H \circ \hat{q}: X \times [0,1] \to Z$ from $X \times [0,1]$ to Z is continuous (where $\hat{q}(x,t) \equiv (q(x), t)$ for all $x \in X$ and $t \in [0,1]$).*

Example Let Z be a topological space, and let $f_0: S^1 \to Z$ and $f_1: S^1 \to Z$ be continuous maps from S^1 to Z, where S^1 denotes the unit circle in $\mathbb{R}^2$. Let $q: [0,1] \to S^1$ be the identification map defined by

$$q(t) = \{(\cos 2\pi t, \sin 2\pi t) \qquad (t \in [0,1]).$$

Then $f_0 \simeq f_1$ if and only if there exists a continuous map $K: [0,1] \times [0,1] \to Z$ such that

$$K(t,0) = f_0(\cos 2\pi t, \sin 2\pi t) \quad \text{and} \quad K(t,1) = f_1(\cos 2\pi t, \sin 2\pi t)$$

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

$$K(0,\tau) = K(1,\tau)$$

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

Example Let S^n denote the n-sphere given by

$$S^n = \{\mathbf{x} \in \mathbb{R}^{n+1} : |\mathbf{x}| = 1\}.$$

Let $\mathbb{R}P^n$ denote real projective n-space. Thus $\mathbb{R}P^n$ is the set whose elements are in bijective correspondence with the lines in $\mathbb{R}^{n+1}$ that pass through the origin. Each point $\mathbf{x}$ of S^n determines a unique element $p(\mathbf{x})$ of $\mathbb{R}P^n$ which represents the line in $\mathbb{R}P^n$ that passes through both the origin and the point $\mathbf{x}$. This defines a map $p\colon S^n \to \mathbb{R}P^n$. Note that the map p is surjective, and that $p(\mathbf{x}) = p(\mathbf{y})$ if and only if $\mathbf{y} = \pm\mathbf{x}$ (where $\mathbf{x}$ and $\mathbf{y}$ are elements of S^n). It follows from Lemma 1.3 that there is a unique topology on $\mathbb{R}P^n$ for which the map $p\colon S^n \to \mathbb{R}P^n$ is an identification map. We regard this topology as the standard topology on $\mathbb{R}P^n$. It follows from Lemma 1.5 that a map $f\colon \mathbb{R}P^n \to Z$ from $\mathbb{R}P^n$ to some topological space Z is continuous if and only if the composition map $f \circ p\colon S^n \to Z$ is continuous. Moreover two continuous maps f_0 and f_1 from $\mathbb{R}P^n$ to Z are homotopic if and only if there exists a continuous map $G\colon S^n \times [0,1] \to Z$ such that

$$G(\mathbf{x}, 0) = f_0(p(\mathbf{x})) \text{ and } G(\mathbf{x}, 1) = f_1(p(\mathbf{x})) \text{ for all } \mathbf{x} \in S^n,$$

$$G(\mathbf{x}, t) = G(-\mathbf{x}, t) \text{ for all } \mathbf{x} \in S^n \text{ and } t \in [0,1].$$

Remark Let X, Y and T be topological spaces, and let $q\colon X \to Y$ be an identification map. Let $q \times 1_T\colon X \times T \to Y \times T$ be the map defined by $(q \times 1_T)(x,t) = (q(x), t)$. The map $q \times 1_T$ is not in general an identification map. However if the topological space T is both locally compact and Hausdorff then it is possible to prove that the map $q \times 1_T\colon X \times T \to Y \times T$ is an identification map. (A topological space T is said to be *locally compact* if and only if every point of T has an open neighbourhood whose closure is compact.) This result generalizes Theorem 1.6.

1.2 Homotopies of Maps onto Product Spaces

Let $X_1, X_2, \ldots, X_r$ be a finite collection of topological spaces. The *product topology* on the Cartesian product $X_1 \times X_2 \times \cdots \times X_r$ of $X_1, X_2, \ldots, X_r$ is characterized by the following property:

> if we are given a topological space Z and continuous maps $f_j\colon Z \to X_j$ for $j = 1, 2, \ldots, r$, then the map $f\colon Z \to X_1 \times X_2 \times \cdots \times X_r$ defined by $f(z) = (f_1(z), f_2(z), \ldots, f_r(z))$ is a continuous map from Z to $X_1 \times X_2 \times \cdots \times X_r$.

We use this characterization of the product topology on $X_1 \times X_2 \times \cdots \times X_r$ in the proof of the following result.

Lemma 1.8 *Let $X_1, X_2, \ldots, X_r$ be a finite collection of topological spaces, let Z be a topological space, and let $f_j\colon Z \to X_j$ and $f'_j\colon Z\colon X_j$ be continuous* *maps from Z into X_j, for $j = 1, 2, \ldots, r$. Let $X = X_1 \times X_2 \times \cdots \times X_r$, and let $f\colon Z \to X$ and $f'\colon Z \to X$ be the continuous maps from Z to X defined by*

$$f(z) = (f_1(z), f_2(z), \ldots, f_r(z)), \qquad f'(z) = (f'_1(z), f'_2(z), \ldots, f'_r(z))$$

for all $z \in Z$. Then

$$f \simeq f' \text{ if and only if } f_j \simeq f'_j \text{ for } j = 1, 2, \ldots, r.$$

Moreover if A is a subset of Z with the property that $f_j|A = f'_j|A$ for $j = 1, 2, \ldots, r$ then

$$f \simeq f' \operatorname{rel} A \text{ if and only if } f_j \simeq f'_j \operatorname{rel} A \text{ for } j = 1, 2, \ldots, r.$$

Proof Let $H_j\colon Z \times [0,1] \to X_j$ be a continuous homotopy between f_j and f'_j for $j = 1, 2, \ldots, r$ (so that $H(z,0) = f(z)$ and $H(z,1) = f'(z)$ for all $z \in Z$). Let $H\colon Z \times [0,1] \to X$ be the continuous map defined by

$$H(z,t) = (H_1(z,t), H_2(z,t), \ldots, H_r(z,t)) \qquad (z \in Z).$$

Then $H(z,0) = f(z)$ and $H(z,1) = f'(z)$ for all $z \in Z$, so that H is a homotopy between the maps f and f'. Conversely if $H\colon Z \times [0,1] \to X$ is a continuous homotopy between f and f' (so that $H(z,0) = f(z)$ and $H(z,1) = f'(z)$ for all $z \in Z$), then the jth component $H_j\colon Z \times [0,1] \to X_j$ of the map H is a homotopy between f_j and f'_j for each j.

Let A be a subset of Z. Suppose that $f_j|A = f'_j|A$ for $j = 1, 2, \ldots, r$. Then $f|A = f'|A$. Moreover $H(a,t) = f(a) = f'(a)$ if and only if $H_j(a,t) = f_j(a) = f'_j(a)$ for all j. Thus $f \simeq f' \operatorname{rel} A$ if and only if $f_j \simeq f'_j \operatorname{rel} A$ for all j, as required. ∎

1.3 Homotopy Equivalences

Definition Let X and Y be topological spaces. A continuous map $f\colon X \to Y$ from X to Y is said to be a *homotopy equivalence* if there exists a continuous map $g\colon Y \to X$ such that $g \circ f$ and $f \circ g$ are homotopic to the identity maps of X and Y respectively. If there exists a homotopy equivalence $f\colon X \to Y$ then the spaces X and Y are said to be *homotopy equivalent*.

Definition Let X be a topological space and let A be a subset of X. Let $i\colon A \to X$ denote the inclusion map of A in X. A continuous map $r\colon X \to A$ with the property that $r|A$ is the identity map of A is said to be a *retraction* of the space X onto the subset A. The retraction $r\colon X \to A$ is said to be a *deformation retraction* if $i \circ r \simeq 1_X$, where 1_X is the identity map of X. In this case A is said to be a *deformation retract of* X. If in addition $i \circ r \simeq 1_X \operatorname{rel} A$, then $r\colon X \to A$ is said to be a *strong deformation retraction* and A is said to be a *strong deformation retract* of X.

Note that if A is a deformation retract of some topological space X then the inclusion map $A \hookrightarrow X$ is a homotopy equivalence.

Example Let S^n denote the unit sphere in $\mathbb{R}^{n+1}$ given by

$$S^n = \{\mathbf{x} \in \mathbb{R}^{n+1} : |\mathbf{x}| = 1\}$$

Then S^n is a strong deformation retract of the following subsets of $\mathbb{R}^{n+1}$:

(i) $\{\mathbf{x} \in \mathbb{R}^{n+1} : |\mathbf{x}| > 0\}$,

(ii) $\{\mathbf{x} \in \mathbb{R}^{n+1} : \frac{1}{2} < |\mathbf{x}| < 2\}$,

(iii) $\{\mathbf{x} \in \mathbb{R}^{n+1} : 0 < |\mathbf{x}| \leq 1\}$,

(iv) $\{\mathbf{x} \in \mathbb{R}^{n+1} : |\mathbf{x}| \geq 1\}$.

However it is possible to show that S^n is not a retract of $\mathbb{R}^{n+1}$ itself.

2 Paths, Loops and the Fundamental Group

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

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

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

Similarly suppose that $x_0, x_1, \ldots, x_r$ are points in the space X and that, for each integer j between 1 and r, we are given a path $\gamma_j\colon [0,1] \to X$ from the point x_{j-1} to x_j. We define $\gamma_1.\gamma_2.\cdots.\gamma_r\colon [0,1] \to X$ to be the path in X from x_0 to x_r defined such that

$$(\gamma_1.\gamma_2.\cdots.\gamma_r)(t) = \gamma_j(rt - j + 1) \text{ for all } t \text{ satisfying } \frac{j-1}{r} \le t \le \frac{j}{r}.$$

Thus the restriction of the path $\gamma_1.\gamma_2.\cdots.\gamma_r$ to the interval $[(j-1)/r, j/r]$ is a reparameterization of the path γ_r from x_{j-1} to x_j: the path $\gamma_1.\gamma_2.\cdots.\gamma_r$ thus represents the path from x_0 to x_r obtained by travelling in succession along the paths $\gamma_1, \gamma_2, \ldots, \gamma_r$. (A straightforward application of Lemma A.7 of Appendix A shows that the path $\gamma_1.\gamma_2.\cdots.\gamma_r$ is indeed continuous.)

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 to the point y then γ^{-1} is the path from y to x obtained by reversing the direction in which the path γ is traversed.)

Lemma 2.1 *Let X be a topological space, and let $\gamma\colon [0,1] \to X$ be a path in X. Then*

$$\varepsilon_0.\gamma \simeq \gamma \simeq \gamma.\varepsilon_1 \text{ rel } \{0,1\}, \qquad \gamma.\gamma^{-1} \simeq \varepsilon_0 \text{ rel } \{0,1\},$$

where ε_0 and ε_1 denote the constant loops at the points $\gamma(0)$ and $\gamma(1)$ respectively (i.e., $\varepsilon_0(t) = \gamma(0)$ and $\varepsilon_1(t) = \gamma(1)$ for all $t \in [0,1]$). Also if γ_1, γ_2 and γ_3 are paths in X for which $\gamma_1(1) = \gamma_2(0)$ and $\gamma_2(1) = \gamma_3(0)$ then

$$(\gamma_1.\gamma_2).\gamma_3 \simeq \gamma_1.(\gamma_2.\gamma_3) \text{ rel } \{0,1\}.$$

Proof First we show that $\varepsilon_0.\gamma \simeq \gamma \simeq \gamma.\varepsilon_1$ rel $\{0,1\}$. Let $F_1\colon [0,1]\times[0,1] \to X$ and $F_2\colon [0,1]\times[0,1] \to X$ be the homotopies defined by

$$F_1(t,\tau) = \begin{cases} \gamma(0) & \text{if } 0 \le t \le \frac{1}{2}(1-\tau); \\ \gamma\left(\dfrac{2t+\tau-1}{1+\tau}\right) & \text{if } \frac{1}{2}(1-\tau) \le t \le 1; \end{cases}$$

$$F_2(t,\tau) = \begin{cases} \gamma\left(\dfrac{2t}{1+\tau}\right) & \text{if } 0 \le t \le \frac{1}{2}(1+\tau); \\ \gamma(1) & \text{if } \frac{1}{2}(1+\tau) \le t \le 1. \end{cases}$$

Then

$$\begin{aligned} F_1(t,0) &= (\varepsilon_0.\gamma)(t), \qquad F_1(t,1) = \gamma(t) \qquad (t \in [0,1]), \\ F_2(t,0) &= (\gamma.\varepsilon_1)(t), \qquad F_2(t,1) = \gamma(t) \qquad (t \in [0,1]), \end{aligned}$$

$$F_1(0,\tau) = F_2(0,\tau) = \gamma(0), \qquad F_1(1,\tau) = F_2(1,\tau) = \gamma(1) \qquad (\tau \in [0,1]).$$

Thus F_1 provides the required homotopy between the paths $\varepsilon_0.\gamma$ and γ, and F_2 provides the required homotopy between the paths $\gamma.\varepsilon_1$ and γ.

Next we show that $\gamma\gamma^{-1} \simeq \varepsilon_0$ rel $\{0,1\}$. Let $G\colon [0,1]\times[0,1] \to X$ be the homotopy defined by

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

Note that

$$G(t,0) = \gamma(0) = \varepsilon_0(t) \text{ and } G(t,1) = (\gamma.\gamma^{-1})(t) \text{ for all } t \in [0,1],$$

$$G(0,\tau) = \gamma(0) = G(1,\tau) \text{ for all } \tau \in [0,1].$$

Thus G provides the required homotopy between the paths $\gamma.\gamma^{-1}$ and ε_0.

Finally we show that if γ_1, γ_2 and γ_3 are paths in X for which $\gamma_1(1) = \gamma_2(0)$ and $\gamma_2(1) = \gamma_3(0)$ then $(\gamma_1.\gamma_2).\gamma_3 \simeq \gamma_1.(\gamma_2.\gamma_3)$ rel $\{0,1\}$. Let $H\colon [0,1]\times [0,1] \to X$ be the homotopy defined by

$$H(t,\tau) = \begin{cases} \gamma_1\left(\dfrac{4t}{\tau+1}\right) & \text{if } 0 \le t \le \frac{1}{4}(\tau+1); \\ \gamma_2(4t-\tau-1) & \text{if } \frac{1}{4}(\tau+1) \le t \le \frac{1}{4}(\tau+2); \\ \gamma_3\left(\dfrac{4t-\tau-2}{2-\tau}\right) & \text{if } \frac{1}{4}(\tau+2) \le t \le 1. \end{cases}$$

Then

$$H(t,0) = ((\gamma_1.\gamma_2).\gamma_3)(t), \qquad H(t,1) = (\gamma_1.(\gamma_2.\gamma_3))(t) \qquad (t \in [0,1]),$$

$$H(0,\tau) = \gamma_1(0), \qquad H(1,\tau) = \gamma_3(1) \qquad (\tau \in [0,1]).$$

Thus H provides the required homotopy between the paths $(\gamma_1.\gamma_2).\gamma_3$ and $\gamma_1.(\gamma_2.\gamma_3)$. ∎

Let X be a topological space, and let $x_0 \in X$ be some chosen basepoint for X. We define an equivalence relation on the set of all (continuous) loops based at the point 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.

We denote by $\pi_1(X, x_0)$ the set of all based homotopy classes of loops based at the point x_0. Thus every element $[\gamma]$ of $\pi_1(X, x_0)$ is represented by some loop $\gamma\colon [0,1] \to X$ based at the point x_0, and two such loops $\gamma_1\colon [0,1] \to X$ and $\gamma_2\colon [0,1] \to X$ based at x_0 represent the same element of $\pi_1(X, x_0)$ if and only if $\gamma_1 \simeq \gamma_2$ rel $\{0,1\}$.

We now describe how one can define a group structure on $\pi_1(X, x_0)$. Let $\gamma_1\colon [0,1] \to X$ and $\gamma_2\colon [0,1] \to X$ be loops based at the point x_0. We define the product $[\gamma_1][\gamma_2]$ of the based homotopy classes $[\gamma_1]$ and $[\gamma_2]$ to be the based homotopy class $[\gamma_1.\gamma_2]$ of the product loop $\gamma_1.\gamma_2$ obtained by concatenating the loops γ_1 and γ_2. We now show that this group operation is indeed well-defined and satisfies the group axioms.

Lemma 2.2 *Let X be a topological space, let x_0 be some chosen point of X, and let $\pi_1(X, x_0)$ be the set of all based homotopy classes of loops based at the point x_0. Then $\pi_1(X, x_0)$ is a group, where the group multiplication on $\pi_1(X, x_0)$ is defined according to the rule $[\gamma_1][\gamma_2] = [\gamma_1.\gamma_2]$, for all loops $\gamma_1\colon [0,1] \to X$ and $\gamma_2\colon [0,1] \to X$ 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 $[\gamma_2] = [\gamma_2']$. We must show that $[\gamma_1.\gamma_2] = [\gamma_1'.\gamma_2']$. Now there exist homotopies $H_1\colon [0,1] \times [0,1] \to X$ and $H_2\colon [0,1] \times [0,1] \to X$ such that $H_1(t,0) = \gamma_1(t)$, $H_1(t,1) = \gamma_1'(t)$, $H_2(t,0) = \gamma_2(t)$ and $H_2(t,1) = \gamma_2'(t)$ for all $t \in [0,1]$, and

$$H_1(0,\tau) = H_1(1,\tau) = H_2(0,\tau) = H_2(1,\tau) = x_0$$

for all $\tau \in [0,1]$. Let $H\colon [0,1] \times [0,1] \to X$ be the continuous map defined by

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

Then $H(t,0) = (\gamma_1.\gamma_2)(t)$ and $H(t,1) = (\gamma_1'.\gamma_2')(t)$ for all $t \in [0,1]$, and also $H(0,\tau) = H(1,\tau) = x_0$ for all $\tau \in [0,1]$. Thus $[\gamma_1.\gamma_2] = [\gamma_1'\gamma_2']$. This shows that the group operation on $\pi_1(X,x_0)$ is well-defined.

If γ_1, γ_2 and γ_3 are loops in X based at the point x_0 then $(\gamma_1.\gamma_2).\gamma_3 \simeq \gamma_1.(\gamma_2.\gamma_3)$ rel $\{0,1\}$ by Lemma 2.1, and hence $([\gamma_1][\gamma_2])[\gamma_3] = [\gamma_1]([\gamma_2][\gamma_3])$. Thus the group operation on $\pi_1(X,x_0)$ is associative.

Let ε_0 denote the constant loop at the basepoint x_0 (i.e., $\varepsilon_0(t) = x_0$ for all $t \in [0,1]$). If γ is a loop in X based at x_0 then $\varepsilon_0.\gamma \simeq \gamma \simeq \gamma.\varepsilon_0$ rel $\{0,1\}$ by Lemma 2.1, and hence $[\varepsilon_0][\gamma] = [\gamma] = [\gamma][\varepsilon_0]$. This shows that $[\varepsilon_0]$ is the identity element of $\pi_1(X,x_0)$.

If $\gamma\colon [0,1] \to X$ is a loop based at x_0 then $\gamma.\gamma^{-1} \simeq \varepsilon_0 \simeq \gamma^{-1}.\gamma$ rel $\{0,1\}$, by Lemma 2.1. Therefore the based homotopy class $[\gamma^{-1}]$ of the loop γ^{-1} is the inverse of the based homotopy class $[\gamma]$ of the loop γ with respect to the group multiplication operation defined on $\pi_1(X,x_0)$. Thus the given group multiplication operation on $\pi_1(X,x_0)$ satisfies all of the the group axioms, as required. ∎

Let X be a topological space and let $x_0 \in X$ be some chosen basepoint for X. The group $\pi_1(X,x_0)$ is referred to as the *fundamental group* of the topological space X based at the point x_0. This group is a topological invariant: if X and Y are topological spaces and if $h\colon X \to Y$ is a homeomorphism then $\pi(X,x_0) \cong \pi(Y,h(x_0))$ for all $x_0 \in X$.

Remark Let S^1 be the standard circle and let $b \in S^1$ be some chosen basepoint. Every loop on the topological space corresponds to some continuous map from S^1 to X. If x_0 is some chosen basepoint in the space X then every loop in X based at x_0 corresponds to some continuous map from S^1 to X which maps the chosen basepoint b of S^1 to x_0. Indeed suppose that we represent S^1 as the unit circle in $\mathbb{R}^2$ and suppose that we choose the point $(1,0)$ on this unit circle to be the basepoint of S^1. If $\gamma\colon [0,1] \to X$ is a (continuous) loop based at the point x_0 of X then the corresponding continuous map from S^1 to X is the map $\hat\gamma\colon S^1 \to X$ defined such that

$$\hat\gamma(\cos 2\pi t, \sin 2\pi t) = \gamma(t)$$

for all $t \in [0,1]$. Observe that $\hat\gamma(1,0) = x_0$ (since $\gamma(0) = x_0 = \gamma(1)$).

Let $\gamma_0\colon [0,1] \to X$ and $\gamma_1\colon [0,1] \to X$ be loops in X based at x_0, and let $\hat\gamma_0\colon S^1 \to X$ and $\hat\gamma_1\colon S^1 \to X$ be the corresponding continuous maps from S^1 to X (where $\hat\gamma_0$ and $\hat\gamma_1$ are determined by the loops γ_0 and γ_1 in the manner described above). Then the loops γ_0 and γ_1 represent the same element of $\pi_1(X,x_0)$ if and only if the maps $\hat\gamma_0\colon S^1 \to X$ and $\hat\gamma_1\colon S^1 \to X$ are homotopic by a basepoint-preserving homotopy (i.e., if and only if there exists a continuous map $H: S^1 \times [0,1] \to X$ such that $H((1,0),t) = x_0$ for all $t \in [0,1]$, and $H(u,0) = \hat{\gamma}_0(u)$ and $H(u,1) = \hat{\gamma}_1(u)$ for all $u \in S^1$). We therefore have a natural identification of the fundamental group $\pi_1(X, x_0)$ of the topological space X at the basepoint x_0 with the set of all based homotopy classes of basepoint-preserving maps from S^1 to X.

Let X be a topological space and let $\alpha: [0,1] \to X$ be a continuous path in X. Let $x = \alpha(0)$ and $y = \alpha(1)$. Let $\tau_\alpha: \pi_1(X, y) \to \pi_1(X, x)$ be the function defined by $\tau_\alpha([\gamma]) = [\alpha.\gamma.\alpha^{-1}]$, for all loops $\gamma: [0,1] \to X$ based at y, where $[\alpha.\gamma.\alpha^{-1}] \in \pi_1(X, x)$ denotes the based homotopy class of the loop $\alpha.\gamma.\alpha^{-1}$ based at x given by

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

(Thus $\alpha.\gamma.\alpha^{-1}$ is the path obtained by following the path α from x to y, then following round the loop γ based at y, and then returning back to x along the path α in the reverse direction.)

Lemma 2.3 *Let X be a topological space and let $\alpha: [0,1] \to X$ be a continuous path in X. Let $x = \alpha(0)$ and $y = \alpha(1)$, and let $\tau_\alpha: \pi_1(X, y) \to \pi_1(X, x)$ be the function defined by $\tau_\alpha([\gamma]) = [\alpha.\gamma.\alpha^{-1}]$. Then τ_α is an isomorphism from $\pi_1(X, y)$ to $\pi_1(X, x)$.*

Proof First we show that τ_α is a homomorphism of groups. Let γ_1 and γ_2 be loops based at the point y. Then

$$\begin{aligned} \tau_\alpha([\gamma_1][\gamma_2]) &= \tau_\alpha([\gamma_1.\gamma_2]) = [\alpha.(\gamma_1.\gamma_2).\alpha^{-1}], \\ \tau_\alpha([\gamma_1])\tau_\alpha([\gamma_2]) &= [\alpha.\gamma_1.\alpha^{-1}.\alpha.\gamma_2.\alpha^{-1}], \end{aligned}$$

where

$$(\alpha.(\gamma_1.\gamma_2).\alpha^{-1})(t) = \begin{cases} \alpha(3t) & \text{if } 0 \leq t \leq \frac{1}{3}; \\ \gamma_1(6t-2) & \text{if } \frac{1}{3} \leq t \leq \frac{1}{2}; \\ \gamma_2(6t-3) & \text{if } \frac{1}{2} \leq t \leq \frac{2}{3}; \\ \alpha(3-3t) & \text{if } \frac{2}{3} \leq t \leq 1; \end{cases}$$

and

$$(\alpha.\gamma_1.\alpha^{-1}.\alpha.\gamma_2.\alpha^{-1})(t) = \begin{cases} \alpha(6t) & \text{if } 0 \leq t \leq \frac{1}{6}; \\ \gamma_1(6t-1) & \text{if } \frac{1}{6} \leq t \leq \frac{1}{3}; \\ \alpha(3-6t) & \text{if } \frac{1}{3} \leq t \leq \frac{1}{2}; \\ \alpha(6t-3) & \text{if } \frac{1}{2} \leq t \leq \frac{2}{3}; \\ \gamma_2(6t-4) & \text{if } \frac{2}{3} \leq t \leq \frac{5}{6}; \\ \alpha(6-6t) & \text{if } \frac{5}{6} \leq t \leq 1. \end{cases}$$

However
$$\alpha.\gamma_1.\alpha^{-1}.\alpha.\gamma_2.\alpha^{-1} \simeq \alpha.(\gamma_1.\gamma_2).\alpha^{-1} \text{ rel } \{0,1\}.$$
Indeed if $H: [0,1] \times [0,1] \to X$ is the continuous map defined by
$$H(t,\tau) = \begin{cases} \alpha\left(\dfrac{6t}{\tau+1}\right) & \text{if } 0 \leq t \leq \frac{1}{6}(\tau+1), \\ \gamma_1(6t-\tau-1) & \text{if } \frac{1}{6}(\tau+1) \leq t \leq \frac{1}{6}(\tau+2), \\ \alpha(3+\tau-6t) & \text{if } \frac{1}{6}(\tau+2) \leq t \leq \frac{1}{2}, \\ \alpha(6t-3+\tau) & \text{if } \frac{1}{2} \leq t \leq \frac{1}{6}(4-\tau), \\ \gamma_2(6t-4+\tau) & \text{if } \frac{1}{6}(4-\tau) \leq t \leq \frac{1}{6}(5-\tau), \\ \alpha\left(\dfrac{6-6t}{\tau+1}\right) & \text{if } \frac{1}{6}(5-\tau) \leq t \leq 1, \end{cases}$$
then
$$H(t,0) = (\alpha.\gamma_1.\alpha^{-1}.\alpha.\gamma_2.\alpha^{-1})(t) \quad \text{and} \quad H(t,1) = (\alpha.(\gamma_1.\gamma_2).\alpha^{-1})(t)$$
for all $t \in [0,1]$, and
$$H(0,\tau) = x = H(1,\tau)$$
for all $\tau \in [0,1]$, and thus H provides the required homotopy between the loops $\alpha.\gamma_1.\alpha^{-1}.\alpha.\gamma_2.\alpha^{-1}$ and $\alpha.(\gamma_1.\gamma_2).\alpha^{-1}$. This shows that $\tau_\alpha([\gamma_1][\gamma_2]) = \tau_\alpha([\gamma_1])\tau_\alpha([\gamma_2])$, so that $\tau_\alpha: \pi_1(X,y) \to \pi_1(X,x)$ is a homomorphism. Moreover τ_α is invertible. Indeed the inverse of τ_α is given by
$$\tau_\alpha^{-1}([\sigma]) = [\alpha^{-1}.\sigma.\alpha]$$
for all loops σ based at the point x. Thus $\tau_\alpha: \pi_1(X,y) \to \pi_1(X,x)$ is an isomorphism. ∎

Definition A topological space X is said to be *path-connected* if and only if, given any two points x_0 and x_1 of X, there exists a continuous path $\alpha: [0,1] \to X$ such that $\alpha(0) = x_0$ and $\alpha(1) = x_1$.

We deduce immediately from Lemma 2.3 the following result.

Corollary 2.4 *Let X be a path-connected topological space, and let x_0 and x_1 be points of X. Then $\pi_1(X,x_0) \cong \pi_1(X,x_1)$.*

Definition A topological space X is said to be *simply-connected* if X is path-connected and $\pi_1(X,x_0)$ is the trivial group (where x_0 is some chosen basepoint for X).

We deduce immediately from Corollary 2.4 that if the topological space X is simply connected then $\pi_1(X, x)$ is trivial for all $x \in X$.

Lemma 2.5 *Let X be a topological space, and let $\gamma_1\colon [0,1] \to X$ and $\gamma_2\colon [0,1] \to X$ be paths in X, where $\gamma_1(0) = \gamma_2(0)$ and $\gamma_1(1) = \gamma_2(1)$. Suppose that the topological space X is simply-connected. Then*

$$\gamma_1 \simeq \gamma_2 \text{ rel } \{0, 1\}.$$

Proof Let $x_0 = \gamma_1(0) = \gamma_2(0)$. Using Lemma 2.1 we see that

$$\gamma_1 \simeq \gamma_1.(\gamma_2^{-1}.\gamma_2) \simeq (\gamma_1.\gamma_2^{-1}).\gamma_2 \text{ rel } \{0, 1\}.$$

But the fundamental group $\pi_1(X, x_0)$ of X at x_0 is trivial, since the topological space X is simply-connected. Therefore

$$\gamma_1.\gamma_2^{-1} \simeq \varepsilon_0 \text{ rel } \{0, 1\},$$

where ε_0 denotes the constant loop at the point x_0. It follows from Lemma 2.1 that

$$\gamma_1 \simeq \varepsilon_0.\gamma_2 \simeq \gamma_2 \text{ rel } \{0, 1\},$$

as required. ∎

Let X and Y be topological spaces, and let x_0 be a point of X. Let $f\colon X \to Y$ be a continuous map. Then f induces a homomorphism $f_\#\colon \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ defined by $f_\#([\gamma]) = [f \circ \gamma]$ for all loops $\gamma\colon [0,1] \to X$ based at x_0.

Lemma 2.6 *Let X, Y and Z be topological spaces, and let x_0, y_0 and z_0 be chosen basepoints of X, Y and Z respectively. Then*

(i) *if $1_X\colon X \to X$ is the identity map of X then $1_{X\#}\colon \pi_1(X, x_0) \to \pi_1(X, x_0)$ is the identity homomorphism of the fundamental group $\pi_1(X, x_0)$ of X,*

(ii) *if $f\colon X \to Y$ and $h\colon Y \to Z$ are continuous maps, and if $f(x_0) = y_0$ and $h(y_0) = z_0$, then $(h \circ f)_\# = h_\# \circ f_\#$,*

(iii) *if $f\colon X \to Y$ and $g\colon X \to Y$ are continuous maps for which $f(x_0) = g(x_0)$, and if $f \simeq g$ rel $\{x_0\}$, then $f_\# = g_\#$.*

We now show that the fundamental group of a product of topological spaces is isomorphic to the direct product of the fundamental groups of those spaces.

Lemma 2.7 *Let $X_1, X_2, \ldots, X_r$ be topological spaces, and let x_j be some chosen basepoint of the space X_j for each integer j between 1 and r. Let x be the point of $X_1 \times X_2 \times \cdots \times X_r$ given by $x = (x_1, x_2, \ldots, x_r)$. Then*

$$\pi_1(X_1 \times X_2 \times \cdots \times X_r, x) \cong \pi_1(X_1, x_1) \times \pi_1(X_2, x_2) \times \cdots \times \pi_1(X_r, x_r).$$

Proof Each element of $\pi_1(X_1 \times X_2 \times \cdots \times X_r, x)$ is represented by a loop of the form

$$t \mapsto (\gamma_1(t), \gamma_2(t), \ldots, \gamma_r(t)),$$

where γ_j is a loop in X_j based at x_j, for $j = 1, 2, \ldots, r$. Moreover it follows from Lemma 1.8 that if γ_j and γ'_j are loops in X_j, for $j = 1, 2, \ldots, r$, and if γ and γ' are the loops in $X_1 \times X_2 \times \cdots \times X_r$ given by

$$\gamma(t) = (\gamma_1(t), \gamma_2(t), \ldots, \gamma_r(t)), \qquad \gamma'(t) = (\gamma'_1(t), \gamma'_2(t), \ldots, \gamma'_r(t)),$$

then $\gamma \simeq \gamma'$ rel $\{0, 1\}$ if and only if $\gamma_j \simeq \gamma'_j$ rel $\{0, 1\}$ for all j. The required result follows easily from these observations. ∎

2.1 The Fundamental Group of a Union of Simply-Connected Spaces

Theorem 2.8 *Let X be a topological space, let x_0 be a point of X, and let U and V be open sets in X containing the point x_0 such that $X = U \cup V$. Suppose that the open sets U and V are simply-connected and that their intersection $U \cap V$ is path-connected. Then the topological space X is simply-connected.*

Proof If x is a point of X then x belongs either to U or to V. Therefore there exists a continuous path in X joining x to x_0 since the sets U and V are both path-connected. Thus X is path-connected. We must prove that the fundamental group $\pi_1(X, x_0)$ of the space X is trivial.

Let $\gamma: [0, 1] \to X$ be a loop in X based at the point x_0. A straightforward application of the Lebesgue Lemma shows that there exist real numbers s_j for $j = 0, 1, 2, \ldots, r$, where

$$0 = s_0 < s_1 < s_2 < \cdots < s_r = 1,$$

such that, for each integer j between 1 and r, either $\gamma([s_{j-1}, s_j]) \subset U$ or else $\gamma([s_{j-1}, s_j]) \subset V$ (see Theorem A.29 of Appendix A).

The sets U, V and $U \cap V$ are all path-connected. Thus we may choose, for $j = 0, 1, \ldots, r$ a path $\eta_j: [0, 1] \to X$ from x_0 to $\gamma(s_j)$, where the choice of η_j is made subject to the following conditions:

(i) $\eta_0(t) = x_0 = \eta_r(t)$ for all $t \in [0,1]$,

(ii) if $\gamma(s_j) \in U$ then $\eta_j([0,1]) \subset U$,

(iii) if $\gamma(s_j) \in V$ then $\eta_j([0,1]) \subset V$,

(Thus if $\gamma(s_j) \in U \cap V$ then the path η_j must be chosen such that $\eta_j([0,1]) \subset U \cap V$, in order that both *(ii)* and *(iii)* are satisfied.)

For each integer j between 1 and r let $\alpha_j\colon [0,1] \to X$ denote the loop in X based at x_0 defined by

$$\alpha_j(t) = \begin{cases} \eta_{j-1}(3t) & \text{if } 0 \leq t \leq \frac{1}{3}, \\ \gamma((2-3t)s_{j-1} + (3t-1)s_j) & \text{if } \frac{1}{3} \leq t \leq \frac{2}{3}, \\ \eta_j(3-3t) & \text{if } \frac{2}{3} \leq t \leq 1, \end{cases}$$

However the loop β is a reparameterization of the loop γ, and therefore β and γ represent the same element of $\pi_1(X, x_0)$. We conclude therefore that the loop γ represents the identity element of $\pi_1(X, x_0)$.

(It is not difficult to construct an explicit homotopy between the loops β and γ. Indeed let $\mu\colon [0,1] \to [0,1]$ be the homeomorphism of the interval $[0,1]$ defined by

$$\mu(t) = (j - rt)s_{j-1} + (rt - j + 1)s_j \text{ for all } t \text{ satisfying } \frac{j-1}{r} \leq t \leq \frac{j}{r}.$$

Then $\beta = \gamma \circ \mu$. Thus the map

$$(t, \tau) \mapsto \gamma((1-\tau)\mu(t) + \tau t)$$

provides the required homotopy between the loops γ and β.)

We have shown that the topological space X is path-connected and that every loop in X based at the point x_0 represents the identity element of $\pi_1(X, x_0)$. Thus $\pi_1(X, x_0)$ is the trivial group. We conclude that X is simply-connected, as required. ∎

Given a non-negative integer n, the *n-sphere* S^n is the topological space given by

$$S^n = \{(x_1, x_2, \ldots, x_{n+1} \in \mathbb{R}^{n+1} : x_1^2 + x_2^2 + \cdots + x_{n+1}^2 = 1\}.$$

If we apply Theorem 2.8 we may deduce the following result.

Corollary 2.9 *The n-sphere S^n is simply-connected for all integers n satisfying $n \geq 2$.*

Proof Let U be the subset of S^n consisting of those points $(x_1, x_2, \ldots, x_{n+1})$ of S^n for which $x_{n+1} > -\frac{1}{2}$, and let V be the subset of S^n consisting of those points $(x_1, x_2, \ldots, x_{n+1})$ of S^n for which $x_{n+1} < \frac{1}{2}$. The sets U and V are open sets with respect to the topology of S^n, and $S^n = U \cup V$. If $n \geq 2$ then the sets U and V are simply-connected and their intersection $U \cap V$ is path-connected. Therefore S^n is simply connected, by Theorem 2.8. ∎

2.2 Homotopy Equivalences and the Fundamental Group

Let X and Y be topological spaces and let $f: X \to Y$ be a continuous map. We recall that the map f is said to be a *homotopy equivalence* if and only if there exists a continuous map $g: Y \to X$ such that $g \circ f$ is homotopic to the identity map of X and $f \circ g$ is homotopic to the identity map of Y. In particular, if X is a topological space, and if A is a subset of X which is a deformation retract of X, then the inclusion map $A \hookrightarrow X$ is a homotopy equivalence. We now prove that a homotopy equivalence between two topological spaces induces an isomorphism between the fundamental groups of those topological spaces.

Theorem 2.10 *Let X and Y be topological spaces and let $f: X \to Y$ be a homotopy equivalence. Then for any given basepoint $x_0 \in X$ the induced homomorphism $f_\#: \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ is an isomorphism.*

Proof Let $g: Y \to X$ be a continuous map such that $g \circ f$ and $f \circ g$ are homotopic to the identity maps of X and Y respectively. (Such a map g exists since f is a homotopy equivalence.)

In the case when $g(f(x_0)) = x_0$ and all homotopies are basepoint-preserving, we see that $g \circ f$ and $f \circ g$ are the identity homomorphisms of $\pi_1(X, x_0)$ and $\pi_1(Y, f(x_0))$ respectively. (This follows directly on applying the results of Lemma 2.6.) Thus $g_\# = f_\#^{-1}$, so that $f_\#$ is an isomorphism from $\pi_1(X, x_0)$ to $\pi_1(Y, f(x_0))$.

In the general case we proceed as follows. Define $y_0 = f(x_0)$ and $x_1 = g(y_0)$. We show that the composition $g_\# \circ f_\#: \pi_1(X, x_0) \to \pi_1(X, x_1)$ of the induced homomorphisms $f_\#: \pi_1(X, x_0) \to \pi_1(Y, y_0)$ and $g_\#: \pi_1(Y, y_0) \to \pi_1(X, x_1)$ is an isomorphism. There exists a continuous homotopy $H: X \times [0, 1] \to X$ such that $H(x, 0) = g(f(x))$ and $H(x, 1) = x$ for all $x \in X$ (since $g \circ f$ is homotopic to the identity map of X). Define $\alpha: [0, 1] \to X$ by $\alpha(t) = H(x_0, t)$ for all $t \in [0, 1]$. Then α is a path from x_1 to x_0. We claim that if $\gamma: [0, 1] \to X$ is a loop based at the basepoint x_0 (so that $\gamma(0) = x_0$ and $\gamma(1) = x_0$) then $g_\#(f_\#([\gamma])) = \tau_\alpha([\gamma])$, where $\tau_\alpha: \pi_1(X, x_1) \to \pi_1(X, x_0)$ is the isomorphism defined by $\tau_\alpha([\gamma]) = [\alpha.\gamma.\alpha^{-1}]$ (see Lemma 2.3).

Consider the map $F: [0,1] \times [0,1] \to X$ defined by

$$F(t,\tau) = \begin{cases} \alpha(3t) & \text{if } 0 \le t \le \frac{1}{3}\tau; \\ H\left(\gamma\left(\dfrac{3t-\tau}{3-2\tau}\right), \tau\right) & \text{if } \frac{1}{3}\tau \le t \le 1 - \frac{1}{3}\tau; \\ \alpha(3-3t) & \text{if } 1 - \frac{1}{3}\tau \le t \le 1. \end{cases}$$

(A straightforward application of Lemma A.7 of Appendix A shows that the map H is well-defined and continuous, since

$$H\left(\gamma\left(\frac{3t-\tau}{3-2\tau}\right), \tau\right) = H(\gamma(0), \tau) = H(x_0, \tau) = \alpha(3t)$$

when $t = \frac{1}{3}\tau$, and

$$H\left(\gamma\left(\frac{3t-\tau}{3-2\tau}\right), \tau\right) = H(\gamma(1), \tau) = H(x_0, \tau) = \alpha(3-3t)$$

when $t = 1 - \frac{1}{3}\tau$.) Note that

$$F(t,0) = H(\gamma(t), 0) = g(f(\gamma(t))), \qquad F(t,1) = (\alpha.\gamma.\alpha^{-1})(t).$$

Therefore F is a homotopy between the loops $g \circ f \circ \gamma$ and $\alpha.\gamma.\alpha^{-1}$. Moreover $F(0,\tau) = \alpha(0) = x_1$ and $F(1,\tau) = \alpha(0) = x_1$ for all $\tau \in [0,1]$. Therefore $g \circ f \circ \gamma \simeq \alpha.\gamma.\alpha^{-1}$ rel $\{0,1\}$, and thus $g_\#(f_\#([\gamma])) = \tau_\alpha[\gamma]$. But $\tau_\alpha: \pi_1(X, x_0) \to \pi_1(X, x_1)$ is an isomorphism, by Lemma 2.3. Thus the composition $g_\# \circ f_\#$ of the homomorphisms $f_\#: \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ and $g_\#: \pi_1(Y, f(x_0)) \to \pi_1(X, x_1)$ is an isomorphism, and hence the homomorphism $f_\#: \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ is injective.

Let $\beta: [0,1] \to Y$ be a loop based at y_0. Then $g_\#([\beta]) = g_\#(f_\#([\gamma]))$ for some loop $\gamma: [0,1] \to X$ based at x_0 (since $g_\# \circ f_\#$ is an isomorphism). But we have just shown that the homomorphism of fundamental groups induced by any homotopy equivalence is injective. This implies that $g_\#: \pi_1(Y, y_0) \to \pi_1(X, g(y_0))$ is injective (since $g: Y \to X$ is a homotopy equivalence). Therefore $[\beta] = f_\#([\gamma])$. This shows that $f_\#: \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ is surjective. We conclude that $f_\#$ is an isomorphism, as required. ∎

Corollary 2.11 *Let X be a topological space, and let A be a deformation retract of X. Let x_0 be a point of A. Then the homomorphism $i_\#: \pi_1(A, x_0) \to \pi_1(X, x_0)$ of fundamental groups induced by the inclusion map $i \hookrightarrow A \to X$ is an isomorphism.*

Proof The inclusion map $i: A \hookrightarrow X$ is a homotopy equivalence, since A is a deformation retract of X. Hence the induced homomorphism $i_\#$ of fundamental groups is an isomorphism, by Theorem 2.10. ∎

Corollary 2.12 *Let X be a contractible topological space and let x_0 be a point of X. Then $\pi_1(X, x_0)$ is trivial.*

Proof If X is contractible then there exists a point p of X such that the set $\{p\}$ consisting of the single element p is a deformation retract of X. It follows from Corollary 2.12 that the fundamental group of the space X is isomorphic to the fundamental group of the topological space consisting of a single point. Thus $\pi_1(X, x_0)$ is trivial. ∎

Example The following spaces are contractible and thus have trivial fundamental group:

(i) n-dimensional Euclidean space $\mathbb{R}^n$,

(ii) the open unit ball B^n in $\mathbb{R}^n$ defined by

$$B^n = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_1^2 + x_2^2 + \cdots + x_n^2 < 1\},$$

(iii) the closed unit ball E^n in $\mathbb{R}^n$ defined by

$$E^n = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_1^2 + x_2^2 + \cdots + x_n^2 \leq 1\},$$

(iv) the open half-space H^n in $\mathbb{R}^n$ defined by

$$H^n = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_n > 0\}.$$

3 Covering Maps

Definition Let X and $\tilde{X}$ be topological spaces and let $p: \tilde{X} \to X$ be a continuous map. The map p is said to be a *local homeomorphism* if and only if every point of $\tilde{X}$ has an open neighbourhood which is mapped homeomorphically by p onto some open set in X.

Example Let S^1 denote the unit circle in $\mathbb{R}^2$, and let $\alpha: (-2, 2) \to S^1$ be the continuous map defined by $\alpha(t) = (\cos 2\pi t, \sin 2\pi t)$. Then the map α is a local homeomorphism from $(-2, 2)$ to S^1.

We shall be considering a particular class of local homeomorphisms known as *covering maps*.

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

Example Let $\alpha: (-2, 2) \to S^1$ be the continuous map defined by $\alpha(t) = (\cos 2\pi t, \sin 2\pi t)$. Let U and V be the open subsets of S^1 defined by

$$\begin{aligned} U &= \{(x, y) \in S^1 : x < 0 \text{ and } x^2 + y^2 = 1\}, \\ V &= \{(x, y) \in S^1 : x > 0 \text{ and } x^2 + y^2 = 1\}. \end{aligned}$$

The set U is evenly covered by the map α, since

$$\begin{aligned} \alpha^{-1}(U) &= \{t \in (-2, 2) : \cos 2\pi t < 0\} \\ &= (-\frac{7}{4}, -\frac{5}{4}) \cup (-\frac{3}{4}, -\frac{1}{4}) \cup (\frac{1}{4}, \frac{3}{4}) \cup (\frac{5}{4}, \frac{7}{4}), \end{aligned}$$

and each of the open intervals $(-\frac{7}{4}, -\frac{5}{4})$, $(-\frac{3}{4}, -\frac{1}{4})$, $(\frac{1}{4}, \frac{3}{4})$ and $(\frac{5}{4}, \frac{7}{4})$ is mapped homeomorphically onto U by the map p. On the other hand the set V is not evenly covered by the map p since

$$\begin{aligned} \alpha^{-1}(V) &= \{t \in (-2, 2) : \cos 2\pi t > 0\} \\ &= (-2, -\frac{7}{4}) \cup (-\frac{5}{4}, -\frac{3}{4}) \cup (-\frac{1}{4}, \frac{1}{4}) \cup (\frac{3}{4}, \frac{5}{4}) \cup (\frac{7}{4}, 2), \end{aligned}$$

where the open intervals $(-2, -\frac{7}{4})$ and $(\frac{7}{4}, 2)$ are not mapped homeomorphiclly onto V by the map p.

Definition Let X and $\tilde{X}$ be topological spaces and let $p: \tilde{X} \to X$ be a continuous map. The map $p: \tilde{X} \to X$ is said to be a *covering map* over the topological space X if and only if the following conditions are satisfied:

(i) the map $p: \tilde{X} \to X$ is surjective,

(ii) every point of X has an open neighbourhood which is evenly covered by the map p.

If $p: \tilde{X} \to X$ is a covering map over a topological space X then the topological space $\tilde{X}$ is said to be a *covering space* of X.

Lemma 3.1 *Every covering map is a local homeomorphism.*

Proof Let $p: \tilde{X} \to X$ be a covering map. Let z be a point of $\tilde{X}$. Then $p(z)$ has an open neighbourhood V which is evenly covered by the map p. Thus $p^{-1}(V)$ is a disjoint union of open sets, each of which is mapped homeomorphically onto V by the map p. One of these open sets contains the point z; let us denote this open set by U. Then U is an open neighbourhood of the point z which is mapped homeomorphically onto the open set V in X by the map p. This shows that p is a local homeomorphism. ∎

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

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

is a covering map. Indeed let $\mathbf{n}$ be a point of S^1. Consider the open neighbourhood U of $\mathbf{n}$ in S^1 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 $e^{-1}(U)$ is the union of the disjoint open sets J_n for all integers n, where

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

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

Example The map $p: \mathbb{C} \to \mathbb{C} \setminus \{0\}$ defined by $p(z) = \exp(2\pi i z)$ is a covering map. Given any $\theta \in [-\pi, \pi]$ let us define

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

Note the U_θ is evenly covered by the map p. Indeed $p^{-1}(U_\theta)$ consists of the union of the open sets

$$\{z \in \mathbb{C} : \frac{\theta}{2\pi} + n - \frac{1}{2} < \operatorname{Im} z < \frac{\theta}{2\pi} + n + \frac{1}{2}\},$$

where each of these open sets is mapped homeomorphically onto U_θ by the map p.

Example Let S^1 denote the unit circle in $\mathbb{R}^2$. Let n be a non-zero integer. Let $\beta_n: S^1 \to S^1$ be defined by

$$\beta_n(\cos\theta, \sin\theta) = (\cos n\theta, \sin n\theta).$$

Then $\beta_n: S^1 \to S^1$ is a covering map.

Example Let $\mathbb{R}P^n$ denote the *real projective n-space*. This is the topological space obtained from the n-sphere S^n by identifying antipodal points on S^n. (We regard S^n as the unit n-sphere in $\mathbb{R}^{n+1}$ consisting of all $\mathbf{x} \in \mathbb{R}^{n-1}$ satisfying $|\mathbf{x}| = 1$. We define an equivalence relation on S^n where distinct points $\mathbf{x}$ and $\mathbf{y}$ of S^n are equivalent if and only if $\mathbf{x} = -\mathbf{y}$. The space $\mathbb{R}P^n$ is then defined to be the set of equivalence classes of points of S^n under this equivalence relation. The topology on $\mathbb{R}P^n$ is the quotient topology induced by the quotient map $\rho: S^n \to \mathbb{R}P^n$. Thus a subset U of $\mathbb{R}P^n$ is open if and only if $\rho^{-1}(U)$ is open in S^n.) It is easily verified that the quotient map $\rho: S^n \to \mathbb{R}P^n$ is a covering map.

The local homeomorphism $\alpha: (-2, 2) \to S^1$ is not a covering map, where S^1 as the unit circle in $\mathbb{R}^2$, and where by $\alpha(t) = (\cos 2\pi t, \sin 2\pi t)$ for all $t \in (-2, 2)$. It can easily be shown that the point $(1, 0)$ of S^1 has no open neighbourhood which is evenly covered by the map α.

Definition Let $p: \tilde{X} \to X$ be a covering map, and let x be a point of $\tilde{X}$. The *fibre* of the map $p: \tilde{X} \to X$ over the point x is defined to be the set $p^{-1}(\{x\})$ consisting of all points of $\tilde{X}$ that are mapped by p onto the point x.

3.1 The Path Lifting and Homotopy Lifting Properties

Let $p: \tilde{X} \to X$ be a covering map over a topological space X. Let Z be a topological space, and let $f: Z \to X$ be a continuous map from Z to X. A continuous map $\tilde{f}: Z \to \tilde{X}$ is said to be a *lift* of the map $f: Z \to X$ if and only if $p \circ \tilde{f} = f$. We shall prove several results concerning the existence and uniqueness of such lifts. These results include the important Path Lifting Property and the Homotopy Lifting Property for covering maps. First we prove a result concerning the uniqueness of lifts of continuous maps from connected topological spaces. (We recall that a topological space X is said to be *connected* if and only if $\emptyset$ and X itself are the only subsets of X that are both open and closed.)

Theorem 3.2 *Let $p: \tilde{X} \to X$ be a covering map over a topological space X. Let Z be a connected topological space, and let $f: Z \to \tilde{X}$ and $g: Z \to \tilde{X}$ be continuous maps. Suppose that $p \circ f = p \circ g$ and that there exists some point z_0 of Z with the property that $f(z_0) = g(z_0)$. Then $f = g$.*

Proof Let Z_0 be the subset of Z defined by

$$Z_0 = \{z \in Z : f(z) = g(z)\}.$$

Note that Z_0 is non-empty, by hypothesis. We show that Z_0 is both open and closed.

Let z be a point of Z_0. There exists an open neighbourhood U of $p(f(z))$ in X which is evenly covered by the map p. Then $p^{-1}(U)$ is a disjoint union of open sets, each of which is mapped homeomorphically onto U by the map p. One of these open sets contains $f(z)$; let this set be denoted by $\tilde{U}$. Let N be the the open neighbourhood of z in Z defined by $N = f^{-1}(\tilde{U}) \cap g^{-1}(\tilde{U})$. Then $f(N) \subset \tilde{U}$ and $g(N) \subset \tilde{U}$. But $p \circ f = p \circ g$, and the restriction $p|\tilde{U}$ of the map p to $\tilde{U}$ maps $\tilde{U}$ homeomorphically onto U. Therefore $f|N = g|N$, and thus $N \subset Z_0$. This shows that Z_0 is open.

We now show that the complement $Z \setminus Z_0$ of Z_0 in Z is open. Let z be a point of $Z \setminus Z_0$. There exists an open neighbourhood U of $p(f(z))$ in X which is evenly covered by the map p. Then $p^{-1}(U)$ is a disjoint union of open sets, each of which is mapped homeomorphically onto U by the map p. One of these open sets contains $f(z)$; let this set be denoted by $\tilde{U}_1$. Another of these open sets contains $g(z)$; let this open set be denoted by $\tilde{U}_2$. Then $\tilde{U}_1 \cap \tilde{U}_2 = \emptyset$. Let N be the open neighbourhood of z in Z defined by $N = f^{-1}(\tilde{U}_1) \cap g^{-1}(\tilde{U}_2)$. Then $f(N) \subset \tilde{U}_1$ and $g(N) \subset \tilde{U}_2$, and hence $f(z') \neq g(z')$ for all $z' \in N$. Thus $N \subset Z \setminus Z_0$. This shows that $Z \setminus Z_0$ is open, so that Z_0 is closed.

The set Z_0 is a non-empty subset of Z that is both open and closed. It follows from the connectedness of Z that $Z_0 = Z$. Therefore $f = g$ throughout Z, as required. ∎

Corollary 3.3 *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 $p \circ c = p \circ f$. Therefore $f = c$ by Theorem 3.2, as required. ∎

Let $p: \tilde{X} \to X$ be a covering map over a topological space X. Let $\gamma: [a, b] \to X$ be a continuous path in X defined on the closed interval $[a, b]$. A continuous path $\tilde{\gamma}: [a, b] \to \tilde{X}$ is said to be a *lift* of γ to $\tilde{X}$ if and only if $p \circ \tilde{\gamma} = \gamma$. We now prove the Path Lifting Property for covering maps.

Theorem 3.4 (Path Lifting Property) *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 $\tilde{\gamma}: [a, b] \to \tilde{X}$ such that $\tilde{\gamma}(a) = w$ and $p \circ \tilde{\gamma} = \gamma$.*

Proof Let S be the subset of $[a, b]$ consisting of all $\tau \in [a, b]$ with the property that there exists a lift of $\gamma|[a, \tau]$ to $\tilde{X}$ starting at w (where $\gamma|[a, \tau]$ denotes the restriction of the path γ to the closed interval $[a, \tau]$). We shall prove that b belongs to S. 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, since s is the supremum of the set S. Choose some element τ_0 of $S \cap J(s, \delta)$. Then there exists a continuous lift $\eta_0: [a, \tau_0] \to \tilde{X}$ of $\gamma|[a, \tau_0]$ to $\tilde{X}$ starting at w. 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 map p. One of these open sets contains the point $\eta_0(\tau_0)$; let this open set be denoted by $\tilde{U}$. Now $\gamma(t) \in U$ for all $t \in J(s, \delta)$ and $\tilde{U}$ is mapped homeomorphically onto U by the map p. Therefore there exists a unique continuous path $\lambda: J(s, \delta) \to \tilde{U}$ such that $\lambda(\tau_0) = \eta_0(\tau_0)$ and $p(\lambda(t)) = \gamma(t)$ for all $t \in J(s, \delta)$. But then, given any $\tau \in J(s, \delta)$, let $\eta: [a, \tau] \to \tilde{X}$ be the continuous path in $\tilde{X}$ given by

$$\eta(t) = \begin{cases} \eta_0(t) & \text{if } a \leq t \leq \tau_0; \\ \lambda(t) & \text{if } \tau_0 \leq t \leq \tau. \end{cases}$$

Then η is a lift of $\gamma|[a, \tau]$. Thus τ belongs to the set S. This shows 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. We conclude therefore that there exists a continuous lift $\tilde{\gamma}: [a, b] \to \tilde{X}$ of γ starting at w. The uniqueness of $\tilde{\gamma}$ follows directly from Theorem 3.2, since the closed interval $[a, b]$ is connected. ∎

Theorem 3.5 (Homotopy Lifting Property) *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 $\gamma_z\colon [0,1] \to Z$ 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 Property (Theorem 3.4) that there exists a unique continuous path $\tilde{\gamma}_z\colon [0,1] \to \tilde{X}$ such that $\tilde{\gamma}_z(0) = g(z)$ for all $z \in Z$ and $p \circ \tilde{\gamma}_z = \gamma_z$. Let the map $G\colon Z \times [0,1] \to \tilde{X}$ 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 map $G\colon Z \times [0,1] \to \tilde{X}$ is continuous and that it is unique.

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

> there exists an open neighbourhood N of z in Z such that the map G is continuous on $N \times [0,t]$.

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 $\tilde{U}$. Thus $\tilde{U}$ is an open set in $\tilde{X}$ which contains the point w and which is mapped homeomorphically onto U by the covering map p. Let $\sigma\colon U \to \tilde{U}$ denote the inverse of the homeomorphism $p|\tilde{U}\colon \tilde{U} \to U$.

Note that $F(z, s_z) = p(w)$. It follows from the continuity of the map F that there exists some $\delta > 0$ and some open neighbourhood N_1 of z in Z such that $F(N_1 \times J(s_z, \delta)) \subset U$, where

$$J(s_z, \delta) = \{t \in \mathbb{R} : 0 \leq t \leq 1 \text{ and } s_z - \delta < t < s_z + \delta\}.$$

Now we can choose some τ belonging to S_z which satisfies $s_z - \delta < \tau \leq s_z$, since 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 neighbourhood N_2 of z in Z such that the map G is continuous on $N_2 \times [0, \tau]$. It follows from this that there exists some open neighbourhood N of z in Z, where N is contained in N_2, such that $G(N \times \{\tau\}) \subset \tilde{U}$. Moreover we can choose N to be a subset of N_1, thus ensuring that $F(N \times J(s_z, \delta)) \subset U$.

Now $p(G(z', t)) = F(z', t) = p(\sigma(F(z', t)))$ for all $z' \in N$ and $t \in J(s_z, \delta)$, and $G(z', \tau) = \sigma(F(z', \tau))$ for all $z' \in N$. It follows from this that $G(z', t) =$

$\sigma(F(z',t))$ for all $z' \in N$ and $t \in J(s_z, \delta)$ (since it follows from Theorem 3.2 that the continuous maps $t \mapsto G(z',t)$ and $t \mapsto \sigma(F(z',t))$ agree on the connected interval $J(s_z, \delta)$). But the maps F and σ are continuous; therefore the map G is continuous on $N \times J(s_z, \delta)$ The map G is also continuous on $N \times [0, \tau]$, for some τ satisfying $s_z - \delta < \tau \leq s_z$ Therefore G is continuous on $N \times [0, t]$ for any $t \in J(s_z, \delta)$, so that $J(s_z, \delta) \subset S_z$. We conclude from this that $s_z = 1$ and that 1 belongs to S_z. Thus we have shown that, given any $z \in Z$ there exists an open neighbourhood N of z such that G is continuous on $N \times [0, 1]$. It follows from this that G is continuous on $Z \times [0, 1]$, as required.

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 $t \in [0, 1]$. ∎

Corollary 3.6 *Let $p: \tilde{X} \to X$ be a covering map over a topological space X. Let $\alpha: [0, 1] \to X$ and $\beta: [0, 1] \to X$ be continuous paths in X, where $\alpha(0) = \beta(0)$ and $\alpha(1) = \beta(1)$. Let $\tilde{\alpha}: [0, 1] \to \tilde{X}$ and $\tilde{\beta}: [0, 1] \to \tilde{X}$ be continuous 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

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

There exists a homotopy $F: [0, 1] \times [0, 1] \to X$ such that

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

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

(since $\alpha \simeq \beta \text{ rel } \{0, 1\}$). It follows from the Homotopy Lifting Property (Theorem 3.5) that there exists a continuous map $G: [0, 1] \times [0, 1] \to \tilde{X}$ such that $p \circ G = F$ and $G(t, 0) = \tilde{\alpha}(t)$ for all $t \in [0, 1]$. Then $p(G(0, \tau)) = x_0$ and $p(G(1, \tau)) = x_1$ for all $\tau \in [0, 1]$. It follows immediately from Corollary 3.3 that $G(0, \tau) = \tilde{x}_0$ and $G(1, \tau) = \tilde{x}_1$ for all $\tau \in [0, 1]$, where

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

However

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

and

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

for all $t \in [0,1]$. Therefore $G(t,1) = \tilde{\beta}(t)$ for all $t \in [0,1]$, by Theorem 3.2. In particular,

$$\tilde{\beta}(1) = G(1,1) = x_1 = \tilde{\alpha}(1).$$

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 fixes the endpoints of these paths, so that $\tilde{\alpha} \simeq \tilde{\beta}$ rel $\{0,1\}$, as required. ∎

Corollary 3.7 *Let $p\colon \tilde{X} \to X$ be a covering map over a topological space X. Let $\tilde{x}_0$ be a point of $\tilde{X}$. Then the homomorphism $p_\#\colon \pi_1(\tilde{X}, \tilde{x}_0) \to \pi_1(X, p(\tilde{x}_0))$ of fundamental groups induced by the covering map p is injective. Moreover if γ is a loop in X based at the point $p(\tilde{x}_0)$ which represents some element of the image $p_\#(\pi_1(\tilde{X}, \tilde{x}_0))$ of the homomorphism $p_\#$, then there exists a loop $\tilde{\gamma}$ in $\tilde{X}$, based at the point $\tilde{x}_0$, such that $p \circ \tilde{\gamma} = \gamma$.*

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 $p(\sigma_0(0) = p(\tilde{x}_0) = p(\sigma_1(0)$. Therefore $\sigma_0 \simeq \sigma_1$ rel $\{0,1\}$, by Corollary 3.6. Thus $[\sigma_0] = [\sigma_1]$. This shows that the homomorphism $p_\#\colon \pi_1(\tilde{X}, \tilde{x}_0) \to \pi_1(X, p(\tilde{x}_0))$ is injective.

Let γ be a loop in X based at the point $p(\tilde{x}_0)$ which represents some element of the image $p_\#(\pi_1(\tilde{X}, \tilde{x}_0))$ of the homomorphism $p_\#$. Then there exists a loop σ in $\tilde{X}$ based at the point $\tilde{x}_0$ such that $\gamma \simeq p \circ \sigma$ rel $\{0,1\}$. Let $\tilde{\gamma}\colon [0,1] \to \tilde{X}$ be the unique path in $\tilde{X}$ for which $\tilde{\gamma}(0) = \tilde{x}_0$ and $p \circ \tilde{\gamma} = \gamma$. (The existence of the path $\tilde{\gamma}$ follows from Theorem 3.4.) Then $\tilde{\gamma}(1) = \sigma(1)$ and $\tilde{\gamma} \simeq \sigma$ rel $\{0,1\}$, by Corollary 3.6. But $\sigma(1) = \tilde{x}_0$. Thus the path $\tilde{\gamma}$ is a loop in $\tilde{X}$ based the point $\tilde{x}_0$ for which $p \circ \tilde{\gamma} = \gamma$. ∎

Let $p\colon \tilde{X} \to X$ be a covering map over a topological space X. Let x_0 and x_1 be points of X, and let α and β be paths in X from x_0 to x_1. Then $\alpha.\beta^{-1}$ is a loop based at the point x_0, given by

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

Thus the loop $\alpha.\beta^{-1}$ represents an element $[\alpha.\beta^{-1}]$ of the fundamental group $\pi_1(X, x_0)$ of X based at the point x_0. Let $\tilde{x}_0$ be a point of $\tilde{X}$ for which $p(\tilde{x}_0) = x_0$, and let $\tilde{\alpha}$ and $\tilde{\beta}$ be the unique lifts of the paths α and β for which $\tilde{\alpha}(0) = \tilde{x}_0$ and $\tilde{\beta}(0) = \tilde{x}_0$. (Thus $p \circ \tilde{\alpha} = \alpha$ and $p \circ \tilde{\beta} = \beta$.) We now prove that $\tilde{\alpha}(1) = \tilde{\beta}(1)$ if and only if $[\alpha.\beta^{-1}]$ belongs to the image $p_\#(\pi_1(\tilde{X}, \tilde{x}_0))$ of the homomorphism $p_\#\colon \pi_1(\tilde{X}, \tilde{x}_0) \to \pi_1(X, x_0)$ of fundamental groups induced by the covering map $p\colon \tilde{X} \to X$).

Lemma 3.8 *Let $p: \tilde{X} \to X$ be a covering map over a topological space X. Let $\alpha: [0,1] \to X$ and $\beta: [0,1] \to X$ be paths in X such that $\alpha(0) = \beta(0)$ and $\alpha(1) = \beta(1)$. Let $\tilde{\alpha}: [0,1] \to \tilde{X}$ and $\tilde{\beta}: [0,1] \to \tilde{X}$ be paths in $\tilde{X}$ such that $p \circ \tilde{\alpha} = \alpha$, and $p \circ \tilde{\beta} = \beta$. Suppose that $\tilde{\alpha}(0) = \tilde{\beta}(0)$. Then $\tilde{\alpha}(1) = \tilde{\beta}(1)$ if and only if $[\alpha.\beta^{-1}] \in p_{\#}(\pi_1(\tilde{X}, \tilde{x}_0))$, where $\tilde{x}_0 = \tilde{\alpha}(0) = \tilde{\beta}(0)$.*

Proof Let $x_0 = p(\tilde{x}_0)$, and $\gamma: [0,1] \to X$ be the loop based at x_0 given by $\gamma = \alpha.\beta^{-1}$. Thus

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

It follows from the Path Lifting Property for covering maps (Theorem 3.4) that there exists a unique path $\tilde{\gamma}: [0,1] \to X$ such that $\tilde{\gamma}(0) = \tilde{x}_0$ and $p \circ \tilde{\gamma} = \gamma$.

Suppose that $[\alpha.\beta^{-1}] \in p_{\#}(\pi_1(\tilde{X}, \tilde{x}_0))$. Then the lift $\tilde{\gamma}$ of γ is a loop in $\tilde{X}$ based at the point $\tilde{x}_0$, by Corollary 3.7. It then follows from the uniqueness of the lifts $t\tilde{\alpha}$ and $\tilde{\beta}$ of the paths α and β that $\tilde{\alpha}(t) = \tilde{\gamma}(\frac{1}{2}t)$ and $\tilde{\beta}(t) = \tilde{\gamma}(\frac{1}{2}(1-t))$ for all $t \in [0,1]$. In particular $\tilde{\alpha}(1) = \tilde{\gamma}(\frac{1}{2} = \tilde{\beta}(1)$, as required.

Conversely suppose that $\tilde{\alpha}(1) = \tilde{\beta}(1)$. Then the unique continuous lift $\tilde{\gamma}$ of the path γ starting at the point $\tilde{x_0}$ is given by

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

But then $\tilde{\gamma}(1) = \tilde{x}_0 = \tilde{\gamma}(0)$, so that $\tilde{\gamma}$ is a loop in $\tilde{X}$ based at the point $\tilde{x}_0$. Thus $[\alpha.\beta^{-1}] = p_{\#}[\tilde{\gamma}]$, and hence $[\alpha.\beta^{-1}]$ belongs to the image of the homomorphism $p_{\#}: \pi_1(\tilde{X}, \tilde{x}_0) \to \pi_1(X, x_0)$ as required. ∎

Theorem 3.9 *Let b be a point of the circle S^1. Then $\pi_1(S^1, b)$ is isomorphic to the additive group $\mathbb{Z}$ of integers.*

Proof We represent S^1 as the unit circle in $\mathbb{R}^2$. Without loss of generality we may suppose that $b = (1, 0)$. Then the map $p: \mathbb{R} \to S^1$ defined by

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

is a covering map. Note that $p(0) = b$ and $p^{-1}(\{b\}) = \mathbb{Z}$.

Let $\gamma: [0,1] \to S^1$ be a loop in S^1 based at the point b which represents some element $[\gamma]$ of $\pi_1(S^1, b)$ (where $b = (1, 0)$). It follows from the Path Lifting Property for covering maps (Theorem 3.4) that there exists a unique path $\tilde{\gamma}: [0,1] \to \mathbb{R}$ in $\mathbb{R}$ such that $\tilde{\gamma}(0) = 0$ and $p \circ \tilde{\gamma} = \gamma$. Then $\tilde{\gamma}(1)$ is an integer (since $p(\tilde{\gamma}(1)) = b$). We claim that there is a well-defined homomorphism which maps the element $[\gamma]$ of $\pi_1(S^1, b)$ represented by the loop γ to the endpoint $\tilde{\gamma}(1)$ of the lift $\tilde{\gamma}$ of γ.

Let α and β be loops in S^1 based at the point b and let $\tilde{\alpha}$ and $\tilde{\beta}$ be the corresponding paths in $\mathbb{R}$ for which $p\circ\tilde{\alpha} = \alpha$, $p\circ\tilde{\beta} = \beta$, and $\tilde{\alpha}(0) = 0 = \tilde{\beta}(0)$. The space $\mathbb{R}$ of real numbers is simply-connected, so that $\pi_1(\mathbb{R}, 0)$ is the trivial group. It follows from Lemma 3.8 that $\tilde{\alpha}(1) = \tilde{\beta}(1)$ if and only if $[\alpha.\beta^{-1}]$ represents the identity element of $\pi_1(S^1, b)$. Thus $\tilde{\alpha}(1) = \tilde{\beta}(1)$ if and only if $[\alpha] = [\beta]$ (where $[\alpha]$ and $[\beta]$ are the elements of $\pi_1(S^1, b)$ represented by the loops α and β respectively. We conclude therefore that there is a well-defined function $\Theta\colon \pi_1(S^1, b) \to \mathbb{Z}$ which maps the element $[\gamma]$ of $\pi_1(S^1, b)$ represented by the loop γ to the endpoint $\tilde{\gamma}(1)$ of the lift $\tilde{\gamma}$ of γ. Moreover this function is injective. It is also surjective, since if $\gamma_n\colon [0,1] \to S^1$ is the loop in S^1 based at B defined by

$$\gamma_n(t) = e(nt) = (\cos 2\pi nt, \sin 2\pi nt)$$

then $\Theta([\gamma_n]) = n$ for each integer n.

We claim that the bijection $\Theta\colon \pi_1(S^1, b) \to \mathbb{Z}$ is a homomorphism. Let α and β be loops in S^1 based at the point b, and let $\tilde{\alpha}$ and $\tilde{\beta}$ be the corresponding paths in $\mathbb{R}$ for which $p \circ \tilde{\alpha} = \alpha$, $p \circ \tilde{\beta} = \beta$, and $\tilde{\alpha}(0) = 0 = \tilde{\beta}(0)$. Let $\gamma\colon [0,1] \to S^1$ be the product loop given by $\gamma = \alpha.\beta$, so that

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

Let $\tilde{\gamma}\colon [0,1] \to \mathbb{R}$ be the unique path in $\mathbb{R}$ for which $p \circ \tilde{\gamma} = \gamma$ and $\tilde{\gamma}(0) = 0$. Then

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

Thus

$$\Theta([\alpha].[\beta]) = \Theta([\alpha.\beta]) = \Theta([\gamma]) = \tilde{\gamma}(1) = \tilde{\alpha}(1) + \tilde{\beta}(1) = \Theta([\alpha]) + \Theta([\beta]).$$

Thus the bijection $\Theta\colon \pi_1(S^1, b) \to \mathbb{Z}$ is a homomorphism. We conclude therefore that $\Theta\colon \pi_1(S^1, b) \to \mathbb{Z}$ is an isomorphism, as required. ∎

Theorem 3.10 *Let $p\colon \tilde{X} \to X$ 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 $p\colon \tilde{X} \to X$ is a homeomorphism.*

Proof We show that the map $p\colon \tilde{X} \to X$ 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 continuous path $\sigma: [0,1] \to \tilde{X}$ with $\sigma(0) = w_0$ and $\sigma(1) = w_1$, since $\tilde{X}$ is path-connected. Then $p \circ \sigma$ 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 3.7 that $p \circ \sigma = p \circ \omega$ for some loop ω in $\tilde{X}$ based at the point w_0. But $\omega(0) = w_0 = \sigma(0)$. Therefore $\omega = \sigma$, by Theorem 3.2. In particular

$$w_1 = \sigma(1) = \omega(1) = w_0.$$

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 $p^{-1}: X \to \tilde{X}$.

Every covering map is a local homeomorphism, by Lemma 3.1. Let x be a point of X. Then there exists an open neighbourhood U $p^{-1}(x)$ which is mapped homeomorphically onto some open set $p(U)$ in X. But then $p(U)$ is an open neighbourhood of the point x, and the inverse p^{-1} of p is continuous on $p(U)$. Thus $p^{-1}: X \to \tilde{X}$ is continuous at each point x of X. Thus $p: \tilde{X} \to X$ is a homeomorphism, as required. ∎

Let $p: \tilde{X} \to X$ be a covering map over some topological space X, and let x_0 be some chosen basepoint of X. We shall investigate the dependence of the subgroup $p_\#\left(\pi_1(\tilde{X}, \tilde{x})\right)$ of $\pi_1(X, x_0)$ on the choice of the point $\tilde{x}$ in $\tilde{X}$, where $\tilde{x}$ is chosen such that $p(\tilde{x}) = x_0$. We first introduce some concepts from group theory.

Let G be a group, and let H be a subgroup of G. Given any $g \in G$, let gHg^{-1} denote the subset of G defined by

$$gHg^{-1} = \{g' \in G : g' = ghg^{-1} \text{ for some } h \in H\}.$$

It is easy to verify that gHg^{-1} is a subgroup of G.

Definition Let G be a group, and let H and H' be subgroups of G. We say that H and H' are *conjugate* if and only if there exists some $g \in G$ for which $H' = gHg^{-1}$.

Note that if $H' = gHg^{-1}$ then $H = g^{-1}H'g$. The relation of conjugacy is an equivalence relation on the set of all subgroups of the group G. Moreover conjugate subgroups of G are isomorphic, since the homomorphism sending $h \in H$ to ghg^{-1} is an isomorphism from H to gHg^{-1}.

Lemma 3.11 *Let $p\colon \tilde{X} \to X$ be a covering map over some topological space X. Let x_0 be a point of X, and let $\tilde{x}_0$ and $\tilde{x}_1$ be points of $\tilde{X}$ for which $p(\tilde{x}_0) = x_0 = p(\tilde{x}_1)$. Let H_0 and H_1 be the subgroups of $\pi_1(X, x_0)$ defined by*

$$H_0 = p_\# \left(\pi_1(\tilde{X}, \tilde{x}_0)\right), H_1 = p_\# \left(\pi_1(\tilde{X}, \tilde{x}_1)\right).$$

Suppose that the covering space $\tilde{X}$ is path-connected. Then the subgroups H_0 and H_1 of $\pi_1(X, x_0)$ are conjugate. Moreover if H is any subgroup of $\pi_1(X, x_0)$ which is conjugate to H_0 then there exists an element $\tilde{x}$ of $\tilde{X}$ for which $p(\tilde{x}) = x$ and $p_\# \left(\pi_1(\tilde{X}, \tilde{x})\right) = H$.

Proof Let $\alpha\colon [0,1] \to \tilde{X}$ be a path in $\tilde{X}$ for which $\alpha(0) = \tilde{x}_0$ and $\alpha(1) = \tilde{x}_1$. (Such a path exists since $\tilde{X}$ is path-connected.) Let

$$\tau_\alpha\colon \pi_1(\tilde{X}, \tilde{x}_1) \to \pi_1(\tilde{X}, \tilde{x}_1)$$

be the isomorphism which sends $[\gamma] \in \pi_1(\tilde{X}, \tilde{x}_1)$ to $[\alpha.\gamma.\alpha^{-1}]$ for all loops γ in $\tilde{X}$ based at $\tilde{x}_1$, where

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

(It follows from Lemma 2.3 that τ_α is well-defined and is an isomorphism from $\pi_1(\tilde{X}, \tilde{x}_1)$ to $\tilde{X}, \tilde{x}_0$.) Let $\eta\colon [0,1] \to X$ be the loop in X based at the point x_0 given by $\eta = \alpha \circ p$. Then

$$p_\#(\tau_\alpha([\gamma])) = [\eta]\,(p_\#([\gamma]))\,[\eta^{-1}]$$

for all $[\gamma] \in \pi_1(\tilde{X}, \tilde{x}_1)$. Therefore $H_0 = [\eta]H_1[\eta]^{-1}$. Thus the subgroups H_0 and H_1 of $\pi_1(X, x_0)$ are conjugate.

Now let H be a subgroup of $\pi_1(X, x_0)$ which is conjugate to H_0. Then $H_0 = [\eta]H[\eta]^{-1}$ for some loop η in X based at the point x_0. It follows from the Path Lifting Property for covering maps (Theorem 3.4) that there exists a path $\alpha\colon [0,1] \to \tilde{X}$ in $\tilde{X}$ for which $\alpha(0) = \tilde{x}_0$ and $p \circ \alpha = \eta$. Let $\tilde{x} = \alpha(1)$. Then

$$p_\# \left(\pi_1(\tilde{X}, \tilde{x})\right) = [\eta]^{-1}H_0[\eta] = H,$$

as required. ∎

4 Winding Numbers

Let $\gamma\colon [a,b] \to \mathbb{C}$ be a continuous closed curve in the complex plane which is defined on some closed interval $[a,b]$ (so that $\gamma(a) = \gamma(b)$), and let w be a complex number which does not belong to the image of the closed curve γ. The map $p_w\colon \mathbb{C} \to \mathbb{C} \setminus \{w\}$ defined by $p_w(z) = w + \exp(2\pi i z)$ is a covering map. Observe that if z_1 and z_2 are complex numbers then $p_w(z_1) = p_w(z_2)$ if and only if $z_1 - z_2$ is an integer. Using the Path Lifting Property for covering maps (Theorem 3.4) we see that there exists a continuous path $\tilde{\gamma}\colon [a,b] \to \mathbb{C}$ in $\mathbb{C}$ such that $p_w \circ \tilde{\gamma} = \gamma$. Let us define

$$n(\gamma, w) = \tilde{\gamma}(b) - \tilde{\gamma}(a).$$

Now $p_w(\tilde{\gamma}(b)) = p_w(\tilde{\gamma}(a))$ (since $\gamma(b) = \gamma(a)$). It follows from this that $n(\gamma, w)$ is an integer. We claim that the value of $n(\gamma, w)$ is independent of the choice of the path $\tilde{\gamma}$ on $\mathbb{C}$.

Let $\sigma\colon [a,b] \to \mathbb{C}$ be a continuous path in $\mathbb{C}$ with the property that $p_w \circ \sigma = \gamma$. Then $p_w(\sigma(t)) = p_w(\tilde{\gamma}(t))$ for all $t \in [a,b]$, and hence $\sigma(t) - \tilde{\gamma}(t)$ is an integer for all $t \in [a,b]$. But the map sending $t \in [a,b]$ to $\sigma(t) - \tilde{\gamma}(t)$ is continuous on $[a,b]$; therefore this map must be a constant map. Thus there exists some integer m with the property that $\sigma(t) = \tilde{\gamma}(t) + m$ for all $t \in [a,b]$, and hence

$$\sigma(b) - \sigma(a) = \tilde{\gamma}(b) - \tilde{\gamma}(a).$$

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

Definition Let $\gamma\colon [a,b] \to \mathbb{C}$ be a continuous closed curve in the complex plane, and let w be a complex number which does not belong to the image of the closed curve γ. Then the *winding number* $n(\gamma, w)$ of the closed curve γ about w is defined by

$$n(\gamma, w) \equiv \tilde{\gamma}(b) - \tilde{\gamma}(a),$$

where $\tilde{\gamma}\colon [a,b] \to \mathbb{C}$ is some continuous path in $\mathbb{C}$ with the property that

$$\gamma(t) = w + \exp(2\pi i \tilde{\gamma}(t)$$

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

Theorem 4.1 *Let w be a complex number and let*

$$\gamma_0\colon [a,b] \to \mathbb{C} \text{ and } \gamma_1\colon [a,b] \to \mathbb{C}$$

be closed curves in $\mathbb{C}$ which do not pass through w. Suppose that there exists some homotopy $F\colon [a,b] \times [0,1] \to \mathbb{C}$ with the following properties:

(i) $F(t,0) = \gamma_0(t)$ *and* $F(t,1) = \gamma_1(t)$ *for all* $t \in [a,b]$,

(ii) $F(a,\tau) = F(b,\tau)$ *for all* $\tau \in [0,1]$,

(iii) *the complex number* w *does not belong to the image* $F([a,b] \times [0,1])$ *of the homotopy* F.

Then $n(\gamma_0, w) = n(\gamma_1, w)$ *(where* $n(\gamma_0, w)$ *and* $n(\gamma_1, w)$ *are the winding numbers of the closed curves* γ_0 *and* γ_1*).*

Proof It follows from the Path Lifting Property for covering maps (Theorem 3.4) that there there exists a continuous path $\eta\colon [0,1] \to \mathbb{C}$ with the property that $p_w(\eta(\tau)) = F(a,\tau)$ for all $\tau \in [0,1]$ (where $p_w(z) = w + \exp(2\pi i z)$ for all $z \in \mathbb{C}$). It then follows from the Homotopy Lifting Property (Theorem 3.5) that there exists a continuous map $G\colon [a,b] \times [0,1] \to \mathbb{C}$ such that $G(a,\tau) = \eta(\tau)$ and $p_w(G(t,\tau)) = F(t,\tau)$ for all $t \in [a,b]$ and $\tau \in [0,1]$. But then

$$p_w(G(a,\tau)) = F(a,\tau) = F(b,\tau) = p_w(G(b,\tau))$$

for all $\tau \in [0,1]$ and hence $G(b,\tau) - G(a,\tau)$ is an integer for all $\tau \in [0,1]$. But the map $t \mapsto G(b,\tau) - G(a,\tau)$ is continuous; therefore this map is a constant map. Thus there exists some integer m with the proprty that $G(b,\tau) - G(a,\tau) = m$ for all $\tau \in [0,1]$. But

$$G(b,0) - G(a,0) = n(\gamma_0, w), \qquad G(b,1) - G(a,1) = n(\gamma_1, w)$$

(since $p_w(G(t,0)) = \gamma_0(t)$ and $p_w(G(t,1)) = \gamma_1(t)$ for all $t \in [a,b]$). Therefore $n(\gamma_0, w) = m = n(\gamma_1, w)$, as required. ∎

Corollary 4.2 *Let* $\gamma_0\colon [a,b] \to \mathbb{C}$ *and* $\gamma_1\colon [a,b] \to \mathbb{C}$ *be continuous closed curves in* $\mathbb{C}$*, and let* w *be a complex number which does not lie on the images of the closed curves* γ_0 *and* γ_1*. Suppose that, for all* $t \in [a,b]$*, the line segment in the complex plane* $\mathbb{C}$ *joining* $\gamma_0(t)$ *to* $\gamma_1(t)$ *does not pass through* w*. Then* $n(\gamma_0, w) = n(\gamma_1, w)$*.*

Proof Let $F\colon [a,b] \times [0,1] \to \mathbb{C}$ be the homotopy defined by

$$F(t,\tau) = (1-\tau)\gamma_0(t) + \tau\gamma_1(t)$$

for all $t \in [a,b]$ and $\tau \in [0,1]$. Note that w does not lie on the image of the homotopy F. We can therefore apply Theorem 4.1 to conclude that $n(\gamma_0, w) = n(\gamma_1, w)$. ∎

Corollary 4.3 (Dog-walking Principle) *Let $\gamma_0\colon [a,b] \to \mathbb{C}$ and $\gamma_1\colon [a,b] \to \mathbb{C}$ be continuous closed curves in $\mathbb{C}$, and let w be a complex number which does not lie on the images of the closed curves γ_0 and γ_1. Suppose that $|\gamma_1(t) - \gamma_0(t)| < |\gamma_0(t) - w|$ for all $t \in [a,b]$. Then $n(\gamma_0, w) = n(\gamma_1, w)$.*

Proof The inequality $|\gamma_1(t) - \gamma_0(t)| < |\gamma_0(t) - w|$ ensures that the line segment in $\mathbb{C}$ joining $\gamma_0(t)$ and $\gamma_1(t)$ does not pass through w. The result therefore follows directly from Corollary 4.2. ∎

Corollary 4.4 *Let $\gamma\colon [a,b] \to \mathbb{C}$ be a continuous closed curve in $\mathbb{C}$, and let $\sigma\colon [0,1] \to \mathbb{C}$ be a continuous path in $\mathbb{C}$. Suppose that $\gamma([a,b]) \cap \sigma([0,1]) = \emptyset$ (so that σ is a continuous path in $\mathbb{C} \setminus \gamma([a,b])$). Then*

$$n(\gamma, \sigma(0)) = n(\gamma, \sigma(1)).$$

Thus the function $w \mapsto n(\gamma, w)$ is constant over each path-component of the set $\mathbb{C} \setminus \gamma([a,b])$.

Proof Let $F\colon [a,b] \times [0,1] \to \mathbb{C}$ be the continuous map defined by $F(t,\tau) = \gamma(t) - \sigma(\tau)$. Then $F(t,\tau) \neq 0$ for all $t \in [a,b]$ and $\tau \in [0,1]$. Given any $\tau \in [0,1]$, let $\gamma_\tau\colon [0,1] \to \mathbb{C}$ be the closed curve defined by

$$\gamma_\tau(t) = F(t,\tau) = \gamma(t) - \sigma(\tau)$$

for all $t \in [a,b]$. Then it follows directly from the definition of winding numbers that $n(\gamma, \sigma(tau)) = n(\gamma_\tau, 0)$ for all $\tau \in [0,1]$. But F is a continuous homotopy between the closed curves γ_0 and γ_1. It follows from Theorem 4.1 that $n(\gamma_0, 0) = n(\gamma_1, 0)$. Hence $n(\gamma, \sigma(0)) = n(\gamma, \sigma(1))$, as required. ∎

Corollary 4.5 *Let $w \in \mathbb{C}$ and $\theta_0 \in \mathbb{R}$ be given, and let $L_{w,\theta}$ denote the half-line in $\mathbb{C}$ defined by*

$$L_{w,\theta} = \{z \in \mathbb{C} : z = w + \rho e^{i\theta} \text{ for some } \rho \in [0, +\infty)\}.$$

Let $\gamma\colon [a,b] \to \mathbb{C}$ be a continuous closed curve in $\mathbb{C}$ whose image is contained in the complement $\mathbb{C} \setminus L_{w,\theta}$ of the half-line $L_{w,\theta}$. Then $n(\gamma, w) = 0$.

Proof Let $\eta\colon [a,b] \to \mathbb{C}$ be the constant curve in $\mathbb{C}$ defined by $\eta(t) = w - e^{i\theta}$ for all $t \in [a,b]$. Then $n(\gamma, w) = n(\eta, w) = 0$ by Corollary 4.2, since the line segment joining $\gamma(t)$ to $w - e^{i\theta}$ does not pass through w for any $t \in [a,b]$. ∎

Let D be the unit disk in $\mathbb{C}$, defined by

$$D = \{z \in \mathbb{C} : |z| \leq 1\}.$$

Let ∂D denote the boundary of D, given by

$$\partial D = \{z \in \mathbb{C} : |z| = 1\}.$$

Thus ∂D is the unit circle in $\mathbb{C}$. Note that the continuous closed curve $\sigma \colon [0,1] \to \mathbb{C}$ defined by $\sigma(t) = e^{2\pi i t}$ traverses the unit circle ∂D once in the anticlockwise direction.

Theorem 4.6 (Kronecker Principle) *Let $f \colon D \to \mathbb{C}$ be a continuous map defined on the closed unit disk D in $\mathbb{C}$. Let w be a complex number which does not lie on the image $f(\partial D)$ of the boundary ∂D of D. Suppose that $n(f \circ \sigma, w) \neq 0$, where $\sigma \colon [0,1] \to \partial D$ is defined by $\sigma(t) = e^{2\pi i t}$. Then there exists some $z \in D \setminus \partial D$ with the property that $f(z) = w$.*

Proof Let w be a complex number which does not belong To the image $f(D)$ of the closed unit disk D under the map f. We show that $n(f \circ \sigma, 0)$. Consider the homotopy $F \colon [0,1] \times [0,1] \to \mathbb{C}$ defined by $F(t,\tau) = f(\tau e^{2\pi i t})$. Note that $F(t,0) = f(0)$ and $F(t,1) = f(\sigma(t))$ for all $t \in [0,1]$. Also $F(0,\tau) = F(1,\tau)$ for all $\tau \in [0,1]$ and the image of the homotopy F is the image $f(D)$ of the closed unit disk D under the map f. Thus if $w \notin f(D)$ then $n(f \circ \sigma, w) = n(\eta, w) = 0$ by Theorem 4.1, where $\eta \colon [0,1] \to \mathbb{C}$ is the constant curve defined by $\eta(t) = f(0)$ for all $t \in [0,1]$. This proves the Kronecker Principle. ∎

We can use the Kronecker Principle in order to prove the 2-dimensional case of the Brouwer Fixed Point Theorem for maps from the closed unit disk into itself.

Theorem 4.7 (The Brouwer Fixed Point Theorem in 2 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 We may assume without loss of generality that $f(z) \neq z$ for all $z \in \partial D$ (since the conclusion of the theorem is clearly satisfied if f has a fixed point on ∂D). Consider the map $g \colon D \to \mathbb{C}$ defined by $g(z) = z - f(z)$. We must show that there exists some $z_0 \in \mathbb{Z}$ such that $g(z_0) = 0$. Let $\sigma \colon [0,1] \to \partial D$ be the continuous closed curve defined by $\sigma(t) = e^{2\pi i t}$ for all $t \in [0,1]$. Define

$$F(t,\tau) = (1-\tau)\sigma(t) + \tau g(\sigma(t)) = \sigma(t) - \tau f(\sigma(t))$$

for all $t, \tau \in [0,1]$. Note that if $0 \leq \tau < 1$ then $|F(t,\tau)| \geq 1 - \tau |f(\sigma(t))| > 0$ (since $f(\sigma(t))$ belongs to D). Also $F(t,1) \neq 0$, since we are assuming that f has no fixed point on ∂D. Thus $F(t,\tau) \neq 0$ for all $t \in [0,1]$ and $\tau \in [0,1]$. Thus the line segment joining $\sigma(t)$ to $g(\sigma(t))$ does not pass through 0, and hence

$$n(g \circ \sigma, 0) = g(\sigma, 0) = 1$$

by Corollary 4.2. It follows from the Kronecker Principle (Theorem 4.6) that there exists some $z_0 \in D$ such that $g(z_0) = 0$. But then $f(z_0) = z_0$. This proves the Brouwer Fixed Point Theorem (in the 2-dimensional case). ∎

Remark One can give a geometrical interpretation of the proof of the Brouwer Fixed Point Theorem given above. let ζ be a point on the boundary ∂D of the closed unit disk in $\mathbb{C}$, and let L_η be the line through 0 defined by

$$L_\zeta = \{z \in \mathbb{C} : \operatorname{Re} z\zeta^{-1} = 0\}.$$

Thus L_ζ is the line through 0 which is parallel to the tangent line to the unit circle ∂D at ζ. Now the unit disk D lies entirely to one side of the tangent line to the unit circle at ζ. It follows from this that $\zeta - f(\zeta)$ lies on the the same side of the line L_ζ as ζ (where $\zeta - f(\zeta)$ represents the displacement vector joining the point $f(\zeta)$ to the point ζ). Therefore the line segment joining $\zeta - f(\zeta)$ to ζ does not pass through 0. It follows from this that the closed curves $g \circ \sigma$ and σ have the same winding number about 0, where $g(z) = z - f(z)$ and where $\sigma\colon [0,1] \to \mathbb{C}$ is the parameterization of the unit circle given by $\sigma(t) = e^{2\pi i t}$. Thus winding number of the closed curve $g \circ \sigma$ about 0 is equal to 1, and so the continuous function $g\colon D \to \mathbb{C}$ has a zero inside D (by the Kronecker Principle) and thus f has a fixed point inside D.

Remark There is an alternative proof of the 2-dimensional case of the Brouwer Fixed Point Theorem which is well-known. Let $f\colon D \to D$ be a continuous map which maps the closed unit disk D into itself. Suppose that it were the case that the map f has no fixed point in D. Then one could define a map $r\colon D \to \partial D$ as follows. Given $z \in D$ 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 $r(z)$. It is not difficult to verify that if $f\colon D \to D$ has no fixed point then $r\colon D \to \partial D$ is continuous. Moreover $r|\partial D$ is the identity map of ∂D. Choose some basepoint $\zeta \in \partial D$. If $i\colon \partial D \hookrightarrow D$ is the inclusion map, and if

$$i_\#\colon \pi_1(\partial D, \zeta) \to \pi_1(D, \zeta), \qquad r_\#\colon \pi_1(D, \zeta) \to \pi_1(\partial D, \zeta)$$

are the homomorphisms of fundamental groups induced by the continuous maps $i\colon \partial D \hookrightarrow D$ and $r\colon D \to \partial D$ respectively then $r_\# \circ i_\#$ is the identity homomorphism $\pi_1(\partial D, \zeta)$, and hence $r_\#$ is surjective. But this is impossible, since $\pi_1(D, \zeta)$ is trivial and $\pi_1(\partial D, \zeta)$ is the infinite cyclic group. This contradiction shows that the continuous map $f \colon D \to D$ must have a fixed point in D.

Remark The Brouwer Fixed Point Theorem in n-dimensions states that every continuous map $f \colon E^n \to E^n$ from the closed unit ball E^n in $\mathbb{R}^n$ into itself has a fixed point $\mathbf{x}_0 \in E^n$ at which $f(\mathbf{x}_0) = \mathbf{x}_0$, where

$$E^n = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| \leq 1\}.$$

This result can be proved using certain topological invariants known as *homology groups.*

Theorem 4.8 (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 Let $P(z) = a_0 + a_1 z + \cdots + a_m z^m$, where $a_1, a_2, \ldots, a_n$ are complex numbers, and where $a_m \neq 0$. We write $P(z) = P_m(z) + Q(z)$, where $P_m(z) = a_m z^m$ and

$$Q(z) = a_0 + a_1 z + \cdots + a_{m-1} z^{m-1}.$$

Let R be defined by $R = |a_m|^{-1}(|a_0| + |a_1| + \cdots + |a_m|)$ If $|z| = R$ then $|z| \geq 1$ and hence

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

Thus if $|z| = R$ then $|P(z) - P_m(z)| < |P_m(z)|$. Let $\sigma_R \colon [0, 1] \to \mathbb{C}$ be defined by $\sigma_R(t) = Re^{2\pi i t}$. It follows from the Dog-Walking Principle (Corollary 4.3) that

$$n(P \circ \sigma_R, 0) = n(P_m \circ \sigma_R, 0) = m > 0$$

(where m is the degree of the non-constant polynomial P). Thus if $f \colon D \to \mathbb{C}$ is defined by $f(z) = P(Rz)$ for all $z \in D$ (where D is the closed unit disk in $\mathbb{C}$) then there exists some $\zeta \in D$ with the property that $f(\zeta) = 0$ by the Kronecker Principle (Theorem 4.6). Thus $P(z_0) = 0$, where $z_0 = R\zeta$. This proves the Fundamental Theorem of Algebra. ∎

Lemma 4.9 *Let $f \colon S^1 \to \mathbb{C}$ be a continuous function defined on the unit circle S^1 in $\mathbb{C}$. Suppose that $f(-z) = -f(z)$ for all $z \in \mathbb{C}$. Then the winding number $n(f \circ \sigma, 0)$ of $f \circ \sigma$ about 0 is odd, where $\sigma \colon [0, 1] \to S^1$ is the parameterization of S^1 given by $\sigma(t) = e^{2\pi i t}$.*

Proof Let $\gamma\colon [0,1] \to \mathbb{C}$ be defined by $\gamma = f \circ \sigma$. Then $\gamma(t+\frac{1}{2}) = -\gamma(t)$ for all $t \in [0, \frac{1}{2}]$. Using the Path Lifting Property for covering maps (Theorem 3.4) (applied to the covering map from $\mathbb{C}$ to $\mathbb{C} \setminus \{0\}$ which sends $z \in \mathbb{C}$ to $e^{2\pi i z}$), we see that there exists a continuous path $\tilde{\gamma}\colon [0,1] \to \mathbb{C}$ such that $\exp(2\pi i \tilde{\gamma}(t)) = \gamma(t)$ for all $t \in [0,1]$. Moreover

$$n(f \circ \sigma, 0) = n(\gamma, 0) = \tilde{\gamma}(1) - \tilde{\gamma}(0).$$

Now $\exp(2\pi i \tilde{\gamma}(t + \frac{1}{2})) = -\exp(2\pi i \tilde{\gamma}(t))$ for all $t \in [0, \frac{1}{2}]$, hence there exists some integer m with the property that $\tilde{\gamma}(t + 1\ over 2) = \tilde{\gamma}(t) + m + \frac{1}{2}$ for all $t \in [0, \frac{1}{2}]$. But then

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

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. Thus we can represent the closed unit disk D either as the closed unit disk

$$\{z \in \mathbb{C} : |z| \leq 1\}$$

in $\mathbb{C}$ or else as the closed unit disk

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

Similarly we regard the circle S^1 as the unit circle about the origin in either $\mathbb{C}$ or $\mathbb{R}^2$. We represent the standard 2-sphere S^2 as the unit sphere about the origin in $\mathbb{R}^3$, defined by

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

Lemma 4.10 *Let $f\colon S^2 \to \mathbb{R}^2$ 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 D be the closed unit disk in $\mathbb{R}^2$, and let $\varphi\colon D \to S^2$ be the map from D to S^2 defined by

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

(Thus the map φ maps D homeomorphically onto the upper hemisphere in $\mathbb{R}^3$.) Let $\sigma\colon [0,1] \to S^2$ 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 $f\colon S^2 \to \mathbb{R}^2$ be a continuous map with the property that $f(-\mathbf{n}) = -f(\mathbf{n})$ for all $\mathbf{n} \in S^2$. It follows from Lemma 4.9 that the winding number $n(f \circ \sigma, 0)$ of the closed curve $f \circ \sigma$ about the origin is an odd integer, and in particular, this winding number is non-zero. Hence there exists some point (u,v) of D such that $f(\sigma(u,v)) = 0$, by the Kronecker Principle (Theorem 4.6) applied to the map $f \circ \sigma\colon D \to \mathbb{R}^2$. Thus $f(\mathbf{n}_0) = 0$, where $\mathbf{n}_0 = \sigma(u,v)$. ∎

We conclude immediately from this result that there are no continuous maps $f\colon S^2 \to S^1$ from the 2-sphere S^2 to the circle S^1 with the property that $f(-\mathbf{n}) = -f(\mathbf{n})$ for all $\mathbf{n} \in S^2$.

Theorem 4.11 (Borsuk-Ulam) *Let $f\colon S^2 \to \mathbb{R}^2$ be a continuous map. Then there exists some point $\mathbf{n}$ of S^2 with the property that $f(-\mathbf{n}) = f(\mathbf{n})$.*

Proof This result follows immediately on applying Lemma 4.11 to the continuous function $g\colon S^2 \to \mathbb{R}^2$ defined by $g(\mathbf{n}) = f(\mathbf{n}) - f(-\mathbf{n})$. ∎

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

5 Properties of Covering Maps over Locally Path-Connected Topological Spaces

Let $p: \tilde{X} \to X$ be a covering map over a topological space X. Let $f: Z \to X$ be a continuous map from some topological space Z into X. If the topological space Z is *locally path-connected* then one can formulate a criterion to determine whether or not there exists a map $\tilde{f}: Z \to \tilde{X}$ for which $p \circ \tilde{f} = f$ (see Theorem 5.4 and Corollary 5.5). This criterion is stated in terms of the homomorphisms of fundamental groups induced by the continuous maps $f: Z \to X$ and $p: \tilde{X} \to X$. We shall use this criterion in order to derive a necessary and sufficient condition for two covering maps over a connected and locally path-connected topological space to be topologically equivalent (see Theorem 5.6). We shall also study the *deck transformations* of a covering space over some connected and locally path-connected topological space. First we must define the concept of a *locally path-connected* topological space.

Definition Let Z be a topological space. The space Z is said to be *locally path-connected* if and only if, for every open subset U of Z and for each point u of U, there exists a path-connected open set N such that $u \in N$ and $N \subset U$ (i.e., the path-connected open subsets of Z form a base for the topology of Z).

Lemma 5.1 *Let X be a locally path-connected topological space and let $p: \tilde{X} \to X$ be a covering map over X. Then the covering space $\tilde{X}$ is also locally path-connected.*

Proof Let U be an open set in $\tilde{X}$, and let u be a point of U. The covering map $p: \tilde{X} \to X$ is a local homeomorphism, by Lemma 3.1. Therefore there exists an open set V containing the point u such that V is mapped homeomorphically onto some open set in X. Thus $p(U \cap V)$ is an open set in X which contains the point $p(u)$. But X is a locally path-connected topological space. Therefore there exists a path-connected open set N in X for which $p(u) \in N$ and $N \subset p(U \cap V)$. Let $\tilde{N} = V \cap p^{-1}(N)$. Then $\tilde{N}$ is an path-connected open subset of $U \cap V$ (since the restriction $p|V: V \to p(V)$ of the map p to V is a homeomorphism from V to $p(V)$). Moreover $u \in \tilde{N}$ and $\tilde{N} \subset U$. This shows that the covering space $\tilde{X}$ is locally path connected. ∎

Theorem 5.2 *Let X be a connected, locally path-connected topological space. Then X is path-connected.*

Proof Choose a point x_0 of X. Let Z be the subset of X consisting of all points x of X with the property that x can be joined to x_0 by a continuous path. We show that the subset Z is both open and closed in X.

Let x be a point of Z. There exists a path-connected open set N which contains the point x, since the topological space X is locally path-connected. Now every point of N can be joined to x by a continuous path, and the point x can be joined to x_0 by a continuous path. Thus every point of N can be joined to x_0 by a continuous path, so that $N \subset Z$. We conclude that Z is an open subset of X.

Now let x be a point of the complement $X \setminus Z$ of Z in X. Then there exists a path-connected open neighbourhood N of x. Now if there were to exist a point y of N which cound be joined to x_0 by a continuous path then we would also be able to join the point x to x_0 by a continuous path (since we could join x to y by a continuous path, and then continue this path to x_0). Thus the point x would belong to Z. But this contradicts the choice of the point x. We conclude therefore that $N \cap Z = \emptyset$. We have therefore shown that every point of the complement $X \setminus Z$ of Z in X has an open neighbourhood which is disjoint from Z. Thus $X \setminus Z$ is open in X, so that Z is closed in X.

Note that x_0 belongs to the set Z. Thus Z is a non-empty subset of X which is both open and closed. But X is connected. Therefore $Z = X$. Thus every point of X can be joined to the point x_0 by a continuous path. This shows that X is path-connected, as required. ∎

Corollary 5.3 *Let X be a locally path-connected topological space, and let $p: \tilde{X} \to X$ be a covering map over X. Suppose that the covering space $\tilde{X}$ is connected. Then $\tilde{X}$ is path-connected.*

Proof The covering space $\tilde{X}$ is locally path-connected, by Lemma 5.1. Therefore $\tilde{X}$ is connected, by Theorem 5.2. ∎

Theorem 5.4 *Let $p: \tilde{X} \to X$ be a covering map over a topological space X, and let $f: Z \to X$ be a continuous map from some topological space Z into X. Suppose that the topological space Z is both connected and locally path-connected. Suppose also that*

$$f_\# \left(\pi_1(Z, z_0)\right) \subset p_\# \left(\pi_1(\tilde{X}, \tilde{x}_0)\right),$$

where z_0 and $\tilde{x}_0$ are points of Z and $\tilde{X}$ respectively which satisfy $f(z_0) = p(\tilde{x}_0)$. Then there exists a unique map $\tilde{f}: Z \to \tilde{X}$ for which $\tilde{f}(z_0) = \tilde{x}_0$ and $p \circ \tilde{f} = f$.

Proof The topological space Z is path-connected, by Theorem 5.2. Let z be a point of Z, and let α and β be paths in Z from z_0 to z. Then $f \circ \alpha$ and $f \circ \beta$ are paths in X from $p(\tilde{x}_0)$ to $f(z)$ (where $p(\tilde{x}_0) = f(z_0)$). It follows from the Path Lifting Property for covering maps (Theorem 3.4) that there exist unique paths $\rho: [0,1] \to \tilde{X}$ and $\sigma: [0,1] \to \tilde{X}$ in $\tilde{X}$ such that $\rho(0) = \sigma(0) = \tilde{x}_0$, $p \circ \rho = f \circ \alpha$ and $p \circ \sigma = f \circ \beta$. Now $[(f \circ \alpha).(f \circ \beta)^{-1}] = f_{\#}[\alpha.\beta^{-1}]$, and $f_{\#}(\pi_1(Z, z_0)) \subset p_{\#}\left(\pi_1(\tilde{X}, \tilde{x}_0)\right)$. It follows from Lemma 3.8 that $\rho(1) = \sigma(1)$. Thus there is a well-defined map $\tilde{f}: Z \to \tilde{X}$ characterized by the following property:

> if $\alpha: [0,1] \to Z$ is a path in Z from z_0 to z, and if $\rho: [0,1] \to \tilde{X}$ is the unique path in $\tilde{X}$ for which $\rho(0) = \tilde{x}_0$ and $p \circ \rho = f \circ \alpha$, then $\tilde{f}(z) = \rho(1)$.

Clearly $p \circ \tilde{f} = f$. Thus it only remains to show that the map $\tilde{f}: Z \to \tilde{X}$ is continuous.

Let z be a point of Z. Then there exists an open neighbourhood V of $f(z)$ in X which is evenly covered by the map p. The inverse image $p^{-1}(V)$ of V is a disjoint union of open sets in $\tilde{X}$, each of which is mapped homeomorphically onto N by p. One of these open sets contains the point $\tilde{f}(z)$. Let us denote this open set by $\tilde{V}$. Thus $\tilde{f}(z) \in \tilde{V}$, and $\tilde{V}$ is mapped homeomorphically onto V by the map p. Let $s: V \to \tilde{V}$ denote the inverse of $(p|\tilde{V}): \tilde{V} \to V$. The map s is continuous, and $p(s(v)) = v$ for all $v \in V$.

Now $f^{-1}(V)$ is an open set in Z containing the point z. But the topological space Z is locally path-connected. Therefore there exists a path-connected open set N in Z such that $z \in N$ and $N \subset f^{-1}(V)$. We claim that $\tilde{f}(N) \subset \tilde{V}$. Let n be a point of N. Let $\gamma: [0,1] \to N$ be a path in N from z to n. Let $\eta: [0,1] \to \tilde{V}$ be the path in $\tilde{V}$ defined by $\eta = s \circ f \circ \gamma$. Then $p \circ \eta = f \circ \gamma$. Moreover $p(\eta(0)) = f(z) = p(\tilde{f}(z))$, and hence $\eta(0) = \tilde{f}(z)$. It follows easily from the definition of the map $\tilde{f}$ that $\tilde{f}(n) = \eta(1)$. Thus $\tilde{f}(n) \in \tilde{V}$ for all $n \in N$, so that $\tilde{f}(N) \subset \tilde{V}$. But $p \circ \tilde{f} = f$. Therefore $\tilde{f}|N = s \circ (f|N)$. Thus the restriction $\tilde{f}|N$ of the map $\tilde{f}$ to the open neighbourhood N of z is a composition of the continuous maps $f|N$ and s, and is thus itself continuous. This shows that the map $\tilde{f}$ is continuous at each point z of Z, as required. ∎

Let $p: \tilde{X} \to X$ be a covering map over some topological space X, and let x_0 be a point of X. Suppose that the covering space $\tilde{X}$ is path-connected. Lemma 3.11 states that if $\tilde{x}_1$ and $\tilde{x}_2$ are points of $\tilde{X}$ for which $p(\tilde{x}_1) = x_0 = p(\tilde{x}_2)$ then the subgroups $p_{\#}\left(\pi_1(\tilde{X}, \tilde{x}_1)\right)$ and $p_{\#}\left(\pi_1(\tilde{X}, \tilde{x}_2)\right)$ of $\pi_1(X, x_0)$ are conjugate. Moreover if H is any subgroup of $\pi_1(X, x_0)$ which is conjugate to

$p_{\#}\left(\pi_1(\tilde{X}, \tilde{x}_1\right)$ then there exists some point $\tilde{x}$ of $\tilde{X}$ for which $p(\tilde{x}) = x_0$ and $p_{\#}\left(\pi_1(\tilde{X}, \tilde{x}\right) = H$. This result can be combined with Theorem 5.4 to yield the following result.

Corollary 5.5 *Let $p: \tilde{X} \to X$ be a covering map over a topological space X, and let $f: Z \to X$ be a continuous map from some topological space Z into X. Suppose that the covering space $\tilde{X}$ is path-connected and that the topological space Z is both connected and locally path-connected. Suppose also that there exist points z_0 and $\tilde{x}_0$ of Z and $\tilde{X}$ respectively for which $f(z_0) = p(\tilde{x}_0)$ and $f_{\#}(\pi_1(Z, z_0)) \subset H$, where H is some subgroup of $\pi_1(X, p(\tilde{x}_0))$ which is conjugate to $p_{\#}\left(\pi_1(\tilde{X}, \tilde{x}_0)\right)$. Then there exists a map $\tilde{f}: Z \to \tilde{X}$ for which $p \circ \tilde{f} = f$.*

Proof Let H be a subgroup conjugate to $p_{\#}\left(\pi_1(\tilde{X}, \tilde{x}_0)\right)$ for which

$$f_{\#}(\pi_1(Z, z_0)) \subset H.$$

It follows from Lemma 3.11 that there exists a point $\tilde{x}$ of $\tilde{X}$ for which $p(\tilde{x}) = p(\tilde{x}_0)$ and $p_{\#}\left(\pi_1(\tilde{X}, \tilde{x})\right) = H$. But then

$$f_{\#}(\pi_1(Z, z_0)) \subset p_{\#}\left(\pi_1(\tilde{X}, \tilde{x})\right).$$

It follows from Theorem 5.4 that there exists a continuous map $\tilde{f}: Z \to \tilde{X}$ for which $p \circ \tilde{f} = f$, as required. ∎

Definition Let $p_1: \tilde{X}_1 \to X$ and $p_2: \tilde{X}_2 \to X$ be covering maps over some topological space X. We say that the covering maps $p_1: \tilde{X}_1 \to X$ and $p_2: \tilde{X}_2 \to X$ are *topologically isomorphic* if and only if there exists a homeomorphism $h: \tilde{X}_1 \to \tilde{X}_2$ from the covering space $\tilde{X}_1$ to the covering space $\tilde{X}_2$ with the property that $p_1 = p_2 \circ h$.

We can apply Theorem 5.4 in deriving a criterion for determining whether or not two covering maps over some connected locally path-connected topological space are isomorphic.

Theorem 5.6 *Let X be a topological space which is both connected and locally path-connected, and let $p_1: \tilde{X}_1 \to X$ and $p_2: \tilde{X}_2 \to X$ be covering maps over X. Let x_0 be a point of X, and let $\tilde{x}_1$ and $\tilde{x}_2$ be elements of $\tilde{X}_1$ and $\tilde{X}_2$ respectively for which $p_1(\tilde{x}_1) = x_0 = p_2(\tilde{x}_2)$. Suppose that the covering spaces $\tilde{X}_1$ and $\tilde{X}_2$ are both connected. Then the covering maps $p_1: \tilde{X}_1 \to X$ and $p_2: \tilde{X}_2 \to X$ are topologically isomorphic if and only if the subgroups $p_{1\#}\left(\pi_1(\tilde{X}_1, \tilde{x}_1)\right)$ and $p_{2\#}\left(\pi_1(\tilde{X}_2, \tilde{x}_2)\right)$ of $\pi_1(X, x_0)$ are conjugate.*

Proof It follows from Lemma 5.3 that the covering spaces $\tilde{X}_1$ and $\tilde{X}_2$ are both locally path-connected, since X is a locally path-connected topological space. But $\tilde{X}_1$ and $\tilde{X}_2$ are both connected. Therefore $\tilde{X}_1$ and $\tilde{X}_2$ are path-connected (see Corollary 5.3).

Suppose that the covering maps $p_1: \tilde{X}_1 \to X$ and $p_2: \tilde{X}_2 \to X$ are topologically isomorphic. Let $h: \tilde{X}_1 \to \tilde{X}_2$ be a homeomorphism for which $p_2 \circ h = p_1$. Then

$$p_{1\#}\left(\pi_1(\tilde{X}_1, \tilde{x}_1)\right) = p_{2\#}\left(\pi_1(\tilde{X}_2, h(\tilde{x}_1))\right).$$

It follows immediately from Lemma 3.11 that the subgroups $p_{1\#}\left(\pi_1(\tilde{X}_1, \tilde{x}_1)\right)$ and $p_{2\#}\left(\pi_1(\tilde{X}_2, \tilde{x}_2)\right)$ of $\pi_1(X, x_0)$ are conjugate.

Conversely, suppose that the subgroups

$$p_{1\#}\left(\pi_1(\tilde{X}_1, \tilde{x}_1)\right) \quad \text{and} \quad p_{2\#}\left(\pi_1(\tilde{X}_2, \tilde{x}_2)\right)$$

of $\pi_1(X, x_0)$ are conjugate. It then follows from Lemma 3.11 that there exists a point w of $\tilde{X}_2$ for which $p_2(w) = x_0$ and

$$p_{1\#}\left(\pi_1(\tilde{X}_1, \tilde{x}_1)\right) = p_{2\#}\left(\pi_1(\tilde{X}_2, w)\right).$$

But the covering spaces $\tilde{X}_1$ and $\tilde{X}_2$ are connected and locally path-connected. It follows from Theorem 5.4 that there exist unique continuous maps $h_1: \tilde{X}_1 \to \tilde{X}_2$ and $h_2: \tilde{X}_2 \to \tilde{X}_1$ for which $p_2 \circ h_1 = p_1$, $p_1 \circ h_2 = p_2$, $h_1(\tilde{x}_1) = w$ and $h_2(w) = \tilde{x}_1$. But then $p_1 \circ h_2 \circ h_1 = p_1$ and $(h_2 \circ h_1)(\tilde{x}_1) = \tilde{x}_1$. It follows from this that the composition map $h_2 \circ h_1$ is the identity map of $\tilde{X}_1$ (since it follows from Theorem 5.4 that the identity map of $\tilde{X}_1$ is the unique map from $\tilde{X}_1$ to itself with the required properties). Similarly the composition map $h_1 \circ h_2$ is the identity map of $\tilde{X}_2$. Thus $h_1: \tilde{X}_1 \to \tilde{X}_2$ is a homeomorphism whose inverse is h_2. Moreover $p_2 \circ h_1 = p_2$. Thus the covering maps $p_1: \tilde{X}_1 \to X$ and $p_2: \tilde{X}_2 \to X$ are topologically isomorphic, as required. ∎

5.1 Deck Transformations

Definition Let $p: \tilde{X} \to X$ be a covering map over a topological space X. A *deck transformation* of the covering space $\tilde{X}$ is a homeomorphism $h: \tilde{X} \to \tilde{X}$ of $\tilde{X}$ with the property that $p \circ h = p$.

The deck transformations of some covering space form a group of homeomorphisms of that covering space (where the group operation is the usual operation of composition of homeomorphisms).

We now define the notion of a *normal subgroup* of a group.

Definition Let G be a group, and let H be a subgroups of G. The subgroup H is said to be a *normal subgroup* of G if and only if ghg^{-1} belongs to H for all $h \in H$ and $g \in G$.

We recall that if H is a subgroup of a group G then gHg^{-1} is a subgroup of G for all $g \in G$, where

$$gHg^{-1} \equiv \{h' \in G : h' = ghg^{-1} \text{ for some } h \in H\}.$$

If H is a normal subgroup of G then $gHg^{-1} \subset H$ for all $g \in G$. We now show that if H is a normal subgroup of G then $gHg^{-1} = H$ for all $g \in G$.

Let H be a normal subgroup of G. Choose an element g of G. If h is an element of H then $h = g(g^{-1}hg)g^{-1}$. But $g^{-1}hg$ is an element of H (since H is a normal subgroup of G). Therefore $h \in gHg^{-1}$. This shows that, for any element g of G, $H \subset gHg^{-1}$. But we have already noted that $gHg^{-1} \subset H$. Therefore $gHg^{-1} = H$, as required. We see therefore that if H is a normal subgroup of G then the only subgroup of G which is conjugate to H is H itself.

Theorem 5.7 *Let X be a topological space which is connected and locally path-connected, let $p\colon \tilde{X} \to X$ be a covering map over X, where the covering space $\tilde{X}$ is connected. Let x_0 be a point of X and let $\tilde{x}_1$ and $\tilde{x}_2$ be points of the covering space $\tilde{X}$ for which $p(\tilde{x}_1) = x_0 = p(\tilde{x}_2)$. Suppose that $p_\#\left(\pi_1(\tilde{X}, \tilde{x}_1)\right)$ is a normal subgroup of $\pi_1(X, x_0)$. Then there exists a unique deck transformation $h\colon \tilde{X} \to \tilde{X}$ such that $h(\tilde{x}_1) = \tilde{x}_2$.*

Proof We see from Lemma 3.11 that the subgroups $p_\#\left(\pi_1(\tilde{X}, \tilde{x}_1)\right)$ and $p_\#\left(\pi_1(\tilde{X}, \tilde{x}_2)\right)$ of $\pi_1(X, x_0)$ are conjugate. But $p_\#\left(\pi_1(\tilde{X}, \tilde{x}_1)\right)$ is a normal subgroup of G, and therefore

$$p_\#\left(\pi_1(\tilde{X}, \tilde{x}_1)\right) = p_\#\left(\pi_1(\tilde{X}, \tilde{x}_2)\right)$$

(see the remarks above). Also the covering space $\tilde{X}$ is locally simply connected (see Lemma 5.1) and is also connected. It follows from Theorem 5.4 that there exists a unique continuous map $h\colon \tilde{X} \to \tilde{X}$ such that $h(\tilde{x}_1) = \tilde{x}_2$. Similarly there exists a unique continuous map $h'\colon \tilde{X} \to \tilde{X}$ such that $h'(\tilde{x}_2) = \tilde{x}_1$, and moreover the composition maps $h' \circ h$ and $h \circ h'$ are both equal to the identity map of $\tilde{X}$ (since the identity map of $\tilde{X}$ is the unique map from $\tilde{X}$ to itself which respects the covering map $p\colon \tilde{X} \to X$ and which fixes some given point of $\tilde{X}$). Thus h' is the inverse of h, so that the map $h\colon \tilde{X} \to \tilde{X}$ is the required homeomorphism of $\tilde{X}$ which sends $\tilde{x}_1$ to $\tilde{x}_2$. ∎

Let $p: \tilde{X} \to X$ be a covering map over some topological space X which is both connected and locally path-connected. Let us denote the group of deck transformations of the covering space $\tilde{X}$ by $\mathrm{Deck}(\tilde{X}|X)$. (Thus each element of the group $\mathrm{Deck}(\tilde{X}|X)$ is a homeomorphism h of $\tilde{X}$ for which $p \circ h = p$.) Let x_0 be a point of X, and let $\tilde{x}_0$ be a point of the covering space $\tilde{X}$ for which $p(\tilde{x}_0) = \tilde{x}$. Suppose that $\tilde{X}$ is connected and that $p_\#\left(\pi_1(\tilde{X}, \tilde{x}_0)\right)$ is a normal subgroup of $\pi_1(X, x_0)$. We now construct a function λ from the fundamental group $\pi_1(X, x_0)$ of X to the group $\mathrm{Deck}(\tilde{X}|X)$ of deck transformations of the covering space $\tilde{X}$.

Let $\gamma: [0,1] \to X$ be a (continuous) loop in X based at the point x_0. Then there exists a unique (continuous) path $\tilde{\gamma}: [0,1] \to \tilde{X}$ in $\tilde{X}$ for which $\tilde{\gamma}(0) = \tilde{x}_0$ and $p \circ \tilde{\gamma} = \gamma$, by the Path Lifting Property for covering maps (Theorem 3.4). Moreover it follows from Corollary 3.6 that the endpoint $\tilde{\gamma}(1)$ of the path $\tilde{\gamma}$ is determined by the element $[\gamma]$ of $\pi_1(X, x_0)$ represented by the loop γ (i.e., if we are given two loops based at the point x_0 which represent the same element of $\pi_1(X, x_0)$ then the lifts of these two loops to the covering space $\tilde{X}$ have the same endpoint). Moreover $p(\tilde{\gamma}(1)) = \gamma(1) = x_0$. It therefore follows from Theorem 5.7 that there exists a unique deck transformation $h_{[\gamma]} \in \mathrm{Deck}(\tilde{X}|X)$ for which $h_{[\gamma]}(\tilde{x}_0) = \tilde{\gamma}(1)$. (Here we have used the fact that $\tilde{X}$ is connected and $p_\#\left(\pi_1(\tilde{X}, \tilde{x}_0)\right)$ is a normal subgroup of $\pi_1(X, x_0)$.) Let

$$\lambda: \pi_1(X, x_0) \to \mathrm{Deck}(\tilde{X}|X)$$

denote the function which sends an element $[\gamma]$ of $\pi_1(X, x_0)$ to the deck transformation $h_{[\gamma]}: \tilde{X} \to \tilde{X}$.

Lemma 5.8 *Let $p: \tilde{X} \to X$ be a covering map over some topological space X which is both connected and locally path-connected. Suppose that $\tilde{X}$ is connected and that $p_\#\left(\pi_1(\tilde{X}, \tilde{x}_0)\right)$ is a normal subgroup of $\pi_1(X, x_0)$ (where $x_0 \in X$ and $\tilde{x}_0 \in \tilde{X}$ are chosen such that $p(\tilde{x}_0) = x_0$). Let*

$$\lambda: \pi_1(X, x_0) \to \mathrm{Deck}(\tilde{X}|X)$$

be the function from the fundamental group $\pi_1(X, x_0)$ of X to the group $\mathrm{Deck}(\tilde{X}|X)$ of deck transformations of the covering space $\tilde{X}$ constructed in the manner described above. Then λ is a homomorphism. Moreover this homomorphism has the following properties:

(i) *the image of the homomorphism λ is the whole of the group $\mathrm{Deck}(\tilde{X}|X)$ (i.e., the homomorphism λ is surjective),*

(ii) *the kernel of the homomorphism λ is $p_\#\left(\pi_1(\tilde{X}, \tilde{x}_0)\right)$.*

Proof Let α and β be loops in X based at the point x_0, and let $[\alpha]$ and $[\beta]$ be the elements of $\pi_1(X, x_0)$ represented by the loops α and β. Then $[\alpha][\beta]$ is represented by the product loop $\alpha.\beta$, where

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

Let $\tilde{\alpha}$ and $\tilde{\beta}$ be the unique paths in $\tilde{X}$ for which $\tilde{\alpha}(0) = \tilde{x}_0 = \tilde{\beta}(0)$, $p \circ \tilde{\alpha} = \alpha$ and $p \circ \tilde{\beta} = \beta$. Then

$$\tilde{\alpha}(1) = h_{[\alpha]}(\tilde{\beta}(0))$$

and $p \circ h_\alpha \circ \tilde{\beta} = \beta$. Let $\sigma\colon [0,1] \to \tilde{X}$ be the product path obtained by concatenating the paths $\tilde{\alpha}$ and $h_{[\alpha]} \circ \tilde{\beta}$. Thus

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

Then $\sigma(0) = \tilde{x}_0$ and $p \circ \sigma = \alpha.\beta$. We conclude that

$$\lambda([\alpha][\beta])(\tilde{x}_0) = h_{[\alpha.\beta]}(\tilde{x}_0) = \sigma(1) = h_{[\alpha]}(\tilde{\beta}(1)) = h_{[\alpha]}(h_{[\beta]}(\tilde{x}_0))$$

But the deck transformation $\lambda([\alpha][\beta])$ is uniquely determined by the value of $\lambda([\alpha][\beta])(\tilde{x}_0)$. We conclude therefore that $\lambda([\alpha][\beta]) = h_{[\alpha]} \circ h_{[\beta]} = \lambda([\alpha]) \circ \lambda([\beta])$ This shows that the function $\lambda\colon \pi_1(X, x_0) \to \operatorname{Deck}(\tilde{X}|X)$ is a homomorphism.

Let $h\colon \tilde{X} \to \tilde{X}$ be a deck transformation of $\tilde{X}$. The covering space $\tilde{X}$ is path-connected. Thus there exists a path $\eta\colon [0,1] \to \tilde{X}$ in $\tilde{X}$ from $\tilde{x}$ to $h(\tilde{x}_0)$. But then $p \circ \eta\colon [0,1] \to X$ is a loop in X based at x_0 and thus represents some element $[p \circ \eta]$ of $\pi_1(X, x_0)$. It then follows from the definition of the homomorphism λ that $h = \lambda([p \circ \eta])$. This shows that the homomorphism λ is surjective. This proves (i).

Let γ be a loop in X based at x_0 which represents an element of the kernel of $\lambda\colon \pi_1(X, x_0) \to \operatorname{Deck}(\tilde{X}|X)$. Then $h_{[\gamma]}$ is the identity map of $\tilde{X}$, and thus $h_{[\gamma]}(\tilde{x}_0) = \tilde{x}_0$. Thus if $\tilde{\gamma}$ is the unique path in $\tilde{X}$ for which $\tilde{\gamma}(0) = \tilde{x}_0$ and $p \circ \tilde{\gamma} = \gamma$ then $\tilde{\gamma}(1) = \tilde{x}_0$, so that $\tilde{\gamma}$ is a loop in $\tilde{X}$ based at the point $\tilde{x}_0$, and $[\gamma] = p_\#[\tilde{\gamma}]$. Thus $\ker \lambda \subset p_\#\left(\pi_1(\tilde{X}, \tilde{x}_0)\right)$. Conversely if σ is a loop in $\tilde{X}$ based at the point $\tilde{x}_0$ then

$$\lambda(p_\#[\sigma])(\tilde{x}_0) = h_{[p \circ \sigma]}(\tilde{x}_0) = \sigma(1) = \tilde{x}_0.$$

It follows that $\lambda(p_\#[\sigma])$ is the identity map of $\tilde{X}$, since the deck transformation $\lambda(p_\#[\sigma])$ is uniquely determined by $\lambda(p_\#[\sigma])(\tilde{x}_0)$. Thus

$$p_\# \left(\pi_1(\tilde{X}, \tilde{x}_0)\right) \subset \ker \lambda.$$

We conclude therefore that $\ker \lambda = p_\# \left(\pi_1(\tilde{X}, \tilde{x}_0)\right)$. This proves (ii). ∎

6 Topological Classification of Covering Maps over Locally Simply-Connected Topological Spaces

Let X be a topological space. We shall show that there is a bijective correspondence between covering maps over X and conjugacy classes of subgroups of the fundamental group of X, provided that the space X satisfies certain topological conditions. (More precisely, we require the topological space X to be *connected* and *locally simply-connected*. (The concept of a locally simply-connected topological space is defined below.) Suppose that X satisfies these conditions. Let x_0 be some chosen basepoint of X, and let H be a subgroup of the fundamental group $\pi_1(X, x_0)$ of X at x_0. We shall construct a path-connected topological space $\tilde{X}_H$ and a covering map $p_H: \tilde{X}_H \to X$ over X with the property that the fundamental group of the covering space $\tilde{X}_H$ is isomorphic to the chosen subgroup H of $\pi_1(X, x_0)$.

In particular, we show that if X is a connected and locally simply-connected topological space then there exists a covering map $p: \tilde{X} \to X$ for which that the covering space $\tilde{X}$ is simply-connected. Such a covering map is said to be a *universal covering map* for the topological space X, and the covering space $\tilde{X}$ is said to be a *universal covering space* for X.

We recall that a topological space X *locally path-connected* if and only if, for each open subset U of X and for each point u of U, there exists a path-connected open set N such that $u \in N$ and $N \subset U$ (i.e., the path-connected open subsets of X form a base for the topology of X). We have shown that if a topological space X is both connected and locally path-connected then it is path-connected (see Theorem 5.2).

We next define the concept of a *locally simply-connected* topological space.

Definition Let X be a topological space. The space X is said to be *locally simply-connected* if and only if, for every open subset U of X and for each point u of U, there exists a simply-connected open set N such that $u \in N$ and $N \subset U$ (i.e., the simply-connected open subsets of X form a base for the topology of X).

A topological space X is said to be *locally Euclidean* of dimension n if and only if every point of X has an open neighbourhood which is homeomorphic to $\mathbb{R}^n$. Every topological space which is locally Euclidean of dimension n for some positive integer n is clearly locally simply-connected.

Let X be a topological space, and let x_0 be point of X. Let H be a subgroup of the fundamental group $\pi_1(X, x_0)$ of X based at the point x_0. Let

α and β be paths in X starting at the point x_0. Suppose that $\alpha(1) = \beta(1)$. Then the path $\alpha.\beta^{-1}$ is a loop in X based at the point x_0; thus it represents an element $[\alpha.\beta^{-1}]$ of $\pi_1(X, x_0)$. (Here $\alpha.\beta^{-1}$ denotes the loop based at x_0 obtained by starting at the point x_0, going out along the path α to the point $\alpha(1)$ and then returning to x_0 along the path β in the reverse direction; it is defined by

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

We define a relation $\sim_H$ on the set of all paths in X starting at the point x_0, where $\alpha \sim_H \beta$ if and only if $\alpha(1) = \beta(1)$ and $[\alpha.\beta^{-1}]$ belongs to the subgroup H of $\pi_1(X, x_0)$. We claim that the relation $\sim_H$ is an equivalence relation.

If α is a path starting at the point x_0 then the loop $\alpha.\alpha^{-1}$ represents the identity element of $\pi_1(X, x_0)$; therefore $\alpha \sim_H \alpha$. Thus the relation $\sim_H$ is reflexive.

If α and β are paths in X starting at the point x_0 and if $\alpha \sim_H \beta$ then $\alpha(1) = \beta(1)$ and $[\alpha.\beta^{-1}]$ belongs to H. But $[\beta.\alpha^{-1}] = [\alpha.\beta^{-1}]^{-1}$, and thus $[\beta.\alpha^{-1}]$ also belongs to H (since H is a subgroup of $\pi_1(X, x_0)$. Hence $\beta \sim_H \alpha$. Thus the relation $\sim_H$ is symmetric.

Let α, β and γ be paths in X starting at x_0 such that $\alpha \sim_H \beta$ and $\beta \sim_H \gamma$. Then $\alpha(1) = \beta(1) = \gamma(1)$, $[\alpha.\beta^{-1}] \in H$ and $[\beta.\gamma^{-1}] \in H$. But

$$\alpha.\gamma^{-1} \simeq \alpha.\beta^{-1}.\beta.\gamma^{-1} \text{ rel } \{0, 1\};$$

thus $[\alpha.\gamma^{-1}] = [\alpha.\beta^{-1}][\beta.\gamma^{-1}]$, and hence $[\alpha.\gamma^{-1}]$ belongs to H. This shows that if $\alpha \sim_H \beta$ and $\beta \sim_H \gamma$ then $\alpha \sim_H \gamma$. Thus the relation $\sim_H$ is transitive.

We have shown that the relation $\sim_H$ is reflexive, symmetric and transitive. We therefore conclude that it is an equivalence relation on the set of all paths in X starting at the point x_0. We denote by $\langle\gamma\rangle_H$ the equivalence class of a path γ in X starting at the point x_0.

Let $\tilde{X}_H$ be the set of all equivalence classes of paths in X starting at the point x_0 (with respect to the equivalence relation $\sim_H$). Thus each element of $\tilde{X}_H$ is of the form $\langle\gamma\rangle_H$ for some path γ in X starting at the point x_0. If α and β are paths in X starting at x_0 then $\langle\alpha\rangle_H = \langle\beta\rangle_H$ if and only if $\alpha(1) = \beta(1)$ and $[\alpha.\beta^{-1}]$ belongs to the subgroup H of $\pi_1(X, x_0)$. Thus there is a well-defined map $p_H: \tilde{X}_H \to X$ defined such that $p_H(\langle\gamma\rangle_H) = \gamma(1)$ for all paths $\gamma: [0, 1] \to X$ in X satisfying $\gamma(0) = x_0$.

We must define a topology on the set $\tilde{X}_H$. Let U be an open set in X, and let $\gamma: [0, 1] \to X$ be a path in X for which $\gamma(0) = x_0$ and $\gamma(1) \in U$. We denote by $\langle\gamma, U\rangle_H$ the set of all elements of $\tilde{X}_H$ that are of the form $\langle\gamma.\alpha\rangle_H$, where $\alpha\colon [0,1] \to U$ is a path in the open set U for which $\alpha(0) = \gamma(1)$. (Here $\gamma.\alpha$ denotes the concatenation of the paths γ and α obtained by first following the path γ from x_0 to $\gamma(1)$ and then following the path α from $\gamma(1)$ to $\alpha(1)$. It is defined by

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

Observe that $p_H(\langle\gamma, U\rangle_H) \subset U$. We define a topology on the set $\tilde{X}_H$ such that a subset V of $\tilde{X}_H$ is open if and only if, for all elements $\langle\gamma\rangle_H$ of V, there exists an open set N in X for which $\gamma(1) \in N$ and $\langle\gamma, N\rangle_H \subset V$. In order to verify that this topology on $\tilde{X}_H$ is well-defined, we must show that the empty set $\emptyset$ and the whole set $\tilde{X}_H$ are both open sets in $\tilde{X}_H$ (with respect to this topology) and that unions of open sets and finite intersections of open sets are themselves open sets.

The empty set $\emptyset$ is an open set in $\tilde{X}_H$ (since the criterion defining open sets becomes vacuous in the case of the empty set). The set $\tilde{X}_H$ itself is an open set, since X is open in X and $\langle\gamma, X\rangle_H \subset \tilde{X}_H$ for each element $\langle\gamma\rangle_H$ of $\tilde{X}_H$ Also it follows directly from the definition of open sets in $\tilde{X}_H$ that any union of open sets in $\tilde{X}_H$ is itself an open set.

Let $V_1, V_2, \ldots, V_r$ be a finite collection of open sets in $\tilde{X}_H$, and let $\langle\gamma\rangle_H$ be an element of the intersection $V_1 \cap V_2 \cap \cdots \cap V_r$ of these open sets. Then $\langle\gamma\rangle_H$ is represented by some path $\gamma\colon [0,1] \to X$ starting at the point x_0. There exist open sets $N_1, N_2, \ldots, N_r$ in X such that $\gamma(1) \in N_j$ and $\langle\gamma, N_j\rangle_H \subset V_j$ for $j = 1, 2, \ldots, r$. Define

$$N = N_1 \cap N_2 \cap \cdots \cap N_r.$$

Then $\gamma(1) \in N$ and $\langle\gamma, N\rangle_H \subset V_1 \cap V_2 \cap \cdots V_r$. This shows that $V_1 \cap V_2 \cap \cdots V_r$ is an open set in $\tilde{X}_H$. We conclude therefore that the topology on $\tilde{X}_H$ described above is indeed well-defined.

Next we show that the map $p_H\colon \tilde{X}_H \to X$ defined above is continuous. Let U be an open set in X. We must show that the inverse image $p_H^{-1}(U)$ of U is an open set in $\tilde{X}_H$. Let $\langle\gamma\rangle_H$ be an element of $p_H^{-1}(U)$ which is represented by some path γ in X starting at the point x_0. Then $p_H(\langle\gamma, U\rangle_H) \subset U$, and hence $\langle\gamma, U\rangle_H \subset p_H^{-1}(U)$. This shows that $p_H^{-1}(U)$ is an open set in $\tilde{X}_H$. Thus the map $p_H\colon \tilde{X}_H \to X$ is continuous.

Lemma 6.1 *Let X be a topological space, let x_0 be a chosen basepoint of X, let H be a subgroup of $\pi_1(X, x_0)$, and let $\tilde{X}_H$ be the topological space constructed in the manner described above. Then $\tilde{X}_H$ is path-connected.*

Proof Let $\tilde{x}_0$ be the element of $\tilde{X}_H$ which represents the constant map at the point x_0 of X. If $\tilde{x}$ is any other element of $\tilde{X}_H$ then $\tilde{x} = \langle \gamma \rangle_H$ for some path $\gamma\colon [0,1] \to X$ for which $\gamma(0) = x_0$ and $\gamma(1) = p_H(\tilde{x})$. For each $\tau \in [0,1]$ let $\gamma_\tau\colon [0,1] \to X$ be the path in X defined by $\gamma_\tau(t) = \gamma(\tau t)$. Let $\tilde{\gamma}\colon [0,1] \to \tilde{X}_H$ be the map from $[0,1]$ to $\tilde{X}_H$ given by $\tilde{\gamma}(\tau) = \langle \gamma_\tau \rangle_H$ for all $\tau \in [0,1]$. It is not difficult to verify that the map $\tilde{\gamma}$ is continuous. Thus $\tilde{\gamma}$ is a path in $\tilde{X}_H$ from $\tilde{x}_0$ to $\tilde{x}$. This shows that the topological space $\tilde{X}_H$ is path-connected. ∎

Lemma 6.2 *Let X be a topological space, let x_0 be a chosen basepoint of X, let H be a subgroup of $\pi_1(X, x_0)$, and let $\tilde{X}_H$ and $p_H\colon \tilde{X}_H \to X$ be the topological space and the continuous map constructed in the manner described above. Let α and β be paths in X starting at x_0, and let U be an open set in X containing the endpoints $\alpha(1)$ and $\beta(1)$ of the paths α and β. Suppose that $\langle \beta \rangle_H$ belongs to $\langle \alpha, U \rangle_H$. Then $\langle \alpha, U \rangle_H = \langle \beta, U \rangle_H$. Thus $\langle \alpha, U \rangle_H$ is an open set in $\tilde{X}_H$.*

Proof Let $\langle \beta \rangle_H$ be an element of $\langle \alpha, N \rangle_H$. Then there exists a path $\eta\colon [0,1] \to U$ in U such that $\eta(0) = \alpha(1)$, $\eta(1) = \beta(1)$ and $\beta \sim_H \alpha.\eta$. Let $\langle \beta.\sigma \rangle_H$ be an element of $\langle \beta, U \rangle_H$, where $\sigma\colon [0,1] \to U$ is some path in U for which $\sigma(0) = \beta(1)$. It follows from Lemma 2.1 that $(\alpha.\eta).\sigma \simeq \alpha.(\eta.\sigma)$ rel $\{0,1\}$. Thus $\beta.\sigma \sim_H \alpha.(\eta.\sigma)$. Moreover $\eta.\sigma$ is a path in U from $\alpha(1)$ to $\sigma(1)$. Thus $\langle \beta.\sigma \rangle_H$ belongs to $\langle \alpha, U \rangle_H$. This shows that $\langle \beta, U \rangle_H \subset \langle \alpha, U \rangle_H$.

Now $\langle \alpha \rangle_H \in \langle \beta, U \rangle_H$, since

$$\alpha \sim_H (\alpha.\eta).\eta^{-1} \sim_H \beta.\eta^{-1}$$

(where $\eta^{-1}\colon [0,1] \to U$ is the path in U defined by $\eta^{-1}(t) = \eta(1-t)$ for all $t \in [0,1]$). It then follows from the result already proved (with the roles of α and β interchanged) that $\langle \alpha, U \rangle_H \subset \langle \beta, U \rangle_H$. Therefore $\langle \alpha, U \rangle_H = \langle \beta, U \rangle_H$, as required. It follows immediately from this result and the definition of the topology of $\tilde{X}_H$ that $\langle \alpha, U \rangle$ is an open set in $\tilde{X}_H$. ∎

We now prove that the map $p_H\colon \tilde{X}_H \to X$ is a covering map, provided that the topological space X is connected and locally simply-connected.

Theorem 6.3 *Let X be a topological space, let x_0 be a chosen basepoint of X, let H be a subgroup of $\pi_1(X, x_0)$, and let $\tilde{X}_H$ and $p_H\colon \tilde{X}_H \to X$ be the topological space and the continuous map constructed in the manner described above. Suppose that the topological space X is connected and locally simply-connected. Then the map $p_H\colon \tilde{X}_H \to X$ is a covering map.*

Proof First we must show that the map $p_H\colon \tilde{X}_H \to X$ is a surjection. The topological space X is connected and path-connected, and thus is path-connected, by Theorem 5.2. Thus if x is a point of X then there exists some path γ in X from x_0 to x. But then

$$x = \gamma(1) = p_H(\langle\gamma\rangle_H).$$

This shows that the map $p_H\colon \tilde{X}_H \to X$ is a surjection.

Let x be a point of X. Then there exists a simply-connected open set N which contains the point x (since X is a locally simply-connected topological space). We shall show that the open set N is evenly covered by the map p_H.

Let $\tilde{y}$ be an element of $p_H^{-1}(N)$. There exists a path $\eta\colon [0,1] \to N$ in N with $\eta(0) = x$ and $\eta(1) = p(\tilde{y})$, since N is path-connected. Also $\tilde{y} = \langle\gamma\rangle_H$ for some path γ in X starting at the point x_0. Let $\alpha\colon [0,1] \to X$ be the path from x_0 to x given by $\alpha = \gamma.\eta^{-1}$. Then $p_H(\langle\alpha\rangle_H) = x$ and and $\langle\gamma\rangle_H \in \langle\alpha, N\rangle_H$. Thus every element of $p_H^{-1}(N)$ belongs to $\langle\alpha, N\rangle$ for some element $\langle\alpha\rangle_H$ of $p_H^{-1}(\{x\})$.

Let y be a point of N and let α be a path in X from x_0 to x. There exists a path η in N from x to y (since the open set N is path-connected). Thus $y = p_H(\tilde{y})$, where $\tilde{y}$ is the element of $\langle\alpha, N\rangle_H$ given by $\tilde{y} = \langle\alpha.\eta\rangle$. Thus $p_H(\langle\alpha, N\rangle_H) = N$ for all $\langle\alpha\rangle_H \in p_H^{-1}(\{x\})$.

Now let $\tilde{y}_1$ and $\tilde{y}_2$ be elements of $\langle\alpha, N\rangle$, where α is some path in X from x_0 to x. Then there exist paths η_1 and η_2 in N starting at the point x such that $\tilde{y}_1 = \langle\alpha.\eta_1\rangle_H$ and $\tilde{y}_2 = \langle\alpha.\eta_2\rangle_H$. Suppose that $p_H(\tilde{x}_1) = p_H(\tilde{x}_2)$. Then $\eta_1(1) = \eta_2(1)$, so that $\eta_1.\eta_2^{-1}$ is a loop in N based at the point x. But this loop represents the identity element of $\pi_1(X, x)$ (since the open set N is simply-connected). It follows from this that the loop $\alpha.\eta_1.\eta_2^{-1}.\alpha^{-1}$ based at the point x_0 represents the identity element of $\pi_1(X, x_0)$, and hence $\langle\alpha.\eta_1\rangle_H = \langle\alpha.\eta_2\rangle_H$ We have therefore shown that if $\tilde{y}_1$ and $\tilde{y}_2$ are elements of $\langle\alpha, N\rangle$ for which $p_H(\tilde{y}_1) = p_H(\tilde{y}_2)$, then $\tilde{y}_1 = \tilde{y}_2$. Since we have already shown that $p_H(\langle\alpha, N\rangle_H) = N$ we conclude that the continuous map p_H maps $\langle\alpha, N\rangle_H$ bijectively onto N for all $\langle\alpha\rangle_H \in p_H^{-1}(\{x\})$. Thus in order to prove that p_H maps $\langle\alpha, N\rangle_H$ homeomorphically onto N it only remains to show that $p_H(V)$ is open in X for every open subset V of $\langle\alpha, N\rangle_H$.

Let V be an open subset of $\langle\alpha, N\rangle$, where α is some path in X from x_0 to x. Let y be an element of $p_H(V)$. There exists a unique element $\tilde{y}$ of V such that $p_H(\tilde{y}) = y$. Moreover $\tilde{y} = \langle\gamma\rangle_H$ for some path γ in X from x_0 to y. It follows from the definition of the topology of $\tilde{X}_H$ that there exists some open set W_1 in X with the property that $\langle\gamma, W_1\rangle_H \subset V$. Moreover there exists a path-connected open set W in X such that $y \in W$ and $W \subset W_1$, since X is a locally path-connected topological space. But then $\langle\gamma, W\rangle_H \subset V$ and $p_H(\langle\gamma, W\rangle_H) = W$ (since W is path-connected). Hence $W \subset p_H(V)$. We conclude from this that $p_H(V)$ is an open set in X. This completes the proof of the fact that, for each path α in X from x_0 to x, the open subset $\langle \alpha, N\rangle$ of $\tilde{X}$ is mapped homeomorphically onto N by the map p_H.

Let α and β be paths in X from x_0 to x. Suppose that $\langle \alpha, N\rangle_H \cap \langle \beta, N\rangle_H$ is non-empty. Let $\langle \gamma\rangle_H$ be an element of $\langle \alpha, N\rangle_H \cap \langle \beta, N\rangle_H$ Then

$$\langle \alpha, N\rangle_H = \langle \gamma, N\rangle_H = \langle \beta, N\rangle_H,$$

by Lemma 6.2. In particular $\langle \beta\rangle_H$ belongs to $\langle \alpha, N\rangle_H$. But $p_H(\langle \beta\rangle_H) = x = p_H(\langle \beta\rangle_H)$, and we have already shown that $\langle \alpha, N\rangle_H$ is mapped bijectively onto N by the map p_H. Therefore $\langle \beta\rangle_H = \langle \alpha\rangle_H$. We conclude therefore that if $\langle \alpha\rangle_H$ and $\langle \beta\rangle_H$ are distinct elements of $p_H^{-1}(\{x\})$ then the open sets $\langle \alpha, N\rangle_H$ and $\langle \beta, N\rangle_H$ are disjoint.

It follows from all the results proved above that, for each point x of X there exists an open neighbourhood N of x such that $p^{-1}(N)$ is a disjoint union of open sets of the form $\langle \alpha, N\rangle$ for some path α in X from x_0 to x. Each of these open sets is mapped homeomorphically onto N by the map p_H. We have also shown that the map $p_H\colon \tilde{X}_H \to X$ is surjective. We conclude therefore that the map $p_H\colon \tilde{X}_H \to X$ is a covering map, as required. ∎

Remark Theorem 6.3 can easily be generalized to cover the case when the topological space X is connected, locally path-connected and semilocally simply-connected. (A topological space X is said to be *semilocally simply-connected* if and only if, for each point x of X, there exists an open neighbourhood N of x with the property that the homomorphism $i_\#\colon \pi_1(N, x) \to \pi_1(X, x)$ induced by the inclusion map $i\colon N \hookrightarrow X$ is the trivial homomorphism which sends each element of $\pi_1(N, x)$ to the identity element of $\pi_1(X, x)$. Every locally simply-connected topological space is both locally path-connected and semilocally simply-connected.) Indeed if X is both locally path-connected and semilocally simply-connected then one can easily show that, for each point x of X, there exists an open neighbourhood N of X with the following properties:

(i) every point of N can be joined to x by a continuous path in N,

(ii) if γ_1 and γ_2 are paths in N from x to some point x' of N then the loop $\gamma_1.\gamma_2^{-1}$ is contractible in X (i.e., this loop represents the identity element of $\pi_1(X, x)$).

One can then prove that the open neighbourhood N of X is evenly covered by the map $p_H\colon \tilde{X}_H \to X$, exactly as in the proof of Theorem 6.3.

Definition Let X and $\tilde{X}$ be topological spaces, and let $p: \tilde{X} \to X$ be a covering map. Suppose that $\tilde{X}$ is simply-connected. Then we say that the space $\tilde{X}$ is a *universal covering space* of X and that the covering map $p: \tilde{X} \to X$ is a *universal covering map* over the space X.

Corollary 6.4 *Let X be a topological space which is connected and locally simply-connected. Then there exists a covering map $p: \tilde{X} \to X$ over X for which the covering space $\tilde{X}$ is simply-connected (i.e., there exists a universal covering map over X).*

Proof This follows directly from Theorem 6.3 in the case when the subgroup H of the fundamental group of X is the trivial subgroup consisting of the identity element of the group. ∎

Let X be a topological space which is connected and locally simply-connected. Theorem 6.3 and Theorem 5.6 provide us with a topological classification of covering maps over the topological space X. Indeed let us choose some basepoint x_0 for the space X. Then to each subgroup H of the fundametal group $\pi_1(X, x_0)$ of X at the point x_0 there corresponds a covering map $p_H: \tilde{X} \to X$ for which

$$p_{H\#}\left(\pi_1(\tilde{X}_H, \tilde{x}_0)\right) = H$$

(where $\tilde{x}_0$ is some appropriate basepoint for the covering space $\tilde{X}_H$). Every covering map over X can be constructed in this fashion, and moreover the covering maps $p_H: \tilde{X}_H \to X$ and $p_K: \tilde{X}_K \to X$ corresponding to subgroups H and K of $\pi_1(X, x_0)$ are topologically isomorphic if and only if the subgroups H and K are conjugate in $\pi_1(X, x_0)$. Thus there is a one-to-one correspondence between covering maps over X and conjugacy classes of subgroups of $\pi_1(X, x_0)$.

A Review of Point Set Topology

Point set topology is the study of topological spaces. We present here a reasonably comprehensive survey of the basic definitions and important theorems of point set topology.

Definition A *topological space* X consists of a set X together with a collection of subsets, referred to as *open sets*, such that the following conditions are satisfied:

(i) the empty set $\emptyset$ and the whole set X are open sets,

(ii) the union of any collection of open sets is itself an open set,

(iii) the intersection of any *finite* collection of open sets is itself an open set.

The collection consisting of all open sets in a topological space X is referred to as a *topology* on the set X, and the elements of X are usually called *points*.

Remark If it is necessary to specify explicitly the topology on a topological space then one denotes by (X, τ) the topological space whose underlying set is X and whose topology is τ. However if no confusion will arise then it is customary to denote this topological space simply by X.

Definition Let X be a topological space. A subset F of X is said to be a *closed set* if and only if its complement $X \setminus F$ is an open set.

Lemma A.1 *Let X be a topological space. Then the collection of closed sets of X has the following properties:*

(i) *the empty set $\emptyset$ and the whole set X are closed sets,*

(ii) *the intersection of any collection of closed sets is itself a closed set,*

(iii) *the union of any finite collection of closed sets is itself a closed set.*

Proof These three properties follow from the corresponding properties of open sets stated in the definition of a topological space, in view of the fact that the intersection of the complements of some collection of subsets of X is equal the complement of the union of that collection of subsets. ∎

Definition Let X be a topological space and let A be a subset of X. The *closure* $\overline{A}$ of A in X is defined to be the intersection of all of the closed subsets of X that contain A. The *interior* A^0 of A in X is defined to be the union of all of the open subsets of X that are contained in A.

Lemma A.2 *Let X be a topological space and let A be a subset of X. Then the closure $\overline{A}$ of A is uniquely chararcterized by the following two properties:*

(i) *$\overline{A}$ is a closed set containing A,*

(ii) *if F is any other closed set containing A then $\overline{A} \subset F$.*

Similarly the interior A^0 of A is uniquely characterized by the following two properties:

(i) *A^0 is an open set contained in A,*

(ii) *if U is any other open set contained in A then $U \subset A^0$.*

Proof Note that the closure $\overline{A}$ of A contains A since it is the intersection of some non-empty collection of subsets of X each of which contains the set A. (One of these subsets is X itself.) Moreover the closure of A is a closed set, since any intersection of closed sets is closed. It follows directly from the definition of the closure $\overline{A}$ of A that if F is a closed subset of X containing A then $\overline{A} \subset F$.

The interior A^0 is an open set contained in A since it is the union of a collection of open sets contained in A, and any union of open sets is itself an open set. It follows directly from the definition of the interior A^0 of A that if U is an open subset of X which is contained in A then $U \subset A^0$. ∎

Lemma A.3 *Let X be a topological space, let A be a subset of X, and let U be an open subset of X. Suppose that $\overline{A} \cap U$ is non-empty (where $\overline{A}$ is the closure of A). Then $A \cap U$ is non-empty.*

Proof Suppose that $A \cap U = \emptyset$. We shall show that this implies that $\overline{A} \cap U = \emptyset$. If $A \cap U = \emptyset$ then $A \subset X \setminus U$ (where $X \setminus U$ is the complement of U in X). But $X \setminus U$ is closed (since U is open). Therefore $\overline{A} \subset X \setminus U$, and hence $\overline{A} \cap U = \emptyset$, as required. We deduce from this that if $\overline{A} \cap U$ is non-empty then $A \cap U$ must also be non-empty. ∎

Definition Let X be a topological space, and let x be a point of X. Let N be a subset of X which contains the point x. Then N is said to be a *neighbourhood* of the point x if and only if there exists an open set U with the property that $x \in U$ and $U \subset N$.

It follows immediately from this definition that if V is an open set in some topological space X and if x is some point of V then V is a neighbourhood of x.

Lemma A.4 *Let X be a topological space, and let V be a subset of X. Then V is open if and only if V is a neighbourhood of every point that belongs to V.*

Proof We have already observed that if V is an open set in X then V is a neighbourhood of every point that belongs to V. Conversely suppose that V is a neighbourhood of every point that belongs to V. We must show that V is an open set.

For each point v of V there exists an open set W_v such that $v \in W_v$ and $W_v \subset V$ (since V is a neighbourhood of V. Then $\bigcup_{v \in V} W_v \subset V$ (where $\bigcup_{v \in V} W_v$ is the union of the open sets W_v for all points v of V. But clearly $V \subset \bigcup_{v \in V} W_v$ Therefore $V = \bigcup_{v \in V} W_v$. Thus V is equal to a union of open sets, and hence V is itself open (since any union of open sets is itself an open set). ∎

A.1 Subspace Topologies

Let X be a topological space and let A be a subset of X. Let τ be the topology of X (i.e., τ is the collection consisting of all open subsets of X). We define a topology τ_A on the set A, where τ_A consists of all subsets of A which are of the form $V \cap A$ for some open subset V of X. The topology τ_A on A is referred to as the *subspace topology* on A. In this way every subset of a topological space can be regarded as a topological space in its own right.

A subset B of A is closed relative to the subspace topology on A if and only if $A \setminus B$ is open relative to the subspace topology on A. Thus B is closed relative to the subspace topology if and only if $B = A \setminus (A \cap V)$ for some open subset V of X. But $A \setminus (A \cap V) = A \cap (X \setminus V)$. It follows that a subset B of A is closed relative to the subspace topology on A if and only if $B = A \cap F$ for some closed subset F of A.

Now suppose that A is an open set in X. Let U be a subset of A which is open with respect to the subspace topology on A. Then there exists an open subset V of X such that $U = V \cap A$. But the intersection of two open subsets of X is itself an open subset of X. Thus U is an open subset of X. Thus if A is itself an open subset of X then a subset U of A is open with respect to the subspace topology on A if and only if U is an open subset of X.

Similarly suppose that A is a closed set in X. Let B be a subset of A which is closed with respect to the subspace topology on A. Then there exists a closed subset F of X such that $B = F \cap A$. But the intersection of two closed subsets of X is itself a closed subset of X. Thus B is a closed subset of X. Thus if A is itself a closed subset of X then a subset B of A is closed with respect to the subspace topology on A if and only if B is a closed subset of X.

A.2 Continuous Maps

Definition Let X and Y be topological spaces. A map $f: X \to Y$ is said to be *continuous* if and only if $f^{-1}(V)$ is an open set in X for every open set V in Y (where

$$f^{-1}(V) \equiv \{x \in X : f(x) \in V\}).$$

Lemma A.5 *Let X, Y and Z be topological spaces, and let $f: X \to Y$ and $g: Y \to Z$ be continuous maps. Then the composition $g \circ f: X \to Z$ of the maps f and g is continuous.*

Proof Let U be an open set in Z. Then $g^{-1}(U)$ is open in Y (since g is continuous), and hence $f^{-1}(g^{-1}(U))$ is open in X (since f is continuous). But $f^{-1}(g^{-1}(U) = (g \circ f)^{-1}(U)$. Thus the composition map $g \circ f$ is continuous. ∎

Lemma A.6 *Let X and Y be topological spaces, and let $f: X \to Y$ be a map from X to Y. The map f is continuous if and only if $f^{-1}(A)$ is closed in X for every closed subset A of Y.*

Proof Let B be a subset of Y. A point x of X belongs to the complement $X \setminus f^{-1}(B)$ in X of $f^{-1}(B)$ if and only if $f(x)$ belongs to the complement $Y \setminus B$ in Y of the subset B. Thus $X \setminus f^{-1}(B) = f^{-1}(Y \setminus B)$.

Suppose that the map $f: X \to Y$ is continuous. If A is a closed subset of Y then $Y \setminus A$ is an open set in Y, and hence $f^{-1}(Y \setminus A)$ is an open subset of X. But the complement in X of this open set is $f^{-1}(A)$. Therefore $f^{-1}(A)$ is closed. We have therefore shown that if the map f is continuous then $f^{-1}(A)$ is closed in X for every closed subset A of Y.

Conversely suppose that $f^{-1}(A)$ is closed in X for every closed subset A of Y. Let U be an open set in Y. Then $Y \setminus U$ is a closed set, and hence $f^{-1}(Y \setminus U)$ is a closed set in X. But $f^{-1}(Y \setminus U) = X \setminus f^{-1}(U)$. Thus $f^{-1}(U)$ is the complement in X of a closed subset of X and is thus open. We have therefore shown that $f^{-1}(U)$ is an open subset of X for every open subset U of Y. Thus the map $f: X \to Y$ is continuous. ∎

Lemma A.7 *Let X be a topological space and let $Z_1, Z_2, \dots, Z_r$ be a finite collection of closed subsets of X such that*

$$Z_1 \cup Z_r \cup \dots \cup Z_r = X.$$

Let Y be a topological space, and let $f: X \to Y$ be a map from X to Y. Suppose that the restriction $f|Z_i: Z_i \to Y$ of f to Z_i is continuous for $i = 1, 2, \dots, r$. Then $f: X \to Y$ is continuous.

Proof It suffices to show that $f^{-1}(A)$ is closed in X for every closed subset A of Y, by Lemma A.6. Let A be a closed subset of Y. Then $f^{-1}(A) \cap Z_i$ is closed in Z_i relative to the subspace topology on Z_i for $i = 1, 2, \ldots, r$, since $f|Z_i$ is continuous for all i. However a subset of Z_i is closed relative to the subspace topology on Z_i if and only if it is a closed subset of X, since Z_i is itself a closed subset of X. Thus $f^{-1} \cap Z_i$ is a closed subset of X for $i = 1, 2, \ldots, r$. But

$$f^{-1}(A) = \bigcup_{i=1}^{r} f^{-1}(A) \cap Z_i.$$

Thus $f^{-1}(A)$ is a finite union of closed sets, so that $f^{-1}(A)$ is itself closed. We conclude from Lemma A.6 that $f\colon X \to Y$ is continuous. ∎

Situations frequently arise where we have a collection of functions $f_i\colon Z_i \to Y$ defined over closed subsets $Z_1, Z_2, \ldots, Z_r$ of some topological space X, where $X = Z_1 \cup Z_2 \cup \cdots \cup Z_r$. It follows from Lemma A.7 that if $f_i\colon Z_i \to Y$ is continuous for $i = 1, 2, \ldots, r$, and if $f_i(z) = f_j(z)$ for all $z \in Z_i \cap Z_j$, then we can piece together the functions f_i in order to obtain a continuous function $f\colon X \to Y$ (where f is defined such that $f|Z_i = f_i$ for $i = 1, 2, \ldots, r$).

Definition Let X and Y be topological spaces. A map $f\colon X \to Y$ is said to be a *homeomorphism* if and only if the following conditions are satisfied:

(i) the map $f\colon X \to Y$ is both injective and surjective (so that the map $f\colon X \to Y$ has a well-defined inverse $f^{-1}\colon Y \to X$),

(ii) the map $f\colon X \to Y$ and its inverse $f^{-1}\colon Y \to X$ are both continuous.

If there exists a homeomorphism $f\colon X \to Y$ from the topological space X to the topological space Y then the topological spaces X and Y are said to be *homeomorphic*.

If $f\colon X \to Y$ is a homeomorphism between topological spaces X and Y then f induces a one-to-one correspondence between the open sets of X and the open sets of Y. Thus the topological spaces X and Y can be regarded as being identical as topological spaces.

A.3 Metric Spaces

Definition A *metric space* (X, d) consists of a set X together with a *distance function* $d\colon X \times X \to [0, +\infty)$ on X, where this distance function satisfies the following axioms:

(i) $d(x, y) \geq 0$ for all $x, y \in X$,

(ii) $d(x, y) = d(y, x)$ for all $x, y \in X$,

(iii) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$,

(iv) $d(x, y) = 0$ if and only if $x = y$.

The quantity $d(x, y)$ should be thought of as measuring the *distance* between the points x and y. The inequality $d(x, z) \leq d(x, y) + d(y, z)$ is referred to as the *Triangle Inequality.*

Let (X, d) be a metric space. A subset V of X is said to be an *open set* if and only if the following condition is satisfied:

> given any point v of V there exists some $\delta > 0$ such that $B(v, \delta) \subset V$, where
> $$B(v, \delta) \equiv \{x \in X : d(x, v) < \delta\}.$$

The set $B(v, \delta)$ is referred to as the *open ball of radius δ about v.*

If (X, d) is a metric space, then the collection of open sets of X constitutes a topology on the set X.

A.4 Hausdorff Spaces

Definition A topological space X is said to be a *Hausdorff space* if and only if it satisfies the following *Hausdorff Axiom*:

> if x and y are distinct points of X then there exist open sets U and V such that $x \in U$, $y \in V$ and $U \cap V = \emptyset$.

Every metric space is a Hausdorff space. Moreover every subset of a Hausdorff space is itself a Hausdorff space (with respect to the subspace topology).

A.5 Discrete and Connected Topological Spaces

Definition A topological space X is said to be *discrete* if and only if every subset of X is an open set.

Definition A topological space X is said to be *connected* if and only if the empty set $\emptyset$ and the whole space X are the only subsets of X that are both open and closed.

Lemma A.8 *A topological space X is connected if and only if the following condition is satisfied:*

> *if U and V are non-empty open sets in X such that $X = U \cup V$ then $U \cap V$ is non-empty.*

Proof First suppose that X is connected. Let U and V be non-empty open sets in X such that $U \cup V = X$. Suppose that it were the case that $U \cap V = \emptyset$. Then $V = X \setminus U$, so that U would be both open and closed. But this is impossible, since U and $X \setminus U$ are non-empty, and the only subsets of X that are both open and closed are $\emptyset$ and X. Thus $U \cap V$ must be non-empty.

Conversely suppose that $U \cap V$ is non-empty whenever U and V are non-empty open sets in X for which $X = U \cup V$. Let A be a non-empty subset of X which is both open and closed. Then $X \setminus A$ is open. Moreover X is the union of the open sets A and $X \setminus A$, and $A \cap (X \setminus A) = \emptyset$. It follows from this that $X \setminus A = \emptyset$, and hence $A = X$. Thus the only subsets of X that are both open and closed are $\emptyset$ and X. Thus X is connected. ∎

Let X be a topological space, and let A be a subset of X. It follows from the definition of connectedness that A is connected (with respect to the subspace topology) if and only if the following condition is satisfied:

> if U and V are open sets in X such that $A \cap U$ and $A \cap V$ are non-empty and $A \subset U \cup V$ then $A \cap U \cap V$ is also non-empty.

Lemma A.9 *Let X be a topological space and let A be a connected subset of A. Then the closure $\overline{A}$ of A is connected.*

Proof Let U and V be open sets in X such that $\overline{A} \cap U$ and $\overline{A} \cap V$ are non-empty and $\overline{A} \subset U \cup V$. Then $A \cap U$ and $A \cap V$ are non-empty, by Lemma A.8. Hence $A \cap U \cap V$ is non-empty, since A is connected and $A \subset U \cup V$. Thus $\overline{A} \cap U \cap V$ is non-empty. This shows that $\overline{A}$ is connected. ∎

Lemma A.10 *Let X and Y be topological spaces, and let $f: X \to Y$ be a continuous map. Let A be a connected subset of X. Then $f(A)$ is a connected subset of Y.*

Proof Let U and V be open sets in Y such that $f(A) \cap U$ and $f(A) \cap V$ are non-empty and $f(A) \subset U \cup V$. Then $A \cap f^{-1}(U)$ and $A \cap f^{-1}(V)$ are non-empty and $A \subset f^{-1}(U) \cup f^{-1}(V)$. But $f^{-1}(U)$ and $f^{-1}(V)$ are open sets in X (since f is continuous), hence $A \cap f^{-1}(U) \cap f^{-1}(V)$ is non-empty, since A is connected. Therefore $f(A) \cap U \cap V$ is non-empty. This shows that $f(A)$ is connected. ∎

Lemma A.11 *Let X be a connected topological space and let Y be a discrete topological space. Then every continuous map $f: X \to Y$ is a constant map (i.e., there exists some element y_0 of Y such that $f(x) = y_0$ for all $x \in X$).*

Proof Let y be a point of Y. Then the set $\{y\}$ is both open and closed (since every subset of a discrete topological space is both open and closed). But then $f^{-1}\{y\}$ is both open and closed (since $f: X \to Y$ is continuous). But the topological space X is connected. Thus either $f^{-1}(\{y\}) = \emptyset$ or else $f^{-1}(\{y\}) = X$. It follows from this that the map $f: X \to Y$ must be a constant map, as required. ∎

In particular, we can consider continuous maps from a topological space X into the the set $\mathbb{Z}$ of integers. This gives us the following criterion for connectedness.

Lemma A.12 *Let X be a topological space. Then X is connected if and only if every continuous map $f: X \to \mathbb{Z}$ from X to $\mathbb{Z}$ is constant.*

Proof The set $\mathbb{Z}$ of integers is a discrete topological space. Thus if X is connected then any continuous map $f: X \to \mathbb{Z}$ is constant, by Lemma A.11. Conversely suppose that every continuous function from X to $\mathbb{Z}$ is constant. Let A be a non-empty subset of X which is both open and closed. Let $f: X \to \mathbb{Z}$ be defined by

$$f(x) = \begin{cases} 1 & \text{if } x \in A; \\ 0 & \text{if } x \in X \setminus A. \end{cases}$$

Then f is continuous, since the sets A and $X \setminus A$ are open. Therefore the function f is constant. But A is non-empty, hence $X \setminus A = \emptyset$, so that $A = X$. Thus $\emptyset$ and X are the only subsets of X that are both open and closed. ∎

Lemma A.13 *Let X be a topological space, and let $\mathcal{V}$ be a collection of connected subsets of X. Suppose that $V \cap W$ is non-empty for all $V, W \in \mathcal{V}$ Then the union of all the sets belonging to $\mathcal{V}$ is connected.*

Proof Let U be the union of all the sets belonging to $\mathcal{V}$, and let $f: U \to \mathbb{Z}$ be a continuous map from U to $\mathbb{Z}$. Given any set V belonging to the collection $\mathcal{V}$ there exists an integer m_V such that $f(v) = m_V$ for all $v \in V$, since V is connected. Moreover if V and W are subsets of X belonging to $\mathcal{V}$ then $m_V = m_W$, since $V \cap W$ is non-empty. Thus there exists some integer m such that $m_V = m$ for all $V \in \mathcal{V}$. Hence $f(u) = m$ for all $u \in U$, where U is the union of all the sets belonging to $\mathcal{U}$. Therefore every continuous function from U to $\mathbb{Z}$ is constant, and hence U is connected, by Lemma A.12. ∎

Definition Let X be a topological space. A subset S of X is said to be a *connected component* of X if and only if the following two conditions are satisfied:

(i) S is connected,

(ii) if A is a connected subset of X which contains S then $A = S$.

Lemma A.14 *Let X be a topological space, and let S be a connected component of X. Then S is closed.*

Proof If S is connected then so is the closure $\overline{S}$ of S. It follows that $S = \overline{S}$. Thus S is closed. ∎

Lemma A.15 *Let X be a topological space and let x be a point of X. Then there exists a unique connected component of the set X which contains the point x.*

Proof Let S be the union of the collection consisting of all connected subsets of X that contain the point x. Then S is itself connected, by Lemma A.13. Moreover if A is a connected set which contains the set S then the point x belongs to A, hence $A \subset S$. Thus S is a connected component of the set S. Suppose that $\tilde{S}$ were another connected component of X which contains the point x. Then $S \cap \tilde{S}$ would be non-empty, hence $S \cup \tilde{S}$ would be connected, by Lemma A.13. But then $S \cup \tilde{S} \subset S$ and $S \cup \tilde{S} \subset \tilde{S}$ and hence $\tilde{S} = S$. This shows that the connected component of X containing the point x is unique, as required. ∎

We see from Lemma A.15 that if X is a topological space then X is the disjoint union of the connected components of X (i.e., X is the union of the connected components of X, and the intersection of any two distinct components of X is the empty set $\emptyset$).

A.6 Compact Topological Spaces

Definition Let X be a topological space. A collection $\mathcal{U}$ of open subsets of X is said to be an *open cover* of X if and only if every point of X belongs to at least one of these open sets. If $\mathcal{U}$ and $\mathcal{V}$ are open covers of the topological space X then $\mathcal{V}$ is said to be a *subcover* of $\mathcal{U}$ if and only if every subset of X which belongs to $\mathcal{V}$ also belongs to $\mathcal{U}$.

Definition A topological space X is said to be *compact* if and only if every open cover of X possesses a finite subcover.

Let X be a topological space and let A be a subset of X. Then A is compact (with respect to the subspace topology) if and only if the following condition is satisfied:

> if $\mathcal{U}$ is a collection of open sets in X such that each point of A belongs to at least one of the open sets in this collection, then there exists a finite collection $V_1, V_2, \ldots, V_r$ of open sets belonging to the collection $\mathcal{U}$ such that
>
> $$A \subset V_1 \cup V_2 \cup \cdots \cup V_r.$$

Lemma A.16 *Let X be a compact topological space, and let A be a closed subset of X. Then A is compact.*

Proof Let $\mathcal{U}$ be a collection of open sets in X such that each point of A belongs to one of the open sets belonging to $\mathcal{U}$. If we adjoin the open set $X \setminus A$ to the collection $\mathcal{U}$ then we obtain an open cover of the space X. This open cover possesses a finite subcover, since X is compact. In particular, there exists a finite collection $V_1, V_2, \ldots, V_r$ of open sets belonging to the collection $\mathcal{U}$ such that $A \subset V_1 \cup V_2 \cup V_r$, as required. ∎

Lemma A.17 *Let X and Y are topological spaces, and $f: X \to Y$ be a continuous map. Let A be a compact subset of X. Then $f(A)$ is a compact subset of Y.*

Proof Let $\mathcal{V}$ be a collection of open sets in Y which covers $f(A)$, and let $\mathcal{W}$ be the collection of open sets in X consisting of all open sets of the form $f^{-1}(V)$ for some $V \in \mathcal{V}$. If x is a point of A then $f(x)$ belongs to some open set V of the collection $\mathcal{V}$, since $\mathcal{V}$ covers $f(A)$. But then x belongs to $f^{-1}(V)$. Thus the collection $\mathcal{W}$ of open sets covers A. It follows from the compactness of A that there exist open sets $V_1, V_2, \ldots, V_r$ belonging to $\mathcal{V}$ such that

$$A \subset f^{-1}(V_1) \cup f^{-1}(V_2) \cup \cdots \cup f^{-1}(V_r).$$

But then

$$f(A) \subset V_1 \cup V_2 \cup \cdots \cup V_r.$$

Thus every open cover of $f(A)$ has a finite subcover, so that $f(A)$ is compact. ∎

Lemma A.18 *Let X be a Hausdorff topological space, and let K be a compact subset of X. Then K is closed. Moreover if x is a point of the complement $X \setminus K$ of K then there exist open subsets V_x and W_x of X such that $x \in V_x$, $K \subset W_x$ and $V_x \cap W_x = \emptyset$.*

Proof Let x be a point of $X \setminus K$. We shall show that there exist open sets V_x and W_x such that $x \in V_x$, $K \subset W_x$ and $V_x \cap W_x = \emptyset$. Now for each point y of K there exist open sets $V_{x,y}$ and $W_{x,y}$ such that $x \in V_{x,y}$, $y \in W_{x,y}$ and $V_{x,y} \cap W_{x,y} = \emptyset$ (since X is a Hausdorff space). But it then follows from the compactness of K that there exists a finite set $\{y_1, y_2, \ldots, y_r\}$ of points of K such that

$$K \subset W_{x,y_1} \cup W_{x,y_2} \cup \cdots \cup W_{x,y_r}.$$

Define

$$V_x = V_{x,y_1} \cap V_{x,y_2} \cap \cdots \cap V_{x,y_r}, \qquad W_x = W_{x,y_1} \cup W_{x,y_2} \cup \cdots \cup W_{x,y_r}.$$

Then V_x and W_x are open sets, x belongs to V_x, $K \subset W_x$ and $V_x \cap W_x = \emptyset$, as required. Note in particular that $V_x \subset X \setminus K$.

We have shown that, for each point x of $X \setminus K$, there exists an open set V_x such that $x \in V_x$ and $V_x \subset X \setminus K$ (i.e., $X \setminus K$ is a neighbourhood of the point x). It follows immediately from Lemma A.4 that $X \setminus K$ is open. Hence K is closed. ∎

Theorem A.19 *Let X and Y be topological spaces, and let $f\colon X \to Y$ be a continuous map which is both injective and surjective. If X is a compact space and if Y is Hausdorff space then $f\colon X \to Y$ is a homeomorphism.*

Proof Let K be a closed subset of the compact space X. Then K is compact, and hence $f(K)$ is compact. But then $f(K)$ is closed, since Y is Hausdorff. Thus the map f sends closed sets in X to closed sets in Y. Let $g\colon Y \to X$ be the inverse of $f\colon X \to Y$. Then $g^{-1}(K)$ is a closed subset of Y for every closed subset K of X (since $g^{-1}(K) = f(K)$). It follows from Lemma A.6 that the inverse g of f is continuous. Thus $f\colon X \to Y$ is a homeomorphism, as required. ∎

Lemma A.20 *Let X be a compact topological space and let $f\colon X \to \mathbb{R}$ be a continuous real-valued function on X. Then f is bounded above and below on X. Moreover f attains its upper and lower bounds on X (i.e., there exists points u and v of X such that $f(u) \leq f(x) \leq f(v)$ for all $x \in X$).*

Proof It follows from Lemma A.17 that the image $f(X)$ of the map f is a compact subset of $\mathbb{R}$. But then $f(X)$ is closed and bounded. (Indeed $f(X)$ can be covered by finitely many open sets of the form $\{t \in \mathbb{R} - m < t < m\}$, so that $f(X)$ must be bounded; $f(X)$ is closed by Lemma A.18, since $\mathbb{R}$ is Hausdorff.) It follows directly from this that the function f is bounded on X and attains its upper and lower bounds on X. ∎

Definition Let (X, d) be a metric space, and let A be a subset of X. The *diameter* of the set A is defined to be the supremum

$$\sup_{u,v \in A} d(u, v)$$

of the distance from the point u to the point v as u and v range over all the points of the set A. (If the distance $d(u, v)$ from u to v is not bounded above as u and v range over the set A then the diameter of A is defined to be $+\infty$.)

We now state and prove the *Lebesgue Lemma*.

Lemma A.21 (Lebesgue Lemma) *Let (X, d) be a compact metric space. Let $\mathcal{U}$ be an open cover of X. Then there exists a positive real number δ such that every subset of X whose diameter is less than δ is contained wholly within one of the open sets belonging to the open cover $\mathcal{U}$.*

Proof Every point of X is contained in at least one of the open sets belonging to the open cover $\mathcal{U}$. It follows from this that, for each point x of X, there exists some $\delta_x > 0$ such that the open ball $B(x, 2\delta_x)$ of radius $2\delta_x$ about the point x is contained wholly within one of the open sets belonging to the open cover $\mathcal{U}$. But then the collection consisting of the open balls $B(x, \delta_x)$ of radius δ_x about the points x of X forms an open cover of the compact space X. Therefore there exists a finite set $x_1, x_2, \ldots, x_r$ of points of X such that

$$B(x_1, \delta_1) \cup B(x_2, \delta_2) \cup \cdots \cup B(x_r, \delta_r) = X,$$

where $\delta_i = \delta_{x_i}$ for $i = 1, 2, \ldots, r$. Let $\delta > 0$ be given by

$$\delta = \text{minimum}(\delta_1, \delta_2, \ldots, \delta_r).$$

Suppose that A is a subset of X whose diameter is less than δ. Let u be a point of A. Then u belongs to $B(x_i, \delta_i)$ for some integer i between 1 and r. But then it follows that $A \subset B(x_i, 2\delta_i)$, since, for each point v of A,

$$d(v, x_i) \leq d(v, u) + d(u, x_i) < \delta + \delta_i \leq 2\delta_i.$$

But $B(x_i, 2\delta_i)$ is contained wholly within one of the open sets belonging to the open cover $\mathcal{U}$. Thus A is contained wholly within one of the open sets belonging to $\mathcal{U}$, as required. ∎

Let $\mathcal{U}$ be an open cover of a compact metric space X. A *Lebesgue number* for the open cover $\mathcal{U}$ is a positive real number δ such that every subset of X whose diameter is less than δ is contained wholly within one of the open sets belonging to the open cover $\mathcal{U}$. The Lebesgue Lemma thus states that there exists a Lebesgue number for every open cover of a compact metric space.

A.7 Product Spaces

Let $X_1, X_2, \ldots, X_r$ be a finite collection of topological spaces, and let $X_1 \times X_2 \times \cdots \times X_r$ be the Cartesian product of the sets $X_1, X_2, \ldots, X_r$. There is a topology on $X_1 \times X_2 \times \cdots \times X_r$ characterized by the following property:

> a subset U of $X_1 \times X_2 \times \cdots \times X_r$ is open if and only if, for each point $(u_1, u_2, \ldots, u_r)$ of U, there exist open sets V_i in X_i containing u_i for $i = 1, 2, \ldots, r$ such that $V_1 \times V_2 \times \cdots \times V_r \subset U$.

(Here $V_1 \times V_2 \times \cdots \times V_r$ denotes the subset of $X_1 \times X_2 \times \cdots \times X_r$ defined by

$$\begin{aligned} V_1 \times V_2 \times \cdots \times V_r &= \{(x_1, x_2, \ldots, x_r) \in X_1 \times X_2 \times \cdots \times X_r : \\ &\qquad x_i \in V_i \text{ for } i = 1, 2, \ldots, r\}.) \end{aligned}$$

One can easily verify that the characterization of open sets does indeed yield a well-defined topology on the product space $X_1 \times X_2 \times \cdots \times X_r$. This topology is referred to as the *product topology* on $X_1 \times X_2 \times \cdots \times X_r$.

Let $p_i: X_1 \times X_2 \times \cdots \times X_r \to X_i$ denote the projection map which sends a point $(x_1, x_2, \ldots, x_r)$ of $X_1 \times X_2 \times \cdots \times X_r$ to x_i. It follows directly from the definition of the product topology that $p_i^{-1}(U)$ is an open set in X_i for each open set U in X_i. Thus the projection map p_i is continuous for $i = 1, 2, \ldots, r$. It follows from this that if Z is a topological space and if $f: Z \to X_1 \times X_2 \times \cdots \times X_r$ is a continuous map then the composition map $p_i \circ f: Z \to X_i$ is continuous for $i = 1, 2, \ldots, r$ (since a composition of continuous maps is continuous by Lemma A.5).

Lemma A.22 *The Cartesian product $X_1 \times X_2 \times \cdots \times X_r$ of the topological spaces $X_1, X_2, \ldots, X_r$ satisfies the following universal property:*

> *given any topological space Z, and given any continuous maps $f_i: Z \to X_i$ from Z to X_i, for $i = 1, 2, \ldots, r$, there exists a unique continuous map $f: Z \to X_1 \times X_2 \times \cdots \times X_r$ such that $p_i \circ f = f_i$ for $i = 1, 2, \ldots, r$.*

Proof Let $X = X_1 \times X_2 \times \cdots \times X_r$. Suppose that we are given a topological space Z continuous maps $f_i: Z \to X_i$ from Z to X_i for $i = 1, 2, \ldots, r$. Let $f: Z \to X$ be the map defined by $f(z) = (f_1(z), f_2(z), \ldots, f_r(z))$. Then f is the unique map from Z to X with the property that $p_i \circ f = f_i$ for $i = 1, 2, \ldots, r$. We must prove that the map f is continuous.

Let U be an open set in X. We must show that $f^{-1}(U)$ is open in Z. Let z be a point of $f^{-1}(U)$, and let $u_i = f_i(z)$ for $i = 1, 2, \ldots, r$. It follows from the definition of the product topology on X that there exists an open set V_i in X_i containing u_i for $i = 1, 2, \ldots, r$ such that

$$V_1 \times V_2 \times \cdots \times V_r \subset U.$$

Let N_z be the subset of Z defined by

$$N_z = f_1^{-1}(V_1) \cap f_2^{-1}(V_2) \cap \cdots \cap f_r^{-1}(V_r).$$

Now $f_i^{-1}(V_i)$ is an open subset of Z for $i = 1, 2, \ldots, r$, since V_i is open in X_i and $f_i\colon Z \to X_i$ is continuous. Thus N_z is a finite intersection of open sets containing the point z, and thus N_z is itself an open set containing z. Moreover

$$f(N_z) \subset V_1 \times V_2 \times \cdots \times V_r \subset U,$$

so that $N_z \subset f^{-1}(U)$. We have thus shown that, for each point z of $f^{-1}(U)$, there exists an open set N_z in Z for which $z \in N_z$ and $N_z \subset f^{-1}(U)$ (i.e., $f^{-1}(U)$ is a neighbourhood of every point belonging to $f^{-1}(U)$). Therefore $f^{-1}(U)$ is an open set in Z (see Lemma A.4). This proves that the map $f\colon Z \to X$ is continuous. Thus the stated universal property is satisfied by the Cartesian product X of the topological spaces $X_1, X_2, \ldots, X_r$ and the projection maps $p_i\colon X \to X_i$ for $i = 1, 2, \ldots, r$. ∎

Lemma A.23 *Let $X_1, X_2, \ldots, X_r$ be Hausdorff spaces. Then the space $X_1 \times X_2 \times \ldots, X_r$ is Hausdorff.*

Proof Let $X = X_1 \times X_2 \times \ldots, X_r$, and let $(x_1, x_2, \ldots, x_r)$ and $(y_1, y_2, \ldots, y_r)$ be distinct points of X. Then $x_i \neq y_i$ for some integer i between 1 and r. But then there exists open sets U and V in X_i such that $x_i \in U$, $y_i \in V$ and $U \cap V = \emptyset$ (since X_i is a Hausdorff space). Let $p_i\colon X \to X_i$ denote the projection map. Then $p_i^{-1}(U)$ and $p_i^{-1}(V)$ are open sets in X, since p_i is continuous. Moreover $(x_1, x_2, \ldots, x_r)$ belongs to $p_i^{-1}(U)$, $(y_1, y_2, \ldots, y_r)$ belongs to $p_i^{-1}(V)$ and $p_i^{-1}(U) \cap p_i^{-1}(V) = \emptyset$. This shows that X is Hausdorff, as required. ∎

We now show that the Cartesian product of two connected topological spaces is connected. To prove this we use the fact that a topological space X is connected if and only if every continuous map from X to the set $\mathbb{Z}$ of integers is constant (see Lemma A.12).

Lemma A.24 *Let X and Y be connected topological spaces. Then $X \times Y$ is connected.*

Proof It suffices to show that every continuous function $f\colon X \times Y \to \mathbb{Z}$ is constant. Let $f\colon X \to \mathbb{Z}$ be a continuous function from X to $\mathbb{Z}$. Choose $x_0 \in X$ and $y_0 \in Y$. Then $f(x, y_0) = f(x_0, y_0)$ for all $x \in X$, since X is connected and the function $x \mapsto f(x, y_0)$ is continuous on X. But then $f(x, y) = f(x, y_0)$ for all $y \in Y$, since Y is connected. Hence $f(x, y) = f(x_0, y_0)$ for all $x \in X$ and $y \in Y$. Thus $f\colon X \times Y \to \mathbb{Z}$ is constant. It follows from Lemma A.12 that $X \times Y$ is connected. ∎

We deduce immediately that a finite Cartesian product of connected topological spaces is connected.

We shall show that a finite Cartesian product of compact spaces is compact. To prove this, we apply the following result, known as the *Tube Lemma*.

Lemma A.25 *Let X and Y be topological spaces, let K be a compact subset of Y, and U be an open set in $X \times Y$. Let V be the subset of X defined by*

$$V = \{x \in X : \{x\} \times K \subset U\}.$$

Then V is an open set in X.

Proof Let x be a point of V. For each point y of K there exist open subsets D_y and E_y of X and Y respectively such that $(x, y) \in D_y \times E_y$ and $D_y \times E_y \subset U$. But K is compact. Therefore there exists a finite set $\{y_1, y_2, \ldots, y_k\}$ of points of K such that

$$K = E_{y_1} \cup E_{y_2} \cup \cdots \cup E_{y_k}.$$

Set

$$N_x = D_{y_1} \cap D_{y_2} \cap \cdots \cap D_{y_k}.$$

Then N_x is an open set in X. Moreover

$$N_x \times K \subset \bigcup_{i=1}^{k} (D_{y_i} \times E_{y_i}) \subset U,$$

so that $N_x \subset V$. It follows that V is the union of the open sets N_x for all $x \in V$. Thus V is itself an open set in X, as required. ∎

Theorem A.26 *Let X and Y be compact topological spaces. Then $X \times Y$ is compact.*

Proof Let $\mathcal{U}$ be an open cover of $X \times Y$. We must show that this open cover possesses a finite subcover.

Let x be a point of X. The set $\{x\} \times Y$ is a compact subset of $X \times Y$, hence there exists a finite collection $U_1, U_2, \ldots, U_r$ of open sets belonging to the open cover $\mathcal{U}$ such that

$$\{x\} \times Y \subset U_1 \cup U_2 \cup \cdots \cup U_r.$$

Let

$$V_x = \{x' \in X : \{x'\} \times Y \subset U_1 \cup U_2 \cup \cdots \cup U_r\}.$$

It follows from Lemma A.25 that V_x is an open set in X. We have therefore shown that, for each point x in X, there exists an open set V_x in X containing the point x such that $V_x \times Y$ is covered by finitely many of the open sets belonging to the open cover $\mathcal{U}$.

Now $(V_x : x \in X)$ is an open cover of the compact space X. This cover possesses a finite subcover. Thus there exists a finite set $\{x_1, x_2, \ldots, x_r\}$ of points of X such that

$$X = V_{x_1} \cup V_{x_2} \cup \cdots \cup V_{x_r}.$$

It follows from this that $X \times Y$ can be covered by finitely many open sets belonging to the open cover $\mathcal{U}$ (since $X \times Y$ is a finite union of sets of the form $V_x \times Y$, and each of these sets can be covered by finitely many of the open sets belonging to $\mathcal{U}$). Therefore $X \times Y$ is compact. ∎

We deduce immediately that a finite product of compact topological spaces is compact.

A.8 Subsets of Euclidean Spaces

We regard the space $\mathbb{R}^n$ of n-tuples of real numbers as a metric space, where the Euclidean distance $|\mathbf{u} - \mathbf{v}|$ between points $\mathbf{u}$ and $\mathbf{v}$ of $\mathbb{R}^n$ is defined such that

$$|\mathbf{u} - \mathbf{v}|^2 = \sum_{j=1}^{n} (u_j - v_j)^2,$$

where $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$. One can verify that all the axioms of a metric space are satisfied. In particular

$$|\mathbf{u} - \mathbf{w}| \leq |\mathbf{u} - \mathbf{v}| + |\mathbf{v} - \mathbf{w}|$$

for all points $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ of $\mathbb{R}^n$. (This inequality is known as the *Triangle Inequality.*) A subset U of $\mathbb{R}^n$ is open if and only if, given any point $\mathbf{u}$ of U, there exists some $\delta > 0$ such that

$$\{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{u}| < \delta\} \subset U.$$

The topological space $\mathbb{R}^n$ is referred to as *n-dimensional Euclidean space.* This space is Hausdorff. Therefore any subset of $\mathbb{R}^n$ is a Hausdorff space (with respect to the subspace topology).

An important theorem of real analysis states that a subset K of $\mathbb{R}^n$ is compact if and only if K is both closed and bounded. We first prove this result in the particular case when K is a closed bounded interval in $\mathbb{R}$. In this case the result is known as the *Heine-Borel Theorem.* The proof of this theorem uses the *least upper bound principle* which states that any set S of real numbers which is bounded above possesses a *least upper bound* (or *supremum*) denoted by $\sup S$.

Theorem A.27 (Heine Borel) *Let a and b be real numbers satisfying $a < b$. Then the closed bounded interval $[a, b]$ is a compact subset of $\mathbb{R}$.*

Proof Let $\mathcal{V}$ be a collection of open sets in $\mathbb{R}$ with the property that each point of the interval $[a, b]$ belongs to one of these open sets. We must show that $[a, b]$ is covered by finitely many of these open sets (i.e., we must prove that there exists a finite collection of open sets $V_1, V_2, \ldots, V_r$ belonging to $\mathcal{V}$ such that

$$[a, b] \subset V_1 \cup V_2 \cup \cdots \cup V_r).$$

Let S be the subset of $[a, b]$ consisting of all $t \in [a, b]$ with the property that the closed interval $[a, t]$ is covered by finitely many of the open sets belonging to $\mathcal{V}$. Let s be the least upper bound $\sup S$ of the set S. Now a belongs to one of the open sets belonging to $\mathcal{V}$, and this open set will contain the interval $[a, t]$ provided that $t - a$ is sufficiently small. Therefore $s > a$.

There exists some open set W belonging to the collection $\mathcal{U}$ such that s belongs to W. Then there exists some $\delta > 0$ such that if $t \in [a, b]$ satisfies $s - \delta < t < s + \delta$ then t belongs to W. Choose some element t_0 of S such that $s - \delta < t_0 < s$. Then there exist open sets $V_1, V_2, \ldots, V_j$ belonging to $\mathcal{V}$ such that

$$[a, t_0] \subset V_1 \cup V_2 \cup \cdots \cup V_j,$$

since t_0 belongs to S. But then if $t \in [a, b]$ satisfies $t < s + \delta$ then

$$[a, t] \subset V_1 \cup V_2 \cup \cdots \cup V_j \cup W,$$

so that t belongs to S. We conclude from this that the least upper bound s of the set S belongs to S and that $s = b$. Thus the closed interval $[a, b]$ can be covered by finitely many of the open sets belonging to $\mathcal{V}$, as required. ∎

Using the Heine-Borel Theorem, we now prove that a subset K of $\mathbb{R}^n$ is compact if and only if it is both closed and bounded.

Theorem A.28 *Let K be a subset of $\mathbb{R}^n$. Then K is compact if and only if K is both closed and bounded.*

Proof Suppose that K is compact. We show that K is closed and bounded. Note that K is closed, since $\mathbb{R}^n$ is Hausdorff, and all compact subsets of Hausdorff spaces are closed, by Lemma A.18. Consider the open cover of $\mathbb{R}^n$ provided by the sets U_j for all positive integers j, where

$$U_j = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| < j\}.$$

The set K must be covered by finitely many of these open sets, since K is compact. Thus K is bounded.

Conversely suppose that K is both closed and bounded. Then there exists some real number L such that K is contained within the closed cube C given by

$$C = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : -L \leq x_j \leq L \text{ for } j = 1, 2, \ldots, n\}.$$

Now the closed interval $[-L, L]$ is compact by the Heine-Borel Theorem (Theorem A.27), and C is the Cartesian product of n-copies of the compact set $[-L, L]$. But repeated application of Theorem A.26 shows that the Cartesian product of a finite collection of compact topological spaces is compact. Thus C is a compact subset of $\mathbb{R}^n$. But K is a closed subset of C. It follows from Lemma A.16 that K is compact, as required. ∎

The following result follows directly from the Heine-Borel Theorem (Theorem A.27) and the Lebesgue Lemma (Lemma A.21).

Theorem A.29 *Let X be a topological space, and let $\mathcal{U}$ be an open cover of X. Let a and b be real numbers satisfying $a < b$, and let $\gamma\colon [a, b] \to X$ be a continuous map from the closed bounded interval $[a, b]$ into X. Then there exist $t_0, t_1, \ldots, t_r \in [a, b]$, where*

$$a = t_0 < t_1 < t_2 < \cdots < t_r = b,$$

such that $\gamma([t_{i-1}, t_i])$ is contained wholly within one of the open sets belonging to the open cover $\mathcal{U}$.

Proof Let $\mathcal{V}$ be the open cover of $[a, b]$ consisting of all the subsets of $[a, b]$ that are of the form $\gamma^{-1}(U)$ for some open set U belonging to $\mathcal{U}$. The closed bounded interval $[a, b]$ is a compact metric space. It follows from the Lebesgue Lemma that one can find $t_0, t_1, \ldots, t_r$ so that each interval $[t_{i-1}, t_i]$ is contained within one of the sets belonging to the open cover $\mathcal{V}$ of $[a, b]$. (If we choose $t_0, t_1, \ldots, t_r$ such that $t_1 - t_{i-1} < \delta$ for all i, where δ is a Lebesgue number for the open cover $\mathcal{V}$ then the required property is satisfied by $t_0, t_1, \ldots, t_r$.) But then $\gamma([t_{i-1}, t_i])$ is contained wholly within one of the open sets belonging to the open cover $\mathcal{U}$ of X, for $i = 1, 2, \ldots, r$, as required. ∎