# Course 421: Michaelmas Term 2002
# Topological Spaces and the Fundamental Group

David R. Wilkins



## Contents

# 1 Topological Spaces

## 1.1 The Concept of Continuity

The concept of continuity plays an important role in mathematics. There is a precise definition of continuity for functions of a real variable. A function $f\colon D \to \mathbb{R}$, defined on a subset $D$ of the real line $\mathbb{R}$, is said to be continuous at a point $p$ of $D$ if, given any real number $\varepsilon$ satisfying $\varepsilon > 0$, there exists a real number $\delta$ satisfying $\delta > 0$ such that $|f(x) - f(p)| < \varepsilon$ for all points $x$ of $D$ satisfying $|x - p| < \delta$. This definition of continuity can easily be adapted so as to apply to functions of a complex variable. It can also be generalized to as to apply to functions of several real or complex variables. We thus obtain a definition of continuity for functions between subsets of Euclidean spaces.

This definition of continuity generalizes directly to functions between metric spaces. A *metric space* is a set provided with a distance function, measuring the distance between any two points of the set. This distance function is required to satisfy certain axioms: the distance between any two points of a metric space is always non-negative, and is zero if and only if those points coincide; the distance from a point $x$ to a point $y$ is the same as the distance from $y$ to $x$; given any three points $x$, $y$ and $z$ of a metric space, the distance from $x$ to $z$ is required to be less than or equal to the sum of the distance from $x$ to $y$ and the distance from $y$ to $z$. A function from a metric space $X$ to a metric space $Y$ is continuous at a point $p$ of $X$ if and only if, given any real number $\varepsilon$ satisfying $\varepsilon > 0$, there exists a real number $\delta$ satisfying $\delta > 0$ such that the distance from $f(x)$ to $f(p)$ is less than $\varepsilon$ for all points $x$ of $X$ whose distance from $p$ is less than $\delta$.

We shall introduce the concept of a *topological space*, and give a definition of continuity for functions from one topological space to another which generalizes the definitions of continuity discussed above for functions of a real variable, for functions of a complex variable, for functions between subsets of Euclidean spaces, and for functions from one metric space to another.

The theory of topological spaces has proved itself to be very useful in many areas of mathematics.

## 1.2 Topological Spaces

**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 the open sets in a topological space $X$ is referred to as a *topology* on the set $X$.

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

## 1.3 Subsets of Euclidean Space

Let $X$ be a subset of $n$-dimensional Euclidean space $\mathbb{R}^n$. The *Euclidean distance* $|\mathbf{x} - \mathbf{y}|$ between two points $\mathbf{x}$ and $\mathbf{y}$ of $X$ is defined as follows:

$$|\mathbf{x} - \mathbf{y}| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2},$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and $\mathbf{y} = (y_1, y_2, \dots, y_n)$. The Euclidean distances between any three points $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ of $X$ satisfy the *Triangle Inequality*:

$$|\mathbf{x} - \mathbf{z}| \leq |\mathbf{x} - \mathbf{y}| + |\mathbf{y} - \mathbf{z}|.$$

A subset $V$ of $X$ is said to be *open* in $X$ if, given any point $\mathbf{v}$ of $V$, there exists some $\delta > 0$ such that

$$\{\mathbf{x} \in X : |\mathbf{x} - \mathbf{v}| < \delta\} \subset V.$$

The empty set is also considered to be open in $X$.

Both $\emptyset$ and $X$ are open sets in $X$. Also it is not difficult to show that any union of open sets in $X$ is open in $X$, and that any finite intersection of open sets in $X$ is open in $X$. (This will be proved in more generality for open sets in metric spaces.) Thus the collection of open sets in a subset $X$ of a Euclidean space $\mathbb{R}^n$ satisfies the topological space axioms. Thus every subset $X$ of $\mathbb{R}^n$ is a topological space with these open sets. This topology on a subset $X$ of $\mathbb{R}^n$ is referred to as the *usual topology* on $X$, generated by the Euclidean distance function.

In particular $\mathbb{R}^n$ is itself a topological space.

## 1.4 Open Sets in Metric Spaces

**Definition** A *metric space* $(X, d)$ consists of a set $X$ together with a *distance function* $d: X \times X \to [0, +\infty)$ on $X$ satisfying 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*. The elements of a metric space are usually referred to as *points* of that metric space.

An $n$-dimensional Euclidean space $\mathbb{R}^n$ is a metric space with with respect to the *Euclidean distance function* $d$, defined by

$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Any subset $X$ of $\mathbb{R}^n$ may be regarded as a metric space whose distance function is the restriction to $X$ of the Euclidean distance function on $\mathbb{R}^n$ defined above.

**Definition** Let $(X, d)$ be a metric space. Given a point $x$ of $X$ and $r \geq 0$, the *open ball* $B_X(x, r)$ of *radius* $r$ about $x$ in $X$ is defined by

$$B_X(x, r) \equiv \{x' \in X : d(x', x) < r\}.$$

**Definition** 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 $B_X(v, \delta) \subset V$.

By convention, we regard the empty set $\emptyset$ as being an open subset of $X$. (The criterion given above is satisfied vacuously in this case.)

**Lemma 1.1** *Let $X$ be a metric space with distance function $d$, and let $x_0$ be a point of $X$. Then, for any $r > 0$, the open ball $B_X(x_0, r)$ of radius $r$ about $x_0$ is an open set in $X$.*

**Proof** Let $x \in B_X(x_0, r)$. We must show that there exists some $\delta > 0$ such that $B_X(x, \delta) \subset B_X(x_0, r)$. Now $d(x, x_0) < r$, and hence $\delta > 0$, where $\delta = r - d(x, x_0)$. Moreover if $x' \in B_X(x, \delta)$ then

$$d(x', x_0) \leq d(x', x) + d(x, x_0) < \delta + d(x, x_0) = r,$$

by the Triangle Inequality, hence $x' \in B_X(x_0, r)$. Thus $B_X(x, \delta) \subset B_X(x_0, r)$, showing that $B_X(x_0, r)$ is an open set, as required. ∎

**Proposition 1.2** *Let $X$ be a metric space. The collection of open sets in $X$ has the following properties:—*

(i) *the empty set $\emptyset$ and the whole set $X$ are both 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.*

**Proof** The empty set $\emptyset$ is an open set by convention. Moreover the definition of an open set is satisfied trivially by the whole set $X$. Thus (i) is satisfied.

Let $\mathcal{A}$ be any collection of open sets in $X$, and let $U$ denote the union of all the open sets belonging to $\mathcal{A}$. We must show that $U$ is itself an open set. Let $x \in U$. Then $x \in V$ for some open set $V$ belonging to the collection $\mathcal{A}$. Therefore there exists some $\delta > 0$ such that $B_X(x, \delta) \subset V$. But $V \subset U$, and thus $B_X(x, \delta) \subset U$. This shows that $U$ is open. Thus (ii) is satisfied.

Finally let $V_1, V_2, V_3, \ldots, V_k$ be a *finite* collection of open sets in $X$, and let $V = V_1 \cap V_2 \cap \cdots \cap V_k$. Let $x \in V$. Now $x \in V_j$ for all $j$, and therefore there exist strictly positive real numbers $\delta_1, \delta_2, \ldots, \delta_k$ such that $B_X(x, \delta_j) \subset V_j$ for $j = 1, 2, \ldots, k$. Let $\delta$ be the minimum of $\delta_1, \delta_2, \ldots, \delta_k$. Then $\delta > 0$. (This is where we need the fact that we are dealing with a finite collection of open sets.) Moreover $B_X(x, \delta) \subset B_X(x, \delta_j) \subset V_j$ for $j = 1, 2, \ldots, k$, and thus $B_X(x, \delta) \subset V$. This shows that the intersection $V$ of the open sets $V_1, V_2, \ldots, V_k$ is itself open. Thus (iii) is satisfied. ∎

Any metric space may be regarded as a topological space. Indeed let $X$ be a metric space with distance function $d$. We recall that a subset $V$ of $X$ is an *open set* if and only if, given any point $v$ of $V$, there exists some $\delta > 0$ such that $\{x \in X : d(x, v) < \delta\} \subset V$. Proposition 1.2 shows that the topological space axioms are satisfied by the collection of open sets in any metric space. We refer to this collection of open sets as the *topology* generated by the distance function $d$ on $X$.

## 1.5 Further Examples of Topological Spaces

**Example** Given any set $X$, one can define a topology on $X$ where every subset of $X$ is an open set. This topology is referred to as the *discrete topology* on $X$.

**Example** Given any set $X$, one can define a topology on $X$ in which the only open sets are the empty set $\emptyset$ and the whole set $X$.

## 1.6 Closed Sets

**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.

We recall that the complement of the union of some collection of subsets of some set $X$ is the intersection of the complements of those sets, and the complement of the intersection of some collection of subsets of $X$ is the union of the complements of those sets. The following result therefore follows directly from the definition of a topological space.

**Proposition 1.3** *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.*

## 1.7 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 $U \cap V = \emptyset$.

**Lemma 1.4** *All metric spaces are Hausdorff spaces.*

**Proof** Let $X$ be a metric space with distance function $d$, and let $x$ and $y$ be points of $X$, where $x \neq y$. Let $\varepsilon = \frac{1}{2}d(x, y)$. Then the open balls $B_X(x, \varepsilon)$ and $B_X(y, \varepsilon)$ of radius $\varepsilon$ centred on the points $x$ and $y$ are open sets (see Lemma 1.1). If $B_X(x, \varepsilon) \cap B_X(y, \varepsilon)$ were non-empty then there would exist $z \in X$ satisfying $d(x, z) < \varepsilon$ and $d(z, y) < \varepsilon$. But this is impossible, since it would then follow from the Triangle Inequality that $d(x,y) < 2\varepsilon$, contrary to the choice of $\varepsilon$. Thus $x \in B_X(x,\varepsilon)$, $y \in B_X(y,\varepsilon)$, $B_X(x,\varepsilon) \cap B_X(y,\varepsilon) = \emptyset$. This shows that the metric space $X$ is a Hausdorff space. ∎

We now give an example of a topological space which is not a Hausdorff space.

**Example** The *Zariski topology* on the set $\mathbb{R}$ of real numbers is defined as follows: a subset $U$ of $\mathbb{R}$ is open (with respect to the Zariski topology) if and only if either $U = \emptyset$ or else $\mathbb{R} \setminus U$ is finite. It is a straightforward exercise to verify that the topological space axioms are satisfied, so that the set $\mathbb{R}$ of real numbers is a topological space with respect to this Zariski topology. Now the intersection of any two non-empty open sets in this topology is always non-empty. (Indeed if $U$ and $V$ are non-empty open sets then $U = \mathbb{R} \setminus F_1$ and $V = \mathbb{R} \setminus F_2$, where $F_1$ and $F_2$ are finite sets of real numbers. But then $U \cap V = \mathbb{R} \setminus (F_1 \cup F_2)$, which is non-empty, since $F_1 \cup F_2$ is finite and $\mathbb{R}$ is infinite.) It follows immediately from this that $\mathbb{R}$, with the Zariski topology, is not a Hausdorff space.

## 1.8 Subspace Topologies

Let $X$ be a topological space with topology $\tau$, and let $A$ be a subset of $X$. Let $\tau_A$ be the collection of all subsets of $A$ that are of the form $V \cap A$ for $V \in \tau$. Then $\tau_A$ is a topology on the set $A$. (It is a straightforward exercise to verify that the topological space axioms are satisfied.) The topology $\tau_A$ on $A$ is referred to as the *subspace topology* on $A$.

Any subset of a Hausdorff space is itself a Hausdorff space (with respect to the subspace topology).

**Lemma 1.5** *Let $X$ be a metric space with distance function $d$, and let $A$ be a subset of $X$. A subset $W$ of $A$ is open with respect to the subspace topology on $A$ if and only if, given any point $w$ of $W$, there exists some $\delta > 0$ such that*

$$\{a \in A : d(a,w) < \delta\} \subset W.$$

*Thus the subspace topology on $A$ coincides with the topology on $A$ obtained on regarding $A$ as a metric space (with respect to the distance function $d$).*

**Proof** Suppose that $W$ is open with respect to the subspace topology on $A$. Then there exists some open set $U$ in $X$ such that $W = U \cap A$. Let $w$ be a point of $W$. Then there exists some $\delta > 0$ such that

$$\{x \in X : d(x,w) < \delta\} \subset U.$$

But then

$$\{a \in A : d(a, w) < \delta\} \subset U \cap A = W.$$

Conversely, suppose that $W$ is a subset of $A$ with the property that, for any $w \in W$, there exists some $\delta_w > 0$ such that

$$\{a \in A : d(a, w) < \delta_w\} \subset W.$$

Define $U$ to be the union of the open balls $B_X(w, \delta_w)$ as $w$ ranges over all points of $W$, where

$$B_X(w, \delta_w) = \{x \in X : d(x, w) < \delta_w\}.$$

The set $U$ is an open set in $X$, since each open ball $B_X(w, \delta_w)$ is an open set in $X$ (Lemma 1.1), and any union of open sets is itself an open set. Moreover

$$B_X(w, \delta_w) \cap A = \{a \in A : d(a, w) < \delta_w\} \subset W$$

for any $w \in W$. Therefore $U \cap A \subset W$. However $W \subset U \cap A$, since, $W \subset A$ and $\{w\} \subset B_X(w, \delta_w) \subset U$ for any $w \in W$. Thus $W = U \cap A$, where $U$ is an open set in $X$. We deduce that $W$ is open with respect to the subspace topology on $A$. ∎

**Example** Let $X$ be any subset of $n$-dimensional Euclidean space $\mathbb{R}^n$. Then the subspace topology on $X$ coincides with the topology on $X$ generated by the Euclidean distance function on $X$. We refer to this topology as the *usual topology* on $X$.

Let $X$ be a topological space, and let $A$ be a subset of $X$. One can readily verify the following:—

- a subset $B$ of $A$ is closed in $A$ (relative to the subspace topology on $A$) if and only if $B = A \cap F$ for some closed subset $F$ of $X$;
- if $A$ is itself open in $X$ then a subset $B$ of $A$ is open in $A$ if and only if it is open in $X$;
- if $A$ is itself closed in $X$ then a subset $B$ of $A$ is closed in $A$ if and only if it is closed in $X$.

## 1.9 Continuous Functions between Topological Spaces

**Definition** A function $f: X \to Y$ from a topological space $X$ to a topological space $Y$ is said to be *continuous* if $f^{-1}(V)$ is an open set in $X$ for every open set $V$ in $Y$, where

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

A continuous function from $X$ to $Y$ is often referred to as a *map* from $X$ to $Y$.

**Lemma 1.6** *Let $X$, $Y$ and $Z$ be topological spaces, and let $f: X \to Y$ and $g: Y \to Z$ be continuous functions. Then the composition $g \circ f: X \to Z$ of the functions $f$ and $g$ is continuous.*

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

**Lemma 1.7** *Let $X$ and $Y$ be topological spaces, and let $f: X \to Y$ be a function from $X$ to $Y$. The function $f$ is continuous if and only if $f^{-1}(G)$ is closed in $X$ for every closed subset $G$ of $Y$.*

**Proof** If $G$ is any subset of $Y$ then $X \setminus f^{-1}(G) = f^{-1}(Y \setminus G)$ (i.e., the complement of the preimage of $G$ is the preimage of the complement of $G$). The result therefore follows immediately from the definitions of continuity and closed sets. ∎

## 1.10 Continuous Functions between Metric Spaces

The following definition of continuity for functions between metric spaces generalizes that for functions of a real or complex variable.

**Definition** Let $X$ and $Y$ be metric spaces with distance functions $d_X$ and $d_Y$ respectively. A function $f: X \to Y$ from $X$ to $Y$ is said to be *continuous* at a point $x$ of $X$ if and only if the following criterion is satisfied:—

- given any real number $\varepsilon$ satisfying $\varepsilon > 0$ there exists some $\delta > 0$ such that $d_Y(f(x), f(x')) < \varepsilon$ for all points $x'$ of $X$ satisfying $d_X(x, x') < \delta$.

The function $f: X \to Y$ is said to be continuous on $X$ if and only if it is continuous at $x$ for every point $x$ of $X$.

This definition can be rephrased in terms of open balls: a function $f\colon X \to Y$ from a metric space $X$ to a metric space $Y$ is continuous at a point $x$ of $X$ if and only if, given any $\varepsilon > 0$, there exists some $\delta > 0$ such that $f$ maps $B_X(x, \delta)$ into $B_Y(f(x), \varepsilon)$ (where $B_X(x, \delta)$ and $B_Y(f(x), \varepsilon)$ denote the open balls of radius $\delta$ and $\varepsilon$ about $x$ and $f(x)$ respectively).

Let $f\colon X \to Y$ be a function from a set $X$ to a set $Y$. Given any subset $V$ of $Y$, we denote by $f^{-1}(V)$ the *preimage* of $V$ under the map $f$, defined by

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

The following result shows that the definition of continuity given above for functions between metric spaces is consistent with the more general definition of continuity for functions between topological spaces.

**Proposition 1.8** *Let $X$ and $Y$ be metric spaces, and let $f\colon X \to Y$ be a function from $X$ to $Y$. The function $f$ is continuous if and only if $f^{-1}(V)$ is an open set in $X$ for every open set $V$ of $Y$.*

**Proof** Suppose that $f\colon X \to Y$ is continuous. Let $V$ be an open set in $Y$. We must show that $f^{-1}(V)$ is open in $X$. Let $x$ be a point belonging to $f^{-1}(V)$. We must show that there exists some $\delta > 0$ with the property that $B_X(x, \delta) \subset f^{-1}(V)$. Now $f(x)$ belongs to $V$. But $V$ is open, hence there exists some $\varepsilon > 0$ with the property that $B_Y(f(x), \varepsilon) \subset V$. But $f$ is continuous at $x$. Therefore there exists some $\delta > 0$ such that $f$ maps the open ball $B_X(x, \delta)$ into $B_Y(f(x), \varepsilon)$ (see the remarks above). Thus $f(x') \in V$ for all $x' \in B_X(x, \delta)$, showing that $B_X(x, \delta) \subset f^{-1}(V)$. We have thus shown that if $f\colon X \to Y$ is continuous then $f^{-1}(V)$ is open in $X$ for every open set $V$ in $Y$.

Conversely suppose that $f\colon X \to Y$ has the property that $f^{-1}(V)$ is open in $X$ for every open set $V$ in $Y$. Let $x$ be any point of $X$. We must show that $f$ is continuous at $x$. Let $\varepsilon > 0$ be given. The open ball $B_Y(f(x), \varepsilon)$ is an open set in $Y$, by Lemma 1.1, hence $f^{-1}\left(B_Y(f(x), \varepsilon)\right)$ is an open set in $X$ which contains $x$. It follows that there exists some $\delta > 0$ such that $B_X(x, \delta) \subset f^{-1}\left(B_Y(f(x), \varepsilon)\right)$. We have thus shown that, given any $\varepsilon > 0$, there exists some $\delta > 0$ such that $f$ maps the open ball $B_X(x, \delta)$ into $B_Y(f(x), \varepsilon)$. We conclude that $f$ is continuous at $x$, as required. ∎

## 1.11 A Criterion for Continuity

We now show that, if a topological space $X$ is the union of a finite collection of closed sets, and if a function from $X$ to some topological space is continuous on each of these closed sets, then that function is continuous on $X$.

**Lemma 1.9** *Let $X$ and $Y$ be topological spaces, let $f: X \to Y$ be a function from $X$ to $Y$, and let $X = A_1 \cup A_2 \cup \cdots \cup A_k$, where $A_1, A_2, \ldots, A_k$ are closed sets in $X$. Suppose that the restriction of $f$ to the closed set $A_i$ is continuous for $i = 1, 2, \ldots, k$. Then $f: X \to Y$ is continuous.*

**Proof** Let $V$ be an open set in $Y$. We must show that $f^{-1}(V)$ is open in $X$. Now the preimage of the open set $V$ under the restriction $f|A_i$ of $f$ to $A_i$ is $f^{-1}(V) \cap A_i$. It follows from the continuity of $f|A_i$ that $f^{-1}(V) \cap A_i$ is relatively open in $A_i$ for each $i$, and hence there exist open sets $U_1, U_2, \ldots, U_k$ in $X$ such that $f^{-1}(V) \cap A_i = U_i \cap A_i$ for $i = 1, 2, \ldots, k$. Let $W_i = U_i \cup (X \setminus A_i)$ for $i = 1, 2, \ldots, k$. Then $W_i$ is an open set in $X$ (as it is the union of the open sets $U_i$ and $X \setminus A_i$), and $W_i \cap A_i = U_i \cap A_i = f^{-1}(V) \cap A_i$ for each $i$. We claim that $f^{-1}(V) = W_1 \cap W_2 \cap \cdots \cap W_k$.

Let $W = W_1 \cap W_2 \cap \cdots \cap W_k$. Then $f^{-1}(V) \subset W$, since $f^{-1}(V) \subset W_i$ for each $i$. Also

$$W = \bigcup_{i=1}^{k} (W \cap A_i) \subset \bigcup_{i=1}^{k} (W_i \cap A_i) = \bigcup_{i=1}^{k} (f^{-1}(V) \cap A_i) \subset f^{-1}(V),$$

since $X = A_1 \cup A_2 \cup \cdots \cup A_k$ and $W_i \cap A_i = f^{-1}(V) \cap A_i$ for each $i$. Therefore $f^{-1}(V) = W$. But $W$ is open in $X$, since it is the intersection of a finite collection of open sets. We have thus shown that $f^{-1}(V)$ is open in $X$ for any open set $V$ in $Y$. Thus $f: X \to Y$ is continuous, as required. ∎

**Alternative Proof** A function $f: X \to Y$ is continuous if and only if $f^{-1}(G)$ is closed in $X$ for every closed set $G$ in $Y$ (Lemma 1.7). Let $G$ be an closed set in $Y$. Then $f^{-1}(G) \cap A_i$ is relatively closed in $A_i$ for $i = 1, 2, \ldots, k$, since the restriction of $f$ to $A_i$ is continuous for each $i$. But $A_i$ is closed in $X$, and therefore a subset of $A_i$ is relatively closed in $A_i$ if and only if it is closed in $X$. Therefore $f^{-1}(G) \cap A_i$ is closed in $X$ for $i = 1, 2, \ldots, k$. Now $f^{-1}(G)$ is the union of the sets $f^{-1}(G) \cap A_i$ for $i = 1, 2, \ldots, k$. It follows that $f^{-1}(G)$, being a finite union of closed sets, is itself closed in $X$. It now follows from Lemma 1.7 that $f: X \to Y$ is continuous. ∎

**Example** Let $Y$ be a topological space, and let $\alpha: [0, 1] \to Y$ and $\beta: [0, 1] \to Y$ be continuous functions defined on the interval $[0, 1]$, where $\alpha(1) = \beta(0)$. Let $\gamma: [0, 1] \to Y$ be defined by

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

Now $\gamma|[0, \frac{1}{2}] = \alpha \circ \rho$ where $\rho: [0, \frac{1}{2}] \to [0, 1]$ is the continuous function defined by $\rho(t) = 2t$ for all $t \in [0, \frac{1}{2}]$. Thus $\gamma|[0, \frac{1}{2}]$ is continuous, being a composition of two continuous functions. Similarly $\gamma|[\frac{1}{2}, 1]$ is continuous. The subintervals $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$ are closed in $[0, 1]$, and $[0, 1]$ is the union of these two subintervals. It follows from Lemma 1.9 that $\gamma: [0, 1] \to Y$ is continuous.

## 1.12 Homeomorphisms

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

- the function $h: X \to Y$ is both injective and surjective (so that the function $h: X \to Y$ has a well-defined inverse $h^{-1}: Y \to X$),
- the function $h: X \to Y$ and its inverse $h^{-1}: Y \to X$ are both continuous.

Two topological spaces $X$ and $Y$ are said to be *homeomorphic* if there exists a homeomorphism $h: X \to Y$ from $X$ to $Y$.

If $h: X \to Y$ is a homeomorphism between topological spaces $X$ and $Y$ then $h$ 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.

## 1.13 Neighbourhoods, Closures and Interiors

**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 $U$ for which $x \in U$ and $U \subset N$.

One can readily verify that this definition of neighbourhoods in topological spaces is consistent with that for neighbourhoods in metric spaces.

**Lemma 1.10** *Let $X$ be a topological space. A subset $V$ of $X$ is open in $X$ if and only if $V$ is a neighbourhood of each point belonging to $V$.*

**Proof** It follows directly from the definition of neighbourhoods that an open set $V$ is a neighbourhood of any point belonging to $V$. Conversely, suppose that $V$ is a subset of $X$ which is a neighbourhood of each $v \in V$. Then, given any point $v$ of $V$, there exists an open set $U_v$ such that $v \in U_v$ and $U_v \subset V$. Thus $V$ is an open set, since it is the union of the open sets $U_v$ as $v$ ranges over all points of $V$. ∎

**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 $A$.

Let $X$ be a topological space and let $A$ be a subset of $X$. It follows directly from the definition of $\overline{A}$ that the closure $\overline{A}$ of $A$ is uniquely characterized by the following two properties:

(i) the closure $\overline{A}$ of $A$ is a closed set containing $A$,

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

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

(i) the interior $A^0$ of $A$ is an open set contained in $A$,

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

Moreover a point $x$ of $A$ belongs to the interior $A^0$ of $A$ if and only if $A$ is a neighbourhood of $x$.

## 1.14 Product Topologies

The *Cartesian product* $X_1 \times X_2 \times \cdots \times X_n$ of sets $X_1, X_2, \ldots, X_n$ is defined to be the set of all ordered $n$-tuples $(x_1, x_2, \ldots, x_n)$, where $x_i \in X_i$ for $i = 1, 2, \ldots, n$.

The sets $\mathbb{R}^2$ and $\mathbb{R}^3$ are the Cartesian products $\mathbb{R} \times \mathbb{R}$ and $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$ respectively.

Cartesian products of sets are employed as the domains of functions of several variables. For example, if $X$, $Y$ and $Z$ are sets, and if an element $f(x, y)$ of $Z$ is determined for each choice of an element $x$ of $X$ and an element $y$ of $Y$, then we have a function $f\colon X \times Y \to Z$ whose domain is the Cartesian product $X \times Y$ of $X$ and $Y$: this function sends the ordered pair $(x, y)$ to $f(x, y)$ for all $x \in X$ and $y \in Y$.

**Definition** Let $X_1, X_2, \ldots, X_n$ be topological spaces. A subset $U$ of the Cartesian product $X_1 \times X_2 \times \cdots \times X_n$ is said to be *open* (with respect to the product topology) if, given any point $p$ of $U$, there exist open sets $V_i$ in $X_i$ for $i = 1, 2, \ldots, n$ such that $\{p\} \subset V_1 \times V_2 \times \cdots \times V_n \subset U$.

**Lemma 1.11** *Let $X_1, X_2, \ldots, X_n$ be topological spaces. Then the collection of open sets in $X_1 \times X_2 \times \cdots \times X_n$ is a topology on $X_1 \times X_2 \times \cdots \times X_n$.*

**Proof** Let $X = X_1 \times X_2 \times \cdots \times X_n$. The definition of open sets ensures that the empty set and the whole set $X$ are open in $X$. We must prove that any union or finite intersection of open sets in $X$ is an open set.

Let $E$ be a union of a collection of open sets in $X$ and let $p$ be a point of $E$. Then $p \in D$ for some open set $D$ in the collection. It follows from this that there exist open sets $V_i$ in $X_i$ for $i = 1, 2, \ldots, n$ such that

$$\{p\} \subset V_1 \times V_2 \times \cdots \times V_n \subset D \subset E.$$

Thus $E$ is open in $X$.

Let $U = U_1 \cap U_2 \cap \cdots \cap U_m$, where $U_1, U_2, \ldots, U_m$ are open sets in $X$, and let $p$ be a point of $U$. Then there exist open sets $V_{ki}$ in $X_i$ for $k = 1, 2, \ldots, m$ and $i = 1, 2, \ldots, n$ such that $\{p\} \subset V_{k1} \times V_{k2} \times \cdots \times V_{kn} \subset U_k$ for $k = 1, 2, \ldots, m$. Let $V_i = V_{1i} \cap V_{2i} \cap \cdots \cap V_{mi}$ for $i = 1, 2, \ldots, n$. Then

$$\{p\} \subset V_1 \times V_2 \times \cdots \times V_n \subset V_{k1} \times V_{k2} \times \cdots \times V_{kn} \subset U_k$$

for $k = 1, 2, \ldots, m$, and hence $\{p\} \subset V_1 \times V_2 \times \cdots \times V_n \subset U$. It follows that $U$ is open in $X$, as required. ∎

**Lemma 1.12** *Let $X_1, X_2, \ldots, X_n$ and $Z$ be topological spaces. Then a function $f\colon X_1 \times X_2 \times \cdots \times X_n \to Z$ is continuous if and only if, given any point $p$ of $X_1 \times X_2 \times \cdots \times X_n$, and given any open set $U$ in $Z$ containing $f(p)$, there exist open sets $V_i$ in $X_i$ for $i = 1, 2, \ldots, n$ such that $p \in V_1 \times V_2 \cdots \times V_n$ and $f(V_1 \times V_2 \times \cdots \times V_n) \subset U$.*

**Proof** Let $V_i$ be an open set in $X_i$ for $i = 1, 2, \ldots, n$, and let $U$ be an open set in $Z$. Then $V_1 \times V_2 \times \cdots \times V_n \subset f^{-1}(U)$ if and only if $f(V_1 \times V_2 \times \cdots \times V_n) \subset U$. It follows that $f^{-1}(U)$ is open in the product topology on $X_1 \times X_2 \times \cdots \times X_n$ if and only if, given any point $p$ of $X_1 \times X_2 \times \cdots \times X_n$ satisfying $f(p) \in U$, there exist open sets $V_i$ in $X_i$ for $i = 1, 2, \ldots, n$ such that $f(V_1 \times V_2 \times \cdots \times V_n) \subset U$. The required result now follows from the definition of continuity. ∎

Let $X_1, X_2, \ldots, X_n$ be topological spaces, and let $V_i$ be an open set in $X_i$ for $i = 1, 2, \ldots, n$. It follows directly from the definition of the product topology that $V_1 \times V_2 \times \cdots \times V_n$ is open in $X_1 \times X_2 \times \cdots \times X_n$.

**Theorem 1.13** *Let $X = X_1 \times X_2 \times \cdots \times X_n$, where $X_1, X_2, \ldots, X_n$ are topological spaces and $X$ is given the product topology, and for each $i$, let $p_i\colon X \to X_i$ denote the projection function which sends $(x_1, x_2, \ldots, x_n) \in X$ to $x_i$. Then the functions $p_1, p_2, \ldots, p_n$ are continuous. Moreover a function $f\colon Z \to X$ mapping a topological space $Z$ into $X$ is continuous if and only if $p_i \circ f\colon Z \to X_i$ is continuous for $i = 1, 2, \ldots, n$.*

**Proof** Let $V$ be an open set in $X_i$. Then

$$p_i^{-1}(V) = X_1 \times \cdots \times X_{i-1} \times V \times X_{i+1} \times \cdots \times X_n,$$

and therefore $p_i^{-1}(V)$ is open in $X$. Thus $p_i\colon X \to X_i$ is continuous for all $i$.

Let $f\colon Z \to X$ be continuous. Then, for each $i$, $p_i \circ f\colon Z \to X_i$ is a composition of continuous functions, and is thus itself continuous.

Conversely suppose that $f\colon Z \to X$ is a function with the property that $p_i \circ f$ is continuous for all $i$. 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 $f(z) = (u_1, u_2, \ldots, u_n)$. Now $U$ is open in $X$, and therefore there exist open sets $V_1, V_2, \ldots, V_n$ in $X_1, X_2, \ldots, X_n$ respectively such that $u_i \in V_i$ for all $i$ and $V_1 \times V_2 \times \cdots \times V_n \subset U$. Let

$$N_z = f_1^{-1}(V_1) \cap f_2^{-1}(V_2) \cap \cdots \cap f_n^{-1}(V_n),$$

where $f_i = p_i \circ f$ for $i = 1, 2, \ldots, n$. Now $f_i^{-1}(V_i)$ is an open subset of $Z$ for $i = 1, 2, \ldots, n$, since $V_i$ is open in $X_i$ and $f_i\colon Z \to X_i$ is continuous. Thus $N_z$, being a finite intersection of open sets, is itself open in $Z$. Moreover

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

so that $N_z \subset f^{-1}(U)$. It follows that $f^{-1}(U)$ is the union of the open sets $N_z$ as $z$ ranges over all points of $f^{-1}(U)$. Therefore $f^{-1}(U)$ is open in $Z$. This shows that $f\colon Z \to X$ is continuous, as required. ∎

**Proposition 1.14** *The usual topology on $\mathbb{R}^n$ coincides with the product topology on $\mathbb{R}^n$ obtained on regarding $\mathbb{R}^n$ as the Cartesian product $\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ of $n$ copies of the real line $\mathbb{R}$.*

**Proof** We must show that a subset $U$ of $\mathbb{R}^n$ is open with respect to the usual topology if and only if it is open with respect to the product topology.

Let $U$ be a subset of $\mathbb{R}^n$ that is open with respect to the usual topology, and let $\mathbf{u} \in U$. Then there exists some $\delta > 0$ such that $B(\mathbf{u}, \delta) \subset U$, where

$$B(\mathbf{u}, \delta) = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{u}| < \delta\}.$$

Let $I_1, I_2, \ldots, I_n$ be the open intervals in $\mathbb{R}$ defined by

$$I_i = \{t \in \mathbb{R} : u_i - \frac{\delta}{\sqrt{n}} < t < u_i + \frac{\delta}{\sqrt{n}}\} \qquad (i = 1, 2, \ldots, n),$$

Then $I_1, I_2, \ldots, I_n$ are open sets in $\mathbb{R}$. Moreover

$$\{\mathbf{u}\} \subset I_1 \times I_2 \times \cdots \times I_n \subset B(\mathbf{u}, \delta) \subset U,$$

since

$$|\mathbf{x} - \mathbf{u}|^2 = \sum_{i=1}^{n} (x_i - u_i)^2 < n \left( \frac{\delta}{\sqrt{n}} \right)^2 = \delta^2$$

for all $\mathbf{x} \in I_1 \times I_2 \times \cdots \times I_n$. This shows that any subset $U$ of $\mathbb{R}^n$ that is open with respect to the usual topology on $\mathbb{R}^n$ is also open with respect to the product topology on $\mathbb{R}^n$.

Conversely suppose that $U$ is a subset of $\mathbb{R}^n$ that is open with respect to the product topology on $\mathbb{R}^n$, and let $\mathbf{u} \in U$. Then there exist open sets $V_1, V_2, \ldots, V_n$ in $\mathbb{R}$ containing $u_1, u_2, \ldots, u_n$ respectively such that $V_1 \times V_2 \times \cdots \times V_n \subset U$. Now we can find $\delta_1, \delta_2, \ldots, \delta_n$ such that $\delta_i > 0$ and $(u_i - \delta_i, u_i + \delta_i) \subset V_i$ for all $i$. Let $\delta > 0$ be the minimum of $\delta_1, \delta_2, \ldots, \ldots, \delta_n$. Then

$$B(\mathbf{u}, \delta) \subset V_1 \times V_2 \times \cdots V_n \subset U,$$

for if $\mathbf{x} \in B(\mathbf{u}, \delta)$ then $|x_i - u_i| < \delta_i$ for $i = 1, 2, \ldots, n$. This shows that any subset $U$ of $\mathbb{R}^n$ that is open with respect to the product topology on $\mathbb{R}^n$ is also open with respect to the usual topology on $\mathbb{R}^n$. ∎

The following result is now an immediate corollary of Proposition 1.14 and Theorem 1.13.

**Corollary 1.15** *Let $X$ be a topological space and let $f: X \to \mathbb{R}^n$ be a function from $X$ to $\mathbb{R}^n$. Let us write*

$$f(x) = (f_1(x), f_2(x), \ldots, f_n(x))$$

*for all $x \in X$, where the components $f_1, f_2, \ldots, f_n$ of $f$ are functions from $X$ to $\mathbb{R}$. The function $f$ is continuous if and only if its components $f_1, f_2, \ldots, f_n$ are all continuous.*

Let $f: X \to \mathbb{R}$ and $g: X \to \mathbb{R}$ be continuous real-valued functions on some topological space $X$. We claim that $f+g$, $f-g$ and $f.g$ are continuous. Now it is a straightforward exercise to verify that the sum and product functions $s: \mathbb{R}^2 \to \mathbb{R}$ and $p: \mathbb{R}^2 \to \mathbb{R}$ defined by $s(x, y) = x + y$ and $p(x, y) = xy$ are continuous, and $f + g = s \circ h$ and $f.g = p \circ h$, where $h: X \to \mathbb{R}^2$ is defined by $h(x) = (f(x), g(x))$. Moreover it follows from Corollary 1.15 that the function $h$ is continuous, and compositions of continuous functions are continuous. Therefore $f + g$ and $f.g$ are continuous, as claimed. Also $-g$ is continuous, and $f - g = f + (-g)$, and therefore $f - g$ is continuous. If in addition the continuous function $g$ is non-zero everywhere on $X$ then $1/g$ is continuous (since $1/g$ is the composition of $g$ with the reciprocal function $t \mapsto 1/t$), and therefore $f/g$ is continuous.

**Lemma 1.16** *The Cartesian product $X_1 \times X_2 \times \ldots X_n$ of Hausdorff spaces $X_1, X_2, \ldots, X_n$ is Hausdorff.*

**Proof** Let $X = X_1 \times X_2 \times \ldots, X_n$, and let $u$ and $v$ be distinct points of $X$, where $u = (x_1, x_2, \ldots, x_n)$ and $v = (y_1, y_2, \ldots, y_n)$. Then $x_i \neq y_i$ for some integer $i$ between 1 and $n$. But then there exist 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 function. Then $p_i^{-1}(U)$ and $p_i^{-1}(V)$ are open sets in $X$, since $p_i$ is continuous. Moreover $u \in p_i^{-1}(U)$, $v \in p_i^{-1}(V)$, and $p_i^{-1}(U) \cap p_i^{-1}(V) = \emptyset$. Thus $X$ is Hausdorff, as required. ∎

## 1.15 Identification Maps and Quotient Topologies

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

- the function $q \colon X \to Y$ is surjective,
- a subset $U$ of $Y$ is open in $Y$ if and only if $q^{-1}(U)$ is open in $X$.

It follows directly from the definition that any identification map is continuous. Moreover, in order to show that a continuous surjection $q \colon X \to Y$ is an identification map, it suffices to prove that if $V$ is a subset of $Y$ with the property that $q^{-1}(V)$ is open in $X$ then $V$ is open in $Y$.

**Lemma 1.17** *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 function $q \colon X \to Y$ is an identification map.*

**Proof** Let $\tau$ 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 $Y \in \tau$. If $\{V_\alpha : \alpha \in A\}$ is any collection of subsets of $Y$ indexed by a set $A$, then it is a straightforward exercise to verify that

$$\bigcup_{\alpha \in A} q^{-1}(V_\alpha) = q^{-1}\left(\bigcup_{\alpha \in A} V_\alpha\right), \qquad \bigcap_{\alpha \in A} q^{-1}(V_\alpha) = q^{-1}\left(\bigcap_{\alpha \in A} V_\alpha\right)$$

(i.e., given any collection of subsets of $Y$, the union of the preimages of the sets is the preimage of the union of those sets, and the intersection of the preimages of the sets is the preimage of the intersection of those sets). It follows easily from this that unions and finite intersections of sets belonging to $\tau$ must themselves belong to $\tau$. Thus $\tau$ is a topology on $Y$, and the function $q \colon X \to Y$ is an identification map with respect to the topology $\tau$. Clearly $\tau$ is the unique topology on $Y$ for which the function $q \colon X \to Y$ is an identification map. ∎

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

**Lemma 1.18** *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 function from $Y$ to $Z$. Then the function $f$ is continuous if and only if the composition function $f \circ q: X \to Z$ is continuous.*

**Proof** Suppose that $f$ is continuous. Then the composition function $f \circ q$ is a composition of continuous functions 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 function $q$ is an identification map). Therefore the function $f$ is continuous, as required. ∎

**Example** Let $S^1$ be the unit circle in $\mathbb{R}^2$, and let $q: [0,1] \to S^1$ be the map that sends $t \in [0,1]$ to $(\cos 2\pi t, \sin 2\pi t)$. Then $q: [0,1] \to S^1$ is an identification map, and therefore a function $f: S^1 \to Z$ from $S^1$ to some topological space $Z$ is continuous if and only if $f \circ q: [0,1] \to Z$ is continuous.

**Example** Let $S^n$ be the $n$-sphere, consisting of all points $\mathbf{x}$ in $\mathbb{R}^{n+1}$ satisfying $|\mathbf{x}| = 1$. Let $\mathbb{R}P^n$ be the set of all lines in $\mathbb{R}^{n+1}$ passing through the origin (i.e., $\mathbb{R}P^n$ is the set of all one-dimensional vector subspaces of $\mathbb{R}^{n+1}$). Let $q: S^n \to \mathbb{R}P^n$ denote the function which sends a point $\mathbf{x}$ of $S^n$ to the element of $\mathbb{R}P^n$ represented by the line in $\mathbb{R}^{n+1}$ that passes through both $\mathbf{x}$ and the origin. Note that each element of $\mathbb{R}P^n$ is the image (under $q$) of exactly two antipodal points $\mathbf{x}$ and $-\mathbf{x}$ of $S^n$. The function $q$ induces a corresponding quotient topology on $\mathbb{R}P^n$ such that $q: S^n \to \mathbb{R}P^n$ is an identification map. The set $\mathbb{R}P^n$, with this topology, is referred to as *real projective $n$-space*. In particular $\mathbb{R}P^2$ is referred to as the *real projective plane*. It follows from Lemma 1.18 that a function $f: \mathbb{R}P^n \to Z$ from $\mathbb{R}P^n$ to any topological space $Z$ is continuous if and only if the composition function $f \circ q: S^n \to Z$ is continuous.

## 1.16 Compact Topological Spaces

Let $X$ be a topological space, and let $A$ be a subset of $X$. A collection of subsets of $X$ in $X$ is said to *cover* $A$ if and only if every point of $A$ belongs to at least one of these subsets. In particular, an *open cover* of $X$ is collection of open sets in $X$ that covers $X$.

If $\mathcal{U}$ and $\mathcal{V}$ are open covers of some topological space $X$ then $\mathcal{V}$ is said to be a *subcover* of $\mathcal{U}$ if and only if every open set belonging 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.

**Lemma 1.19** *Let $X$ be a topological space. A subset $A$ of $X$ is compact (with respect to the subspace topology on $A$) if and only if, given any collection $\mathcal{U}$ of open sets in $X$ covering $A$, there exists a finite collection $V_1, V_2, \ldots, V_r$ of open sets belonging to $\mathcal{U}$ such that $A \subset V_1 \cup V_2 \cup \cdots \cup V_r$.*

**Proof** A subset $B$ of $A$ is open in $A$ (with respect to the subspace topology on $A$) if and only if $B = A \cap V$ for some open set $V$ in $X$. The desired result therefore follows directly from the definition of compactness. ∎

We now show that any closed bounded interval in the real line is compact. This result is known as the *Heine-Borel Theorem*. The proof of this theorem uses the *least upper bound principle* which states that, given any non-empty set $S$ of real numbers which is bounded above, there exists a *least upper bound* (or *supremum*) $\sup S$ for the set $S$.

**Theorem 1.20** (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 $\mathcal{U}$ be a collection of open sets in $\mathbb{R}$ with the property that each point of the interval $[a, b]$ belongs to at least one of these open sets. We must show that $[a, b]$ is covered by finitely many of these open sets.

Let $S$ be the set of all $\tau \in [a, b]$ with the property that $[a, \tau]$ is covered by some finite collection of open sets belonging to $\mathcal{U}$, and let $s = \sup S$. Now $s \in W$ for some open set $W$ belonging to $\mathcal{U}$. Moreover $W$ is open in $\mathbb{R}$, and therefore there exists some $\delta > 0$ such that $(s - \delta, s + \delta) \subset W$. Moreover $s - \delta$ is not an upper bound for the set $S$, hence there exists some $\tau \in S$ satisfying $\tau > s - \delta$. It follows from the definition of $S$ that $[a, \tau]$ is covered by some finite collection $V_1, V_2, \ldots, V_r$ of open sets belonging to $\mathcal{U}$.

Let $t \in [a, b]$ satisfy $\tau \leq t < s + \delta$. Then

$$[a, t] \subset [a, \tau] \cup (s - \delta, s + \delta) \subset V_1 \cup V_2 \cup \cdots \cup V_r \cup W,$$

and thus $t \in S$. In particular $s \in S$, and moreover $s = b$, since otherwise $s$ would not be an upper bound of the set $S$. Thus $b \in S$, and therefore $[a, b]$ is covered by a finite collection of open sets belonging to $\mathcal{U}$, as required. ∎

**Lemma 1.21** *Let $A$ be a closed subset of some compact topological space $X$. Then $A$ is compact.*

**Proof** Let $\mathcal{U}$ be any collection of open sets in $X$ covering $A$. On adjoining the open set $X \setminus A$ to $\mathcal{U}$, we obtain an open cover of $X$. This open cover of $X$ possesses a finite subcover, since $X$ is compact. Moreover $A$ is covered by the open sets in the collection $\mathcal{U}$ that belong to this finite subcover. It follows from Lemma 1.19 that $A$ is compact, as required. ∎

**Lemma 1.22** *Let $f\colon X \to Y$ be a continuous function between topological spaces $X$ and $Y$, and 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 $f(A)$. Then $A$ is covered by the collection of all open sets of the form $f^{-1}(V)$ for some $V \in \mathcal{V}$. It follows from the compactness of $A$ that there exists a finite collection $V_1, V_2, \dots, V_k$ of open sets belonging to $\mathcal{V}$ such that

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

But then $f(A) \subset V_1 \cup V_2 \cup \cdots \cup V_k$. This shows that $f(A)$ is compact. ∎

**Lemma 1.23** *Let $f\colon X \to \mathbb{R}$ be a continuous real-valued function on a compact topological space $X$. Then $f$ is bounded above and below on $X$.*

**Proof** The range $f(X)$ of the function $f$ is covered by some finite collection $I_1, I_2, \dots, I_k$ of open intervals of the form $(-m, m)$, where $m \in \mathbb{N}$, since $f(X)$ is compact (Lemma 1.22) and $\mathbb{R}$ is covered by the collection of all intervals of this form. It follows that $f(X) \subset (-M, M)$, where $(-M, M)$ is the largest of the intervals $I_1, I_2, \dots, I_k$. Thus the function $f$ is bounded above and below on $X$, as required. ∎

**Proposition 1.24** *Let $f\colon X \to \mathbb{R}$ be a continuous real-valued function on a compact topological space $X$. Then there exist points $u$ and $v$ of $X$ such that $f(u) \leq f(x) \leq f(v)$ for all $x \in X$.*

**Proof** Let $m = \inf\{f(x) : x \in X\}$ and $M = \sup\{f(x) : x \in X\}$. There must exist $v \in X$ satisfying $f(v) = M$, for if $f(x) < M$ for all $x \in X$ then the function $x \mapsto 1/(M - f(x))$ would be a continuous real-valued function on $X$ that was not bounded above, contradicting Lemma 1.23. Similarly there must exist $u \in X$ satisfying $f(u) = m$, since otherwise the function $x \mapsto 1/(f(x) - m)$ would be a continuous function on $X$ that was not bounded above, again contradicting Lemma 1.23. But then $f(u) \leq f(x) \leq f(v)$ for all $x \in X$, as required. ∎

**Proposition 1.25** *Let $A$ be a compact subset of a metric space $X$. Then $A$ is closed in $X$.*

**Proof** Let $p$ be a point of $X$ that does not belong to $A$, and let $f(x) = d(x,p)$, where $d$ is the distance function on $X$. It follows from Proposition 1.24 that there is a point $q$ of $A$ such that $f(a) \geq f(q)$ for all $a \in A$, since $A$ is compact. Now $f(q) > 0$, since $q \neq p$. Let $\delta$ satisfy $0 < \delta \leq f(q)$. Then the open ball of radius $\delta$ about the point $p$ is contained in the complement of $A$, since $f(x) < f(q)$ for all points $x$ of this open ball. It follows that the complement of $A$ is an open set in $X$, and thus $A$ itself is closed in $X$. ∎

**Proposition 1.26** *Let $X$ be a Hausdorff topological space, and let $K$ be a compact subset of $X$. Let $x$ be a point of $X \setminus K$. Then there exist open sets $V$ and $W$ in $X$ such that $x \in V$, $K \subset W$ and $V \cap W = \emptyset$.*

**Proof** For each point $y \in K$ 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 then there exists a finite set $\{y_1, y_2, \ldots, y_r\}$ of points of $K$ such that $K$ is contained in $W_{x,y_1} \cup W_{x,y_2} \cup \cdots \cup W_{x,y_r}$, since $K$ is compact. Define

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

Then $V$ and $W$ are open sets, $x \in V$, $K \subset W$ and $V \cap W = \emptyset$, as required. ∎

**Corollary 1.27** *A compact subset of a Hausdorff topological space is closed.*

**Proof** Let $K$ be a compact subset of a Hausdorff topological space $X$. It follows immediately from Proposition 1.26 that, for each $x \in X \setminus K$, there exists an open set $V_x$ such that $x \in V_x$ and $V_x \cap K = \emptyset$. But then $X \setminus K$ is equal to the union of the open sets $V_x$ as $x$ ranges over all points of $X \setminus K$, and any set that is a union of open sets is itself an open set. We conclude that $X \setminus K$ is open, and thus $K$ is closed. ∎

**Proposition 1.28** *Let $X$ be a Hausdorff topological space, and let $K_1$ and $K_2$ be compact subsets of $X$, where $K_1 \cap K_2 = \emptyset$. Then there exist open sets $U_1$ and $U_2$ such that $K_1 \subset U_1$, $K_2 \subset U_2$ and $U_1 \cap U_2 = \emptyset$.*

**Proof** It follows from Proposition 1.26 that, for each point $x$ of $K_1$, there exist open sets $V_x$ and $W_x$ such that $x \in V_x$, $K_2 \subset W_x$ and $V_x \cap W_x = \emptyset$. But then there exists a finite set $\{x_1, x_2, \ldots, x_r\}$ of points of $K_1$ such that

$$K_1 \subset V_{x_1} \cup V_{x_2} \cup \cdots \cup V_{x_r},$$

since $K_1$ is compact. Define

$$U_1 = V_{x_1} \cup V_{x_2} \cup \cdots \cup V_{x_r}, \qquad U_2 = W_{x_1} \cap W_{x_2} \cap \cdots \cap W_{x_r}.$$

Then $U_1$ and $U_2$ are open sets, $K_1 \subset U_1$, $K_2 \subset U_2$ and $U_1 \cap U_2 = \emptyset$, as required. ∎

**Lemma 1.29** *Let $f: X \to Y$ be a continuous function from a compact topological space $X$ to a Hausdorff space $Y$. Then $f(K)$ is closed in $Y$ for every closed set $K$ in $X$.*

**Proof** If $K$ is a closed set in $X$, then $K$ is compact (Lemma 1.21), and therefore $f(K)$ is compact (Lemma 1.22). But any compact subset of a Hausdorff space is closed (Corollary 1.27). Thus $f(K)$ is closed in $Y$, as required. ∎

**Remark** If the Hausdorff space $Y$ in Lemma 1.29 is a metric space, then Proposition 1.25 may be used in place of Corollary 1.27 in the proof of the lemma.

**Theorem 1.30** *A continuous bijection $f: X \to Y$ from a compact topological space $X$ to a Hausdorff space $Y$ is a homeomorphism.*

**Proof** Let $g: Y \to X$ be the inverse of the bijection $f: X \to Y$. If $U$ is open in $X$ then $X \setminus U$ is closed in $X$, and hence $f(X \setminus U)$ is closed in $Y$, by Lemma 1.29. But $f(X \setminus U) = g^{-1}(X \setminus U) = Y \setminus g^{-1}(U)$. It follows that $g^{-1}(U)$ is open in $Y$ for every open set $U$ in $X$. Therefore $g: Y \to X$ is continuous, and thus $f: X \to Y$ is a homeomorphism. ∎

We recall that a function $f: X \to Y$ from a topological space $X$ to a topological space $Y$ is said to be an *identification map* if it is surjective and satisfies the following condition: a subset $U$ of $Y$ is open in $Y$ if and only if $f^{-1}(U)$ is open in $X$.

**Proposition 1.31** *A continuous surjection $f: X \to Y$ from a compact topological space $X$ to a Hausdorff space $Y$ is an identification map.*

**Proof** Let $U$ be a subset of $Y$. We claim that $Y \setminus U = f(K)$, where $K = X \setminus f^{-1}(U)$. Clearly $f(K) \subset Y \setminus U$. Also, given any $y \in Y \setminus U$, there exists $x \in X$ satisfying $y = f(x)$, since $f: X \to Y$ is surjective. Moreover $x \in K$, since $f(x) \notin U$. Thus $Y \setminus U \subset f(K)$, and hence $Y \setminus U = f(K)$, as claimed.

We must show that the set $U$ is open in $Y$ if and only if $f^{-1}(U)$ is open in $X$. First suppose that $f^{-1}(U)$ is open in $X$. Then $K$ is closed in $X$, and hence $f(K)$ is closed in $Y$, by Lemma 1.29. It follows that $U$ is open in $Y$. Conversely if $U$ is open in $Y$ then $f^{-1}(U)$ is open in $X$, since $f: X \to Y$ is continuous. Thus the surjection $f: X \to Y$ is an identification map. ∎

**Example** Let $S^1$ be the unit circle in $\mathbb{R}^2$, defined by $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, and let $q\colon [0, 1] \to S^1$ be defined by $q(t) = (\cos 2\pi t, \sin 2\pi t)$ for all $t \in [0, 1]$. It has been shown that the map $q$ is an identification map. This also follows directly from the fact that $q\colon [0, 1] \to S^1$ is a continuous surjection from the compact space $[0, 1]$ to the Hausdorff space $S^1$.

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 1.32** *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 $V = \{x \in X : \{x\} \times K \subset U\}$. Then $V$ is an open set in $X$.*

**Proof** Let $x \in V$. For each $y \in K$ 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 $D_y \times E_y \subset U$. Now there exists a finite set $\{y_1, y_2, \ldots, y_k\}$ of points of $K$ such that $K \subset E_{y_1} \cup E_{y_2} \cup \cdots \cup E_{y_k}$, since $K$ is compact. 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

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

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 1.33** *A Cartesian product of a finite number of compact spaces is itself compact.*

**Proof** It suffices to prove that the product of two compact topological spaces $X$ and $Y$ is compact, since the general result then follows easily by induction on the number of compact spaces in the product.

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 $X \times Y$, since it is the image of the compact space $Y$ under the continuous map from $Y$ to $X \times Y$ which sends $y \in Y$ to $(x, y)$, and the image of any compact set under a continuous map is itself compact (Lemma 1.22). Therefore there exists a finite collection $U_1, U_2, \ldots, U_r$ of open sets belonging to the open cover $\mathcal{U}$ such that $\{x\} \times Y$ is contained in $U_1 \cup U_2 \cup \cdots \cup U_r$. Let $V_x$ denote the set of all points $x'$ of $X$ for which $\{x'\} \times Y$ is contained in $U_1 \cup U_2 \cup \cdots \cup U_r$. Then $x \in V_x$, and Lemma 1.32 ensures That $V_x$ is an open set in $X$. Note 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 space $X$. It follows from the compactness of $X$ that there exists a finite set $\{x_1, x_2, \ldots, x_r\}$ of points of $X$ such that $X = V_{x_1} \cup V_{x_2} \cup \cdots \cup V_{x_r}$. Now $X \times Y$ is the union of the sets $V_{x_j} \times Y$ for $j = 1, 2, \ldots, r$, and each of these sets can be covered by a finite collection of open sets belonging to the open cover $\mathcal{U}$. On combining these finite collections, we obtain a finite collection of open sets belonging to $\mathcal{U}$ which covers $X \times Y$. This shows that $X \times Y$ is compact. ∎

**Theorem 1.34** *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. Then $K$ is closed, since $\mathbb{R}^n$ is Hausdorff, and a compact subset of a Hausdorff space is closed (by Corollary 1.27). For each natural number $m$, let $B_m$ be the open ball of radius $m$ about the origin, given by $B_m = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| < m\}$. Then $\{B_m : m \in \mathbb{N}\}$ is an open cover of $\mathbb{R}^n$. It follows from the compactness of $K$ that there exist natural numbers $m_1, m_2, \ldots, m_k$ such that $K \subset B_{m_1} \cup B_{m_2} \cup \cdots \cup B_{m_k}$. But then $K \subset B_M$, where $M$ is the maximum of $m_1, m_2, \ldots, m_k$, and 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 1.20), and $C$ is the Cartesian product of $n$ copies of the compact set $[-L, L]$. It follows from Theorem 1.33 that $C$ is compact. But $K$ is a closed subset of $C$, and a closed subset of a compact topological space is itself compact, by Lemma 1.21. Thus $K$ is compact, as required. ∎

## 1.17 The Lebesgue Lemma and Uniform Continuity

**Definition** Let $X$ be a metric space with distance function $d$. A subset $A$ of $X$ is said to be *bounded* if there exists a non-negative real number $K$ such that $d(x, y) \leq K$ for all $x, y \in A$. The smallest real number $K$ with this property is referred to as the *diameter* of $A$, and is denoted by $\operatorname{diam} A$. (Note that $\operatorname{diam} A$ is the supremum of the values of $d(x, y)$ as $x$ and $y$ range over all points of $A$.)

**Lemma 1.35** (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 $\delta$ such that* *every subset of $X$ whose diameter is less than $\delta$ 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 $\delta_x$ 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 $\delta$. 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 $\delta$ such that every subset of $X$ whose diameter is less than $\delta$ 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.

Let $X$ and $Y$ be metric spaces with distance functions $d_X$ and $d_Y$ respectively, and let $f: X \to Y$ be a function from $X$ to $Y$. The function $f$ is said to be *uniformly continuous* on $X$ if and only if, given $\varepsilon > 0$, there exists some $\delta > 0$ such that $d_Y(f(x), f(x')) < \varepsilon$ for all points $x$ and $x'$ of $X$ satisfying $d_X(x, x') < \delta$. (The value of $\delta$ should be independent of both $x$ and $x'$.)

**Theorem 1.36** *Let $X$ and $Y$ be metric spaces. Suppose that $X$ is compact. Then every continuous function from $X$ to $Y$ is uniformly continuous.*

**Proof** Let $d_X$ and $d_Y$ denote the distance functions for the metric spaces $X$ and $Y$ respectively. Let $f: X \to Y$ be a continuous function from $X$ to $Y$. We must show that $f$ is uniformly continuous.

Let $\varepsilon > 0$ be given. For each $y \in Y$, define

$$V_y = \{x \in X : d_Y(f(x), y) < \tfrac{1}{2}\varepsilon\}.$$

Note that $V_y = f^{-1}\left(B_Y(y, \frac{1}{2}\varepsilon)\right)$, where $B_Y(y, \frac{1}{2}\varepsilon)$ denotes the open ball of radius $\frac{1}{2}\varepsilon$ about $y$ in $Y$. Now the open ball $B_Y(y, \frac{1}{2}\varepsilon)$ is an open set in $Y$, and $f$ is continuous. Therefore $V_y$ is open in $X$ for all $y \in Y$. Note that $x \in V_{f(x)}$ for all $x \in X$.

Now $\{V_y : y \in Y\}$ is an open cover of the compact metric space $X$. It follows from the Lebesgue Lemma (Lemma 1.35) that there exists some $\delta > 0$ such that every subset of $X$ whose diameter is less than $\delta$ is a subset of some set $V_y$. Let $x$ and $x'$ be points of $X$ satisfying $d_X(x, x') < \delta$. The diameter of the set $\{x, x'\}$ is $d_X(x, x')$, which is less than $\delta$. Therefore there exists some $y \in Y$ such that $x \in V_y$ and $x' \in V_y$. But then $d_Y(f(x), y) < \frac{1}{2}\varepsilon$ and $d_Y(f(x'), y) < \frac{1}{2}\varepsilon$, and hence

$$d_Y(f(x), f(x')) \leq d_Y(f(x), y) + d_Y(y, f(x')) < \varepsilon.$$

This shows that $f\colon X \to Y$ is uniformly continuous, as required. ∎

Let $K$ be a closed bounded subset of $\mathbb{R}^n$. It follows from Theorem 1.34) and Theorem 1.36 that any continuous function $f\colon K \to \mathbb{R}^k$ is uniformly continuous.

## 1.18 Connected Topological Spaces

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

**Lemma 1.37** *A topological space $X$ is connected if and only if it has the following property: 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** If $U$ is a subset of $X$ that is both open and closed, and if $V = X \setminus U$, then $U$ and $V$ are both open, $U \cup V = X$ and $U \cap V = \emptyset$. Conversely if $U$ and $V$ are open subsets of $X$ satisfying $U \cup V = X$ and $U \cap V = \emptyset$, then $U = X \setminus V$, and hence $U$ is both open and closed. Thus a topological space $X$ is connected if and only if there do not exist non-empty open sets $U$ and $V$ such that $U \cup V = X$ and $U \cap V = \emptyset$. The result follows. ∎

Let $\mathbb{Z}$ be the set of integers with the usual topology (i.e., the subspace topology on $\mathbb{Z}$ induced by the usual topology on $\mathbb{R}$). Then $\{n\}$ is open for all $n \in \mathbb{Z}$, since

$$\{n\} = \mathbb{Z} \cap \{t \in \mathbb{R} : |t - n| < \tfrac{1}{2}\}.$$

It follows that every subset of $\mathbb{Z}$ is open (since it is a union of sets consisting of a single element, and any union of open sets is open). It follows that a function $f \colon X \to \mathbb{Z}$ on a topological space $X$ is continuous if and only if $f^{-1}(V)$ is open in $X$ for any subset $V$ of $\mathbb{Z}$. We use this fact in the proof of the next theorem.

**Proposition 1.38** *A topological space $X$ is connected if and only if every continuous function $f \colon X \to \mathbb{Z}$ from $X$ to the set $\mathbb{Z}$ of integers is constant.*

**Proof** Suppose that $X$ is connected. Let $f \colon X \to \mathbb{Z}$ be a continuous function. Choose $n \in f(X)$, and let

$$U = \{x \in X : f(x) = n\}, \qquad V = \{x \in X : f(x) \neq n\}.$$

Then $U$ and $V$ are the preimages of the open subsets $\{n\}$ and $\mathbb{Z} \setminus \{n\}$ of $\mathbb{Z}$, and therefore both $U$ and $V$ are open in $X$. Moreover $U \cap V = \emptyset$, and $X = U \cup V$. It follows that $V = X \setminus U$, and thus $U$ is both open and closed. Moreover $U$ is non-empty, since $n \in f(X)$. It follows from the connectedness of $X$ that $U = X$, so that $f \colon X \to \mathbb{Z}$ is constant, with value $n$.

Conversely suppose that every continuous function $f \colon X \to \mathbb{Z}$ is constant. Let $S$ be a subset of $X$ which is both open and closed. Let $f \colon X \to \mathbb{Z}$ be defined by

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

Now the preimage of any subset of $\mathbb{Z}$ under $f$ is one of the open sets $\emptyset$, $S$, $X \setminus S$ and $X$. Therefore the function $f$ is continuous. But then the function $f$ is constant, so that either $S = \emptyset$ or $S = X$. This shows that $X$ is connected. ∎

**Lemma 1.39** *The closed interval $[a, b]$ is connected, for all real numbers $a$ and $b$ satisfying $a \leq b$.*

**Proof** Let $f \colon [a, b] \to \mathbb{Z}$ be a continuous integer-valued function on $[a, b]$. We show that $f$ is constant on $[a, b]$. Indeed suppose that $f$ were not constant. Then $f(\tau) \neq f(a)$ for some $\tau \in [a, b]$. But the Intermediate Value Theorem would then ensure that, given any real number $c$ between $f(a)$ and $f(\tau)$, there would exist some $t \in [a, \tau]$ for which $f(t) = c$, and this is clearly impossible, since $f$ is integer-valued. Thus $f$ must be constant on $[a, b]$. We now deduce from Proposition 1.38 that $[a, b]$ is connected. ∎

**Example** Let $X = \{(x, y) \in \mathbb{R}^2 : x \neq 0\}$. The topological space $X$ is not connected. Indeed if $f\colon X \to \mathbb{Z}$ is defined by

$$f(x, y) = \begin{cases} 1 & \text{if } x > 0, \\ -1 & \text{if } x < 0, \end{cases}$$

then $f$ is continuous on $X$ but is not constant.

A concept closely related to that of connectedness is *path-connectedness*. Let $x_0$ and $x_1$ be points in a topological space $X$. A *path* in $X$ from $x_0$ to $x_1$ is defined to be a continuous function $\gamma\colon [0, 1] \to X$ such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$. 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 path in $X$ from $x_0$ to $x_1$.

**Proposition 1.40** *Every path-connected topological space is connected.*

**Proof** Let $X$ be a path-connected topological space, and let $f\colon X \to \mathbb{Z}$ be a continuous integer-valued function on $X$. If $x_0$ and $x_1$ are any two points of $X$ then there exists a path $\gamma\colon [0, 1] \to X$ such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$. But then $f \circ \gamma\colon [0, 1] \to \mathbb{Z}$ is a continuous integer-valued function on $[0, 1]$. But $[0, 1]$ is connected (Lemma 1.39), therefore $f \circ \gamma$ is constant (Proposition 1.38). It follows that $f(x_0) = f(x_1)$. Thus every continuous integer-valued function on $X$ is constant. Therefore $X$ is connected, by Proposition 1.38. ∎

The topological spaces $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{R}^n$ are all path-connected. Indeed, given any two points of one of these spaces, the straight line segment joining these two points is a continuous path from one point to the other. Also the $n$-sphere $S^n$ is path-connected for all $n > 0$. We conclude that these topological spaces are connected.

Let $A$ be a subset of a topological space $X$. Using Lemma 1.37 and the definition of the subspace topology, we see that $A$ is connected 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 1.41** *Let $X$ be a topological space and let $A$ be a connected subset of $X$. Then the closure $\overline{A}$ of $A$ is connected.*

**Proof** It follows from the definition of the closure of $A$ that $\overline{A} \subset F$ for any closed subset $F$ of $X$ for which $A \subset F$. On taking $F$ to be the complement of some open set $U$, we deduce that $\overline{A} \cap U = \emptyset$ for any open set $U$ for which

$A \cap U = \emptyset$. Thus if $U$ is an open set in $X$ and if $\overline{A} \cap U$ is non-empty then $A \cap U$ must also be non-empty.

Now 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, and $A \subset U \cup V$. But $A$ is connected. Therefore $A \cap U \cap V$ is non-empty, and thus $\overline{A} \cap U \cap V$ is non-empty. This shows that $\overline{A}$ is connected. ∎

**Lemma 1.42** *Let $f: X \to Y$ be a continuous function between topological spaces $X$ and $Y$, and let $A$ be a connected subset of $X$. Then $f(A)$ is connected.*

**Proof** Let $g: f(A) \to \mathbb{Z}$ be any continuous integer-valued function on $f(A)$. Then $g \circ f: A \to \mathbb{Z}$ is a continuous integer-valued function on $A$. It follows from Proposition 1.38 that $g \circ f$ is constant on $A$. Therefore $g$ is constant on $f(A)$. We deduce from Proposition 1.38 that $f(A)$ is connected. ∎

**Lemma 1.43** *The Cartesian product $X \times Y$ of connected topological spaces $X$ and $Y$ is itself connected.*

**Proof** Let $f: X \times Y \to \mathbb{Z}$ be a continuous integer-valued function from $X \times Y$ to $Z$. Choose $x_0 \in X$ and $y_0 \in Y$. The function $x \mapsto f(x, y_0)$ is continuous on $X$, and is thus constant. Therefore $f(x, y_0) = f(x_0, y_0)$ for all $x \in X$. Now fix $x$. The function $y \mapsto f(x, y)$ is continuous on $Y$, and is thus constant. Therefore

$$f(x, y) = f(x, y_0) = f(x_0, y_0)$$

for all $x \in X$ and $y \in Y$. We deduce from Proposition 1.38 that $X \times Y$ is connected. ∎

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

**Proposition 1.44** *Let $X$ be a topological space. For each $x \in X$, let $S_x$ be the union of all connected subsets of $X$ that contain $x$. Then*

(i) *$S_x$ is connected,*

(ii) *$S_x$ is closed,*

(iii) *if $x, y \in X$, then either $S_x = S_y$, or else $S_x \cap S_y = \emptyset$.*

**Proof** Let $f: S_x \to \mathbb{Z}$ be a continuous integer-valued function on $S_x$, for some $x \in X$. Let $y$ be any point of $S_x$. Then, by definition of $S_x$, there exists some connected set $A$ containing both $x$ and $y$. But then $f$ is constant on $A$, and thus $f(x) = f(y)$. This shows that the function $f$ is constant on $S_x$. We deduce that $S_x$ is connected. This proves (i). Moreover the closure $\overline{S_x}$ is connected, by Lemma 1.41. Therefore $\overline{S_x} \subset S_x$. This shows that $S_x$ is closed, proving (ii).

Finally, suppose that $x$ and $y$ are points of $X$ for which $S_x \cap S_y \neq \emptyset$. Let $f: S_x \cup S_y \to \mathbb{Z}$ be any continuous integer-valued function on $S_x \cup S_y$. Then $f$ is constant on both $S_x$ and $S_y$. Moreover the value of $f$ on $S_x$ must agree with that on $S_y$, since $S_x \cap S_y$ is non-empty. We deduce that $f$ is constant on $S_x \cup S_y$. Thus $S_x \cup S_y$ is a connected set containing both $x$ and $y$, and thus $S_x \cup S_y \subset S_x$ and $S_x \cup S_y \subset S_y$, by definition of $S_x$ and $S_y$. We conclude that $S_x = S_y$. This proves (iii). ∎

Given any topological space $X$, the connected subsets $S_x$ of $X$ defined as in the statement of Proposition 1.44 are referred to as the *connected components* of $X$. We see from Proposition 1.44, part (iii) that the topological space $X$ is the disjoint union of its connected components.

**Example** The connected components of $\{(x, y) \in \mathbb{R}^2 : x \neq 0\}$ are

$$\{(x, y) \in \mathbb{R}^2 : x > 0\} \text{ and } \{(x, y) \in \mathbb{R}^2 : x < 0\}.$$

**Example** The connected components of

$$\{t \in \mathbb{R} : |t - n| < \tfrac{1}{2} \text{ for some integer } n\}.$$

are the sets $J_n$ for all $n \in \mathbb{Z}$, where $J_n = (n - \frac{1}{2}, n + \frac{1}{2})$.

# 2 Homotopies and Covering Maps

## 2.1 Homotopies

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

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

**Lemma 2.1** *Let $X$ and $Y$ be topological spaces. The homotopy relation $\simeq$ is an equivalence relation on the set of all continuous maps from $X$ to $Y$.*

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

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

Now $H|X \times [0, \frac{1}{2}]$ and $H|X \times [\frac{1}{2}, 1]$ are continuous. It follows from elementary point set topology that $H$ is continuous on $X \times [0,1]$. Moreover $H(x,0) = f(x)$ and $H(x,1) = h(x)$ for all $x \in X$. Thus $f \simeq h$. Thus the relation is transitive. The relation $\simeq$ is therefore an equivalence relation. ∎

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

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

## 2.2 Covering Maps

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

**Lemma 2.2** *Let $p: \tilde{X} \to X$ be a covering map. Then $p(V)$ is open in $X$ for every open set $V$ in $\tilde{X}$. In particular, a covering map $p: \tilde{X} \to X$ is a homeomorphism if and only if it is a bijection.*

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

## 2.3 Path Lifting and the Monodromy Theorem

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

**Proposition 2.3** *Let $p: \tilde{X} \to X$ be a covering map, let $Z$ be a connected topological space, and let $g: Z \to \tilde{X}$ and $h: Z \to \tilde{X}$ be continuous maps. Suppose that $p \circ g = p \circ h$ and that $g(z) = h(z)$ for some $z \in Z$. Then $g = h$.*

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

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

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

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

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

**Lemma 2.4** *Let $p: \tilde{X} \to X$ be a covering map, let $Z$ be a topological space, let $A$ be a connected subset of $Z$, and let $f: Z \to X$ and $g: A \to \tilde{X}$ be continuous maps with the property that $p \circ g = f|A$. Suppose that $f(Z) \subset U$, where $U$ is an open subset of $X$ that is evenly covered by the covering map $p$. Then there exists a continuous map $\tilde{f}: Z \to \tilde{X}$ such that $\tilde{f}|A = g$ and $p \circ \tilde{f} = f$.*

**Proof** The open set $U$ is evenly covered by the covering map $p$, and therefore $p^{-1}(U)$ is a disjoint union of open sets, each of which is mapped homeomorphically onto $U$ by the covering map $p$. One of these open sets contains $g(a)$ for some $a \in A$; let this set be denoted by $\tilde{U}$. Let $\sigma: U \to \tilde{U}$ be the inverse of the homeomorphism $p|\tilde{U}: \tilde{U} \to U$, and let $\tilde{f} = \sigma \circ f$. Then $p \circ \tilde{f} = f$. Also $p \circ \tilde{f}|A = p \circ g$ and $\tilde{f}(a) = g(a)$. It follows from Proposition 2.3 that $\tilde{f}|A = g$, since $A$ is connected. Thus $\tilde{f}: Z \to \tilde{X}$ is the required map. ∎

**Theorem 2.5** (Path Lifting Theorem) *Let $p: \tilde{X} \to X$ be a covering map, let $\gamma: [0, 1] \to X$ be a continuous path in $X$, and let $w$ be a point of $\tilde{X}$ satisfying $p(w) = \gamma(0)$. Then there exists a unique continuous path $\tilde{\gamma}: [0, 1] \to \tilde{X}$ such that $\tilde{\gamma}(0) = w$ and $p \circ \tilde{\gamma} = \gamma$.*

**Proof** The map $p: \tilde{X} \to X$ is a covering map; therefore there exists an open cover $\mathcal{U}$ of $X$ such that each open set $U$ belonging to $X$ is evenly covered by the map $p$. Now the collection consisting of the preimages $\gamma^{-1}(U)$ of the open sets $U$ belonging to $\mathcal{U}$ is an open cover of the interval $[0, 1]$. But $[0, 1]$ is compact, by the Heine-Borel Theorem. It follows from the Lebesgue Lemma that there exists some $\delta > 0$ such that every subinterval of length less than $\delta$ is mapped by $\gamma$ into one of the open sets belonging to $\mathcal{U}$. Partition the interval $[0, 1]$ into subintervals $[t_{i-1}, t_i]$, where $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = 1$, and where the length of each subinterval is less than $\delta$. Then each subinterval $[t_{i-1}, t_i]$ is mapped by $\gamma$ into some open set in $X$ that is evenly covered by the map $p$. It follows from Lemma 2.4 that once $\tilde{\gamma}(t_{i-1})$ has been determined, we can extend $\tilde{\gamma}$ continuously over the $i$th subinterval $[t_{i-1}, t_i]$. Thus by extending $\tilde{\gamma}$ successively over $[t_0, t_1]$, $[t_1, t_2]$,. . ., $[t_{n-1}, t_n]$, we can lift the path $\gamma: [0, 1] \to X$ to a path $\tilde{\gamma}: [0, 1] \to \tilde{X}$ starting at $w$. The uniqueness of $\tilde{\gamma}$ follows from Proposition 2.3. ∎

**Theorem 2.6** (The Monodromy Theorem) *Let $p: \tilde{X} \to X$ be a covering map, let $H: [0, 1] \times [0, 1] \to X$ be a continuous map, and let $w$ be a point of $\tilde{X}$ satisfying $p(w) = H(0, 0)$. Then there exists a unique continuous map $\tilde{H}: [0, 1] \times [0, 1] \to \tilde{X}$ such that $\tilde{H}(0, 0) = w$ and $p \circ \tilde{H} = H$.*

**Proof** The unit square $[0, 1] \times [0, 1]$ is compact. By applying the Lebesgue Lemma to an open cover of the square by preimages of evenly covered open sets in $X$ (as in the proof of Theorem 2.5), we see that there exists some $\delta > 0$ with the property that any square contained in $[0, 1] \times [0, 1]$ whose sides have length less than $\delta$ is mapped by $H$ into some open set in $X$ which is evenly covered by the covering map $p$. It follows from Lemma 2.4 that if the lift $\tilde{H}$ of $H$ has already been determined over a corner, or along one side, or along two adjacent sides of a square whose sides have length less than $\delta$, then $\tilde{H}$ can be extended over the whole of that square. Thus if we subdivide $[0, 1] \times [0, 1]$ into squares $S_{j,k}$, where

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

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

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

The uniqueness of $\tilde{H}$ follows from Proposition 2.3. ∎

# 3 The Fundamental Group

## 3.1 The Fundamental Group of a Topological Space

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

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

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

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

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

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

**Theorem 3.1** *Let $X$ be a topological space, let $x_0$ be some chosen point of $X$, and let $\pi_1(X, x_0)$ be the set of all based homotopy classes of loops based at the point $x_0$. Then $\pi_1(X, x_0)$ is a group, the group multiplication on $\pi_1(X, x_0)$ being defined according to the rule $[\gamma_1][\gamma_2] = [\gamma_1.\gamma_2]$ for all loops $\gamma_1$ and $\gamma_2$ based at $x_0$.*

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

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

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

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

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

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

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

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

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

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

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

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

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

Let $f: X \to Y$ be a continuous map between topological spaces $X$ and $Y$, and let $x_0$ be a point of $X$. Then $f$ induces a homomorphism $f_\#: \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$, where $f_\#([\gamma]) = [f \circ \gamma]$ for all loops $\gamma: [0,1] \to X$ based at $x_0$. If $x_0$, $y_0$ and $z_0$ are points belonging to topological spaces $X$, $Y$ and $Z$, and if $f: X \to Y$ and $g: Y \to Z$ are continuous maps satisfying $f(x_0) = y_0$ and

$g(y_0) = z_0$, then the induced homomorphisms $f_\#: \pi_1(X, x_0) \to \pi_1(Y, y_0)$ and $g_\#: \pi_1(Y, x_0) \to \pi_1(Z, z_0)$ satisfy $g_\# \circ f_\# = (g \circ f)_\#$. It follows easily from this that any homeomorphism of topological spaces induces a corresponding isomorphism of fundamental groups, and thus the fundamental group is a topological invariant.

## 3.2 Simply-Connected Topological Spaces

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

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

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

**Theorem 3.2** *A path-connected topological space $X$ is simply-connected if and only if $\pi_1(X, x)$ is trivial for all $x \in X$.*

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

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

and $G(t,1) = x$ for all $t \in [0,1]$ and $G(0,\tau) = G(1,\tau) = x$ for all $\tau \in [0,1]$, since $\pi_1(X,x)$ is trivial. Moreover $G(t_1,\tau_1) = G(t_2,\tau_2)$ whenever $q(t_1,\tau_1) = q(t_2,\tau_2)$, where

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

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

One can show that, if two points $x_1$ and $x_2$ in a topological space $X$ can be joined by a path in $X$ then $\pi_1(X,x_1)$ and $\pi_1(X,x_2)$ are isomorphic. On combining this result with Theorem 3.2, we see that a path-connected topological space $X$ is simply-connected if and only if $\pi_1(X,x)$ is trivial for some $x \in X$.

**Theorem 3.3** *Let $X$ be a topological space, and let $U$ and $V$ be open subsets of $X$, with $U \cup V = X$. Suppose that $U$ and $V$ are simply-connected, and that $U \cap V$ is non-empty and path-connected. Then $X$ is itself simply-connected.*

**Proof** We must show that any continuous function $f\colon \partial D \to X$ defined on the unit circle $\partial D$ can be extended continuously over the closed unit disk $D$. Now the preimages $f^{-1}(U)$ and $f^{-1}(V)$ of $U$ and $V$ are open in $\partial D$ (since $f$ is continuous), and $\partial D = f^{-1}(U) \cup f^{-1}(V)$. It follows from the Lebesgue Lemma that there exists some $\delta > 0$ such that any arc in $\partial D$ whose length is less than $\delta$ is entirely contained in one or other of the sets $f^{-1}(U)$ and $f^{-1}(V)$. Choose points $z_1, z_2, \ldots, z_n$ around $\partial D$ such that the distance from $z_i$ to $z_{i+1}$ is less than $\delta$ for $i = 1, 2, \ldots, n-1$ and the distance from $z_n$ to $z_1$ is also less than $\delta$. Then, for each $i$, the short arc joining $z_{i-1}$ to $z_i$ is mapped by $f$ into one or other of the open sets $U$ and $V$.

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

Now, for each $i$, consider the sector $T_i$ of the closed unit disk bounded by the line segments joining the centre of the disk to the points $z_{i-1}$ and $z_i$ and by the short arc joining $z_{i-1}$ to $z_i$. Now this sector is homeomorphic to the closed unit disk, and therefore any continuous function mapping the boundary $\partial T_i$

of $T_i$ into a simply-connected space can be extended continuously over the whole of $T_i$. In particular, let $F_i$ be the function on $\partial T_i$ defined by

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

Note that $F_i(\partial T_i) \subset U$ whenever the short arc joining $z_{i-1}$ to $z_i$ is mapped by $f$ into $U$, and $F_i(\partial T_i) \subset V$ whenever this short arc is mapped into $V$. But $U$ and $V$ are both simply-connected. It follows that each of the functions $F_i$ can be extended continuously over the whole of the sector $T_i$. Moreover the functions defined in this fashion on each of the sectors $T_i$ agree with one another wherever the sectors intersect, and can therefore be pieced together to yield a continuous map defined over the the whole of the closed disk $D$ which extends the map $f$, as required. ∎

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

## 3.3 The Fundamental Group of the Circle

**Theorem 3.4** *$\pi_1(S^1, b) \cong \mathbb{Z}$ for any $b \in S^1$.*

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

Let $\alpha$ and $\beta$ be loops in $S^1$ based at $b$, and let $\tilde{\alpha}$ and $\tilde{\beta}$ be paths in $\mathbb{R}$ that satisfy $p \circ \tilde{\alpha} = \alpha$ and $p \circ \tilde{\beta} = \beta$. Suppose that $\alpha$ and $\beta$ represent the same element of $\pi_1(S^1, b)$. Then there exists a homotopy $F\colon [0,1] \times [0,1] \to S^1$ such that $F(t, 0) = \alpha(t)$ and $F(t, 1) = \beta(t)$ for all $t \in [0,1]$, and $F(0, \tau) = F(1, \tau) = b$ for all $\tau \in [0,1]$. It follows from the Monodromy Theorem (Theorem 2.6) that this homotopy lifts to a continuous map $G\colon [0,1] \times [0,1] \to \mathbb{R}$ satisfying $p \circ G = F$. Moreover $G(0, \tau)$ and $G(1, \tau)$ are integers for all $\tau \in [0,1]$, since $p(G(0, \tau)) = b = p(G(1, \tau))$. Also $G(t, 0) - \tilde{\alpha}(t)$ and $G(t, 1) - \tilde{\beta}(t)$ are integers for all $t \in [0,1]$, since $p(G(t, 0)) = \alpha(t) = p(\tilde{\alpha}(t))$ and $p(G(t, 1)) = \beta(t) = p(\tilde{\beta}(t))$. Now any continuous integer-valued function on $[0,1]$ is constant, by the Intermediate Value Theorem. In particular the functions sending $\tau \in [0,1]$ to $G(0,\tau)$ and $G(1,\tau)$ are constant, as are the functions sending $t \in [0,1]$ to $G(t,0) - \tilde{\alpha}(t)$ and $G(t,1) - \tilde{\beta}(t)$. Thus

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

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

On combining these results, we see that

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

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

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

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

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

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

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

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

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

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

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

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

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

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

**Theorem 3.5** *Let $f\colon D \to D$ be a continuous map which maps the closed disk $D$ into itself. Then $f(\mathbf{x}_0) = \mathbf{x}_0$ for some $\mathbf{x}_0 \in D$.*

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

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

Let $r_\#\colon \pi_1(D, \mathbf{b}) \to \pi_1(\partial D, \mathbf{b})$ be the homomorphism of fundamental groups induced by $r\colon D \to \partial D$. Now $(r \circ i)_\#\colon \pi_1(\partial D, \mathbf{b}) \to \pi_1(\partial D, \mathbf{b})$ is the identity isomorphism of $\pi_1(\partial D, \mathbf{b})$, since $r \circ i\colon \partial D \to \partial D$ is the identity map. But it follows directly from the definition of induced homomorphisms that $(r \circ i)_\# = r_\# \circ i_\#$. Therefore $i_\#\colon \pi_1(\partial D, \mathbf{b}) \to \pi_1(D, \mathbf{b})$ is injective, and $r_\#\colon \pi_1(D, \mathbf{b}) \to \pi_1(\partial D, \mathbf{b})$ is surjective. But this is impossible, since $\pi_1(\partial D, \mathbf{b}) \cong \mathbb{Z}$ (Theorem 3.4) and $\pi_1(D, \mathbf{b})$ is the trivial group. This contradiction shows that the continuous map $f\colon D \to D$ must have at least one fixed point. ∎

# 4 Simplicial Complexes

## 4.1 Geometrical Independence

**Definition** Points $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$ in some Euclidean space $\mathbb{R}^k$ are said to be *geometrically independent* (or *affine independent*) if the only solution of the linear system

$$\left\{ \begin{array}{rcl} \sum_{j=0}^{q} \lambda_j \mathbf{v}_j & = & \mathbf{0}, \\ \sum_{j=0}^{q} \lambda_j & = & 0 \end{array} \right.$$

is the trivial solution $\lambda_0 = \lambda_1 = \cdots = \lambda_q = 0$.

It is straightforward to verify that $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$ are geometrically independent if and only if the vectors $\mathbf{v}_1 - \mathbf{v}_0, \mathbf{v}_2 - \mathbf{v}_0, \ldots, \mathbf{v}_q - \mathbf{v}_0$ are linearly independent. It follows from this that any set of geometrically independent points in $\mathbb{R}^k$ has at most $k + 1$ elements. Note also that if a set consists of geometrically independent points in $\mathbb{R}^k$, then so does every subset of that set.

**Definition** A *q-simplex* in $\mathbb{R}^k$ is defined to be a set of the form

$$\left\{ \sum_{j=0}^{q} t_j \mathbf{v}_j : 0 \leq t_j \leq 1 \text{ for } j = 0, 1, \ldots, q \text{ and } \sum_{j=0}^{q} t_j = 1 \right\},$$

where $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$ are geometrically independent points of $\mathbb{R}^k$. The points $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$ are referred to as the *vertices* of the simplex. The non-negative integer $q$ is referred to as the *dimension* of the simplex.

Note that a 0-simplex in $\mathbb{R}^k$ is a single point of $\mathbb{R}^k$, a 1-simplex in $\mathbb{R}^k$ is a line segment in $\mathbb{R}^k$, a 2-simplex is a triangle, and a 3-simplex is a tetrahedron.

Let $\sigma$ be a $q$-simplex in $\mathbb{R}^k$ with vertices $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$. If $\mathbf{x}$ is a point of $\sigma$ then there exist real numbers $t_0, t_1, \ldots, t_q$ such that

$$\sum_{j=0}^{q} t_j \mathbf{v}_j = \mathbf{x}, \quad \sum_{j=0}^{q} t_j = 1 \text{ and } 0 \leq t_j \leq 1 \text{ for } j = 0, 1, \ldots, q.$$

Moreover $t_0, t_1, \ldots, t_q$ are uniquely determined: if $\sum\limits_{j=0}^{q} s_j \mathbf{v}_j = \sum\limits_{j=0}^{q} t_j \mathbf{v}_j$ and $\sum\limits_{j=0}^{q} s_j = 1 = \sum\limits_{j=0}^{q} t_j$, then $\sum\limits_{j=0}^{q} (t_j - s_j)\mathbf{v}_j = \mathbf{0}$ and $\sum\limits_{j=0}^{q} (t_j - s_j) = 0$, hence $t_j - s_j = 0$ for all $j$, since $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$ are geometrically independent. We refer to $t_0, t_1, \ldots, t_q$ as the *barycentric coordinates* of the point $\mathbf{x}$ of $\sigma$.

**Lemma 4.1** *Let $q$ be a non-negative integer, let $\sigma$ be a $q$-simplex in $\mathbb{R}^m$, and let $\tau$ be a $q$-simplex in $\mathbb{R}^n$, where $m \geq q$ and $n \geq q$. Then $\sigma$ and $\tau$ are homeomorphic.*

**Proof** Let $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$ be the vertices of $\sigma$, and let $\mathbf{w}_0, \mathbf{w}_1, \ldots, \mathbf{w}_q$ be the vertices of $\tau$. The required homeomorphism $h\colon \sigma \to \tau$ is given by

$$h\left( \sum_{j=0}^{q} t_j \mathbf{v}_j \right) = \sum_{j=0}^{q} t_j \mathbf{w}_j$$

for all $t_0, t_1, \ldots, t_q$ satisfying $0 \leq t_j \leq 1$ for $j = 0, 1, \ldots, q$ and $\sum\limits_{j=0}^{q} t_j = 1$. ∎

A homeomorphism between two $q$-simplices defined as in the above proof is referred to as a *simplicial homeomorphism*.

## 4.2 Simplicial Complexes in Euclidean Spaces

**Definition** Let $\sigma$ and $\tau$ be simplices in $\mathbb{R}^k$. We say that $\tau$ is a *face* of $\sigma$ if the set of vertices of $\tau$ is a subset of the set of vertices of $\sigma$. A face of $\sigma$ is said to be a *proper face* if it is not equal to $\sigma$ itself. An $r$-dimensional face of $\sigma$ is referred to as an *$r$-face* of $\sigma$. A 1-dimensional face of $\sigma$ is referred to as an *edge* of $\sigma$.

Note that any simplex is a face of itself. Also the vertices and edges of any simplex are by definition faces of the simplex.

**Definition** A finite collection $K$ of simplices in $\mathbb{R}^k$ is said to be a *simplicial complex* if the following two conditions are satisfied:—

- if $\sigma$ is a simplex belonging to $K$ then every face of $\sigma$ also belongs to $K$,
- if $\sigma_1$ and $\sigma_2$ are simplices belonging to $K$ then either $\sigma_1 \cap \sigma_2 = \emptyset$ or else $\sigma_1 \cap \sigma_2$ is a common face of both $\sigma_1$ and $\sigma_2$.

The *dimension* of a simplicial complex $K$ is the greatest non-negative integer $n$ with the property that $K$ contains an $n$-simplex. The union of all the simplices of $K$ is a compact subset $|K|$ of $\mathbb{R}^k$ referred to as the *polyhedron* of $K$. (The polyhedron is compact since it is both closed and bounded in $\mathbb{R}^k$.)

**Example** Let $K_\sigma$ consist of some $n$-simplex $\sigma$ together with all of its faces. Then $K_\sigma$ is a simplicial complex of dimension $n$, and $|K_\sigma| = \sigma$.

**Lemma 4.2** *Let $K$ be a simplicial complex, and let $X$ be a topological space. A function $f\colon |K| \to X$ is continuous on the polyhedron $|K|$ of $K$ if and only if the restriction of $f$ to each simplex of $K$ is continuous on that simplex.*

**Proof** If a topological space can be expressed as a finite union of closed subsets, then a function is continuous on the whole space if and only if its restriction to each of the closed subsets is continuous on that closed set. The required result is a direct application of this general principle. ∎

We shall denote by $\operatorname{Vert} K$ the set of vertices of a simplicial complex $K$ (i.e., the set consisting of all vertices of all simplices belonging to $K$). A collection of vertices of $K$ is said to *span* a simplex of $K$ if these vertices are the vertices of some simplex belonging to $K$.

**Definition** Let $K$ be a simplicial complex in $\mathbb{R}^k$. A *subcomplex* of $K$ is a collection $L$ of simplices belonging to $K$ with the following property:—

- if $\sigma$ is a simplex belonging to $L$ then every face of $\sigma$ also belongs to $L$.

Note that every subcomplex of a simplicial complex $K$ is itself a simplicial complex.

**Definition** Let $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$ be the vertices of a $q$-simplex $\sigma$ in some Euclidean space $\mathbb{R}^k$. We define the *interior* of the simplex $\sigma$ to be the set of all points of $\sigma$ that are of the form $\sum_{j=0}^{q} t_j \mathbf{v}_j$, where $t_j > 0$ for $j = 0, 1, \ldots, q$ and $\sum_{j=0}^{q} t_j = 1$. One can readily verify that the interior of the simplex $\sigma$ consists of all points of $\sigma$ that do not belong to any proper face of $\sigma$. (Note that, if $\sigma \in \mathbb{R}^k$, then the interior of a simplex defined in this fashion will not coincide with the topological interior of $\sigma$ unless $\dim \sigma = k$.)

Note that any point of a simplex $\sigma$ belongs to the interior of a unique face of $\sigma$. Indeed let $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$ be the vertices of $\sigma$, and let $\mathbf{x} \in \sigma$. Then $\mathbf{x} = \sum_{j=0}^{q} t_j \mathbf{v}_j$, where $0 \leq t_j \leq 1$ for $j = 0, 1, \ldots, q$ and $\sum_{j=0}^{q} t_j = 1$. The unique face of $\sigma$ containing $\mathbf{x}$ in its interior is then the face spanned by those vertices $\mathbf{v}_j$ for which $t_j > 0$.

**Lemma 4.3** *Let $K$ be a finite collection of simplices in some Euclidean space $\mathbb{R}^k$, and let $|K|$ be the union of all the simplices in $K$. Then $K$ is a simplicial complex (with polyhedron $|K|$) if and only if the following two conditions are satisfied:—*

- *$K$ contains the faces of its simplices,*
- *every point of $|K|$ belongs to the interior of a unique simplex of $K$.*

**Proof** Suppose that $K$ is a simplicial complex. Then $K$ contains the faces of its simplices. We must show that every point of $|K|$ belongs to the interior of a unique simplex of $K$. Let $\mathbf{x} \in |K|$. Then $\mathbf{x}$ belongs to the interior of a face $\sigma$ of some simplex of $K$ (since every point of a simplex belongs to the interior of some face). But then $\sigma \in K$, since $K$ contains the faces of all its simplices. Thus $\mathbf{x}$ belongs to the interior of at least one simplex of $K$.

Suppose that $\mathbf{x}$ were to belong to the interior of two distinct simplices $\sigma$ and $\tau$ of $K$. Then $\mathbf{x}$ would belong to some common face $\sigma \cap \tau$ of $\sigma$ and $\tau$ (since $K$ is a simplicial complex). But this common face would be a proper face of one or other of the simplices $\sigma$ and $\tau$ (since $\sigma \neq \tau$), contradicting the fact that $\mathbf{x}$ belongs to the interior of both $\sigma$ and $\tau$. We conclude that the simplex $\sigma$ of $K$ containing $\mathbf{x}$ in its interior is uniquely determined, as required.

Conversely, we must show that any collection of simplices satisfying the given conditions is a simplicial complex. Since $K$ contains the faces of all its simplices, it only remains to verify that if $\sigma$ and $\tau$ are any two simplices of $K$ with non-empty intersection then $\sigma \cap \tau$ is a common face of $\sigma$ and $\tau$.

Let $\mathbf{x} \in \sigma \cap \tau$. Then $\mathbf{x}$ belongs to the interior of a unique simplex $\omega$ of $K$. However any point of $\sigma$ or $\tau$ belongs to the interior of a unique face of that simplex, and all faces of $\sigma$ and $\tau$ belong to $K$. It follows that $\omega$ is a common face of $\sigma$ and $\tau$, and thus the vertices of $\omega$ are vertices of both $\sigma$ and $\tau$. We deduce that the simplices $\sigma$ and $\tau$ have vertices in common, and that every point of $\sigma \cap \tau$ belongs to the common face $\rho$ of $\sigma$ and $\tau$ spanned by these common vertices. But this implies that $\sigma \cap \tau = \rho$, and thus $\sigma \cap \tau$ is a common face of both $\sigma$ and $\tau$, as required. ∎

**Definition** A *triangulation* $(K, h)$ of a topological space $X$ consists of a simplicial complex $K$ in some Euclidean space, together with a homeomorphism $h\colon |K| \to X$ mapping the polyhedron $|K|$ of $K$ onto $X$.

The polyhedron of a simplicial complex is a compact Hausdorff space. Thus if a topological space admits a triangulation then it must itself be a compact Hausdorff space.

**Lemma 4.4** *Let $X$ be a Hausdorff topological space, let $K$ be a simplicial complex, and let $h\colon |K| \to X$ be a bijection mapping $|K|$ onto $X$. Suppose that the restriction of $h$ to each simplex of $K$ is continuous on that simplex. Then the map $h\colon |K| \to X$ is a homeomorphism, and thus $(K, h)$ is a triangulation of $X$.*

**Proof** Each simplex of $K$ is a closed subset of $|K|$, and the number of simplices of $K$ is finite. It follows from Lemma 4.2 that $h\colon |K| \to X$ is continuous. Also the polyhedron $|K|$ of $K$ is a compact topological space. But every continuous bijection from a compact topological space to a Hausdorff space is a homeomorphism. Thus $(K, h)$ is a triangulation of $X$. ∎

## 4.3 Simplicial Maps

**Definition** A *simplicial map* $\varphi\colon K \to L$ between simplicial complexes $K$ and $L$ is a function $\varphi\colon \operatorname{Vert} K \to \operatorname{Vert} L$ from the vertex set of $K$ to that of $L$ such that $\varphi(\mathbf{v}_0), \varphi(\mathbf{v}_1), \ldots, \varphi(\mathbf{v}_q)$ span a simplex belonging to $L$ whenever $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$ span a simplex of $K$.

Note that a simplicial map $\varphi: K \to L$ between simplicial complexes $K$ and $L$ can be regarded as a function from $K$ to $L$: this function sends a simplex $\sigma$ of $K$ with vertices $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$ to the simplex $\varphi(\sigma)$ of $L$ spanned by the vertices $\varphi(\mathbf{v}_0), \varphi(\mathbf{v}_1), \ldots, \varphi(\mathbf{v}_q)$.

A simplicial map $\varphi: K \to L$ also induces in a natural fashion a continuous map $\varphi: |K| \to |L|$ between the polyhedra of $K$ and $L$, where

$$\varphi\left(\sum_{j=0}^{q} t_j \mathbf{v}_j\right) = \sum_{j=0}^{q} t_j \varphi(\mathbf{v}_j)$$

whenever $0 \leq t_j \leq 1$ for $j = 0, 1, \ldots, q$, $\sum_{j=0}^{q} t_j = 1$, and $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$ span a simplex of $K$. The continuity of this map follows immediately from a straightforward application of Lemma 4.2. Note that the interior of a simplex $\sigma$ of $K$ is mapped into the interior of the simplex $\varphi(\sigma)$ of $L$.

There are thus three equivalent ways of describing a simplicial map: as a function between the vertex sets of two simplicial complexes, as a function from one simplicial complex to another, and as a continuous map between the polyhedra of two simplicial complexes. In what follows, we shall describe a simplicial map using the representation that is most appropriate in the given context.

## 4.4 Barycentric Subdivision of a Simplicial Complex

Let $\sigma$ be a $q$-simplex in $\mathbb{R}^k$ with vertices $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$. The *barycentre* of $\sigma$ is defined to be the point

$$\hat{\sigma} = \frac{1}{q+1}(\mathbf{v}_0 + \mathbf{v}_1 + \cdots + \mathbf{v}_q).$$

Let $\sigma$ and $\tau$ be simplices in some Euclidean space. If $\sigma$ is a proper face of $\tau$ then we denote this fact by writing $\sigma < \tau$.

A simplicial complex $K_1$ is said to be a *subdivision* of a simplicial complex $K$ if $|K_1| = |K|$ and each simplex of $K_1$ is contained in a simplex of $K$.

**Definition** Let $K$ be a simplicial complex in some Euclidean space $\mathbb{R}^k$. The *first barycentric subdivision* $K'$ of $K$ is defined to be the collection of simplices in $\mathbb{R}^k$ whose vertices are $\hat{\sigma}_0, \hat{\sigma}_1, \ldots, \hat{\sigma}_r$ for some sequence $\sigma_0, \sigma_1, \ldots, \sigma_r$ of simplices of $K$ with $\sigma_0 < \sigma_1 < \cdots < \sigma_r$. Thus the set of vertices of $K'$ is the set of all the barycentres of all the simplices of $K$.

Note that every simplex of $K'$ is contained in a simplex of $K$. Indeed if $\sigma_0, \sigma_1, \ldots, \sigma_r \in K$ satisfy $\sigma_0 < \sigma_1 < \cdots < \sigma_r$ then the simplex of $K'$ spanned by $\hat{\sigma}_0, \hat{\sigma}_1, \ldots, \hat{\sigma}_r$, is contained in the simplex $\sigma_r$ of $K$.

**Proposition 4.5** *Let $K$ be a simplicial complex in some Euclidean space, and let $K'$ be the first barycentric subdivision of $K$. Then $K'$ is itself a simplicial complex, and $|K'| = |K|$.*

**Proof** We prove the result by induction on the number of simplices in $K$. The result is clear when $K$ consists of a single simplex, since that simplex must then be a point and therefore $K' = K$. We prove the result for a simplicial complex $K$, assuming that it holds for all complexes with fewer simplices.

It is clear from the definition of the barycentric subdivision $K'$ that any face of a simplex of $K'$ must itself belong to $K'$. We must verify that any two simplices of $K'$ are disjoint or else intersect in a common face.

Choose a simplex $\sigma$ of $K$ for which $\dim \sigma = \dim K$, and let $L = K \setminus \{\sigma\}$. Then $L$ is a subcomplex of $K$, since $\sigma$ is not a proper face of any simplex of $K$. Now $L$ has fewer simplices than $K$. It follows from the induction hypothesis that $L'$ is a simplicial complex and $|L'| = |L|$. Also it follows from the definition of $K'$ that $K'$ consists of the following simplices:—

- the simplices of $L'$,
- the barycentre $\hat{\sigma}$ of $\sigma$,
- simplices $\hat{\sigma}\rho$ whose vertex set is obtained by adjoining $\hat{\sigma}$ to the vertex set of some simplex $\rho$ of $L'$, where the vertices of $\rho$ are barycentres of proper faces of $\sigma$.

By checking all possible intersections of simplices of the above types, it is easy to verify that any two simplices of $K'$ intersect in a common face. Indeed any two simplices of $L'$ intersect in a common face, since $L'$ is a simplicial complex. If $\rho_1$ and $\rho_2$ are simplices of $L'$ whose vertices are barycentres of proper faces of $\sigma$, then $\rho_1 \cap \rho_2$ is a common face of $\rho_1$ and $\rho_2$ which is of this type, and $\hat{\sigma}\rho_1 \cap \hat{\sigma}\rho_2 = \hat{\sigma}(\rho_1 \cap \rho_2)$. Thus $\hat{\sigma}\rho_1 \cap \hat{\sigma}\rho_2$ is a common face of $\hat{\sigma}\rho_1$ and $\hat{\sigma}\rho_2$. Also any simplex $\tau$ of $L'$ is disjoint from the barycentre $\hat{\sigma}$ of $\sigma$, and $\hat{\sigma}\rho \cap \tau = \rho \cap \tau$. We conclude that $K'$ is indeed a simplicial complex.

It remains to verify that $|K'| = |K|$. Now $|K'| \subset |K|$, since every simplex of $K'$ is contained in a simplex of $K$. Let $\mathbf{x}$ be a point of the chosen simplex $\sigma$. Then there exists a point $\mathbf{y}$ belonging to a proper face of $\sigma$ and some $t \in [0, 1]$ such that $\mathbf{x} = (1-t)\hat{\sigma} + t\mathbf{y}$. But then $\mathbf{y} \in |L|$, and $|L| = |L'|$ by the induction hypothesis. It follows that $\mathbf{y} \in \rho$ for some simplex $\rho$ of $L'$ whose vertices are barycentres of proper faces of $\sigma$. But then $\mathbf{x} \in \hat{\sigma}\rho$, and therefore $\mathbf{x} \in |K'|$. Thus $|K| \subset |K'|$, and hence $|K'| = |K|$, as required. ∎

We define (by induction on $j$) the $j$th barycentric subdivision $K^{(j)}$ of $K$ to be the first barycentric subdivision of $K^{(j-1)}$ for each $j > 1$.

**Lemma 4.6** *Let $\sigma$ be a $q$-simplex and let $\tau$ be a face of $\sigma$. Let $\hat{\sigma}$ and $\hat{\tau}$ be the barycentres of $\sigma$ and $\tau$ respectively. If all the 1-simplices (edges) of $\sigma$ have length not exceeding $d$ for some $d > 0$ then*

$$|\hat{\sigma} - \hat{\tau}| \leq \frac{qd}{q+1}.$$

**Proof** Let $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$ be the vertices of $\sigma$. Let $\mathbf{x}$ and $\mathbf{y}$ be points of $\sigma$. We can write $\mathbf{y} = \sum_{j=0}^{q} t_j \mathbf{v}_j$, where $0 \leq t_i \leq 1$ for $i = 0, 1, \ldots, q$ and $\sum_{j=0}^{q} t_j = 1$. Now

$$\begin{aligned} |\mathbf{x} - \mathbf{y}| &= \left| \sum_{i=0}^{q} t_i (\mathbf{x} - \mathbf{v}_i) \right| \leq \sum_{i=0}^{q} t_i |\mathbf{x} - \mathbf{v}_i| \\ &\leq \text{maximum} \left( |\mathbf{x} - \mathbf{v}_0|, |\mathbf{x} - \mathbf{v}_1|, \ldots, |\mathbf{x} - \mathbf{v}_q| \right). \end{aligned}$$

Applying this result with $\mathbf{x} = \hat{\sigma}$ and $\mathbf{y} = \hat{\tau}$, we find that

$$|\hat{\sigma} - \hat{\tau}| \leq \text{maximum} \left( |\hat{\sigma} - \mathbf{v}_0|, |\hat{\sigma} - \mathbf{v}_1|, \ldots, |\hat{\sigma} - \mathbf{v}_q| \right).$$

But

$$\hat{\sigma} = \frac{1}{q+1} \mathbf{v}_i + \frac{q}{q+1} \mathbf{z}_i$$

for $i = 0, 1, \ldots, q$, where $\mathbf{z}_i$ is the barycentre of the $(q-1)$-face of $\sigma$ opposite to $\mathbf{v}_i$, given by

$$\mathbf{z}_i = \frac{1}{q} \sum_{j \neq i} \mathbf{v}_j.$$

Moreover $\mathbf{z}_i \in \sigma$. It follows that

$$|\hat{\sigma} - \mathbf{v}_i| = \frac{q}{q+1} |\mathbf{z}_i - \mathbf{v}_i| \leq \frac{qd}{q+1}$$

for $i = 1, 2, \ldots, q$, and thus

$$|\hat{\sigma} - \hat{\tau}| \leq \text{maximum} \left( |\hat{\sigma} - \mathbf{v}_0|, |\hat{\sigma} - \mathbf{v}_1|, \ldots, |\hat{\sigma} - \mathbf{v}_q| \right) \leq \frac{qd}{q+1},$$

as required. ∎

The *mesh* $\mu(K)$ of a simplicial complex $K$ is the length of the longest edge of $K$.

**Lemma 4.7** *Let $K$ be a simplicial complex in $\mathbb{R}^k$ for some $k$, and let $n$ be the dimension of $K$. Let $K'$ be the first barycentric subdivision of $K$. Then*

$$\mu(K') \leq \frac{n}{n+1}\mu(K).$$

**Proof** A 1-simplex of $K'$ is of the form $(\hat{\tau}, \hat{\sigma})$, where $\sigma$ is a $q$-simplex of $K$ for some $q \leq n$ and $\tau$ is a proper face of $\sigma$. Then

$$|\hat{\tau} - \hat{\sigma}| \leq \frac{q}{q+1}\mu(K) \leq \frac{n}{n+1}\mu(K)$$

by Lemma 4.6, as required. ∎

It follows directly from the above lemma that $\lim_{j \to +\infty} \mu(K^{(j)}) = 0$, where $K^{(j)}$ is the $j$th barycentric subdivision of $K$.

## 4.5 The Simplicial Approximation Theorem

**Definition** Let $f\colon |K| \to |L|$ be a continuous map between the polyhedra of simplicial complexes $K$ and $L$. A simplicial map $s\colon K \to L$ is said to be a *simplicial approximation* to $f$ if, for each $\mathbf{x} \in |K|$, $s(\mathbf{x})$ is an element of the unique simplex of $L$ which contains $f(\mathbf{x})$ in its interior.

Note that if $s\colon K \to L$ is a simplicial approximation to $f\colon |K| \to |L|$ then $s$ and $f$ are homotopic. Indeed the map from $|K| \times [0,1]$ to $|L|$ sending $(\mathbf{x}, t)$ to $(1-t)f(\mathbf{x}) + ts(\mathbf{x})$ is a well-defined homotopy between $f$ and $s$.

**Definition** Let $K$ be a simplicial complex, and let $\mathbf{x} \in |K|$. The *star* $\mathrm{st}_K(\mathbf{x})$ of $\mathbf{x}$ in $K$ is the union of the interiors of all simplices of $K$ that contain the point $\mathbf{x}$.

**Lemma 4.8** *Let $K$ be a simplicial complex and let $\mathbf{x} \in |K|$. Then the star $\mathrm{st}_K(\mathbf{x})$ of $\mathbf{x}$ is open in $|K|$, and $\mathbf{x} \in \mathrm{st}_K(\mathbf{x})$.*

**Proof** Every point of $|K|$ belongs to the interior of a unique simplex of $K$ (Lemma 4.3). It follows that the complement $|K| \setminus \mathrm{st}_K(\mathbf{x})$ of $\mathrm{st}_K(\mathbf{x})$ in $|K|$ is the union of the interiors of those simplices of $K$ that do not contain the point $\mathbf{x}$. But if a simplex of $K$ does not contain the point $\mathbf{x}$, then the same is true of its faces. Moreover the union of the interiors of all the faces of some simplex is the simplex itself. It follows that $|K| \setminus \mathrm{st}_K(\mathbf{x})$ is the union of all simplices of $K$ that do not contain the point $\mathbf{x}$. But each simplex of $K$ is closed in $|K|$. It follows that $|K| \setminus \mathrm{st}_K(\mathbf{x})$ is a finite union of closed sets, and is thus itself closed in $|K|$. We deduce that $\mathrm{st}_K(\mathbf{x})$ is open in $|K|$. Also $\mathbf{x} \in \mathrm{st}_K(\mathbf{x})$, since $\mathbf{x}$ belongs to the interior of at least one simplex of $K$. ∎

**Proposition 4.9** *A function $s\colon \mathrm{Vert}\, K \to \mathrm{Vert}\, L$ between the vertex sets of simplicial complexes $K$ and $L$ is a simplicial map, and a simplicial approximation to some continuous map $f\colon |K| \to |L|$, if and only if $f\left(\mathrm{st}_K(\mathbf{v})\right) \subset \mathrm{st}_L\left(s(\mathbf{v})\right)$ for all vertices $\mathbf{v}$ of $K$.*

**Proof** Let $s\colon K \to L$ be a simplicial approximation to $f\colon |K| \to |L|$, let $\mathbf{v}$ be a vertex of $K$, and let $\mathbf{x} \in \mathrm{st}_K(\mathbf{v})$. Then $\mathbf{x}$ and $f(\mathbf{x})$ belong to the interiors of unique simplices $\sigma \in K$ and $\tau \in L$. Moreover $\mathbf{v}$ must be a vertex of $\sigma$, by definition of $\mathrm{st}_K(\mathbf{v})$. Now $s(\mathbf{x})$ must belong to $\tau$ (since $s$ is a simplicial approximation to the map $f$), and therefore $s(\mathbf{x})$ must belong to the interior of some face of $\tau$. But $s(\mathbf{x})$ must belong to the interior of $s(\sigma)$, since $\mathbf{x}$ is in the interior of $\sigma$. It follows that $s(\sigma)$ must be a face of $\tau$, and therefore $s(\mathbf{v})$ must be a vertex of $\tau$. Thus $f(\mathbf{x}) \in \mathrm{st}_L(s(\mathbf{v}))$. We conclude that if $s\colon K \to L$ is a simplicial approximation to $f\colon |K| \to |L|$, then $f\left(\mathrm{st}_K(\mathbf{v})\right) \subset \mathrm{st}_L\left(s(\mathbf{v})\right)$.

Conversely let $s\colon \mathrm{Vert}\, K \to \mathrm{Vert}\, L$ be a function with the property that $f\left(\mathrm{st}_K(\mathbf{v})\right) \subset \mathrm{st}_L\left(s(\mathbf{v})\right)$ for all vertices $\mathbf{v}$ of $K$. Let $\mathbf{x}$ be a point in the interior of some simplex of $K$ with vertices $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q$. Then $\mathbf{x} \in \mathrm{st}_K(\mathbf{v}_j)$ and hence $f(\mathbf{x}) \in \mathrm{st}_L\left(s(\mathbf{v}_j)\right)$ for $j = 0, 1, \ldots, q$. It follows that each vertex $s(\mathbf{v}_j)$ must be a vertex of the unique simplex $\tau \in L$ that contains $f(\mathbf{x})$ in its interior. In particular, $s(\mathbf{v}_0), s(\mathbf{v}_1), \ldots, s(\mathbf{v}_q)$ span a face of $\tau$, and $s(\mathbf{x}) \in \tau$. We conclude that the function $s\colon \mathrm{Vert}\, K \to \mathrm{Vert}\, L$ represents a simplicial map which is a simplicial approximation to $f\colon |K| \to |L|$, as required. ∎

**Corollary 4.10** *If $s\colon K \to L$ and $t\colon L \to M$ are simplicial approximations to continuous maps $f\colon |K| \to |L|$ and $g\colon |L| \to |M|$, where $K$, $L$ and $M$ are simplicial complexes, then $t \circ s\colon K \to M$ is a simplicial approximation to $g \circ f\colon |K| \to |M|$.*

**Theorem 4.11** (Simplicial Approximation Theorem) *Let $K$ and $L$ be simplicial complexes, and let $f\colon |K| \to |L|$ be a continuous map. Then, for some sufficiently large integer $j$, there exists a simplicial approximation $s\colon K^{(j)} \to L$ to $f$ defined on the $j$th barycentric subdivision $K^{(j)}$ of $K$.*

**Proof** The collection consisting of the stars $\mathrm{st}_L(\mathbf{w})$ of all vertices $\mathbf{w}$ of $L$ is an open cover of $|L|$, since each star $\mathrm{st}_L(\mathbf{w})$ is open in $|L|$ (Lemma 4.8) and the interior of any simplex of $L$ is contained in $\mathrm{st}_L(\mathbf{w})$ whenever $\mathbf{w}$ is a vertex of that simplex. It follows from the continuity of the map $f\colon |K| \to |L|$ that the collection consisting of the preimages $f^{-1}(\mathrm{st}_L(\mathbf{w}))$ of the stars of all vertices $\mathbf{w}$ of $L$ is an open cover of $|K|$. It then follows from the Lebesgue Lemma that there exists some $\delta > 0$ with the property that every subset of $|K|$ whose diameter is less than $\delta$ is mapped by $f$ into $\mathrm{st}_L(\mathbf{w})$ for some vertex $\mathbf{w}$ of $L$.

Now the mesh $\mu(K^{(j)})$ of the $j$th barycentric subdivision of $K$ tends to zero as $j \to +\infty$, since

$$\mu(K^{(j)}) \leq \left(\frac{\dim K}{\dim K + 1}\right)^j \mu(K)$$

for all $j$ (Lemma 4.7). Thus we can choose $j$ such that $\mu(K^{(j)}) < \frac{1}{2}\delta$. If $\mathbf{v}$ is a vertex of $K^{(j)}$ then each point of $\mathrm{st}_{K^{(j)}}(\mathbf{v})$ is within a distance $\frac{1}{2}\delta$ of $\mathbf{v}$, and hence the diameter of $\mathrm{st}_{K^{(j)}}(\mathbf{v})$ is at most $\delta$. We can therefore choose, for each vertex $\mathbf{v}$ of $K^{(j)}$ a vertex $s(\mathbf{v})$ of $L$ such that $f(\mathrm{st}_{K^{(j)}}(\mathbf{v})) \subset \mathrm{st}_L(s(\mathbf{v}))$. In this way we obtain a function $s\colon \mathrm{Vert}\, K^{(j)} \to \mathrm{Vert}\, L$ from the vertices of $K^{(j)}$ to the vertices of $L$. It follows directly from Proposition 4.9 that this is the desired simplicial approximation to $f$. ∎

## 4.6 The Brouwer Fixed Point Theorem

**Definition** Let $K$ be a simplicial complex which is a subdivision of some $n$-dimensional simplex $\Delta$. We define a *Sperner labelling* of the vertices of $K$ to be a function, labelling each vertex of $K$ with an integer between 0 and $n$, with the following properties:—

- for each $j \in \{0, 1, \ldots, n\}$, there is exactly one vertex of $\Delta$ labelled by $j$,
- if a vertex $\mathbf{v}$ of $K$ belongs to some face of $\Delta$, then some vertex of that face has the same label as $\mathbf{v}$.

**Lemma 4.12** (Sperner's Lemma) *Let $K$ be a simplicial complex which is a subdivision of an $n$-simplex $\Delta$. Then, for any Sperner labelling of the vertices of $K$, the number of $n$-simplices of $K$ whose vertices are labelled by $0, 1, \ldots, n$ is odd.*

**Proof** Given integers $i_0, i_1, \ldots, i_q$ between 0 and $n$, let $N(i_0, i_1, \ldots, i_q)$ denote the number of $q$-simplices of $K$ whose vertices are labelled by $i_0, i_1, \ldots, i_q$ (where an integer occurring $k$ times in the list labels exactly $k$ vertices of the simplex). We must show that $N(0, 1, \ldots, n)$ is odd.

We prove the result by induction on the dimension $n$ of the simplex $\Delta$; it is clearly true when $n = 0$. Suppose that the result holds in dimensions less than $n$. For each simplex $\sigma$ of $K$ of dimension $n$, let $p(\sigma)$ denote the number of $(n-1)$-faces of $\sigma$ labelled by $0, 1, \ldots, n-1$. If $\sigma$ is labelled by $0, 1, \ldots, n$ then $p(\sigma) = 1$; if $\sigma$ is labelled by $0, 1, \ldots, n-1, j$, where $j < n$, then $p(\sigma) = 2$; in all other cases $p(\sigma) = 0$. Therefore

$$\sum_{\substack{\sigma \in K \\ \dim \sigma = n}} p(\sigma) = N(0, 1, \ldots, n) + 2 \sum_{j=0}^{n-1} N(0, 1, \ldots, n-1, j).$$

Now the definition of Sperner labellings ensures that the only $(n-1)$-face of $\Delta$ containing simplices of $K$ labelled by $0, 1, \ldots, n-1$ is that with vertices labelled by $0, 1, \ldots, n-1$. Thus if $M$ is the number of $(n-1)$-simplices of $K$ labelled by $0, 1, \ldots, n-1$ that are contained in this face, then $N(0, 1, \ldots, n-1) - M$ is the number of $(n-1)$-simplices labelled by $0, 1, \ldots, n-1$ that intersect the interior of $\Delta$. It follows that

$$\sum_{\substack{\sigma \in K \\ \dim \sigma = n}} p(\sigma) = M + 2(N(0, 1, \ldots, n-1) - M),$$

since any $(n-1)$-simplex of $K$ that is contained in a proper face of $\Delta$ must be a face of exactly one $n$-simplex of $K$, and any $(n-1)$-simplex that intersects the interior of $\Delta$ must be a face of exactly two $n$-simplices of $K$. On combining these equalities, we see that $N(0, 1, \ldots, n) - M$ is an even integer. But the induction hypothesis ensures that Sperner's Lemma holds in dimension $n-1$, and thus $M$ is odd. It follows that $N(0, 1, \ldots, n)$ is odd, as required. ∎

**Proposition 4.13** *Let $\Delta$ be an $n$-simplex with boundary $\partial\Delta$. Then there does not exist any continuous map $r\colon \Delta \to \partial\Delta$ with the property that $r(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \partial\Delta$.*

**Proof** Suppose that such a map $r\colon \Delta \to \partial\Delta$ were to exist. It would then follow from the Simplicial Approximation Theorem (Theorem 4.11) that there would exist a simplicial approximation $s\colon K \to L$ to the map $r$, where $L$ is the simplicial complex consisting of all of the proper faces of $\Delta$, and $K$ is the $j$th barycentric subdivision, for some sufficiently large $j$, of the simplicial complex consisting of the simplex $\Delta$ together with all of its faces.

If $\mathbf{v}$ is a vertex of $K$ belonging to some proper face $\Sigma$ of $\Delta$ then $r(\mathbf{v}) = \mathbf{v}$, and hence $s(\mathbf{v})$ must be a vertex of $\Sigma$, since $s\colon K \to L$ is a simplicial approximation to $r\colon \Delta \to \partial\Delta$. In particular $s(\mathbf{v}) = \mathbf{v}$ for all vertices $\mathbf{v}$ of $\Delta$. Thus if $\mathbf{v} \mapsto m(\mathbf{v})$ is a labelling of the vertices of $\Delta$ by the integers $0, 1, \ldots, n$, then $\mathbf{v} \mapsto m(s(\mathbf{v}))$ is a Sperner labelling of the vertices of $K$. Thus Sperner's Lemma (Lemma 4.12) guarantees the existence of at least one $n$-simplex $\sigma$ of $K$ labelled by $0, 1, \ldots, n$. But then $s(\sigma) = \Delta$, which is impossible, since $\Delta$ is not a simplex of $L$. We conclude therefore that there cannot exist any continuous map $r\colon \Delta \to \partial\Delta$ satisfying $r(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \partial\Delta$. ∎

**Theorem 4.14** (Brouwer Fixed Point Theorem) *A continuous map $f\colon E^n \to E^n$ sending a closed $n$-dimensional ball $E^n$ into itself has at least one fixed point (i.e., there exists $\mathbf{x} \in E^n$ for which $f(\mathbf{x}) = \mathbf{x}$).*

**Proof** Suppose that the map $f\colon E^n \to E^n$ had no fixed point. For each $\mathbf{x} \in E$, let $q(\mathbf{x})$ be the point at which the half line starting at $f(\mathbf{x})$ and passing through $\mathbf{x}$ intersects the boundary sphere $S^{n-1}$ of $E^n$. Then $q\colon E^n \to S^{n-1}$ would be continuous, and $q(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in S^{n-1}$. But the closed $n$-dimensional ball $E^n$ is homeomorphic to an $n$-simplex $\Delta$. Therefore the map $q\colon E^n \to S^{n-1}$ would correspond under some homeomorphism $h\colon \Delta \to E^n$ to a continuous map $r\colon \Delta \to \partial\Delta$ mapping $\Delta$ onto its boundary $\partial\Delta$, where $h(r(\mathbf{y})) = q(h(\mathbf{y}))$ for all $\mathbf{y} \in \Delta$. Moreover $r(\mathbf{y}) = \mathbf{y}$ for all $\mathbf{y} \in \partial\Delta$. However Proposition 4.13 shows that there does not exist any continuous map $r\colon \Delta \to \partial\Delta$ with this property. Therefore the map $f$ must have at least one fixed point, as required. ∎

# 5 An Application of Topology to Mathematical Economics

## 5.1 The Existence of Equilibria in an Exchange Economy

We consider an exchange economy consisting of a finite number of commodities and a finite number of households, each provided with an initial endowment of each of the commodities. The commodities are required to be *infinitely divisible*: this means that a household can hold an amount $x$ of that commodity for any non-negative real number $x$. (Thus salt, for example, could be regarded as an 'infinitely divisible' quantity whereas cars cannot: it makes little sense to talk about a particular household owning 2.637 of a car, for example, though such a household may well own 2.637 kilograms of salt.) Now the households may well wish to exchange commodities with one another so as improve on their initial endowment. They might for example seek to barter commodities with one another: however this method of redistribution would not work very efficiently in a large economy. Alternatively they might attempt to set up a price mechanism to simplify the task of redistributing the commodities. Thus suppose that each commodity is assigned a given price. Then each household could sell its initial endowment to the market, receiving in return the value of its initial endowment at the given prices. The household could then purchase from the market a quantity of each commodity so as to maximize its own preference, subject to the constraint that the total value of the commodities purchased by any household cannot exceed the value of its initial endowment at the given prices. The problem of redistribution then becomes one of fixing prices so that there is exactly enough of each commodity to go around: if the price of any commodity is too low then the demand for that commodity is likely to outstrip supply, whereas if the price is too high then supply will exceed demand. A *Walras equilibrium* is achieved if prices can be found so that the supply of each commodity matches its demand. We shall use the *Brouwer fixed point theorem* to prove the existence of a Walras equilibrium in this idealized economy.

Let our exchange economy consist of $n$ commodities and $m$ households. We suppose that household $h$ is provided with an initial endowment $\overline{x}_{hi}$ of commodity $i$, where $\overline{x}_{hi} \geq 0$. Thus the initial endowment of household $h$ can be represented by a vector $\overline{\mathbf{x}}_h$ in $\mathbb{R}^n$ whose $i$th component is $\overline{x}_{hi}$. The prices of the commodities are given by a price vector $\mathbf{p}$ whose $i$th component $p_i$ specifies the price of a unit of the $i$th commodity: a price vector $\mathbf{p}$ is required to satisfy $p_i \geq 0$ for all $i$. Then the value of the initial endowment of household $h$ at the given prices is $\mathbf{p}.\overline{\mathbf{x}}_h$. Let $x_{hi}(\mathbf{p})$ be the quantity of commodity $i$ that household $h$ seeks to purchase at prices $\mathbf{p}$, and let $\mathbf{x}_h(\mathbf{p}) \in \mathbb{R}^n$ be the vector whose $i$th component is $x_{hi}(\mathbf{p})$. The budget constraint certainly ensures that $\mathbf{p}.(\mathbf{x}_h(\mathbf{p}) - \overline{\mathbf{x}}_h) \leq 0$ (i.e., the value of the goods purchased cannot exceed the value of the initial endowment at the given prices). We assume that the value of the commodities that each household seeks to purchase is equal to the value of its initial endowment, and thus $\mathbf{p}.\mathbf{x}_h(\mathbf{p}) = \mathbf{p}.\overline{\mathbf{x}}_h$. Also the preferences of the household will only depend on the relative prices of the commodities, and therefore $\mathbf{x}_h(\lambda\mathbf{p}) = \mathbf{x}_h(\mathbf{p})$ for all $\lambda > 0$.

Now the total supply of each commodity in the economy is represented by the vector $\sum_h \overline{\mathbf{x}}_h$, and the total demand at prices $\mathbf{p}$ is represented by $\sum_h \mathbf{x}_h(\mathbf{p})$. The *excess demand* in the economy at prices $\mathbf{p}$ is therefore represented by the vector $\mathbf{z}(\mathbf{p})$, where $\mathbf{z}(\mathbf{p}) = \sum_h (\mathbf{x}_h(\mathbf{p}) - \overline{\mathbf{x}}_h)$. Let $z_i(\mathbf{p})$ be the $i$th component of $\mathbf{z}(\mathbf{p})$. Then $z_i(\mathbf{p}) > 0$ when the demand for the $i$th commodity exceeds supply, whereas $z_i(\mathbf{p}) < 0$ when the supply exceeds demand. Note that $\mathbf{p}.\mathbf{z}(\mathbf{p}) = 0$ for any price vector $\mathbf{p}$. This identity, known as *Walras' Law*, follows immediately on summing the budget constraint $\mathbf{p}.\mathbf{x}_h(\mathbf{p}) = \mathbf{p}.\overline{\mathbf{x}}_h$

over all households.

**Theorem 5.1** *Consider an exchange economy consisting of a finite number of infinitely divisible commodities and a finite number of households. Let the excess demand in the economy at prices* $\mathbf{p}$ *be given by* $\mathbf{z}(\mathbf{p})$*, where*

(i) *the excess demand vector* $\mathbf{z}(\mathbf{p})$ *is well-defined for any price vector* $\mathbf{p}$*, and depends continuously on* $\mathbf{p}$*,*

(ii) $\mathbf{p}.\mathbf{z}(\mathbf{p}) = 0$ *for any price vector* $\mathbf{p}$ *(Walras' Law).*

(iii) *Then there exist equilibrium prices* $\mathbf{p}^*$ *at which* $z_i(\mathbf{p}^*) \leq 0$ *for all* $i$*.*

**Proof** Let $\Delta$ be the $(n-1)$-dimensional simplex in $\mathbb{R}^n$ consisting of all points $(p_1, p_2, \ldots, p_n)$ in $\mathbb{R}^n$ satisfying $0 \leq p_i \leq 1$ for $i = 1, 2, \ldots, n$ and $\sum_{i=1}^n p_i = 1$, and let $\mathbf{v}\colon \Delta \to \mathbb{R}^n$ be the function with $i$th component $v_i$ given by

$$v_i(\mathbf{p}) = \begin{cases} p_i + z_i(\mathbf{p}) & \text{if } z_i(\mathbf{p}) > 0; \\ p_i & \text{if } z_i(\mathbf{p}) \leq 0. \end{cases}$$

Note that $\mathbf{v}(\mathbf{p}) \neq \mathbf{0}$ and the components of $\mathbf{v}(\mathbf{p})$ are non-negative for all $\mathbf{p} \in \Delta$. It follows that there is a well-defined map $\varphi\colon \Delta \to \Delta$ given by

$$\varphi(\mathbf{p}) = \frac{1}{\sum\limits_{i=1}^{n} v_i(\mathbf{p})} \mathbf{v}(\mathbf{p}),$$

The Brouwer Fixed Point Theorem (Theorem 4.14) ensures that there exists $\mathbf{p}^* \in \Delta$ satisfying $\varphi(\mathbf{p}^*) = \mathbf{p}^*$. Then $\mathbf{v}(\mathbf{p}^*) = \lambda \mathbf{p}^*$ for some $\lambda \geq 1$. We claim that $\lambda = 1$.

Suppose that it were the case that $\lambda > 1$. Then $v_i(\mathbf{p}^*) > p_i^*$, and thus $z_i(\mathbf{p}^*) > 0$ whenever $p_i^* > 0$. But $p_i^* \geq 0$ for all $i$, and $p_i^* > 0$ for at least one value of $i$, since $\mathbf{p}^* \in \Delta$. It would follow that $\mathbf{p}^*.\mathbf{z}(\mathbf{p}^*) > 0$, contradicting Walras' Law. We conclude that $\lambda = 1$, and thus $v_i = p_i^*$ and $z_i(\mathbf{p}^*) \leq 0$ for all $i$, as required. ∎

Note that if $z_i(\mathbf{p}^*) \leq 0$ for all $i$ and $\mathbf{p}^*.\mathbf{z}(\mathbf{p}^*) = 0$, then $p_i^* z_i(\mathbf{p}^*) \leq 0$ for all $i$, and $z_i(\mathbf{p}^*) = 0$ whenever $p_i > 0$. Thus, at equilibrium prices, supply always equals or exceeds demand, and supply equals demand for those commodities with positive prices.

The proof of the existence of Walras equilibria can readily be generalized to *Arrow-Debreu* models where economic activity is carried out by both households and firms. The problem of existence of equilibria was studied by

L. Walras in the 1870s, though a rigorous proof of the existence of equilibria was not found till the 1930s, when A. Wald proved existence for a limited range of economic models. Proofs of existence using the Brouwer Fixed Point Theorem, or a more general fixed point theorem due to Katukani, were first published in 1954 by K. J. Arrow and G. Debreu and by L. McKenzie. Subsequent research has centred on problems of uniqueness and stability, and the existence theorems have been generalized to economies with an infinite number of commodities and economic agents (households and firms). An alternative approach to the existence theorems using techniques of differential topology was pioneered by G. Debreu and by S. Smale.

More detailed accounts of the theory of 'general equilibrium' can be found in, for example, *The theory of value*, by G. Debreu, *General competitive analysis*, by K. J. Arrow and F. H. Hahn, or *Economics for mathematicians* by J. W. S. Cassels.