Course 221: Analysis
Academic year 2007-08, First Semester

David R. Wilkins

Contents

1 Basic Theorems of Real Analysis

1.1 The Least Upper Bound Principle

A widely-used basic principle of analysis, from which many important theorems ultimately derive, is the Least Upper Bound Principle.

Let D be a subset of the set $\mathbb{R}$ of real numbers. A real number u is said to be an *upper bound* of the set D of $x \leq u$ for all $x \in D$. The set D is said to be *bounded above* if such an upper bound exists.

Definition Let D be some set of real numbers which is bounded above. A real number s is said to be the *least upper bound* (or *supremum*) of D (denoted by $\sup D$) if s is an upper bound of D and $s \leq u$ for all upper bounds u of D.

Example The real number 2 is the least upper bound of the sets $\{x \in \mathbb{R} : x \leq 2\}$ and $\{x \in \mathbb{R} : x < 2\}$. Note that the first of these sets contains its least upper bound, whereas the second set does not.

The *Least Upper Bound Principle* may be stated as follows:

> if D is any non-empty subset of $\mathbb{R}$ which is bounded above then there exists a *least upper bound* $\sup D$ for the set D.

A *lower bound* of a set D of real numbers is a real number l with the property that $l \leq x$ for all $x \in D$. A set D of real numbers is said to be *bounded below* if such a lower bound exists. If D is bounded below, then there exists a greatest lower bound (or *infimum*) $\inf D$ of the set D. Indeed $\inf D = -\sup\{x \in \mathbb{R} : -x \in D\}$.

1.2 Monotonic Sequences

An infinite sequence $a_1, a_2, a_3, \ldots$ of real numbers is said to be *strictly increasing* if $a_{n+1} > a_n$ for all n, *strictly decreasing* if $a_{n+1} < a_n$ for all n, *non-decreasing* if $a_{n+1} \geq a_n$ for all n, or *non-increasing* if $a_{n+1} \leq a_n$ for all n. A sequence satisfying any one of these conditions is said to be *monotonic*; thus a monotonic sequence is either non-decreasing or non-increasing.

Theorem 1.1 *Any non-decreasing sequence of real numbers that is bounded above is convergent. Similarly any non-increasing sequence of real numbers that is bounded below is convergent.*

Proof Let $a_1, a_2, a_3, \ldots$ be a non-decreasing sequence of real numbers that is bounded above. It follows from the Least Upper Bound Principle that there exists a least upper bound l for the set $\{a_n : n \in \mathbb{N}\}$. We claim that the sequence converges to l.

Let $\varepsilon > 0$ be given. We must show that there exists some natural number N such that $|a_n - l| < \varepsilon$ whenever $n \geq N$. Now $l - \varepsilon$ is not an upper bound for the set $\{a_n : n \in \mathbb{N}\}$ (since l is the least upper bound), and therefore there must exist some natural number N such that $a_N > l - \varepsilon$. But then $l - \varepsilon < a_n \leq l$ whenever $n \geq N$, since the sequence is non-decreasing and bounded above by l. Thus $|a_n - l| < \varepsilon$ whenever $n \geq N$. Therefore $a_n \to l$ as $n \to +\infty$, as required.

If the sequence $a_1, a_2, a_3, \ldots$ is non-increasing and bounded below then the sequence $-a_1, -a_2, -a_3, \ldots$ is non-decreasing and bounded above, and is therefore convergent. It follows that the sequence $a_1, a_2, a_3, \ldots$ is also convergent. ∎

1.3 Upper and Lower Limits of Bounded Sequences of Real Numbers

Let $a_1, a_2, a_3, \ldots$ be a bounded infinite sequence of real numbers, and, for each positive integer j, let

$$S_j = \{a_j, a_{j+1}, a_{j+2}, \ldots\} = \{a_k : k \geq j\}.$$

The sets $S_1, S_2, S_3, \ldots$ are all bounded. It follows that there exist well-defined infinite sequences $u_1, u_2, u_3, \ldots$ and $l_1, l_2, l_3, \ldots$ of real numbers, where $u_j = \sup S_j$ and $l_j = \inf S_j$ for all positive integers j. Now S_{j+1} is a subset of S_j for each positive integer j, and therefore $u_{j+1} \leq u_j$ and $l_{j+1} \geq l_j$ for each positive integer j. It follows that the bounded infinite sequence $(u_j : j \in \mathbb{N})$ is a non-increasing sequence, and is therefore convergent (Theorem 1.1). Similarly the bounded infinite sequence $(l_j : j \in \mathbb{N})$ is a non-decreasing sequence, and is therefore convergent. We define

$$\begin{aligned} \limsup_{j \to +\infty} a_j &= \lim_{j \to +\infty} u_j = \lim_{j \to +\infty} \sup\{a_j, a_{j+1}, a_{j+2}, \ldots\}, \\ \liminf_{j \to +\infty} a_j &= \lim_{j \to +\infty} l_j = \lim_{j \to +\infty} \inf\{a_j, a_{j+1}, a_{j+2}, \ldots\}. \end{aligned}$$

The quantity $\limsup_{j \to +\infty} a_j$ is referred to as the *upper limit* of the sequence $a_1, a_2, a_3, \ldots$. The quantity $\liminf_{j \to +\infty} a_j$ is referred to as the *lower limit* of the sequence $a_1, a_2, a_3, \ldots$.

Note that every bounded infinite sequence $a_1, a_2, a_3, \ldots.$ of real numbers has a well-defined upper limit $\limsup_{j \to +\infty} a_j$ and a well-defined lower limit $\liminf_{j \to +\infty} a_j$.

Proposition 1.2 *A bounded infinite sequence $a_1, a_2, a_3, \ldots.$ of real numbers is convergent if and only if $\liminf_{j \to +\infty} a_j = \limsup_{j \to +\infty} a_j$, in which case the limit of the sequence is equal to the common value of its upper and lower limits.*

Proof For each positive integer j, let $u_j = \sup S_j$ and $l_j = \inf S_j$, where

$$S_j = \{a_j, a_{j+1}, a_{j+2}, \ldots\} = \{a_k : k \geq j\}.$$

Then $\liminf_{j \to +\infty} a_j = \lim_{j \to +\infty} l_j$ and $\limsup_{j \to +\infty} a_j = \lim_{j \to +\infty} u_j$.

Suppose that $\liminf_{j \to +\infty} a_j = \limsup_{j \to +\infty} a_j = c$ for some real number c. Then, given any positive real number ε, there exist natural numbers N_1 and N_2 such that $c - \varepsilon < l_j \leq c$ whenever $j \geq N_1$, and $c \leq u_j \leq c + \varepsilon$ whenever $j \geq N_2$. Let N be the maximum of N_1 and N_2. If $j \geq N$ then $a_j \in S_N$, and therefore

$$c - \varepsilon < l_N \leq a_j \leq u_N < c + \varepsilon.$$

Thus $|a_j - c| < \varepsilon$ whenever $j \geq N$. This proves that the infinite sequence $a_1, a_2, a_3, \ldots$ converges to the limit c.

Conversely let $a_1, a_2, a_3, \ldots$ be a bounded sequence of real numbers that converges to some value c. Let $\varepsilon > 0$ be given. Then there exists some natural number N such that $c - \frac{1}{2}\varepsilon < a_j < c + \frac{1}{2}\varepsilon$ whenever $j \geq N$. It follows that $S_j \subset (c - \frac{1}{2}\varepsilon, c + \frac{1}{2}\varepsilon)$ whenever $j \geq N$. But then

$$c - \tfrac{1}{2}\varepsilon \leq l_j \leq u_j \leq c + \tfrac{1}{2}\varepsilon$$

whenever $j \geq N$, where $u_j = \sup S_j$ and $l_j = \inf S_j$. We see from this that, given any positive real number ε, there exists some natural number N such that $|l_j - c| < \varepsilon$ and $|u_j - c| < \varepsilon$ whenever $j \geq N$. It follows from this that

$$\limsup_{j \to +\infty} a_j = \lim_{j \to +\infty} u_j = c \text{ and } \liminf_{j \to +\infty} a_j = \lim_{j \to +\infty} l_j = c,$$

as required. ∎

1.4 Cauchy's Criterion for Convergence

Definition An infinite sequence $a_1, a_2, a_3, \ldots$ of real numbers said to be a *Cauchy sequence* if, given any positive real number ε, there exists some positive integer N such that $|a_j - a_k| < \varepsilon$ for all j and k satisfying $j \geq N$ and $k \geq N$.

Theorem 1.3 (Cauchy's Criterion for Convergence) *A sequence of real numbers is convergent if and only if it is a Cauchy sequence.*

Proof Let $a_1, a_2, a_3, \ldots$ be a sequence of real numbers. Suppose that this sequence converges to some limit c. Let some positive real number ε be given. Then there exists some natural number N such that $|a_j - c| < \frac{1}{2}\varepsilon$ whenever $j \geq N$. If j and k are positive integers satisfying $j \geq N$ and $k \geq N$ then

$$|a_j - a_k| \leq |a_j - c| + |c - a_k| < \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon.$$

This shows that any convergent sequence of real numbers is a Cauchy sequence.

Next let $a_1, a_2, a_3, \ldots$ be a Cauchy sequence of real numbers. We must prove that this sequence is convergent. First we show that it is bounded. Now there exists some natural number M such that $|a_j - a_k| < 1$ for all positive integers j and k satisfying $j > M$ and $k > M$. Let R be the maximum of the real numbers

$$|a_1|, |a_2|, \ldots, |a_{M-1}|, |a_M| + 1.$$

It is clear that $|a_j| \leq R$ when $j < M$. If $j \geq M$ then $|a_j - a_M| < 1$, and therefore $|a_j| < |a_M| + 1 \leq R$. Thus $|a_j| \leq R$ for all positive integers j. This proves that the Cauchy sequence is bounded.

For each positive integer j, let

$$u_j = \sup\{a_k : k \geq j\} \text{ and } l_j = \inf\{a_k : k \geq j\}.$$

Then $u_1, u_2, u_3, \ldots$ is a non-increasing sequence which converges to $\limsup_{j\to+\infty} a_j$, and $l_1, l_2, l_3, \ldots$ is a non-decreasing sequence which converges to $\liminf_{j\to+\infty} a_j$.

Let ε be some given positive real number. Then there exists some natural number N such that $|a_j - a_k| < \varepsilon$ for all positive integers j and k satisfying $j \geq N$ and $k \geq N$. It follows from this that $a_N - \varepsilon < a_j < a_N + \varepsilon$ for all positive integers j satisfying $j \geq N$. It then follow from the definitions of u_N and l_N that $a_N - \varepsilon \leq l_N \leq u_N \leq a_N + \varepsilon$. Now $0 \leq u_j - l_j \leq u_N - l_N$ whenever $j \geq N$. It follows that

$$\limsup_{j\to+\infty} a_j - \liminf_{j\to+\infty} a_j = \lim_{j\to+\infty} (u_j - l_j) \leq u_N - l_N \leq 2\varepsilon.$$

Thus if $d = \limsup_{j\to+\infty} a_j - \liminf_{j\to+\infty} a_j$ then $0 \leq d \leq 2\varepsilon$ for all positive real numbers ε. It must therefore be the case that $d = 0$. Thus $\limsup_{j\to+\infty} a_j = \liminf_{j\to+\infty} a_j$. It now follows from Proposition 1.2 that the Cauchy sequence $a_1, a_2, a_3, \ldots$ is convergent, as required. ∎

1.5 The Bolzano-Weierstrass Theorem

Let $a_1, a_2, a_3, \ldots$ be an infinite sequence of real numbers. A *subsequence* of this sequence is a sequence that is of the form $a_{m_1}, a_{m_2}, a_{m_3}, \ldots$, where $m_1, m_2, m_3, \ldots$ are positive integers satisfying $m_1 < m_2 < m_3 < \cdots$. Thus, for example, $a_2, a_4, a_6, \ldots$ and $a_1, a_4, a_9, \ldots$ are subsequences of the given sequence.

Lemma 1.4 *Let $a_1, a_2, a_3, \ldots$ be a bounded infinite sequence of real numbers, and let c be a real number satisfying $c < \limsup_{j\to+\infty} a_j$. Then there exist infinitely many positive integers j such that $a_j > c$.*

Proof Let N be a positive integer. Then

$$c < \limsup_{j\to+\infty} a_j \leq \sup\{a_j : j \geq N\},$$

It follows that c is not an upper bound for the set $\{a_j : j \geq N\}$, and therefore there exists some positive integer satisfying $j \geq N$ for which $a_j > c$. We conclude from this that there does not exist any positive integer N with the property that $a_j \leq c$ whenever $j \geq N$. Therefore $\{j \in \mathbb{N} : a_j > c\}$ is not a finite set. The result follows. ∎

Proposition 1.5 *Any bounded infinite sequence $a_1, a_2, a_3, \ldots$ of real numbers has a subsequence which converges to the upper limit $\limsup_{j\to+\infty} a_j$ of the given sequence.*

Proof Let $s = \limsup_{j\to+\infty} a_j$, and let

$$u_N = \sup\{a_N, a_{N+1}, a_{N+2}, \ldots\} = \sup\{a_j : j \geq N\}$$

for all positive integers N. The upper limit s of the sequence $a_1, a_2, a_3, \ldots$ is then the limit of the non-increasing sequence $u_1, u_2, u_3, \ldots$.

Let ε be positive real number. The convergence of the infinite sequence $u_1, u_2, u_3, \ldots$ to s ensures that there exists some positive integer N such that

$u_N < s + \varepsilon$. But then $a_j < s + \varepsilon$ whenever $j \geq N$. It follows that the number of positive integers j for which $a_j \geq s + \varepsilon$ is finite. Also it follows from Lemma 1.4 that the number of positive integers j for which $a_j > s - \varepsilon$ is infinite. Putting these two facts together, we see that the number of positive integers j for which $s - \varepsilon < a_j < s + \varepsilon$ is infinite. (Indeed let $S_1 = \{j \in \mathbb{N} : a_j > s - \varepsilon\}$ and $S_2 = \{j \in \mathbb{N} : a_j \geq s + \varepsilon\}$. Then S_1 is an infinite set, S_2 is a finite set, and therefore $S_1 \setminus S_2$ is an infinite set. Moreover $s - \varepsilon < a_j < s + \varepsilon$ for all $j \in S_1 \setminus S_2$.)

Now given any positive integer j, and given any positive number m_j such that $|a_{m_j} - s| < j^{-1}$, there exists some positive integer m_{j+1} such that $m_{j+1} > m_j$ and $|a_{m_{j+1}} - s| < (j+1)^{-1}$. It follow from this that there exists a subsequence $a_{m_1}, a_{m_2}, a_{m_3}, \ldots$ of the infinite sequence $a_1, a_2, a_3, \ldots$, where $m_1 < m_2 < m_3 < \cdots$, which has the property that $|a_{m_j} - s| < j^{-1}$ for all positive integers j. This subsequence converges to s as required. ∎

The following theorem, known as the *Bolzano-Weierstrass Theorem*, is an immediate consequence of Proposition 1.5.

Theorem 1.6 (Bolzano-Weierstrass) *Every bounded sequence of real numbers has a convergent subsequence.*

1.6 The Intermediate Value Theorem

Proposition 1.7 *Let $f\colon [a,b] \to \mathbb{Z}$ continuous integer-valued function defined on a closed interval $[a,b]$. Then the function f is constant.*

Proof Let

$$S = \{x \in [a,b] : f \text{ is constant on the interval } [a,x]\},$$

and let $s = \sup S$. Now $s \in [a,b]$, and therefore the function f is continuous at s. Therefore there exists some real number δ satisfying $\delta > 0$ such that $|f(x) - f(s)| < \frac{1}{2}$ for all $x \in [a,b]$ satisfying $|x - s| < \delta$. But the function f is integer-valued. It follows that $f(x) = f(s)$ for all $x \in [a,b]$ satisfying $|x - s| < \delta$. Now $s - \delta$ is not an upper bound for the set S. Therefore there exists some element x_0 of S satisfying $s - \delta < x_0 \leq s$. But then $f(s) = f(x_0) = f(a)$, and therefore the function f is constant on the interval $[a,x]$ for all $x \in [a,b]$ satisfying $s \leq x < s+\delta$. Thus $x \in [a,b] \cap [s, s+\delta) \subset S$. In particular $s \in S$. Now S cannot contain any elements x of $[a,b]$ satisfying $x > s$. Therefore $[a,b] \cap [s, s+\delta) = \{s\}$, and therefore $s = b$. This shows that $b \in S$, and thus the function f is constant on the interval $[a,b]$, as required. ∎

Theorem 1.8 (The Intermediate Value Theorem) *Let a and b be real numbers satisfying $a < b$, and let $f\colon [a,b] \to \mathbb{R}$ be a continuous function defined on the interval $[a,b]$. Let c be a real number which lies between $f(a)$ and $f(b)$ (so that either $f(a) \leq c \leq f(b)$ or else $f(a) \geq c \geq f(b)$.) Then there exists some $s \in [a,b]$ for which $f(s) = c$.*

Proof Let c be a real number which lies between $f(a)$ and $f(b)$, and let $g_c\colon \mathbb{R} \setminus \{c\} \to \mathbb{Z}$ be the continuous integer-valued function on $\mathbb{R} \setminus \{c\}$ defined such that $g_c(x) = 0$ whenever $x < c$ and $g_c(x) = 1$ if $x > c$. Suppose that c were not in the range of the function f. Then the composition function $g_c \circ f\colon [a,b] \to \mathbb{R}$ would be a continuous integer-valued function defined throughout the interval $[a,b]$. This function would not be constant, since $g_c(f(a)) \neq g_c(f(b))$. But every continuous integer-valued function on the interval $[a,b]$ is constant (Proposition 1.7). It follows that every real number c lying between $f(a)$ and $f(b)$ must belong to the range of the function f, as required. ∎

Corollary 1.9 *Let $f\colon [a,b] \to [c,d]$ be a strictly increasing continuous function mapping an interval $[a,b]$ into an interval $[c,d]$, where a, b, c and d are real numbers satisfying $a < b$ and $c < d$. Suppose that $f(a) = c$ and $f(b) = d$. Then the function f has a continuous inverse $f^{-1}\colon [c,d] \to [a,b]$.*

Proof Let x_1 and x_2 be distinct real numbers belonging to the interval $[a,b]$ then either $x_1 < x_2$, in which case $f(x_1) < f(x_2)$ or $x_1 > x_2$, in which case $f(x_1) > f(x_2)$. Thus $f(x_1) \neq f(x_2)$ whenever $x_1 \neq x_2$. It follows that the function f is injective. The Intermediate Value Theorem (Theorem 1.8) ensures that f is surjective. It follows that the function f has a well-defined inverse $f^{-1}\colon [c,d] \to [a,b]$. It only remains to show that this inverse function is continuous.

Let y be a real number satisfying $c < y < d$, and let x be the unique real number such that $a < x < b$ and $f(x) = y$. Let $\varepsilon > 0$ be given. We can then choose $x_1, x_2 \in [a,b]$ such that $x - \varepsilon < x_1 < x < x_2 < x + \varepsilon$. Let $y_1 = f(x_1)$ and $y_2 = f(x_2)$. Then $y_1 < y < y_2$. Choose $\delta > 0$ such that $\delta < y - y_1$ and $\delta < y_2 - y$. If $v \in [c,d]$ satisfies $|v - y| < \delta$ then $y_1 < v < y_2$ and therefore $x_1 < f^{-1}(v) < x_2$. But then $|f^{-1}(v) - f^{-1}(y)| < \varepsilon$. We conclude that the function $f^{-1}\colon [c,d] \to [a,b]$ is continuous at all points in the interior of the interval $[a,b]$. A similar argument shows that it is continuous at the endpoints of this interval. Thus the function f has a continuous inverse, as required. ∎

2 Analysis in Euclidean Spaces

2.1 Euclidean Spaces

We denote by $\mathbb{R}^n$ the set consisting of all n-tuples $(x_1, x_2, \ldots, x_n)$ of real numbers. The set $\mathbb{R}^n$ represents n-dimensional *Euclidean space* (with respect to the standard Cartesian coordinate system). Let $\mathbf{x}$ and $\mathbf{y}$ be elements of $\mathbb{R}^n$, where

$$\mathbf{x} = (x_1, x_2, \ldots, x_n), \quad \mathbf{y} = (y_1, y_2, \ldots, y_n),$$

and let λ be a real number. We define

$$\begin{aligned}
\mathbf{x} + \mathbf{y} &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n), \\
\mathbf{x} - \mathbf{y} &= (x_1 - y_1, x_2 - y_2, \ldots, x_n - y_n), \\
\lambda \mathbf{x} &= (\lambda x_1, \lambda x_2, \ldots, \lambda x_n), \\
\mathbf{x} \cdot \mathbf{y} &= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n, \\
|\mathbf{x}| &= \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.
\end{aligned}$$

The quantity $\mathbf{x} \cdot \mathbf{y}$ is the *scalar product* (or *inner product*) of $\mathbf{x}$ and $\mathbf{y}$, and the quantity $|\mathbf{x}|$ is the *Euclidean norm* of $\mathbf{x}$. Note that $|\mathbf{x}|^2 = \mathbf{x} \cdot \mathbf{x}$. The *Euclidean distance* between two points $\mathbf{x}$ and $\mathbf{y}$ of $\mathbb{R}^n$ is defined to be the Euclidean norm $|\mathbf{y} - \mathbf{x}|$ of the vector $\mathbf{y} - \mathbf{x}$.

Lemma 2.1 (Schwarz' Inequality) *Let $\mathbf{x}$ and $\mathbf{y}$ be elements of $\mathbb{R}^n$. Then $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|$.*

Proof We note that $|\lambda \mathbf{x} + \mu \mathbf{y}|^2 \geq 0$ for all real numbers λ and μ. But

$$|\lambda \mathbf{x} + \mu \mathbf{y}|^2 = (\lambda \mathbf{x} + \mu \mathbf{y}).(\lambda \mathbf{x} + \mu \mathbf{y}) = \lambda^2 |\mathbf{x}|^2 + 2\lambda\mu \mathbf{x} \cdot \mathbf{y} + \mu^2 |\mathbf{y}|^2.$$

Therefore $\lambda^2 |\mathbf{x}|^2 + 2\lambda\mu \mathbf{x} \cdot \mathbf{y} + \mu^2 |\mathbf{y}|^2 \geq 0$ for all real numbers λ and μ. In particular, suppose that $\lambda = |\mathbf{y}|^2$ and $\mu = -\mathbf{x} \cdot \mathbf{y}$. We conclude that

$$|\mathbf{y}|^4 |\mathbf{x}|^2 - 2|\mathbf{y}|^2 (\mathbf{x} \cdot \mathbf{y})^2 + (\mathbf{x} \cdot \mathbf{y})^2 |\mathbf{y}|^2 \geq 0,$$

so that $\left(|\mathbf{x}|^2 |\mathbf{y}|^2 - (\mathbf{x} \cdot \mathbf{y})^2\right) |\mathbf{y}|^2 \geq 0$. Thus if $\mathbf{y} \neq \mathbf{0}$ then $|\mathbf{y}| > 0$, and hence

$$|\mathbf{x}|^2 |\mathbf{y}|^2 - (\mathbf{x} \cdot \mathbf{y})^2 \geq 0.$$

But this inequality is trivially satisfied when $\mathbf{y} = \mathbf{0}$. Thus $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|$, as required. ∎

It follows easily from Schwarz' Inequality that $|\mathbf{x}+\mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. For

$$\begin{aligned} |\mathbf{x}+\mathbf{y}|^2 &= (\mathbf{x}+\mathbf{y}).(\mathbf{x}+\mathbf{y}) = |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2\mathbf{x}\cdot\mathbf{y} \\ &\leq |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2|\mathbf{x}||\mathbf{y}| = (|\mathbf{x}| + |\mathbf{y}|)^2. \end{aligned}$$

It follows that

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

for all points $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ of $\mathbb{R}^n$. This important inequality is known as the *Triangle Inequality.* It expresses the geometric fact the the length of any triangle in a Euclidean space is less than or equal to the sum of the lengths of the other two sides.

2.2 Convergence of Sequences in Euclidean Spaces

Definition Let n be a positive integer, and let $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \ldots$ be an infinite sequence of points in n-dimensional Euclidean space $\mathbb{R}^n$. This sequence of points is said to *converge* to some point $\mathbf{r}$ of $\mathbb{R}^n$ if, given any real number ε satisfying $\varepsilon > 0$, there exists some positive integer N such that $|\mathbf{p}_j - \mathbf{r}| < \varepsilon$ whenever $j \geq N$.

Lemma 2.2 *Let $\mathbf{p}$ be a point of $\mathbb{R}^n$, where $\mathbf{p} = (p_1, p_2, \ldots, p_n)$. Then a sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ of points in $\mathbb{R}^n$ converges to $\mathbf{p}$ if and only if the ith components of the elements of this sequence converge to p_i for $i = 1, 2, \ldots, n$.*

Proof Let x_{ji} and p_i denote the ith components of $\mathbf{x}_j$ and $\mathbf{p}$. Then $|x_{ji} - p_i| \leq |\mathbf{x}_j - \mathbf{p}|$ for all j. It follows directly from the definition of convergence that if $\mathbf{x}_j \to \mathbf{p}$ as $j \to +\infty$ then $x_{ji} \to p_i$ as $j \to +\infty$.

Conversely suppose that, for each i, $x_{ji} \to p_i$ as $j \to +\infty$. Let $\varepsilon > 0$ be given. Then there exist natural numbers $N_1, N_2, \ldots, N_n$ such that $|x_{ji} - p_i| < \varepsilon/\sqrt{n}$ whenever $j \geq N_i$. Let N be the maximum of $N_1, N_2, \ldots, N_n$. If $j \geq N$ then

$$|\mathbf{x}_j - \mathbf{p}|^2 = \sum_{i=1}^{n} (x_{ji} - p_i)^2 < n(\varepsilon/\sqrt{n})^2 = \varepsilon^2,$$

so that $\mathbf{x}_j \to \mathbf{p}$ as $j \to +\infty$. ∎

The following result is the analogue of the Bolzano-Weierstrass Theorem for sequences in n-dimensional Euclidean space.

Theorem 2.3 *Every bounded sequence of points in $\mathbb{R}^n$ has a convergent subsequence.*

Proof Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ be a bounded sequence of points in $\mathbb{R}^n$. Let us denote by $x_j^{(i)}$ the ith component of the point $\mathbf{x}_j$, so that

$$\mathbf{x}_j = (x_j^{(1)}, x_j^{(2)}, \ldots, x_j^{(n)})$$

for all positive integers j. Suppose that, for some integer s between 1 and $n-1$, the sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ has a subsequence $\mathbf{x}_{p_1}, \mathbf{x}_{p_2}, \mathbf{x}_{p_3}, \ldots$ with the property that, for each integer i satisfying $1 \leq i \leq s$, the ith components of the members of this subsequence constitute a convergent sequence $x_{p_1}^{(i)}, x_{p_2}^{(i)}, x_{p_3}^{(i)}, \ldots$ of real numbers. Let $a_j = x_{p_j}^{(s+1)}$ for each positive integer j. Then $a_1, a_2, a_3, \ldots$ is a bounded sequence of real numbers. It follows from the Bolzano-Weierstrass Theorem (Theorem 1.6) that this sequence has a convergent subsequence $a_{m_1}, a_{m_2}, a_{m_3}, \ldots$, where $m_1 < m_2 < m_3 < \cdots$. Let $q_j = p_{m_j}$ for each positive integer j. Then $\mathbf{x}_{q_1}, \mathbf{x}_{q_2}, \mathbf{x}_{q_3}, \ldots$ is a subsequence of the original bounded sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ which has the property that, for each integer i satisfying $1 \leq i \leq s+1$, the ith components of the members of the subsequence constitute a convergent sequence $x_{q_1}^{(i)}, x_{q_2}^{(i)}, x_{q_3}^{(i)}, \ldots$ of real numbers.

Repeated applications of this result show that the bounded sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ has a subsequence $\mathbf{x}_{r_1}, \mathbf{x}_{r_2}, \mathbf{x}_{r_3}, \ldots$ with the property that, for each integer i satisfying $1 \leq i \leq n$, the ith components of the members of the subsequence constitute a convergent sequence of real numbers. Let $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ where, for each value of i between 1 and n, the ith component z_i of $\mathbf{z}$ is the limit of the sequence $x_{r_1}^{(i)}, x_{r_2}^{(i)}, x_{r_3}^{(i)}, \ldots$ of ith components of the members of the subsequence $\mathbf{x}_{r_1}, \mathbf{x}_{r_2}, \mathbf{x}_{r_3}, \ldots$. Then this subsequence converges to the point $\mathbf{z}$, as required. ∎

2.3 Cauchy Sequences of Points in Euclidean Spaces

Definition A sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ of points in a Euclidean space is said to be a *Cauchy sequence* if, given any $\varepsilon > 0$, there exists some natural number N such that $|\mathbf{x}_j - \mathbf{x}_k| < \varepsilon$ for all integers j and k satisfying $j \geq N$ and $k \geq N$.

Lemma 2.4 *Every convergent sequence in a Euclidean space is a Cauchy sequence.*

Proof Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ be a sequence of points in a Euclidean space $\mathbb{R}^n$ which converges to some point $\mathbf{p}$ of $\mathbb{R}^n$. Given any $\varepsilon > 0$, there exists some natural number N such that $|\mathbf{x}_j - \mathbf{p}| < \varepsilon/2$ whenever $j \geq N$. But then it follows from the Triangle Inequality that

$$|\mathbf{x}_j - \mathbf{x}_k| \leq |\mathbf{x}_j - \mathbf{p}| + |\mathbf{p} - \mathbf{x}_k| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

whenever $j \geq N$ and $k \geq N$. ∎

Theorem 2.5 *Every Cauchy sequence in $\mathbb{R}^n$ converges to some point of $\mathbb{R}^n$.*

Proof Let $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \ldots$ be a Cauchy sequence in $\mathbb{R}^n$. Then, given any real number ε satisfying $\varepsilon > 0$, there exists some natural number N such that $|\mathbf{p}_j - \mathbf{p}_k| < \varepsilon$ whenever $j \geq N$ and $k \geq N$. In particular, there exists some natural number L such that $|\mathbf{p}_j - \mathbf{p}_k| < 1$ whenever $j \geq L$ and $k \geq L$. Let R be the maximum of the numbers $|\mathbf{p}_1|, |\mathbf{p}_2|, \ldots, |\mathbf{p}_{L-1}|$ and $|\mathbf{p}_L| + 1$. Then $|\mathbf{p}_j| \leq R$ whenever $j < L$. Moreover if $j \geq L$ then

$$|\mathbf{p}_j| \leq |\mathbf{p}_L| + |\mathbf{p}_j - \mathbf{p}_L| < |\mathbf{p}_L| + 1 \leq R.$$

Thus $|\mathbf{p}_j| \leq R$ for all positive integers j. We conclude that the Cauchy sequence $(\mathbf{p}_j : j \in N)$ is bounded.

Now every bounded sequence of points in a Euclidean space has a convergent subsequence (Theorem 2.3). In particular, the Cauchy sequence $(\mathbf{p}_j : j \in N)$ has a convergent subsequence $(\mathbf{p}_{k_j} : j \in N)$, where $k_1, k_2, k_3, \ldots$ are positive integers satisfying $k_1 < k_2 < k_3 < \cdots$. Let the point $\mathbf{q}$ of $\mathbb{R}^n$ be the limit of this subsequence. Then, given any positive number ε satisfying $\varepsilon > 0$, there exists some positive integer M such that $|\mathbf{p}_{k_m} - \mathbf{q}| < \frac{1}{2}\varepsilon$ whenever $m \geq M$. Also there exists some positive integer N such that $|\mathbf{p}_j - \mathbf{p}_k| < \frac{1}{2}\varepsilon$ whenever $j \geq N$ and $k \geq N$. Choose m large enough to ensure that $m \geq M$ and $k_m \geq N$. If $j \geq N$ then

$$|\mathbf{p}_j - \mathbf{q}| \leq |\mathbf{p}_j - \mathbf{p}_{k_m}| + |\mathbf{p}_{k_m} - \mathbf{q}| < \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon.$$

It follows that the Cauchy sequence $(\mathbf{p}_j : j \in N)$ converges to the point $\mathbf{q}$. Thus every Cauchy sequence is convergent, as required. ∎

2.4 Continuity

Definition Let X and Y be a subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. A function $f\colon X \to Y$ from X to Y is said to be *continuous* at a point $\mathbf{p}$ of X if and only if the following criterion is satisfied:—

> given any real number ε satisfying $\varepsilon > 0$ there exists some $\delta > 0$ such that $|f(\mathbf{x}) - f(\mathbf{p})| < \varepsilon$ for all points $\mathbf{x}$ of X satisfying $|\mathbf{x} - \mathbf{p}| < \delta$.

The function $f\colon X \to Y$ is said to be continuous on X if and only if it is continuous at every point $\mathbf{p}$ of X.

Lemma 2.6 *The 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.*

Proof Let $(u,v) \in \mathbb{R}^2$. We first show that $s: \mathbb{R}^2 \to \mathbb{R}$ is continuous at (u,v). Let $\varepsilon > 0$ be given. Let $\delta = \frac{1}{2}\varepsilon$. If (x,y) is any point of $\mathbb{R}^2$ whose distance from (u,v) is less than δ then $|x-u| < \delta$ and $|y-v| < \delta$, and hence

$$|s(x,y) - s(u,v)| = |x+y-u-v| \leq |x-u| + |y-v| < 2\delta = \varepsilon.$$

This shows that $s: \mathbb{R}^2 \to \mathbb{R}$ is continuous at (u,v).

Next we show that $p: \mathbb{R}^2 \to \mathbb{R}$ is continuous at (u,v). Now

$$p(x,y) - p(u,v) = xy - uv = (x-u)(y-v) + u(y-v) + (x-u)v$$

for all points (x,y) of $\mathbb{R}^2$. Thus if the distance from (x,y) to (u,v) is less than δ then $|x-u| < \delta$ and $|y-v| < \delta$, and hence $|p(x,y) - p(u,v)| < \delta^2 + (|u| + |v|)\delta$. Let $\varepsilon > 0$ is given. If $\delta > 0$ is chosen to be the minimum of 1 and $\varepsilon/(1+|u|+|v|)$ then $\delta^2 + (|u|+|v|)\delta < (1+|u|+|v|)\delta < \varepsilon$, and thus $|p(x,y) - p(u,v)| < \varepsilon$ for all points (x,y) of $\mathbb{R}^2$ whose distance from (u,v) is less than δ. This shows that $p: \mathbb{R}^2 \to \mathbb{R}$ is continuous at (u,v). ∎

Lemma 2.7 *Let X, Y and Z be subsets of $\mathbb{R}^m$, $\mathbb{R}^n$ and $\mathbb{R}^k$ respectively, let $f: X \to Y$ be a function mapping X into Y, and let $g: Y \to Z$ be a function mapping Y into Z. Let $\mathbf{p}$ be a point of X. Suppose that f is continuous at $\mathbf{p}$ and g is continuous at $f(\mathbf{p})$. Then the composition function $g \circ f: X \to Z$ is continuous at $\mathbf{p}$.*

Proof Let $\varepsilon > 0$ be given. Then there exists some $\eta > 0$ such that $|g(\mathbf{y}) - g(f(\mathbf{p}))| < \varepsilon$ for all $\mathbf{y} \in Y$ satisfying $|\mathbf{y} - f(\mathbf{p})| < \eta$. But then there exists some $\delta > 0$ such that $|f(\mathbf{x}) - f(\mathbf{p})| < \eta$ for all $\mathbf{x} \in X$ satisfying $|\mathbf{x} - \mathbf{p}| < \delta$. It follows that $|g(f(\mathbf{x})) - g(f(\mathbf{p}))| < \varepsilon$ for all $\mathbf{x} \in X$ satisfying $|\mathbf{x} - \mathbf{p}| < \delta$, and thus $g \circ f$ is continuous at $\mathbf{p}$, as required. ∎

Lemma 2.7 guarantees that a composition of continuous functions between subsets of Euclidean spaces is continuous.

2.5 Convergent Sequences and Continuous Functions

Lemma 2.8 *Let X and Y be a subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively, and let $f: X \to Y$ be a continuous function from X to Y. Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ be a sequence of points of X which converges to some point $\mathbf{p}$ of X. Then the sequence $f(\mathbf{x}_1), f(\mathbf{x}_2), f(\mathbf{x}_3), \ldots$ converges to $f(\mathbf{p})$.*

Proof Let $\varepsilon > 0$ be given. Then there exists some $\delta > 0$ such that $|f(\mathbf{x}) - f(\mathbf{p})| < \varepsilon$ for all $\mathbf{x} \in X$ satisfying $|\mathbf{x} - \mathbf{p}| < \delta$, since the function f is continuous at $\mathbf{p}$. Also there exists some natural number N such that $|\mathbf{x}_j - \mathbf{p}| < \delta$ whenever $j \geq N$, since the sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ converges to $\mathbf{p}$. Thus if $j \geq N$ then $|f(\mathbf{x}_j) - f(\mathbf{p})| < \varepsilon$. Thus the sequence $f(\mathbf{x}_1), f(\mathbf{x}_2), f(\mathbf{x}_3), \ldots$ converges to $f(\mathbf{p})$, as required. ∎

Proposition 2.9 *Let $a_1, a_2, a_3, \ldots$ and $b_1, b_2, b_3, \ldots$ be convergent infinite sequences of real numbers. Then the sum, difference and product of these sequences are convergent, and*

$$\begin{aligned}\lim_{j\to+\infty}(a_j + b_j) &= \lim_{j\to+\infty} a_j + \lim_{j\to+\infty} b_j,\\ \lim_{j\to+\infty}(a_j - b_j) &= \lim_{j\to+\infty} a_j - \lim_{j\to+\infty} b_j,\\ \lim_{j\to+\infty}(a_j b_j) &= \left(\lim_{j\to+\infty} a_j\right)\left(\lim_{j\to+\infty} b_j\right).\end{aligned}$$

If in addition $b_j \neq 0$ for all n and $\lim_{j\to+\infty} b_j \neq 0$, then the quotient of the sequences (a_j) and (b_j) is convergent, and

$$\lim_{j\to+\infty}\frac{a_j}{b_j} = \frac{\lim\limits_{j\to+\infty} a_j}{\lim\limits_{j\to+\infty} b_j}.$$

Proof Throughout this proof let $l = \lim_{j\to+\infty} a_j$ and $m = \lim_{j\to+\infty} b_j$.

Now $a_j + b_j = s(a_j, b_j)$ and $a_j b_j = p(a_j, b_j)$ for all positive integers j, where $s\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ and $p\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ are the functions given by $s(x, y) = x + y$ and $p(x, y) = xy$ for all real numbers x and y. Also the sequence $((a_j, b_j) : j \in \mathbb{N})$ is a sequence of points in $\mathbb{R}^2$ which converges to the point (l, m), since its components are sequences of real numbers converging to the limits l and m (Lemma 2.2). Moreover the functions s and p are continuous (Lemma 2.6). It now follows from Lemma 2.8 that

$$\lim_{j\to+\infty}(a_j + b_j) = \lim_{j\to+\infty} s(a_j, b_j) = s\left(\lim_{j\to+\infty}(a_j, b_j)\right) = s(l, m) = l + m,$$

and

$$\lim_{j\to+\infty}(a_j b_j) = \lim_{j\to+\infty} p(a_j, b_j) = p\left(\lim_{j\to+\infty}(a_j, b_j)\right) = p(l, m) = lm.$$

Also the sequence $(-b_j : j \in \mathbb{N})$ converges to $-m$, and therefore $\lim_{j\to+\infty}(a_j - b_j) = l - m$. Now the reciprocal function $r\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$ is continuous on

$\mathbb{R} \setminus \{0\}$, where $r(x) = 1/x$ for all non-zero real numbers x. It follows from Lemma 2.8 that if $b_j \neq 0$ for all positive integers j then $1/b_j$ converges to $1/m$ as $j \to +\infty$. But then $\lim_{j \to +\infty} (a_j/b_j) = l/m$. This completes the proof of Proposition 2.9. ∎

2.6 Components of Continuous Functions

Let $f: X \to \mathbb{R}^n$ be a function mapping a mapping a set X into n-dimensional Euclidean space $\mathbb{R}^n$. Then

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

for all $x \in X$, where $f_1, f_2, \ldots, f_n$ are functions from X to $\mathbb{R}$, referred to as the *components* of the function f.

Proposition 2.10 *Let X be a subset of some Euclidean space, and let $\mathbf{p}$ be a point of X. A function $f: X \to \mathbb{R}^n$ mapping X into the Euclidean space $\mathbb{R}^n$ is continuous at $\mathbf{p}$ if and only if its components are continuous at $\mathbf{p}$.*

Proof Note that the ith component f_i of f is given by $f_i = p_i \circ f$, where $p_i: \mathbb{R}^n \to \mathbb{R}$ is the continuous function which maps $(y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$ onto its ith coordinate y_i. It therefore follows immediately from Lemma 2.7 that if f is continuous the point $\mathbf{p}$, then so are the components of f.

Conversely suppose that the components of f are continuous at $\mathbf{p} \in X$. Let $\varepsilon > 0$ be given. Then there exist positive real numbers $\delta_1, \delta_2, \ldots, \delta_n$ such that $|f_i(\mathbf{x}) - f_i(\mathbf{p})| < \varepsilon/\sqrt{n}$ for $\mathbf{x} \in X$ satisfying $|\mathbf{x} - \mathbf{p}| < \delta_i$. Let δ be the minimum of $\delta_1, \delta_2, \ldots, \delta_n$. If $\mathbf{x} \in X$ satisfies $|\mathbf{x} - \mathbf{p}| < \delta$ then

$$|f(\mathbf{x}) - f(\mathbf{p})|^2 = \sum_{i=1}^{n} |f_i(\mathbf{x}) - f_i(\mathbf{p})|^2 < \varepsilon^2,$$

and hence $|f(\mathbf{x}) - f(\mathbf{p})| < \varepsilon$. Thus the function f is continuous at $\mathbf{p}$, as required. ∎

Proposition 2.11 *Let $f: X \to \mathbb{R}$ and $g: X \to \mathbb{R}$ be real-valued functions defined on some subset X of a Euclidean space, and let $\mathbf{p}$ be a point of X. Suppose that the functions f and g are continuous at the point $\mathbf{p}$. Then so are the functions $f + g$, $f - g$ and $f \cdot g$. If in addition $g(x) \neq 0$ for all $x \in X$ then the quotient function f/g is continuous at $\mathbf{p}$.*

Proof Note that $f + g = s \circ h$ and $f \cdot g = p \circ h$, where $h: X \to \mathbb{R}^2$, $s: \mathbb{R}^2 \to \mathbb{R}$ and $p: \mathbb{R}^2 \to \mathbb{R}$ are given by $h(x) = (f(x), g(x))$, $s(u, v) = u + v$ and $p(u, v) = uv$ for all $x \in X$ and $u, v \in \mathbb{R}$. If the functions f and g are continuous at $\mathbf{p}$ then so is the function h (Proposition 2.10). The functions s and p are continuous on $\mathbb{R}^2$. It therefore follows from Lemma 2.7 that the composition functions $s \circ h$ and $p \circ h$ are continuous at $\mathbf{p}$. Thus the functions $f + g$ and $f \cdot g$ are continuous at $\mathbf{p}$. Now $f - g = f + (-g)$, and both f and $-g$ are continuous. Therefore $f - g$ is continuous.

Now suppose that $g(x) \neq 0$ for all $x \in X$. Note that $1/g = r \circ g$, where $r: \mathbb{R} \setminus \{0\} \to \mathbb{R}$ is the reciprocal function, defined by $r(t) = 1/t$. Now the reciprocal function r is continuous. It now follows on applying Lemma 2.7 that the function $1/g$ is continuous at $\mathbf{p}$. The function f/g, being the product of the functions f and $1/g$ is therefore continuous at $\mathbf{p}$. ∎

2.7 Limits of Functions

Let X be a subset of some Euclidean space $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of $\mathbb{R}^n$. We say that the point $\mathbf{p}$ is a *limit point* of X if, given any real number δ satisfying $\delta > 0$, there exists a point $\mathbf{x}$ of X satisfying $0 < |\mathbf{x} - \mathbf{p}| < \delta$. It follows easily from this that a point $\mathbf{p}$ of $\mathbb{R}^n$ is a limit point of X if and only if there exists a sequence of points of $X \setminus \{\mathbf{p}\}$ which converges to the point $\mathbf{p}$.

Definition Let X be a subset of a Euclidean space $\mathbb{R}^m$, let $f: X \to \mathbb{R}^n$ mapping X into a Euclidean space $\mathbb{R}^n$, let $\mathbf{p}$ be a limit point of X, and let $\mathbf{q}$ be a point of $\mathbb{R}^n$. We say that $\mathbf{q}$ is the *limit* of $f(\mathbf{x})$ as $\mathbf{x}$ tends to $\mathbf{p}$ in X if, given any real number ε satisfying $\varepsilon > 0$, there exists some real number δ satisfying $\delta > 0$ such that $|f(\mathbf{x}) - \mathbf{q}| < \varepsilon$ for all points $\mathbf{x}$ of X satisfying $0 < |\mathbf{x} - \mathbf{p}| < \delta$. If the point $\mathbf{q}$ is the limit of $f(\mathbf{x})$ as $\mathbf{x}$ tends to $\mathbf{p}$ in X, then we denote this fact by writing: $\mathbf{q} = \lim_{\mathbf{x} \to \mathbf{p}} f(\mathbf{x})$.

Lemma 2.12 *Let X and Y be subsets of Euclidean spaces, let $f: X \to Y$ be a function from X to Y, and let $\mathbf{p}$ be a point of X that is also limit point of X. Then the function f is continuous at $\mathbf{p}$ if and only if $\lim_{\mathbf{x} \to \mathbf{p}} f(\mathbf{x}) = f(\mathbf{p})$.*

Proof The result follows immediately on comparing the definitions of convergence and of limits of functions. ∎

Let X be a subset of some Euclidean space. A point $\mathbf{p}$ of X is said to be an *isolated point* of X if it is not a limit point of X. A point $\mathbf{p}$ of X is an isolated point of X if and only if there exists some real number δ satisfying $\delta > 0$ such that the only point of X whose distance from $\mathbf{p}$ is less than δ is the point $\mathbf{p}$ itself. It follows directly from the definition of continuity that any function between subsets of Euclidean space is continuous at all the isolated points of its domain.

Lemma 2.13 *Let X, Y and Z be subsets of Euclidean spaces, let $\mathbf{p}$ be a limit point of X, and let $f: X \to Y$ and $g: Y \to Z$ be functions. Suppose that $\lim_{x \to p} f(x) = \mathbf{q}$. Suppose also that the function g is defined and is continuous at $\mathbf{q}$. Then $\lim_{x \to p} g(f(\mathbf{x})) = g(\mathbf{q})$.*

Proof The function g is continuous at $\mathbf{q}$. Therefore there exists some $\eta > 0$ such that $|g(\mathbf{y}) - g(\mathbf{q})| < \varepsilon$ for all $\mathbf{y} \in Y$ satisfying $|\mathbf{y} - \mathbf{q}| < \eta$. But then there exists some $\delta > 0$ such that $|f(\mathbf{x}) - \mathbf{q}| < \eta$ for all $\mathbf{x} \in X$ satisfying $0 < |\mathbf{x} - \mathbf{p}| < \delta$. Thus $|g(f(\mathbf{x})) - g(\mathbf{q})| < \varepsilon$ for all $\mathbf{x} \in X$ satisfying $0 < |\mathbf{x} - \mathbf{p}| < \delta$, showing that $\lim_{\mathbf{x} \to \mathbf{p}} g(f(\mathbf{x})) = g(\mathbf{q})$, as required. ∎

Let X be a subset of some Euclidean space, let $f: X \to \mathbb{R}^n$ be a function mapping X into n-dimensional Euclidean space $\mathbb{R}^n$, let $\mathbf{p}$ be a limit point of the set X, and let $\mathbf{q}$ be a point in $\mathbb{R}^n$. Let $\tilde{f}: X \cup \{\mathbf{p}\} \to \mathbb{R}^n$ be the function on $X \cup \{\mathbf{p}\}$ defined such that

$$\tilde{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \in X \setminus \{\mathbf{p}\}; \\ \mathbf{q} & \text{if } \mathbf{x} = \mathbf{p}. \end{cases}$$

Then $\lim_{\mathbf{x} \to \mathbf{p}} f(\mathbf{x}) = \mathbf{q}$ if and only if the function $\tilde{f}$ is continuous at $\mathbf{p}$. This enables one to deduce basic results concerning limits of functions from the corresponding results concerning continuity of functions.

The following result is thus a conseqence of Proposition 2.10.

Proposition 2.14 *Let X be a subset of some Euclidean space, let $f: X \to \mathbb{R}^n$ be a function mapping X into n-dimensional Euclidean space $\mathbb{R}^n$, let $\mathbf{p}$ be a limit point of the set X, and let $\mathbf{q}$ be a point in $\mathbb{R}^n$. Let the real-valued functions $f_1, f_2, \ldots, f_n$ be the components of f, so that*

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

for all $\mathbf{x} \in X$, and let $\mathbf{q} = (q_1, q_2, \ldots, q_n)$. Then $\lim_{\mathbf{x} \to \mathbf{p}} f(\mathbf{x}) = \mathbf{q}$ if and only if $\lim_{\mathbf{x} \to \mathbf{p}} f_i(\mathbf{x}) = q_i$ for $i = 1, 2, \ldots, n$.

The following result is a consequence of Proposition 2.11.

Proposition 2.15 *Let X be a subset of some Euclidean space, let $\mathbf{p}$ be a limit point of X, and let $f\colon X \to \mathbb{R}^n$ and $g\colon X \to \mathbb{R}^n$ be functions on X taking values in some Euclidean space $\mathbb{R}^n$. Suppose that the limits $\lim_{\mathbf{x}\to\mathbf{p}} f(\mathbf{x})$ and $\lim_{\mathbf{x}\to\mathbf{p}} g(\mathbf{x})$ exist. Then*

$$\begin{aligned}
\lim_{\mathbf{x}\to\mathbf{p}} (f(\mathbf{x}) + g(\mathbf{x})) &= \lim_{\mathbf{x}\to\mathbf{p}} f(\mathbf{x}) + \lim_{\mathbf{x}\to\mathbf{p}} g(\mathbf{x}), \\
\lim_{\mathbf{x}\to\mathbf{p}} (f(\mathbf{x}) - g(\mathbf{x})) &= \lim_{\mathbf{x}\to\mathbf{p}} f(\mathbf{x}) - \lim_{\mathbf{x}\to\mathbf{p}} g(\mathbf{x}), \\
\lim_{\mathbf{x}\to\mathbf{p}} (f(\mathbf{x})g(\mathbf{x})) &= \left(\lim_{\mathbf{x}\to\mathbf{p}} f(\mathbf{x})\right)\left(\lim_{\mathbf{x}\to\mathbf{p}} g(\mathbf{x})\right).
\end{aligned}$$

If moreover $\lim_{\mathbf{x}\to\mathbf{p}} g(\mathbf{x}) \neq 0$ and the function g is non-zero throughout its domain X then

$$\lim_{\mathbf{x}\to\mathbf{p}} \frac{f(\mathbf{x})}{g(\mathbf{x})} = \frac{\lim_{\mathbf{x}\to\mathbf{p}} f(\mathbf{x})}{\lim_{\mathbf{x}\to\mathbf{p}} g(\mathbf{x})}.$$

2.8 Uniform Convergence

Definition Let X be a subset of some Euclidean spaces, and let $f_1, f_2, f_3, \ldots$ be a sequence of functions mapping X into some Euclidean space $\mathbb{R}^n$. The sequence (f_j) is said to converge *uniformly* to a function $f\colon X \to \mathbb{R}^n$ on X as $j \to +\infty$ if, given any real number ε satisfying $\varepsilon > 0$, there exists some positive integer N such that $|f_j(\mathbf{x}) - f(\mathbf{x})| < \varepsilon$ for all $\mathbf{x} \in X$ and for all integers j satisfying $j \geq N$ (where the value of N is independent of $\mathbf{x}$).

Theorem 2.16 *Let $f_1, f_2, f_3, \ldots$ be a sequence of continuous functions mapping some subset X of a Euclidean space into $\mathbb{R}^n$. Suppose that this sequence converges uniformly on X to some function $f\colon X \to \mathbb{R}^n$. Then this limit function f is continuous.*

Proof Let $\mathbf{p}$ be an element of X, and let $\varepsilon > 0$ be given. If j is chosen sufficiently large then $|f(\mathbf{x}) - f_j(\mathbf{x})| < \frac{1}{3}\varepsilon$ for all $\mathbf{x} \in X$, since $f_j \to f$ uniformly on X as $j \to +\infty$. It then follows from the continuity of f_j that there exists some $\delta > 0$ such that $|f_j(\mathbf{x}) - f_j(\mathbf{p})| < \frac{1}{3}\varepsilon$ for all $\mathbf{x} \in X$ satisfying $|\mathbf{x} - \mathbf{p}| < \delta$. But then

$$\begin{aligned}
|f(\mathbf{x}) - f(\mathbf{p})| &\leq |f(\mathbf{x}) - f_j(\mathbf{x})| + |f_j(\mathbf{x}) - f_j(\mathbf{p})| + |f_j(\mathbf{p}) - f(\mathbf{p})| \\
&< \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon = \varepsilon
\end{aligned}$$

whenever $|\mathbf{x} - \mathbf{p}| < \delta$. Thus the function f is continuous at $\mathbf{p}$, as required. ∎

2.9 Open Sets in Euclidean Spaces

Given a point $\mathbf{p}$ of $\mathbb{R}^n$ and a non-negative real number r, the *open ball* $B(\mathbf{p}, r)$ of *radius* r about $\mathbf{p}$ is defined to be the subset of $\mathbb{R}^n$ given by

$$B(\mathbf{p}, r) \equiv \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{p}| < r\}.$$

(Thus $B(\mathbf{p}, r)$ is the set consisting of all points of $\mathbb{R}^n$ that lie within a sphere of radius r centred on the point $\mathbf{p}$.)

Definition A subset V of $\mathbb{R}^n$ is said to be *open* in $\mathbb{R}^n$ if and only if, given any point $\mathbf{p}$ of V, there exists some $\delta > 0$ such that $B(\mathbf{p}, \delta) \subset V$.

By convention, we regard the empty set $\emptyset$ as being an open subset of $\mathbb{R}^n$. (The criterion given above is satisfied vacuously in the case when V is the empty set.)

Example Let $H = \{(x, y, z) \in \mathbb{R}^3 : z > c\}$, where c is some real number. Then H is an open set in $\mathbb{R}^3$. Indeed let $\mathbf{p}$ be a point of H. Then $\mathbf{p} = (u, v, w)$, where $w > c$. Let $\delta = w - c$. If the distance from a point (x, y, z) to the point (u, v, w) is less than δ then $|z - w| < \delta$, and hence $z > c$, so that $(x, y, z) \in H$. Thus $B(\mathbf{p}, \delta) \subset H$, and therefore H is an open set.

The previous example can be generalized. Given any integer i between 1 and n, and given any real number c_i, the sets

$$\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_i > c_i\}, \qquad \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_i < c_i\}$$

are open sets in $\mathbb{R}^n$.

Lemma 2.17 *Let* $\mathbf{p}$ *be a point of* n*-dimensional Euclidean space* $\mathbb{R}^n$*. Then, for any positive real number* r*, the open ball* $B(\mathbf{p}, r)$ *of radius* r *about* $\mathbf{p}$ *is an open set in* $\mathbb{R}^n$*.*

Proof Let $\mathbf{x}$ be an element of $B(\mathbf{p}, r)$. We must show that there exists some $\delta > 0$ such that $B(\mathbf{x}, \delta) \subset B(\mathbf{p}, r)$. Let $\delta = r - |\mathbf{x} - \mathbf{p}|$. Then $\delta > 0$, since $|\mathbf{x} - \mathbf{p}| < r$. Moreover if $\mathbf{y} \in B(\mathbf{x}, \delta)$ then

$$|\mathbf{y} - \mathbf{p}| \leq |\mathbf{y} - \mathbf{x}| + |\mathbf{x} - \mathbf{p}| < \delta + |\mathbf{x} - \mathbf{p}| = r,$$

by the Triangle Inequality, and hence $\mathbf{y} \in B(\mathbf{p}, r)$. Thus $B(\mathbf{x}, \delta) \subset B(\mathbf{p}, r)$. This shows that $B(\mathbf{p}, r)$ is an open set, as required. ∎

Lemma 2.18 *Let* $\mathbf{p}$ *be a point of* n*-dimensional Euclidean space* $\mathbb{R}^n$*. Then, for any non-negative real number* r*, the set* $\{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{p}| > r\}$ *is an open set in* $\mathbb{R}^n$*.*

Proof Let $\mathbf{x}$ be a point of $\mathbb{R}^n$ satisfying $|\mathbf{x} - \mathbf{p}| > r$, and let $\mathbf{y}$ be any point of $\mathbb{R}^n$ satisfying $|\mathbf{y} - \mathbf{x}| < \delta$, where $\delta = |\mathbf{x} - \mathbf{p}| - r$. Then

$$|\mathbf{x} - \mathbf{p}| \leq |\mathbf{x} - \mathbf{y}| + |\mathbf{y} - \mathbf{p}|,$$

by the Triangle Inequality, and therefore

$$|\mathbf{y} - \mathbf{p}| \geq |\mathbf{x} - \mathbf{p}| - |\mathbf{y} - \mathbf{x}| > |\mathbf{x} - \mathbf{p}| - \delta = r.$$

Thus $B(\mathbf{x}, \delta)$ is contained in the given set. The result follows. ∎

Proposition 2.19 *The collection of open sets in n-dimensional Euclidean space $\mathbb{R}^n$ has the following properties:—*

(i) *the empty set $\emptyset$ and the whole space $\mathbb{R}^n$ are both open in $\mathbb{R}^n$;*

(ii) *the union of any collection of open sets in $\mathbb{R}^n$ is itself open in $\mathbb{R}^n$;*

(iii) *the intersection of any* finite *collection of open sets in $\mathbb{R}^n$ is itself open in $\mathbb{R}^n$.*

Proof The empty set $\emptyset$ is an open set by convention. Moreover the definition of an open set is satisfied trivially by the whole space $\mathbb{R}^n$. This proves (i).

Let $\mathcal{A}$ be any collection of open sets in $\mathbb{R}^n$, and let U denote the union of all the open sets belonging to $\mathcal{A}$. We must show that U is itself open in $\mathbb{R}^n$. Let $\mathbf{x} \in U$. Then $\mathbf{x} \in V$ for some set V belonging to the collection $\mathcal{A}$. It follows that there exists some $\delta > 0$ such that $B(\mathbf{x}, \delta) \subset V$. But $V \subset U$, and thus $B(\mathbf{x}, \delta) \subset U$. This shows that U is open in $\mathbb{R}^n$. This proves (ii).

Finally let $V_1, V_2, V_3, \ldots, V_k$ be a *finite* collection of subsets of $\mathbb{R}^n$ that are open in $\mathbb{R}^n$, and let V denote the intersection $V_1 \cap V_2 \cap \cdots \cap V_k$ of these sets. Let $\mathbf{x} \in V$. Now $\mathbf{x} \in V_j$ for $j = 1, 2, \ldots, k$, and therefore there exist strictly positive real numbers $\delta_1, \delta_2, \ldots, \delta_k$ such that $B(\mathbf{x}, \delta_j) \subset V_j$ for $j = 1, 2, \ldots, k$. Let δ 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 sets.) Now $B(\mathbf{x}, \delta) \subset B(\mathbf{x}, \delta_j) \subset V_j$ for $j = 1, 2, \ldots, k$, and thus $B(\mathbf{x}, \delta) \subset V$. Thus the intersection V of the sets $V_1, V_2, \ldots, V_k$ is itself open in $\mathbb{R}^n$. This proves (iii). ∎

Example The set $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < 4 \text{ and } z > 1\}$ is an open set in $\mathbb{R}^3$, since it is the intersection of the open ball of radius 2 about the origin with the open set $\{(x, y, z) \in \mathbb{R}^3 : z > 1\}$.

Example The set $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < 4 \text{ or } z > 1\}$ is an open set in $\mathbb{R}^3$, since it is the union of the open ball of radius 2 about the origin with the open set $\{(x, y, z) \in \mathbb{R}^3 : z > 1\}$.

Example The set

$$\{(x, y, z) \in \mathbb{R}^3 : (x-n)^2 + y^2 + z^2 < \tfrac{1}{4} \text{ for some } n \in \mathbb{Z}\}$$

is an open set in $\mathbb{R}^3$, since it is the union of the open balls of radius $\frac{1}{2}$ about the points $(n, 0, 0)$ for all integers n.

Example For each natural number k, let

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

Now each set V_k is an open ball of radius $1/k$ about the origin, and is therefore an open set in $\mathbb{R}^3$. However the intersection of the sets V_k for all natural numbers k is the set $\{(0, 0, 0)\}$, and thus the intersection of the sets V_k for all natural numbers k is not itself an open set in $\mathbb{R}^3$. This example demonstrates that infinite intersections of open sets need not be open.

Lemma 2.20 *A sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ of points in $\mathbb{R}^n$ converges to a point $\mathbf{p}$ if and only if, given any open set U which contains $\mathbf{p}$, there exists some natural number N such that $\mathbf{x}_j \in U$ for all j satisfying $j \geq N$.*

Proof Suppose that the sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ has the property that, given any open set U which contains $\mathbf{p}$, there exists some natural number N such that $\mathbf{x}_j \in U$ whenever $j \geq N$. Let $\varepsilon > 0$ be given. The open ball $B(\mathbf{p}, \varepsilon)$ of radius ε about $\mathbf{p}$ is an open set by Lemma 2.17. Therefore there exists some natural number N such that $\mathbf{x}_j \in B(\mathbf{p}, \varepsilon)$ whenever $j \geq N$. Thus $|\mathbf{x}_j - \mathbf{p}| < \varepsilon$ whenever $j \geq N$. This shows that the sequence converges to $\mathbf{p}$.

Conversely, suppose that the sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ converges to $\mathbf{p}$. Let U be an open set which contains $\mathbf{p}$. Then there exists some $\varepsilon > 0$ such that the open ball $B(\mathbf{p}, \varepsilon)$ of radius ε about $\mathbf{p}$ is a subset of U. Thus there exists some $\varepsilon > 0$ such that U contains all points $\mathbf{x}$ of $\mathbb{R}^n$ that satisfy $|\mathbf{x} - \mathbf{p}| < \varepsilon$. But there exists some natural number N with the property that $|\mathbf{x}_j - \mathbf{p}| < \varepsilon$ whenever $j \geq N$, since the sequence converges to $\mathbf{p}$. Therefore $\mathbf{x}_j \in U$ whenever $j \geq N$, as required. ∎

2.10 Open Sets in Subsets of Euclidean Spaces

Let X be a subset of $\mathbb{R}^n$. Given a point $\mathbf{p}$ of X and a non-negative real number r, the *open ball* $B_X(\mathbf{p}, r)$ in X of *radius* r about $\mathbf{p}$ is defined to be the subset of X given by

$$B_X(\mathbf{p}, r) \equiv \{\mathbf{x} \in X : |\mathbf{x} - \mathbf{p}| < r\}.$$

(Thus $B_X(\mathbf{p}, r)$ is the set consisting of all points of X that lie within a sphere of radius r centred on the point $\mathbf{p}$.)

Definition Let X be a subset of $\mathbb{R}^n$. A subset V of X is said to be *open* in X if and only if, given any point $\mathbf{p}$ of V, there exists some $\delta > 0$ such that $B_X(\mathbf{p}, \delta) \subset V$.

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

Example Let U be an open set in $\mathbb{R}^n$. Then for any subset X of $\mathbb{R}^n$, the intersection $U \cap X$ is open in X. (This follows directly from the definitions.) Thus for example, let S^2 be the unit sphere in $\mathbb{R}^3$, given by

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

and let N be the subset of S^2 given by

$$N = \{(x, y, z) \in \mathbb{R}^n : x^2 + y^2 + z^2 = 1 \text{ and } z > 0\}.$$

Then N is open in S^2, since $N = H \cap S^2$, where H is the open set in $\mathbb{R}^3$ given by

$$H = \{(x, y, z) \in \mathbb{R}^3 : z > 0\}.$$

Note that N is not itself an open set in $\mathbb{R}^3$. Indeed the point $(0, 0, 1)$ belongs to N, but, for any $\delta > 0$, the open ball (in $\mathbb{R}^3$ of radius δ about $(0, 0, 1)$ contains points (x, y, z) for which $x^2 + y^2 + z^2 \neq 1$. Thus the open ball of radius δ about the point $(0, 0, 1)$ is not a subset of N.

2.11 Closed Sets in Euclidean Spaces

Definition A subset F of n-dimensional Euclidean space $\mathbb{R}^n$ is said to be *closed* in $\mathbb{R}^n$ if and only if its complement $\mathbb{R}^n \setminus F$ in $\mathbb{R}^n$ is open in $\mathbb{R}^n$. (Recall that $\mathbb{R}^n \setminus F = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \notin F\}$.)

Example The sets $\{(x, y, z) \in \mathbb{R}^3 : z \geq c\}$, $\{(x, y, z) \in \mathbb{R}^3 : z \leq c\}$, and $\{(x, y, z) \in \mathbb{R}^3 : z = c\}$ are closed sets in $\mathbb{R}^3$ for each real number c, since the complements of these sets are open in $\mathbb{R}^3$.

Example Let $\mathbf{p}$ be a point of n-dimensional Euclidean space $\mathbb{R}^n$. Then the sets $\{\mathbf{x} \in X : |\mathbf{x} - \mathbf{p}| \leq r\}$ and $\{\mathbf{x} \in X : |\mathbf{x} - \mathbf{p}| \geq r\}$ are closed for each non-negative real number r. In particular, the set $\{\mathbf{p}\}$ consisting of the single point $\mathbf{p}$ is a closed set in X. (These results follow immediately using Lemma 2.17 and Lemma 2.18 and the definition of closed sets.)

Let $\mathcal{A}$ be some collection of subsets of a set X. Then

$$X \setminus \bigcup_{S \in \mathcal{A}} S = \bigcap_{S \in \mathcal{A}} (X \setminus S), \qquad X \setminus \bigcap_{S \in \mathcal{A}} S = \bigcup_{S \in \mathcal{A}} (X \setminus S)$$

(i.e., the complement of the union of some collection of subsets of 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 Proposition 2.19.

Proposition 2.21 *The collection of closed sets in n-dimensional Euclidean space $\mathbb{R}^n$ has the following properties:—*

(i) *the empty set $\emptyset$ and the whole space $\mathbb{R}^n$ are both closed in $\mathbb{R}^n$;*

(ii) *the intersection of any collection of closed sets in $\mathbb{R}^n$ is itself closed in $\mathbb{R}^n$;*

(iii) *the union of any* finite *collection of closed sets in $\mathbb{R}^n$ is itself closed in $\mathbb{R}^n$.*

Lemma 2.22 *Let F be a closed subset of n-dimensional Euclidean space $\mathbb{R}^n$, and let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ be a sequence of points of F which converges to a point $\mathbf{p}$ of $\mathbb{R}^n$. Then $\mathbf{p} \in F$.*

Proof The complement $\mathbb{R}^n \setminus F$ of F in $\mathbb{R}^n$ is open, since F is closed. Suppose that $\mathbf{p}$ were a point belonging to $\mathbb{R}^n \setminus F$. It would then follow from Lemma 2.20 that $\mathbf{x}_j \in \mathbb{R}^n \setminus F$ for all values of j greater than some positive integer N, contradicting the fact that $\mathbf{x}_j \in F$ for all j. This contradiction shows that $\mathbf{p}$ must belong to F, as required. ∎

Lemma 2.23 *Let F be a closed bounded set in $\mathbb{R}^n$, and let U be an open set in $\mathbb{R}^n$. Suppose that $F \subset U$. Then there exists positive real number δ such that $|\mathbf{x} - \mathbf{y}| \geq \delta > 0$ for all $\mathbf{x} \in F$ and $\mathbf{y} \in \mathbb{R}^n \setminus U$.*

Proof Suppose that such a positive real number δ did not exist. Then there would exist an infinite sequence $(\mathbf{x}_j : j \in \mathbb{N})$ of points of F and a correspondinding infinite sequence $(\mathbf{y}_j : j \in \mathbb{N})$ of points of $\mathbb{R}^n \setminus U$ such that $|\mathbf{x}_j - \mathbf{y}_j| < 1/j$ for all positive integers j. The sequence $(\mathbf{x}_j : j \in \mathbb{N})$ would be a bounded sequence of points of $\mathbb{R}^n$, and would therefore have a convergent subsequence $(\mathbf{x}_{m_j} : j \in \mathbb{N})$ (Theorem 2.3). Let $\mathbf{p} = \lim_{j \to +\infty} \mathbf{x}_{m_j}$. Then $\mathbf{p} = \lim_{j \to +\infty} \mathbf{y}_{m_j}$, because $\lim_{j \to +\infty} (\mathbf{x}_{m_j} - \mathbf{y}_{m_j}) = \mathbf{0}$. But then $\mathbf{p} \in F$ and $\mathbf{p} \in \mathbb{R}^n \setminus U$, because the sets F and $\mathbb{R}^n \setminus U$ are closed (Lemma 2.22). But this is impossible, as $F \subset U$. It follows that there must exist some positive real number δ with the required properties. ∎

3 Metric Spaces

3.1 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.

Note that if X is a metric space with distance function d and if A is a subset of X then the restriction $d|A \times A$ of d to pairs of points of A defines a distance function on A satisfying the axioms for a metric space.

The set $\mathbb{R}$ of real numbers becomes a metric space with distance function d given by $d(x, y) = |x - y|$ for all $x, y \in \mathbb{R}$. Similarly the set $\mathbb{C}$ of complex numbers becomes a metric space with distance function d given by $d(z, w) = |z - w|$ for all $z, w \in \mathbb{C}$, and n-dimensional Euclidean space $\mathbb{R}^n$ is a metric space with with respect to the *Euclidean distance function* d, given 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 $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{R}^n$ may be regarded as a metric space whose distance function is the restriction to X of the distance function on $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{R}^n$ defined above.

Example The *n-sphere* S^n is defined to be the subset of $(n+1)$-dimensional Euclidean space $\mathbb{R}^{n+1}$ consisting of all elements $\mathbf{x}$ of $\mathbb{R}^{n+1}$ for which $|\mathbf{x}| = 1$. Thus

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

(Note that S^2 is the standard (2-dimensional) unit sphere in 3-dimensional Euclidean space.) The *chordal distance* between two points $\mathbf{x}$ and $\mathbf{y}$ of S^n is defined to be the length $|\mathbf{x} - \mathbf{y}|$ of the line segment joining $\mathbf{x}$ and $\mathbf{y}$. The n-sphere S^n is a metric space with respect to the chordal distance function.

3.2 Convergence of Sequences in a Metric Space

Definition Let X be a metric space with distance function d. A sequence $x_1, x_2, x_3, \ldots$ of points in X is said to *converge* to a point p in X if, given any strictly positive real number ε, there exists some natural number N such that $d(x_j, p) < \varepsilon$ whenever $j \geq N$.

We refer to p as the *limit* $\lim\limits_{j\to+\infty} x_j$ of the sequence $x_1, x_2, x_3, \ldots$.

This definition of convergence for infinite sequence of points in a metric space generalizes the standard definition of convergence for sequences of real numbers, and that for sequences of points in a Euclidean space.

If a sequence of points in a metric space is convergent then the limit of that sequence is unique. Indeed let $x_1, x_2, x_3, \ldots$ be a sequence of points in a metric space (X, d) which converges to points p and p' of X. We show that $p = p'$. Now, given any $\varepsilon > 0$, there exist natural numbers N_1 and N_2 such that $d(x_j, p) < \varepsilon$ whenever $j \geq N_1$ and $d(x_j, p') < \varepsilon$ whenever $j \geq N_2$. On choosing j so that $j \geq N_1$ and $j \geq N_2$ we see that

$$0 \leq d(p, p') \leq d(p, x_j) + d(x_j, p') < 2\varepsilon$$

by a straightforward application of the metric space axioms (i)–(iii). Thus $0 \leq d(p, p') < 2\varepsilon$ for every $\varepsilon > 0$, and hence $d(p, p') = 0$, so that $p = p'$ by Axiom (iv).

Lemma 3.1 *Let (X, d) be a metric space, and let $x_1, x_2, x_3, \ldots$ be a sequence of points of X which converges to some point p of X. Then, for any point y of X, $d(x_j, y) \to d(p, y)$ as $j \to +\infty$.*

Proof Let $\varepsilon > 0$ be given. We must show that there exists some natural number N such that $|d(x_j, y) - d(p, y)| < \varepsilon$ whenever $j \geq N$. However N can be chosen such that $d(x_j, p) < \varepsilon$ whenever $j \geq N$. But

$$d(x_j, y) \leq d(x_j, p) + d(p, y), \qquad d(p, y) \leq d(p, x_j) + d(x_j, y)$$

for all j, hence

$$-d(x_j, p) \leq d(x_j, y) - d(p, y) \leq d(x_j, p)$$

for all j, and hence $|d(x_j, y) - d(p, y)| < \varepsilon$ whenever $j \geq N$, as required. ∎

3.3 Continuity of Functions between Metric Spaces

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

- given any strictly positive real number ε, there exists some strictly positive real number δ such that $d_Y(f(x), f(p)) < \varepsilon$ for all points x of X satisfying $d_X(x, p) < \delta$.

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

This definition of continuity for functions between metric spaces generalizes the standard definition of continuity for functions between subsets of Euclidean spaces.

Lemma 3.2 *Let X, Y and Z be metric spaces, and let $f\colon X \to Y$ and $g\colon Y \to Z$ be continuous functions. Then the composition function $g \circ f\colon X \to Z$ is continuous.*

Proof We denote by d_X, d_Y and d_Z the distance functions on X, Y and Z respectively. Let p be any point of X. We show that $g \circ f$ is continuous at p. Let $\varepsilon > 0$ be given. Now the function g is continuous at $f(p)$. Hence there exists some $\eta > 0$ such that $d_Z(g(y), g(f(p))) < \varepsilon$ for all $y \in Y$ satisfying $d_Y(y, f(p)) < \eta$. But then there exists some $\delta > 0$ such that $d_Y(f(x), f(p)) < \eta$ for all $x \in X$ satisfying $d_X(x, p) < \delta$. Thus $d_Z(g(f(x)), g(f(p))) < \varepsilon$ for all $x \in X$ satisfying $d_X(x, p) < \delta$, showing that $g \circ f$ is continuous at p, as required. ∎

Lemma 3.3 *Let $f\colon X \to Y$ be a continuous function between metric spaces X and Y, and let $x_1, x_2, x_3, \ldots$ be a sequence of points in X which converges to some point p of X. Then the sequence $f(x_1), f(x_2), f(x_3), \ldots$ converges to $f(p)$.*

Proof We denote by d_X and d_Y the distance functions on X and Y respectively. Let $\varepsilon > 0$ be given. We must show that there exists some natural number N such that $d_Y(f(x_n), f(p)) < \varepsilon$ whenever $n \geq N$. However there exists some $\delta > 0$ such that $d_Y(f(x), f(p)) < \varepsilon$ for all $x \in X$ satisfying $d_X(x, p) < \delta$, since the function f is continuous at p. Also there exists some natural number N such that $d_X(x_n, p) < \delta$ whenever $n \geq N$, since the sequence $x_1, x_2, x_3, \ldots$ converges to p. Thus if $n \geq N$ then $d_Y(f(x_n), f(p)) < \varepsilon$, as required. ∎

3.4 Open Sets in Metric Spaces

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

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

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

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

$$d(x, p) \leq d(x, q) + d(q, p) < \delta + d(q, p) = r,$$

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

Lemma 3.5 *Let X be a metric space with distance function d, and let p be a point of X. Then, for any $r \geq 0$, the set $\{x \in X : d(x, p) > r\}$ is an open set in X.*

Proof Let q be a point of X satisfying $d(q, p) > r$, and let x be any point of X satisfying $d(x, q) < \delta$, where $\delta = d(q, p) - r$. Then

$$d(q, p) \leq d(q, x) + d(x, p),$$

by the Triangle Inequality, and therefore

$$d(x, p) \geq d(q, p) - d(x, q) > d(q, p) - \delta = r.$$

Thus $B_X(x, \delta) \subset \{x \in X : d(x, p) > r\}$, as required. ∎

Proposition 3.6 *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 δ 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. ∎

Lemma 3.7 *Let* X *be a metric space. A sequence* $x_1, x_2, x_3, \ldots$ *of points in* X *converges to a point* p *if and only if, given any open set* U *which contains* p*, there exists some natural number* N *such that* $x_j \in U$ *for all* $j \geq N$.

Proof Let $x_1, x_2, x_3, \ldots$ be a sequence satisfying the given criterion, and let $\varepsilon > 0$ be given. The open ball $B_X(p, \varepsilon)$ of radius ε about p is an open set (see Lemma 3.4). Therefore there exists some natural number N such that, if $j \geq N$, then $x_j \in B_X(p, \varepsilon)$, and thus $d(x_j, p) < \varepsilon$. Hence the sequence (x_j) converges to p.

Conversely, suppose that the sequence (x_j) converges to p. Let U be an open set which contains p. Then there exists some $\varepsilon > 0$ such that $B_X(p, \varepsilon) \subset U$. But $x_j \to p$ as $j \to +\infty$, and therefore there exists some natural number N such that $d(x_j, p) < \varepsilon$ for all $j \geq N$. If $j \geq N$ then $x_j \in B_X(p, \varepsilon)$ and thus $x_j \in U$, as required. ∎

Definition Let (X, d) be a metric space, and let x be a point of X. A subset N of X is said to be a *neighbourhood* of x (in X) if and only if there exists some $\delta > 0$ such that $B_X(x, \delta) \subset N$, where $B_X(x, \delta)$ is the open ball of radius δ about x.

It follows directly from the relevant definitions that a subset V of a metric space X is an open set if and only if V is a neighbourhood of v for all $v \in V$.

3.5 Closed Sets in a Metric Space

A subset F of a metric space X is said to be a *closed set* in X if and only if its complement $X \setminus F$ is open. (Recall that the *complement* $X \setminus F$ of F in X is, by definition, the set of all points of the metric space X that do not belong to F.) The following result follows immediately from Lemma 3.4 and Lemma 3.5.

Lemma 3.8 *Let X be a metric space with distance function d, and let $x_0 \in X$. Given any $r \geq 0$, the sets*

$$\{x \in X : d(x, x_0) \leq r\}, \qquad \{x \in X : d(x, x_0) \geq r\}$$

are closed. In particular, the set $\{x_0\}$ consisting of the single point x_0 is a closed set in X.

Let $\mathcal{A}$ be some collection of subsets of a set X. Then

$$X \setminus \bigcup_{S \in \mathcal{A}} S = \bigcap_{S \in \mathcal{A}} (X \setminus S), \qquad X \setminus \bigcap_{S \in \mathcal{A}} S = \bigcup_{S \in \mathcal{A}} (X \setminus S)$$

(i.e., the complement of the union of some collection of subsets of 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, so that the operation of taking complements converts unions into intersections and intersections into unions). The following result therefore follows directly from Proposition 3.6.

Proposition 3.9 *Let X be a metric space. The collection of closed sets in X has the following properties:—*

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

(ii) *the intersection of any collection of closed sets in X is itself a closed set;*

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

Lemma 3.10 *Let F be a closed set in a metric space X and let $(x_j : j \in \mathbb{N})$ be a sequence of points of F. Suppose that $x_j \to p$ as $j \to +\infty$. Then p also belongs to F.*

Proof Suppose that the limit p of the sequence were to belong to the complement $X \setminus F$ of the closed set F. Now $X \setminus F$ is open, and thus it would follow from Lemma 3.7 that there would exist some natural number N such that $x_j \in X \setminus F$ for all $j \geq N$, contradicting the fact that $x_j \in F$ for all j. This contradiction shows that p must belong to F, as required. ∎

Definition Let A be a subset of a metric space X. The *closure* $\overline{A}$ of A is the intersection of all closed subsets of X containing A.

Let A be a subset of the metric space X. Note that the closure $\overline{A}$ of A is itself a closed set in X, since the intersection of any collection of closed subsets of X is itself a closed subset of X (see Proposition 3.9). Moreover if F is any closed subset of X, and if $A \subset F$, then $\overline{A} \subset F$. Thus the closure $\overline{A}$ of A is the smallest closed subset of X containing A.

Lemma 3.11 *Let X be a metric space with distance function d, let A be a subset of X, and let x be a point of X. Then x belongs to the closure $\overline{A}$ of A if and only if, given any $\varepsilon > 0$, there exists some point a of A such that $d(x, a) < \varepsilon$.*

Proof Let x be a point of X with the property that, given any $\varepsilon > 0$, there exists some $a \in A$ satisfying $d(x, a) < \varepsilon$. Let F be any closed subset of X containing A. If x did not belong to F then there would exist some $\varepsilon > 0$ with the property that $B_X(x, \varepsilon) \cap F = \emptyset$, where $B_X(x, \varepsilon)$ denotes the open ball of radius ε about x. But this would contradict the fact that $B_X(x, \varepsilon) \cap A$ is non-empty for all $\varepsilon > 0$. Thus the point x belongs to every closed subset F of X that contains A, and therefore $x \in \overline{A}$, by definition of the closure $\overline{A}$ of A.

Conversely let $x \in \overline{A}$, and let $\varepsilon > 0$ be given. Let F be the complement $X \setminus B_X(x, \varepsilon)$ of $B_X(x, \varepsilon)$. Then F is a closed subset of X, and the point x does not belong to F. If $B_X(x, \varepsilon) \cap A = \emptyset$ then A would be contained in F, and hence $x \in F$, which is impossible. Therefore there exists $a \in A$ satisfying $d(x, a) < \varepsilon$, as required. ∎

3.6 Continuous Functions and Open and Closed Sets

Let X and Y be metric spaces, and let $f: X \to Y$ be a function from X to Y. We recall that the function f is continuous at a point p of X if and only if, given any $\varepsilon > 0$, there exists some $\delta > 0$ such that $d_Y(f(x), f(p)) < \varepsilon$ for all points x of X satisfying $d_X(x, p) < \delta$, where d_X and d_Y denote the distance functions on X and Y respectively. Expressed in terms of open balls, this means that the function $f: X \to Y$ is continuous at p if and only if, given any $\varepsilon > 0$, there exists some $\delta > 0$ such that f maps $B_X(p, \delta)$ into $B_Y(f(p), \varepsilon)$ (where $B_X(p, \delta)$ and $B_Y(f(p), \varepsilon)$ denote the open balls of radius δ and ε about p and $f(p)$ respectively).

Let $f: X \to Y$ 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

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

Proposition 3.12 *Let X and Y be metric spaces, and let $f: X \to Y$ be a function from X to Y. The function f is continuous if and only if $f^{-1}(V)$ is an open set in X for every open set V of Y.*

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

Conversely suppose that $f: X \to Y$ has the property that $f^{-1}(V)$ is open in X for every open set V in Y. Let p be any point of X. We must show that f is continuous at p. Let $\varepsilon > 0$ be given. The open ball $B_Y(f(p), \varepsilon)$ is an open set in Y, by Lemma 3.4, hence $f^{-1}(B_Y(f(p), \varepsilon))$ is an open set in X which contains p. It follows that there exists some $\delta > 0$ such that $B_X(p, \delta) \subset f^{-1}(B_Y(f(p), \varepsilon))$. We have thus shown that, given any $\varepsilon > 0$, there exists some $\delta > 0$ such that f maps the open ball $B_X(p, \delta)$ into $B_Y(f(p), \varepsilon)$. We conclude that f is continuous at p, as required. ∎

Let $f: X \to Y$ be a function between metric spaces X and Y. Then the preimage $f^{-1}(Y \setminus G)$ of the complement $Y \setminus G$ of any subset G of Y is equal to the complement $X \setminus f^{-1}(G)$ of the preimage $f^{-1}(G)$ of G. Indeed

$$x \in f^{-1}(Y \setminus G) \iff f(x) \in Y \setminus G \iff f(x) \notin G \iff x \notin f^{-1}(G).$$

Also a subset of a metric space is closed if and only if its complement is open. The following result therefore follows directly from Proposition 3.12.

Corollary 3.13 *Let X and Y be metric 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 a closed set in X for every closed set G in Y.*

Let $f: X \to Y$ be a continuous function from a metric space X to a metric space Y. Then, for any point y of Y, the set $\{x \in X : f(x) = y\}$ is a closed subset of X. This follows from Corollary 3.13, together with the fact that the set $\{y\}$ consisting of the single point y is a closed subset of the metric space Y.

Let X be a metric space, and let $f: X \to \mathbb{R}$ be a continuous function from X to $\mathbb{R}$. Then, given any real number c, the sets

$$\{x \in X : f(x) > c\}, \qquad \{x \in X : f(x) < c\}$$

are open subsets of X, and the sets

$$\{x \in X : f(x) \geq c\}, \qquad \{x \in X : f(x) \leq c\}, \qquad \{x \in X : f(x) = c\}$$

are closed subsets of X. Also, given real numbers a and b satisfying $a < b$, the set

$$\{x \in X : a < f(x) < b\}$$

is an open subset of X, and the set

$$\{x \in X : a \leq f(x) \leq b\}$$

is a closed subset of X.

Similar results hold for continuous functions $f: X \to \mathbb{C}$ from X to $\mathbb{C}$. Thus, for example,

$$\{x \in X : |f(x)| < R\}, \qquad \{x \in X : |f(x)| > R\}$$

are open subsets of X and

$$\{x \in X : |f(x)| \leq R\}, \qquad \{x \in X : |f(x)| \geq R\}, \qquad \{x \in X : |f(x)| = R\}$$

are closed subsets of X, for any non-negative real number R.

3.7 Homeomorphisms

Let X and Y be metric spaces. A function $h: X \to Y$ from X to Y is said to be a *homeomorphism* if it is a bijection and both $h: X \to Y$ and its inverse $h^{-1}: Y \to X$ are continuous. If there exists a homeomorphism $h: X \to Y$ from a metric space X to a metric space Y, then the metric spaces X and Y are said to be *homeomorphic*.

The following result follows directly on applying Proposition 3.12 to $h: X \to Y$ and to $h^{-1}: Y \to X$.

Lemma 3.14 *Any homeomorphism $h: X \to Y$ between metric spaces X and Y induces a one-to-one correspondence between the open sets of X and the open sets of Y: a subset V of Y is open in Y if and only if $h^{-1}(V)$ is open in X.*

Let X and Y be metric spaces, and let $h: X \to Y$ be a homeomorphism. A sequence $x_1, x_2, x_3, \ldots$ of points in X is convergent in X if and only if the corresponding sequence $h(x_1), h(x_2), h(x_3), \ldots$ is convergent in Y. (This follows directly on applying Lemma 3.3 to $h: X \to Y$ and its inverse $h^{-1}: Y \to X$.) Let Z and W be metric spaces. A function $f: Z \to X$ is continuous if and only if $h \circ f: Z \to Y$ is continuous, and a function $g: Y \to W$ is continuous if and only if $g \circ h: X \to W$ is continuous.

4 Complete Metric Spaces, Normed Vector Spaces and Banach Spaces

4.1 Complete Metric Spaces

Definition Let X be a metric space with distance function d. A sequence $x_1, x_2, x_3, \ldots$ of points of X is said to be a *Cauchy sequence* in X if and only if, given any $\varepsilon > 0$, there exists some positive integer N such that $d(x_j, x_k) < \varepsilon$ for all j and k satisfying $j \geq N$ and $k \geq N$.

Every convergent sequence in a metric space is a Cauchy sequence. Indeed let X be a metric space with distance function d, and let $x_1, x_2, x_3, \ldots$ be a sequence of points in X which converges to some point p of X. Given any positive real number ε, there exists some positive integer N such that $d(x_n, p) < \varepsilon/2$ whenever $n \geq N$. But then it follows from the Triangle Inequality that

$$d(x_j, x_k) \leq d(x_j, p) + d(p, x_k) < \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon$$

whenever $j \geq N$ and $k \geq N$.

Definition A metric space (X, d) is said to be *complete* if every Cauchy sequence in X converges to some point of X.

It follows immediately from Theorem 2.5 that n-dimensional Euclidean space $\mathbb{R}^n$ is a complete metric space. In particular, the set $\mathbb{R}$ of real numbers is a complete metric space.

Example The space $\mathbb{Q}$ of rational numbers (with distance function $d(q, r) = |q - r|$) is not complete. Indeed one can construct an infinite sequence $q_1, q_2, q_3, \ldots$ of rational numbers which converges (in $\mathbb{R}$) to $\sqrt{2}$. Such a sequence of rational numbers is a Cauchy sequence in both $\mathbb{R}$ and $\mathbb{Q}$. However this Cauchy sequence does not converge to an point of the metric space $\mathbb{Q}$ (since $\sqrt{2}$ is an irrational number). Thus the metric space $\mathbb{Q}$ is not complete.

Lemma 4.1 *Let X be a complete metric space, and let A be a subset of X. Then A is complete if and only if A is closed in X.*

Proof Suppose that A is closed in X. Let $a_1, a_2, a_3, \ldots$ be a Cauchy sequence in A. This Cauchy sequence must converge to some point p of X, since X is complete. But the limit of every sequence of points of A must belong to A, since A is closed. In particular $p \in A$. We deduce that A is complete.

Conversely, suppose that A is complete. Suppose that A were not closed. Then the complement $X \setminus A$ of A would not be open, and therefore there would exist a point p of $X \setminus A$ with the property that $B_X(p, \delta) \cap A$ is non-empty for all $\delta > 0$, where $B_X(p, \delta)$ denotes the open ball in X of radius δ centred at p. We could then find a sequence $a_1, a_2, a_3, \ldots$ of points of A satisfying $d(a_j, p) < 1/j$ for all positive integers j. This sequence would be a Cauchy sequence in A which did not converge to a point of A, contradicting the completeness of A. Thus A must be closed, as required. ∎

The following result follows directly from Theorem 2.5 and Lemma 4.1.

Corollary 4.2 *A subset X of $\mathbb{R}^n$ is complete if and only if it is closed.*

Example The n-sphere S^n (with the chordal distance function given by $d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$) is a complete metric space, where

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

4.2 Normed Vector Spaces

A set X is a *vector space* over some field $\mathbb{F}$ if

- given any $x, y \in X$ and $\lambda \in \mathbb{F}$, there are well-defined elements $x + y$ and λx of X,
- X is an Abelian group with respect to the operation $+$ of addition,
- the identities

  $$\lambda(x + y) = \lambda x + \lambda y, \qquad (\lambda + \mu)x = \lambda x + \mu x,$$

  $$(\lambda\mu)x = \lambda(\mu x), \qquad 1x = x$$

  are satisfied for all $x, y \in X$ and $\lambda, \mu \in \mathbb{F}$.

Elements of the field $\mathbb{F}$ are referred to as *scalars*. We consider here only *real vector spaces* and *complex vector spaces*: these are vector spaces over the fields of real numbers and complex numbers respectively.

Definition A *norm* $\|.\|$ on a real or complex vector space X is a function, associating to each element x of X a corresponding real number $\|x\|$, such that the following conditions are satisfied:—

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

(ii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$,

(iii) $\|\lambda x\| = |\lambda| \, \|x\|$ for all $x \in X$ and for all scalars λ,

(iv) $\|x\| = 0$ if and only if $x = 0$.

A *normed vector space* $(X, \|.\|)$ consists of a a real or complex vector space X, together with a norm $\|.\|$ on X.

Note that any normed complex vector space can also be regarded as a normed real vector space.

Example The field $\mathbb{R}$ is a one-dimensional normed vector space over itself: the norm $|t|$ of $t \in \mathbb{R}$ is the absolute value of t.

Example The field $\mathbb{C}$ is a one-dimensional normed vector space over itself: the norm $|z|$ of $z \in \mathbb{C}$ is the modulus of z. The field $\mathbb{C}$ is also a two-dimensional normed vector space over $\mathbb{R}$.

Example Let $\|.\|_1$, $\|.\|_2$ and $\|.\|_\infty$ be the real-valued functions on $\mathbb{C}^n$ defined by

$$\begin{aligned}
\|\mathbf{z}\|_1 &= \sum_{j=1}^{n} |z_j|, \\
\|\mathbf{z}\|_2 &= \left(\sum_{j=1}^{n} |z_j|^2 \right)^{\frac{1}{2}}, \\
\|\mathbf{z}\|_\infty &= \max(|z_1|, |z_2|, \ldots, |z_n|),
\end{aligned}$$

for each $\mathbf{z} \in \mathbb{C}^n$, where $\mathbf{z} = (z_1, z_2, \ldots, z_n)$. Then $\|.\|_1$, $\|.\|_2$ and $\|.\|_\infty$ are norms on $\mathbb{C}^n$. In particular, if we regard $\mathbb{C}^n$ as a $2n$-dimensional real vector space naturally isomorphic to $\mathbb{R}^{2n}$ (via the isomorphism

$$(z_1, z_2, \ldots, z_n) \mapsto (x_1, y_1, x_2, y_2, \ldots, x_n, y_n),$$

where x_j and y_j are the real and imaginary parts of z_j for $j = 1, 2, \ldots, n$) then $\|.\|_2$ represents the Euclidean norm on this space. The inequality $\|\mathbf{z}+\mathbf{w}\|_2 \leq \|\mathbf{z}\|_2 + \|\mathbf{w}\|_2$ satisfied for all $\mathbf{z}, \mathbf{w} \in \mathbb{C}^n$ is therefore just the standard Triangle Inequality for the Euclidean norm.

Example The space $\mathbb{R}^n$ is also an n-dimensional real normed vector space with respect to the norms $\|.\|_1$, $\|.\|_2$ and $\|.\|_\infty$ defined above. Note that $\|.\|_2$ is the standard Euclidean norm on $\mathbb{R}^n$.

Example Let

$$
\begin{aligned}
\ell_1 &= \{(z_1, z_2, z_3, \ldots) \in \mathbb{C}^\infty : |z_1| + |z_2| + |z_3| + \cdots \text{ converges}\}, \\
\ell_2 &= \{(z_1, z_2, z_3, \ldots) \in \mathbb{C}^\infty : |z_1|^2 + |z_2|^2 + |z_3|^2 + \cdots \text{ converges}\}, \\
\ell_\infty &= \{(z_1, z_2, z_3, \ldots) \in \mathbb{C}^\infty : \text{the sequence } |z_1|, |z_2|, |z_3|, \ldots \text{ is bounded}\}.
\end{aligned}
$$

where $\mathbb{C}^\infty$ denotes the set of all sequences $(z_1, z_2, z_3, \ldots)$ of complex numbers. Then ℓ_1, ℓ_2 and ℓ_∞ are infinite-dimensional normed vector spaces, with norms $\|.\|_1$, $\|.\|_2$ and $\|.\|_\infty$ respectively, where

$$
\begin{aligned}
\|\mathbf{z}\|_1 &= \sum_{j=1}^{+\infty} |z_j|, \\
\|\mathbf{z}\|_2 &= \left(\sum_{j=1}^{+\infty} |z_j|^2 \right)^{\frac{1}{2}}, \\
\|\mathbf{z}\|_\infty &= \sup\{|z_1|, |z_2|, |z_3|, \ldots\}.
\end{aligned}
$$

(For example, to show that $\|\mathbf{z} + \mathbf{w}\|_2 \leq \|\mathbf{z}\|_2 + \|\mathbf{w}\|_2$ for all $\mathbf{z}, \mathbf{w} \in \ell_2$, we note that

$$
\left(\sum_{j=1}^{n} |z_j + w_j|^2 \right)^{\frac{1}{2}} \leq \left(\sum_{j=1}^{n} |z_j|^2 \right)^{\frac{1}{2}} + \left(\sum_{j=1}^{n} |w_j|^2 \right)^{\frac{1}{2}} \leq \|\mathbf{z}\|_2 + \|\mathbf{w}\|_2
$$

for all positive integers n, by the Triangle Inequality in $\mathbb{C}^n$. Taking limits as $n \to +\infty$, we deduce that $\|\mathbf{z} + \mathbf{w}\|_2 \leq \|\mathbf{z}\|_2 + \|\mathbf{w}\|_2$, as required.)

If $x_1, x_2, \ldots, x_m$ are elements of a normed vector space X then

$$
\left\| \sum_{k=1}^{m} x_k \right\| \leq \sum_{k=1}^{m} \|x_k\|,
$$

where $\|.\|$ denotes the norm on X. (This can be verified by induction on m, using the inequality $\|x + y\| \leq \|x\| + \|y\|$.)

A norm $\|.\|$ on a vector space X induces a corresponding distance function on X: the distance $d(x, y)$ between elements x and y of X is defined by $d(x, y) = \|x - y\|$. This distance function satisfies the metric space axioms. Thus any vector space with a given norm can be regarded as a metric space.

Lemma 4.3 *Let X be a normed vector space over the field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let (x_j) and (y_j) be convergent sequences in X, and let (λ_j) be* *a convergent sequence in $\mathbb{F}$. Then the sequences $(x_j + y_j)$ and $(\lambda_j x_j)$ are convergent in X, and*

$$\begin{aligned} \lim_{j\to+\infty}(x_j + y_j) &= \lim_{j\to+\infty} x_j + \lim_{j\to+\infty} y_j, \\ \lim_{j\to+\infty}(\lambda_j x_j) &= \left(\lim_{j\to+\infty} \lambda_j\right)\left(\lim_{j\to+\infty} x_j\right). \end{aligned}$$

Proof First we prove that $\lim_{j\to+\infty}(x_j + y_j) = x + y$, where Let $x = \lim_{j\to+\infty} x_j$, $y = \lim_{j\to+\infty} y_j$. Let $\varepsilon > 0$ be given. Then there exist natural numbers N_1 and N_2 such that $\|x_j - x\| < \frac{1}{2}\varepsilon$ whenever $j \geq N_1$ and $\|y_j - y\| < \frac{1}{2}\varepsilon$ whenever $j \geq N_2$. Let N be the maximum of N_1 and N_2. If $j \geq N$ then

$$\|(x_j + y_j) - (x + y)\| \leq \|x_j - x\| + \|y_j - y\| < \varepsilon.$$

It follows from this that $\lim_{j\to+\infty}(x_j + y_j) = x + y$.

Next we prove that $\lim_{j\to+\infty}(\lambda_j x_j) = \lambda x$, where $\lambda = \lim_{j\to+\infty} \lambda_j$. Let $\varepsilon > 0$ be given. Then there exist natural numbers N_3 and N_4 such that

$$\|x_j - x\| < \frac{\varepsilon}{2(|\lambda| + 1)}$$

whenever $j \geq N_3$, and

$$|\lambda_j - \lambda| < \frac{\varepsilon}{2(\|x\| + 1)} \text{ and } |\lambda_j - \lambda| < 1$$

whenever $j \geq N_4$. Let N be the maximum of N_3 and N_4. if $j \geq N$ then

$$\begin{aligned} \|\lambda_j x_j - \lambda x\| &= \|\lambda_j(x_j - x) + (\lambda_j - \lambda)x\| \leq |\lambda_j|\,\|x_j - x\| + |\lambda_j - \lambda|\|x\| \\ &\leq (|\lambda| + 1)\,\|x_j - x\| + |\lambda_j - \lambda|\|x\| < \varepsilon. \end{aligned}$$

It follows from this that $\lim_{j\to+\infty}(\lambda_j x_j) = \lambda x$, as required. ∎

Let X be a normed vector space, and let $x_1, x_2, x_3, \ldots$ be elements of X. The infinite series $\sum_{n=1}^{+\infty} x_n$ is said to *converge* to some element s of X if, given any positive real number ε, there exists some positive integer N such that

$$\|s - \sum_{n=1}^{m} x_n\| < \varepsilon$$

for all $m \geq N$ (where $\|.\|$ denotes the norm on X).

We say that a normed vector space X is *complete* A normed vector space is complete if and only if every Cauchy sequence in X is convergent. A complete normed vector space is referred to as a *Banach space*. (The basic theory of such spaces was extensively developed by the famous Polish mathematician Stefan Banach and his colleagues.)

Lemma 4.4 *Let X be a Banach space, and let $x_1, x_2, x_3, \ldots$ be elements of X. Suppose that $\sum_{n=1}^{+\infty} \|x_n\|$ is convergent. Then $\sum_{n=1}^{+\infty} x_n$ is convergent, and*

$$\left\| \sum_{n=1}^{+\infty} x_n \right\| \leq \sum_{n=1}^{+\infty} \|x_n\|.$$

Proof For each positive integer n, let

$$s_n = x_1 + x_2 + \cdots + x_n.$$

Let $\varepsilon > 0$ be given. We can find N such that $\sum_{n=N}^{+\infty} \|x_n\| < \varepsilon$, since $\sum_{n=1}^{+\infty} \|x_n\|$ is convergent. Let $s_n = x_1 + x_2 + \cdots + x_n$. If $j \geq N$, $k \geq N$ and $j < k$ then

$$\|s_k - s_j\| = \left\| \sum_{n=j+1}^{k} x_n \right\| \leq \sum_{n=j+1}^{k} \|x_n\| \leq \sum_{n=N}^{+\infty} \|x_n\| < \varepsilon.$$

Thus $s_1, s_2, s_3, \ldots$ is a Cauchy sequence in X, and therefore converges to some element s of X, since X is complete. But then $s = \sum_{j=1}^{+\infty} x_j$. Moreover, on choosing m large enough to ensure that $\|s - s_m\| < \varepsilon$, we deduce that

$$\|s\| \leq \left\| \sum_{n=1}^{m} x_n \right\| + \left\| s - \sum_{n=1}^{m} x_n \right\| \leq \sum_{n=1}^{m} \|x_n\| + \left\| s - \sum_{n=1}^{m} x_n \right\| < \sum_{n=1}^{+\infty} \|x_n\| + \varepsilon.$$

Since this inequality holds for all $\varepsilon > 0$, we conclude that

$$\|s\| \leq \sum_{n=1}^{+\infty} \|x_n\|,$$

as required. ∎

4.3 Bounded Linear Transformations

Let X and Y be real or complex vector spaces. A function $T: X \to Y$ is said to be a *linear transformation* if $T(x + y) = Tx + Ty$ and $T(\lambda x) = \lambda Tx$ for all elements x and y of X and scalars λ. A linear transformation mapping X into itself is referred to as a *linear operator* on X.

Definition Let X and Y be normed vector spaces. A linear transformation $T: X \to Y$ is said to be *bounded* if there exists some non-negative real number C with the property that $\|Tx\| \leq C\|x\|$ for all $x \in X$. If T is bounded, then the smallest non-negative real number C with this property is referred to as the *operator norm* of T, and is denoted by $\|T\|$.

Lemma 4.5 *Let X and Y be normed vector spaces, and let $S: X \to Y$ and $T: X \to Y$ be bounded linear transformations. Then $S + T$ and λS are bounded linear transformations for all scalars λ, and*

$$\|S + T\| \leq \|S\| + \|T\|, \qquad \|\lambda S\| = |\lambda|\|S\|.$$

Moreover $\|S\| = 0$ if and only if $S = 0$. Thus the vector space $B(X, Y)$ of bounded linear transformations from X to Y is a normed vector space (with respect to the operator norm).

Proof $\|(S+T)x\| \leq \|Sx\|+\|Tx\| \leq (\|S\|+\|T\|)\|x\|$ for all $x \in X$. Therefore $S+T$ is bounded, and $\|S+T\| \leq \|S\|+\|T\|$. Using the fact that $\|(\lambda S)x\| = |\lambda|\,\|Sx\|$ for all $x \in X$, we see that λS is bounded, and $\|\lambda S\| = |\lambda|\,\|S\|$. If $S = 0$ then $\|S\| = 0$. Conversely if $\|S\| = 0$ then $\|Sx\| \leq \|S\|\,\|x\| = 0$ for all $x \in X$, and hence $S = 0$. The result follows. ∎

Lemma 4.6 *Let X, Y and Z be normed vector spaces, and let $S: X \to Y$ and $T: Y \to Z$ be bounded linear transformations. Then the composition TS of S and T is also bounded, and $\|TS\| \leq \|T\|\,\|S\|$.*

Proof $\|TSx\| \leq \|T\|\,\|Sx\| \leq \|T\|\,\|S\|\,\|x\|$ for all $x \in X$. The result follows. ∎

Proposition 4.7 *Let X and Y be normed vector spaces, and let $T: X \to Y$ be a linear transformation from X to Y. Then the following conditions are equivalent:—*

(i) *$T: X \to Y$ is continuous,*

(ii) *$T: X \to Y$ is continuous at 0,*

(iii) $T: X \to Y$ *is bounded.*

Proof Obviously (i) implies (ii). We show that (ii) implies (iii) and (iii) implies (i). The equivalence of the three conditions then follows immediately.

Suppose that $T: X \to Y$ is continuous at 0. Then there exists $\delta > 0$ such that $\|Tx\| < 1$ for all $x \in X$ satisfying $\|x\| < \delta$. Let C be any positive real number satisfying $C > 1/\delta$. If x is any non-zero element of X then $\|\lambda x\| < \delta$, where $\lambda = 1/(C\|x\|)$, and hence

$$\|Tx\| = C\|x\| \, \|\lambda Tx\| = C\|x\| \, \|T(\lambda x)\| < C\|x\|.$$

Thus $\|Tx\| \leq C\|x\|$ for all $x \in X$, and hence $T: X \to Y$ is bounded. Thus (ii) implies (iii).

Finally suppose that $T: X \to Y$ is bounded. Let x be a point of X, and let $\varepsilon > 0$ be given. Choose $\delta > 0$ satisfying $\|T\|\delta < \varepsilon$. If $x' \in X$ satisfies $\|x' - x\| < \delta$ then

$$\|Tx' - Tx\| = \|T(x' - x)\| \leq \|T\| \, \|x' - x\| < \|T\|\delta < \varepsilon.$$

Thus $T: X \to Y$ is continuous. Thus (iii) implies (i), as required. ∎

Proposition 4.8 *Let X be a normed vector space and let Y be a Banach space. Then the space $B(X, Y)$ of bounded linear transformations from X to Y is also a Banach space.*

Proof We have already shown that $B(X, Y)$ is a normed vector space (see Lemma 4.5). Thus it only remains to show that $B(X, Y)$ is complete.

Let $S_1, S_2, S_3, \ldots$ be a Cauchy sequence in $B(X, Y)$. Let $x \in X$. We claim that $S_1x, S_2x, S_3x, \ldots$ is a Cauchy sequence in Y. This result is trivial if $x = 0$. If $x \neq 0$, and if $\varepsilon > 0$ is given then there exists some positive integer N such that $\|S_j - S_k\| < \varepsilon/\|x\|$ whenever $j \geq N$ and $k \geq N$. But then $\|S_jx - S_kx\| \leq \|S_j - S_k\| \, \|x\| < \varepsilon$ whenever $j \geq N$ and $k \geq N$. This shows that $S_1x, S_2x, S_3x, \ldots$ is indeed a Cauchy sequence. It therefore converges to some element of Y, since Y is a Banach space.

Let the function $S: X \to Y$ be defined by $Sx = \lim_{n \to +\infty} S_nx$. Then

$$S(x + y) = \lim_{n \to +\infty} (S_nx + S_ny) = \lim_{n \to +\infty} S_nx + \lim_{n \to +\infty} S_ny = Sx + Sy,$$

(see Lemma 4.3), and

$$S(\lambda x) = \lim_{n \to +\infty} S_n(\lambda x) = \lambda \lim_{n \to +\infty} S_nx = \lambda Sx,$$

Thus $S: X \to Y$ is a linear transformation.

Next we show that $S_n \to S$ in $B(X,Y)$ as $n \to +\infty$. Let $\varepsilon > 0$ be given. Then there exists some positive integer N such that $\|S_j - S_n\| < \frac{1}{2}\varepsilon$ whenever $j \geq N$ and $n \geq N$, since the sequence $S_1, S_2, S_3, \ldots$ is a Cauchy sequence in $B(X,Y)$. But then $\|S_j x - S_n x\| \leq \frac{1}{2}\varepsilon \|x\|$ for all $j \geq N$ and $n \geq N$, and thus

$$\begin{aligned} \|Sx - S_n x\| &= \left\| \lim_{j \to +\infty} (S_j x - S_n x) \right\| \leq \lim_{j \to +\infty} \|S_j x - S_n x\| \\ &\leq \lim_{j \to +\infty} \|S_j - S_n\| \, \|x\| \leq \tfrac{1}{2}\varepsilon \|x\| \end{aligned}$$

for all $n \geq N$ (since the norm is a continuous function on Y). But then

$$\|Sx\| \leq \|S_n x\| + \|Sx - S_n x\| \leq \left(\|S_n\| + \tfrac{1}{2}\varepsilon \right) \|x\|$$

for any $n \geq N$, showing that $S\colon X \to Y$ is a bounded linear transformation, and $\|S - S_n\| \leq \frac{1}{2}\varepsilon < \varepsilon$ for all $n \geq N$, showing that $S_n \to S$ in $B(X,Y)$ as $n \to +\infty$. Thus the Cauchy sequence $S_1, S_2, S_3, \ldots$ is convergent in $B(X,Y)$, as required. ∎

Corollary 4.9 *Let X and Y be Banach spaces, and let $T_1, T_2, T_3, \ldots$ be bounded linear transformations from X to Y. Suppose that $\sum_{n=0}^{+\infty} \|T_n\|$ is convergent. Then $\sum_{n=0}^{+\infty} T_n$ is convergent, and*

$$\left\| \sum_{n=0}^{+\infty} T_n \right\| \leq \sum_{n=0}^{+\infty} \|T_n\|.$$

Proof The space $B(X,Y)$ of bounded linear maps from X to Y is a Banach space by Proposition 4.8. The result therefore follows immediately on applying Lemma 4.4. ∎

Example Let T be a bounded linear operator on a Banach space X (i.e., a bounded linear transformation from X to itself). The infinite series

$$\sum_{n=0}^{+\infty} \frac{\|T\|^n}{n!}$$

converges to $\exp(\|T\|)$. It follows immediately from Lemma 4.6 (using induction on n) that $\|T^n\| \leq \|T\|^n$ for all $n \geq 0$ (where T^0 is the identity operator on X). It therefore follows from Corollary 4.9 that there is a well-defined bounded linear operator $\exp T$ on X, defined by

$$\exp T = \sum_{n=0}^{+\infty} \frac{1}{n!} T^n$$

(where T^0 is the identity operator I on X).

Proposition 4.10 *Let T be a bounded linear operator on a Banach space X. Suppose that $\|T\| < 1$. Then the operator $I - T$ has a bounded inverse $(I-T)^{-1}$ (where I denotes the identity operator on X). Moreover*

$$(I-T)^{-1} = I + T + T^2 + T^3 + \cdots.$$

Proof $\|T^n\| \leq \|T\|^n$ for all n, and the geometric series

$$1 + \|T\| + \|T\|^2 + \|T\|^3 + \cdots$$

is convergent (since $\|T\| < 1$). It follows from Corollary 4.9 that the infinite series

$$I + T + T^2 + T^3 + \cdots$$

converges to some bounded linear operator S on X. Now

$$\begin{aligned} (I-T)S &= \lim_{n\to+\infty} (I-T)(I+T+T^2+\cdots+T^n) = \lim_{n\to+\infty} (I - T^{n+1}) \\ &= I - \lim_{n\to+\infty} T^{n+1} = I, \end{aligned}$$

since $\|T\|^{n+1} \to 0$ and therefore $T^{n+1} \to 0$ as $n \to +\infty$. Similarly $S(I-T) = I$. This shows that $I - T$ is invertible, with inverse S, as required. ∎

4.4 Spaces of Bounded Continuous Functions

Let X be a metric space. We say that a function $f: X \to \mathbb{R}^n$ from X to $\mathbb{R}^n$ is *bounded* if there exists some non-negative constant K such that $|f(x)| \leq K$ for all $x \in X$. If f and g are bounded continuous functions from X to $\mathbb{R}^n$, then so is $f + g$. Also λf is bounded and continuous for any real number λ. It follows from this that the space $C(X, \mathbb{R}^n)$ of bounded continuous functions from X to $\mathbb{R}^n$ is a vector space over $\mathbb{R}$. Given $f \in C(X, \mathbb{R}^n)$, we define the *supremum norm* $\|f\|$ of f by the formula

$$\|f\| = \sup_{x\in X} |f(x)|.$$

One can readily verify that $\|.\|$ is a norm on the vector space $C(X, \mathbb{R}^n)$. We shall show that $C(X, \mathbb{R}^n)$, with the supremum norm, is a Banach space (i.e., the supremum norm on $C(X, \mathbb{R}^n)$ is complete). The proof of this result will make use of the following characterization of continuity for functions whose range is $\mathbb{R}^n$.

Theorem 4.11 *The normed vector space $C(X, \mathbb{R}^n)$ of all bounded continuous functions from some metric space X to $\mathbb{R}^n$, with the supremum norm, is a Banach space.*

Proof Let $f_1, f_2, f_3, \ldots$ be a Cauchy sequence in $C(X, \mathbb{R}^n)$. Then, for each $x \in X$, the sequence $f_1(x), f_2(x), f_3(x), \ldots$ is a Cauchy sequence in $\mathbb{R}^n$ (since $|f_j(x) - f_k(x)| \leq \|f_j - f_k\|$ for all positive integers j and k), and $\mathbb{R}^n$ is a complete metric space. Thus, for each $x \in X$, the sequence $f_1(x), f_2(x), f_3(x), \ldots$ converges to some point $f(x)$ of $\mathbb{R}^n$. We must show that the limit function f defined in this way is bounded and continuous.

Let $\varepsilon > 0$ be given. Then there exists some positive integer N with the property that $\|f_j - f_k\| < \frac{1}{3}\varepsilon$ for all $j \geq N$ and $k \geq N$, since $f_1, f_2, f_3, \ldots$ is a Cauchy sequence in $C(X, \mathbb{R}^n)$. But then, on taking the limit of the left hand side of the inequality $|f_j(x) - f_k(x)| < \frac{1}{3}\varepsilon$ as $k \to +\infty$, we deduce that $|f_j(x) - f(x)| \leq \frac{1}{3}\varepsilon$ for all $x \in X$ and $j \geq N$. In particular $|f_N(x) - f(x)| \leq \frac{1}{3}\varepsilon$ for all $x \in X$. It follows that $|f(x)| \leq \|f_N\| + \frac{1}{3}\varepsilon$ for all $x \in X$, showing that the limit function f is bounded.

Next we show that the limit function f is continuous. Let $p \in X$ and $\varepsilon > 0$ be given. Let N be chosen large enough to ensure that $|f_N(x) - f(x)| \leq \frac{1}{3}\varepsilon$ for all $x \in X$. Now f_N is continuous. It follows from the definition of continuity for functions between metric spaces that there exists some real number δ satisfying $\delta > 0$ such that $|f_N(x) - f_N(p)| < \frac{1}{3}\varepsilon$ for all elements x of X satisfying $d_X(x, p) < \delta$, where d_X denotes the distance function on X. Thus if $x \in X$ satisfies $d_X(x, p) < \delta$ then

$$\begin{aligned} |f(x) - f(p)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(p)| + |f_N(p) - f(p)| \\ &< \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon = \varepsilon. \end{aligned}$$

Therefore the limit function f is continuous. Thus $f \in C(X, \mathbb{R}^n)$.

Finally we observe that $f_j \to f$ in $C(X, \mathbb{R}^n)$ as $j \to +\infty$. Indeed we have already seen that, given $\varepsilon > 0$ there exists some positive integer N such that $|f_j(x) - f(x)| \leq \frac{1}{3}\varepsilon$ for all $x \in X$ and for all $j \geq N$. Thus $\|f_j - f\| \leq \frac{1}{3}\varepsilon < \varepsilon$ for all $j \geq N$, showing that $f_j \to f$ in $C(X, \mathbb{R}^n)$ as $j \to +\infty$. This shows that $C(X, \mathbb{R}^n)$ is a complete metric space, as required. ∎

Corollary 4.12 *Let X be a metric space and let F be a closed subset of $\mathbb{R}^n$. Then the space $C(X, F)$ of bounded continuous functions from X to F is a complete metric space with respect to the distance function ρ, where*

$$\rho(f, g) = \|f - g\| = \sup_{x \in X} |f(x) - g(x)|$$

for all $f, g \in C(X, F)$.

Proof Let $f_1, f_2, f_3, \ldots$ be a Cauchy sequence in $C(X, F)$. Then $f_1, f_2, f_3, \ldots$ is a Cauchy sequence in $C(X, \mathbb{R}^n)$ and therefore converges in $C(X, \mathbb{R}^n)$ to some function $f: X \to \mathbb{R}^n$. Let x be some point of X. Then $f_j(x) \to f(x)$ as $j \to +\infty$. But then $f(x) \in F$, since $f_j(x) \in F$ for all j, and F is closed in $\mathbb{R}^n$. This shows that $f \in C(X, F)$, and thus the Cauchy sequence $f_1, f_2, f_3, \ldots$ converges in $C(X, F)$. We conclude that $C(X, F)$ is a complete metric space, as required. ∎

4.5 The Contraction Mapping Theorem and Picard's Theorem

Let X be a metric space with distance function d. A function $T: X \to X$ mapping X to itself is said to be a *contraction mapping* if there exists some constant λ satisfying $0 \leq \lambda < 1$ with the property that $d(T(x), T(x')) \leq \lambda d(x, x')$ for all $x, x' \in X$.

One can readily check that any contraction map $T: X \to X$ on a metric space (X, d) is continuous. Indeed let x be a point of X, and let $\varepsilon > 0$ be given. Then $d(T(x), T(x')) < \varepsilon$ for all points x' of X satisfying $d(x, x') < \varepsilon$.

Theorem 4.13 (Contraction Mapping Theorem) *Let X be a complete metric space, and let $T: X \to X$ be a contraction mapping defined on X. Then T has a unique fixed point in X (i.e., there exists a unique point x of X for which $T(x) = x$).*

Proof Let λ be chosen such that $0 \leq \lambda < 1$ and $d(T(u), T(u')) \leq \lambda d(u, u')$ for all $u, u' \in X$, where d is the distance function on X. First we show the existence of the fixed point x. Let x_0 be any point of X, and define a sequence $x_0, x_1, x_2, x_3, x_4, \ldots$ of points of X by the condition that $x_n = T(x_{n-1})$ for all positive integers n. It follows by induction on n that $d(x_{n+1}, x_n) \leq \lambda^n d(x_1, x_0)$. Using the Triangle Inequality, we deduce that if j and k are positive integers satisfying $k > j$ then

$$d(x_k, x_j) \leq \sum_{n=j}^{k-1} d(x_{n+1}, x_n) \leq \frac{\lambda^j - \lambda^k}{1 - \lambda} d(x_1, x_0) \leq \frac{\lambda^j}{1 - \lambda} d(x_1, x_0).$$

(Here we have used the identity

$$\lambda^j + \lambda^{j+1} + \cdots + \lambda^{k-1} = \frac{\lambda^j - \lambda^k}{1 - \lambda}.)$$

Using the fact that $0 \leq \lambda < 1$, we deduce that the sequence (x_n) is a Cauchy sequence in X. This Cauchy sequence must converge to some point x of X, since X is complete. But then we see that

$$T(x) = T\left(\lim_{n \to +\infty} x_n\right) = \lim_{n \to +\infty} T(x_n) = \lim_{n \to +\infty} x_{n+1} = x,$$

since $T: X \to X$ is a continuous function, and thus x is a fixed point of T.

If x' were another fixed point of T then we would have

$$d(x', x) = d(T(x'), T(x)) \leq \lambda d(x', x).$$

But this is impossible unless $x' = x$, since $\lambda < 1$. Thus the fixed point x of the contraction map T is unique. ∎

We use the Contraction Mapping Theorem in order to prove the following existence theorem for solutions of ordinary differential equations.

Theorem 4.14 (Picard's Theorem) *Let $F: U \to \mathbb{R}$ be a continuous function defined over some open set U in the plane $\mathbb{R}^2$, and let (x_0, t_0) be an element of U. Suppose that there exists some non-negative constant M such that*

$$|F(u,t) - F(v,t)| \leq M|u - v| \text{ for all } (u,t) \in U \text{ and } (v,t) \in U.$$

Then there exists a continuous function $\varphi: [t_0 - \delta, t_0 + \delta] \to \mathbb{R}$ defined on the interval $[t_0 - \delta, t_0 + \delta]$ for some $\delta > 0$ such that $x = \varphi(t)$ is a solution to the differential equation

$$\frac{dx(t)}{dt} = F(x(t), t)$$

with initial condition $x(t_0) = x_0$.

Proof Solving the differential equation with the initial condition $x(t_0) = x_0$ is equivalent to finding a continuous function $\varphi: I \to \mathbb{R}$ satisfying the integral equation

$$\varphi(t) = x_0 + \int_{t_0}^{t} F(\varphi(s), s)\, ds.$$

where I denotes the closed interval $[t_0 - \delta, t_0 + \delta]$. (Note that any continuous function φ satisfying this integral equation is automatically differentiable, since the indefinite integral of a continuous function is always differentiable.)

Let $K = |F(x_0, t_0)| + 1$. Using the continuity of the function F, together with the fact that U is open in $\mathbb{R}^2$, one can find some $\delta_0 > 0$ such that the open disk of radius δ_0 about (x_0, t_0) is contained in U and $|F(x,t)| \leq K$ for all points (x,t) in this open disk. Now choose $\delta > 0$ such that

$$\delta\sqrt{1 + K^2} < \delta_0 \text{ and } M\delta < 1.$$

Note that if $|t - t_0| \leq \delta$ and $|x - x_0| \leq K\delta$ then (x,t) belongs to the open disk of radius δ_0 about (x_0, t_0), and hence $(x,t) \in U$ and $|F(x,t)| \leq K$.

Let J denote the closed interval $[x_0 - K\delta, x_0 + K\delta]$. The space $C(I, J)$ of continuous functions from the interval I to the interval J is a complete metric space, by Corollary 4.12. Define $T: C(I, J) \to C(I, J)$ by

$$T(\varphi)(t) = x_0 + \int_{t_0}^{t} F(\varphi(s), s)\, ds.$$

We claim that T does indeed map $C(I, J)$ into itself and is a contraction mapping.

Let $\varphi: I \to J$ be an element of $C(I, J)$. Note that if $|t - t_0| \leq \delta$ then

$$|(\varphi(t), t) - (x_0, t_0)|^2 = (\varphi(t) - x_0)^2 + (t - t_0)^2 \leq \delta^2 + K^2\delta^2 < \delta_0^2,$$

hence $|F(\varphi(t), t)| \leq K$. It follows from this that

$$|T(\varphi)(t) - x_0| \leq K\delta$$

for all t satisfying $|t - t_0| < \delta$. The function $T(\varphi)$ is continuous, and is therefore a well-defined element of $C(I, J)$ for all $\varphi \in C(I, J)$.

We now show that T is a contraction mapping on $C(I, J)$. Let φ and ψ be elements of $C(I, J)$. The hypotheses of the theorem ensure that

$$|F(\varphi(t), t) - F(\psi(t), t)| \leq M|\varphi(t) - \psi(t)| \leq M\rho(\varphi, \psi)$$

for all $t \in I$, where $\rho(\varphi, \psi) = \sup_{t \in I} |\varphi(t) - \psi(t)|$. Therefore

$$\begin{aligned} |T(\varphi)(t) - T(\psi)(t)| &= \left| \int_{t_0}^{t} \left(F(\varphi(s), s) - F(\psi(s), s) \right) ds \right| \\ &\leq M|t - t_0|\rho(\varphi, \psi) \end{aligned}$$

for all t satisfying $|t - t_0| \leq \delta$. Therefore $\rho(T(\varphi), T(\psi)) \leq M\delta\rho(\varphi, \psi)$ for all $\varphi, \psi \in C(I, J)$. But δ has been chosen such that $M\delta < 1$. This shows that $T: C(I, J) \to C(I, J)$ is a contraction mapping on $C(I, J)$. It follows from the Contraction Mapping Theorem (Theorem 4.13) that there exists a unique element φ of $C(I, J)$ satisfying $T(\varphi) = \varphi$. This function φ is the required solution to the differential equation. ∎

A straightforward, but somewhat technical, least upper bound argument can be used to show that if $x = \psi(t)$ is any other continuous solution to the differential equation

$$\frac{dx}{dt} = F(x, t)$$

on the interval $[t_0 - \delta, t_0 + \delta]$ satisfying the initial condition $\psi(t_0) = x_0$, then $|\psi(t) - x_0| \leq K\delta$ for all t satisfying $|t - t_0| \leq \delta$. Thus such a solution to the differential equation must belong to the space $C(I,J)$ defined in the proof of Theorem 4.14. The uniqueness of the fixed point of the contraction mapping $T\colon C(I,J) \to C(I,J)$ then shows that $\psi = \varphi$, where $\varphi\colon [t_0 - \delta, t_0 + \delta] \to \mathbb{R}$ is the solution to the differential equation whose existence was proved in Theorem 4.14. This shows that the solution to the differential equation is in fact unique on the interval $[t_0 - \delta, t_0 + \delta]$.

4.6 The Completion of a Metric Space

We describe below a construction whereby any metric space can be embedded in a complete metric space.

Lemma 4.15 *Let X be a metric space with distance function d, let (x_j) and (y_j) be Cauchy sequences of points in X, and let $d_j = d(x_j, y_j)$ for all positive integers j. Then (d_j) is a Cauchy sequence of real numbers.*

Proof It follows from the Triangle Inequality that

$$d_j \leq d(x_j, x_k) + d_k + d(y_k, y_j)$$

and thus $d_j - d_k \leq d(x_j, x_k) + d(y_j, y_k)$ for all integers j and k. Similarly $d_k - d_j \leq d(x_j, x_k) + d(y_j, y_k)$. It follows that

$$|d_j - d_k| \leq d(x_j, x_k) + d(y_j, y_k)$$

for all integers j and k.

Let $\varepsilon > 0$ be given. Then there exists some positive integer N such that $d(x_j, x_k) < \frac{1}{2}\varepsilon$ and $d(y_j, y_k) < \frac{1}{2}\varepsilon$ whenever $j \geq N$ and $k \geq N$, since the sequences (x_j) and (y_j) are Cauchy sequences in X. But then $|d_j - d_k| < \varepsilon$ whenever $j \geq N$ and $k \geq N$. Thus the sequence (d_j) is a Cauchy sequence of real numbers, as required. ∎

Let X be a metric space with distance function d. It follows from Cauchy's Criterion for Convergence and Lemma 4.15 that $\lim\limits_{j \to +\infty} d(x_j, y_j)$ exists for all Cauchy sequences (x_j) and (y_j) in X.

Lemma 4.16 *Let X be a metric space with distance function d, and let (x_j), (y_j) and (z_j) be Cauchy sequences of points in X. Then*

$$0 \leq \lim_{j \to +\infty} d(x_j, z_j) \leq \lim_{j \to +\infty} d(x_j, y_j) + \lim_{j \to +\infty} d(y_j, z_j).$$

Proof This follows immediately on taking limits of both sides of the Triangle Inequality. ∎

Lemma 4.17 *Let X be a metric space with distance function d, and let (x_j), (y_j) and (z_j) be Cauchy sequences of points in X. Suppose that*

$$\lim_{j\to+\infty} d(x_j, y_j) = 0 \text{ and } \lim_{j\to+\infty} d(y_j, z_j) = 0.$$

Then $\lim_{j\to+\infty} d(x_j, z_j) = 0$.

Proof This is an immediate consequence of Lemma 4.16. ∎

Lemma 4.18 *Let X be a metric space with distance function d, and let (x_j), (x'_j), (y_j) and (y'_j) be Cauchy sequences of points in X. Suppose that*

$$\lim_{j\to+\infty} d(x_j, x'_j) = 0 \text{ and } \lim_{j\to+\infty} d(y_j, y'_j) = 0.$$

Then $\lim_{j\to+\infty} d(x_j, y_j) = \lim_{j\to+\infty} d(x'_j, y'_j)$.

Proof It follows from Lemma 4.16 that

$$\begin{aligned} \lim_{j\to+\infty} d(x_j, y_j) &\leq \lim_{j\to+\infty} d(x_j, x'_j) + \lim_{j\to+\infty} d(x'_j, y'_j) + \lim_{j\to+\infty} d(y'_j, y_j) \\ &= \lim_{j\to+\infty} d(x'_j, y'_j). \end{aligned}$$

Similarly $\lim_{j\to+\infty} d(x'_j, y'_j) \leq \lim_{j\to+\infty} d(x_j, y_j)$. It follows that $\lim_{j\to+\infty} d(x_j, y_j) = \lim_{j\to+\infty} d(x'_j, y'_j)$, as required. ∎

Let X be a metric space with distance function d. Then there is an equivalence relation on the set of Cauchy sequences of points in X, where two Cauchy sequences (x_j) and (x'_j) in X are equivalent if and only if $\lim_{j\to+\infty} d(x_j, x'_j) = 0$. Let $\tilde{X}$ denote the set of equivalence classes of Cauchy sequences in X with respect to this equivalence relation. Let $\tilde{x}$ and $\tilde{y}$ be elements of $\tilde{X}$, and let (x_j) and (y_j) be Cauchy sequences belonging to the equivalence classes represented by $\tilde{x}$ and $\tilde{y}$. We define

$$d(\tilde{x}, \tilde{y}) = \lim_{j\to+\infty} d(x_j, y_j).$$

It follows from Lemma 4.18 that the value of $d(\tilde{x}, \tilde{y})$ does not depend on the choice of Cauchy sequences (x_j) and (y_j) representing $\tilde{x}$ and $\tilde{y}$. We obtain in this way a distance function on the set $\tilde{X}$. This distance function satisfies the Triangle Inequality (Lemma 4.16) and the other metric space axioms. Therefore $\tilde{X}$ with this distance function is a metric space. We refer to the space $\tilde{X}$ as the *completion* of the metric space X.

We can regard the metric space X as being embedded in its completion $\tilde{X}$, where a point x of X is represented in $\tilde{X}$ by the equivalence class of the constant sequence $x, x, x, \ldots$.

Example The completion of the space $\mathbb{Q}$ of rational numbers is the space $\mathbb{R}$ of real numbers.

Theorem 4.19 *The completion $\tilde{X}$ of a metric space X is a complete metric space.*

Proof Let $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \ldots$ be a Cauchy sequence in the completion $\tilde{X}$ of X. For each positive integer m let $x_{m,1}, x_{m,2}, x_{m,3}, \ldots$ be a Cauchy sequence in X belonging to the equivalence class that represents the element $\tilde{x}_m$ of $\tilde{X}$. Then, for each positive integer m there exists a positive integer $N(m)$ such that $d(x_{m,j}, x_{m,k}) < 1/m$ whenever $j \geq N(m)$ and $k \geq N(m)$. Let $y_m = x_{m,N(m)}$. We claim that the sequence $y_1, y_2, y_3, \ldots$ is a Cauchy sequence in X, and that the element $\tilde{y}$ of $\tilde{X}$ corresponding to this Cauchy sequence is the limit in $\tilde{X}$ of the sequence $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \ldots$.

Let $\varepsilon > 0$ be given. Then there exists some positive integer M such that $M > 3/\varepsilon$ and $d(\tilde{x}_p, \tilde{x}_q) < \frac{1}{3}\varepsilon$ whenever $p \geq M$ and $q \geq M$. It follows from the definition of the distance function on $\tilde{X}$ that if $p \geq M$ and $q \geq M$ then $d(x_{p,k}, x_{q,k}) < \frac{1}{3}\varepsilon$ for all sufficiently large positive integers k. If $p \geq M$ and $k \geq N(p)$ then

$$d(y_p, x_{p,k}) = d(x_{p,N(p)}, x_{p,k}) < 1/p \leq 1/M < \tfrac{1}{3}\varepsilon$$

It follows that if $p \geq M$ and $q \geq M$, and if k is sufficiently large, then $d(y_p, x_{p,k}) < \frac{1}{3}\varepsilon$, $d(y_q, x_{q,k}) < \frac{1}{3}\varepsilon$, and $d(x_{p,k}, x_{q,k}) < \frac{1}{3}\varepsilon$, and hence $d(y_p, y_q) < \varepsilon$. We conclude that the sequence $y_1, y_2, y_3, \ldots$ of points of X is indeed a Cauchy sequence.

Let $\tilde{y}$ be the element of $\tilde{X}$ which is represented by the Cauchy sequence $y_1, y_2, y_3, \ldots$ of points of X, and, for each positive integer m, let $\tilde{y}_m$ be the element of $\tilde{X}$ represented by the constant sequence $y_m, y_m, y_m, \ldots$ in X. Now

$$d(\tilde{y}, \tilde{y_m}) = \lim_{p \to +\infty} d(y_p, y_m),$$

and therefore $d(\tilde{y}, \tilde{y_m}) \to 0$ as $m \to +\infty$. Also

$$d(\tilde{y}_m, \tilde{x}_m) = \lim_{j \to +\infty} d(x_{m,N(m)}, x_{m,j}) \leq \frac{1}{m}$$

and hence $d(\tilde{y}_m, \tilde{x}_m) \to 0$ as $m \to +\infty$. It follows from this that $d(\tilde{y}, \tilde{x}_m) \to 0$ as $m \to +\infty$, and therefore the Cauchy sequence $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \ldots$ in $\tilde{X}$ converges to the point $\tilde{y}$ of $\tilde{X}$. We conclude that $\tilde{X}$ is a complete metric space, since we have shown that every Cauchy sequence in $\tilde{X}$ is convergent. ∎

Remark In a paper published in 1872, Cantor gave a construction of the real number system in which real numbers are represented as Cauchy sequences of rational numbers. The real numbers represented by two Cauchy sequences of rational numbers are equal if and only if the difference of the Cauchy sequences converges to zero. Thus the construction of the completion of a metric space, described above, generalizes Cantor's construction of the system of real numbers from the system of rational numbers.

5 Topological Spaces

The theory of topological spaces provides a setting for the notions of continuity and convergence which is more general than that provided by the theory of metric spaces. In the theory of metric spaces one can find necessary and sufficient conditions for convergence and continuity that do not refer explicitly to the distance function on a metric space but instead are expressed in terms of open sets. Thus a sequence of points in a metric space X converges to a point p of X if and only if every open set which contains the point p also contains all but finitely many members of the sequence. Also a function $f\colon X \to Y$ between metric spaces X and Y is continuous if and only if the preimage $f^{-1}(V)$ of every open set V in Y is an open set in X. It follows from this that we can generalize the notions of convergence and continuity by introducing the concept of a *topological space*: a topological space consists of a set together with a collection of subsets termed *open sets* that satisfy appropriate axioms. The axioms for open sets in a topological space are satisfied by the open sets in any metric space.

5.1 Topological Spaces: Definitions and Examples

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, τ) the topological space whose underlying set is X and whose topology is τ. However if no confusion will arise then it is customary to denote this topological space simply by X.

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 $\{x \in X : d(x, v) < \delta\} \subset V$. The empty set $\emptyset$ and the whole space X are open sets. Also any union of open sets in a metric space is an open set, and any finite intersection of open sets in a metric space is an open set. Thus 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.

Any subset X of n-dimensional Euclidean space $\mathbb{R}^n$ is a topological space: a subset V of X is open in X if and only if, given any point $\mathbf{v}$ of V, there exists some $\delta > 0$ such that

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

In particular $\mathbb{R}^n$ is itself a topological space whose topology is generated by the Euclidean distance function on $\mathbb{R}^n$. This topology on $\mathbb{R}^n$ is referred to as the *usual topology* on $\mathbb{R}^n$. One defines the usual topologies on $\mathbb{R}$ and $\mathbb{C}$ in an analogous fashion.

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.

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 5.1 *Let X be a topological space. Then the collection of closed sets of X has the following properties:—*

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

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

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

5.2 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 5.2 *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 ε centred on the points x and y are open sets. 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 ε. 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.

5.3 Subspace Topologies

Let X be a topological space with topology τ, and let A be a subset of X. Let τ_A be the collection of all subsets of A that are of the form $V \cap A$ for $V \in \tau$. Then τ_A is a topology on the set A. (It is a straightforward exercise to verify that the topological space axioms are satisfied.) The topology τ_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 5.3 *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, 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.

5.4 Continuous Functions between Topological Spaces

Definition A function $f\colon X \to Y$ 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 5.4 *Let X, Y and Z be topological spaces, and let $f\colon X \to Y$ and $g\colon Y \to Z$ be continuous functions. Then the composition $g \circ f\colon X \to Z$ 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 5.5 *Let X and Y be topological spaces, and let $f\colon X \to Y$ be a function from X to Y. The function f is continuous if and only if $f^{-1}(G)$ 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. ∎

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 5.6 *Let X and Y be topological spaces, let $f\colon X \to Y$ 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 $f\colon X \to Y$ 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 $f\colon X \to Y$ is continuous, as required. ∎

Alternative Proof A function $f\colon X \to Y$ is continuous if and only if $f^{-1}(G)$ is closed in X for every closed set G in Y (Lemma 5.5). 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 5.5 that $f\colon X \to Y$ is continuous. ∎

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

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

Now $\gamma|[0, \frac{1}{2}] = \alpha \circ \rho$ where $\rho\colon [0, \frac{1}{2}] \to [0, 1]$ 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 5.6 that $\gamma: [0, 1] \to Y$ is continuous.

5.5 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.

5.6 Sequences and Convergence

Definition A sequence $x_1, x_2, x_3, \ldots$ of points in a topological space X is said to *converge* to a point p of X if, given any open set U containing the point p, there exists some natural number N such that $x_j \in U$ for all $j \geq N$. If the sequence (x_j) converges to p then we refer to p as a *limit* of the sequence.

This definition of convergence generalizes the definition of convergence for a sequence of points in a metric space.

It can happen that a sequence of points in a topological space can have more than one limit. For example, consider the set $\mathbb{R}$ of real numbers with the Zariski topology. (The open sets of $\mathbb{R}$ in the Zariski topology are the empty set and those subsets of $\mathbb{R}$ whose complements are finite.) Let $x_1, x_2, x_3, \ldots$ be the sequence in $\mathbb{R}$ defined by $x_j = j$ for all natural numbers j. One can readily check that this sequence converges to every real number p (with respect to the Zariski topology on $\mathbb{R}$).

Lemma 5.7 *A sequence $x_1, x_2, x_3, \ldots$ of points in a Hausdorff space X converges to at most one limit.*

Proof Suppose that p and q were limits of the sequence (x_j), where $p \neq q$. Then there would exist open sets U and V such that $p \in U$, $q \in V$ and $U \cap V = \emptyset$, since X is a Hausdorff space. But then there would exist natural numbers N_1 and N_2 such that $x_j \in U$ for all j satisfying $j \geq N_1$ and $x_j \in V$ for all j satisfying $j \geq N_2$. But then $x_j \in U \cap V$ for all j satisfying $j \geq N_1$ and $j \geq N_2$, which is impossible, since $U \cap V = \emptyset$. This contradiction shows that the sequence (x_j) has at most one limit. ∎

Lemma 5.8 *Let X be a topological space, and let F be a closed set in X. Let $(x_j : j \in \mathbb{N})$ be a sequence of points in F. Suppose that the sequence (x_j) converges to some point p of X. Then $p \in F$.*

Proof Suppose that p were a point belonging to the complement $X \setminus F$ of F. Now $X \setminus F$ is open (since F is closed). Therefore there would exist some natural number N such that $x_j \in X \setminus F$ for all values of j satisfying $j \geq N$, contradicting the fact that $x_j \in F$ for all j. This contradiction shows that p must belong to F, as required. ∎

Lemma 5.9 *Let $f: X \to Y$ be a continuous function between topological spaces X and Y, and let $x_1, x_2, x_3, \ldots$ be a sequence of points in X which converges to some point p of X. Then the sequence $f(x_1), f(x_2), f(x_3), \ldots$ converges to $f(p)$.*

Proof Let V be an open set in Y which contains the point $f(p)$. Then $f^{-1}(V)$ is an open set in X which contains the point p. It follows that there exists some natural number N such that $x_j \in f^{-1}(V)$ whenever $j \geq N$. But then $f(x_j) \in V$ whenever $j \geq N$. We deduce that the sequence $f(x_1), f(x_2), f(x_3), \ldots$ converges to $f(p)$, as required. ∎

5.7 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 5.10 *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.

Lemma 5.11 *Let X be a topological space, and let A be a subset of X. Suppose that a sequence $x_1, x_2, x_3, \ldots$ of points of A converges to some point p of X. Then p belongs to the closure $\overline{A}$ of A.*

Proof If F is any closed set containing A then $x_j \in F$ for all j, and therefore $p \in F$, by Lemma 5.8. Therefore $p \in \overline{A}$ by definition of $\overline{A}$. ∎

Definition Let X be a topological space, and let A be a subset of X. We say that A is *dense* in X if $\overline{A} = X$.

Example The set of all rational numbers is dense in $\mathbb{R}$.

5.8 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 $f: X \times Y \to Z$ 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 5.12 *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 5.13 *Let $X_1, X_2, \ldots, X_n$ and Z be topological spaces. Then a function $f: X_1 \times X_2 \times \cdots \times X_n \to Z$ 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 5.14 *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 $p_i: X \to X_i$ 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 $f: Z \to X$ mapping a topological space Z into X is continuous if and only if $p_i \circ f: Z \to X_i$ 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 $p_i: X \to X_i$ is continuous for all i.

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

Conversely suppose that $f: Z \to X$ 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 $f_i: Z \to X_i$ 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 5.15 *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

$$I_i = \{t \in \mathbb{R} : u_i - \frac{\delta}{\sqrt{n}} < t < u_i + \frac{\delta}{\sqrt{n}}\}$$

for $i = 1, 2, \ldots, n$. 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 5.15 and Theorem 5.14.

Corollary 5.16 *Let X be a topological space and let $f\colon X \to \mathbb{R}^n$ 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 $f\colon X \to \mathbb{R}$ and $g\colon X \to \mathbb{R}$ 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 $s\colon \mathbb{R}^2 \to \mathbb{R}$ and $p\colon \mathbb{R}^2 \to \mathbb{R}$ 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 $h\colon X \to \mathbb{R}^2$ is defined by $h(x) = (f(x), g(x))$. Moreover it follows from Corollary 5.16 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 5.17 *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. ∎

5.9 Cut and Paste Constructions

Suppose we start out with a square of paper. If we join together two opposite edges of this square we obtain a cylinder. The boundary of the cylinder consists of two circles. If we join together the two boundary circles we obtain a torus (which corresponds to the surface of a doughnut).

Let the square be represented by the set $[0,1] \times [0,1]$ consisting of all ordered pairs (s,t) where s and t are real numbers between 0 and 1. There is an equivalence relation on the square $[0,1] \times [0,1]$, where points (s,t) and

(u, v) of the square are related if and only if at least one of the following conditions is satisfied:

- $s = u$ and $t = v$;
- $s = 0$, $u = 1$ and $t = v$;
- $s = 1$, $u = 0$ and $t = v$;
- $t = 0$, $v = 1$ and $s = u$;
- $t = 1$, $v = 0$ and $s = u$;
- (s, t) and (u, v) both belong to $\{(0, 0), (0, 1), (1, 0), (1, 1)\}$.

Note that if $0 < s < 1$ and $0 < t < 1$ then the equivalence class of the point (s, t) is the set $\{(s, t)\}$ consisting of that point. If $s = 0$ or 1 and if $0 < t < 1$ then the equivalence class of (s, t) is the set $\{(0, t), (1, t)\}$. Similarly if $t = 0$ or 1 and if $0 < s < 1$ then the equivalence class of (s, t) is the set $\{(s, 0), (s, 1)\}$. The equivalence class of each corner of the square is the set $\{(0, 0), (1, 0), (0, 1), (1, 1)\}$ consisting of all four corners. Thus each equivalence class contains either one point in the interior of the square, or two points on opposite edges of the square, or four points at the four corners of the square. Let T^2 denote the set of these equivalence classes. We have a map $q: [0, 1] \times [0, 1] \to T^2$ which sends each point (s, t) of the square to its equivalence class. Each element of the set T^2 is the image of one, two or four points of the square. The elements of T^2 represent points on the torus obtained from the square by first joining together two opposite sides of the square to form a cylinder and then joining together the boundary circles of this cylinder as described above. We say that the torus T^2 is obtained from the square $[0, 1] \times [0, 1]$ by *identifying* the points $(0, t)$ and $(1, t)$ for all $t \in [0, 1]$ and identifying the points $(s, 0)$ and $(s, 1)$ for all $s \in [0, 1]$.

The topology on the square $[0, 1] \times [0, 1]$ induces a corresponding topology on the set T^2, where a subset U of T^2 is open in T^2 if and only if $q^{-1}(U)$ is open in the square $[0, 1] \times [0, 1]$. (The fact that these open sets in T^2 constitute a topology on the set T^2 is a consequence of Lemma 5.18.) The function $q: [0, 1] \times [0, 1] \to T^2$ is then a continuous surjection. We say that the topological space T^2 is the *identification space* obtained from the square $[0, 1] \times [0, 1]$ by identifying points on the sides to the square as described above. The continuous map q from the square to the torus is an example of an *identification map*, and the topology on the torus T^2 is referred to as the *quotient topology* on T^2 induced by the identification map $q: [0, 1] \times [0, 1] \to T^2$.

Another well-known identification space obtained from the square is the *Klein bottle* (*Kleinsche Flasche*). The Klein bottle K^2 is obtained from the square $[0,1] \times [0,1]$ by identifying $(0,t)$ with $(1,1-t)$ for all $t \in [0,1]$ and identifying $(s,0)$ with $(s,1)$ for all $s \in [0,1]$. These identifications correspond to an equivalence relation on the square, where points (s,t) and (u,v) of the square are equivalent if and only if one of the following conditions is satisfied:

- $s = u$ and $t = v$;
- $s = 0$, $u = 1$ and $t = 1 - v$;
- $s = 1$, $u = 0$ and $t = 1 - v$;
- $t = 0$, $v = 1$ and $s = u$;
- $t = 1$, $v = 0$ and $s = u$;
- (s,t) and (u,v) both belong to $\{(0,0), (0,1), (1,0), (1,1)\}$.

The corresponding set of equivalence classes is the Klein bottle K^2. Thus each point of the Klein bottle K^2 represents an equivalence class consisting of either one point in the interior of the square, or two points $(0,t)$ and $(1,1-t)$ with $0 < t < 1$ on opposite edges of the square, or two points $(s,0)$ and $(s,1)$ with $0 < s < 1$ on opposite edges of the square, or the four corners of the square. There is a surjection $r\colon [0,1] \times [0,1] \to K^2$ from the square to the Klein bottle that sends each point of the square to its equivalence class. The identifications used to construct the Klein bottle ensure that $r(0,t) = r(1,1-t)$ for all $t \in [0,1]$ and $r(s,0) = r(s,1)$ for all $s \in [0,1]$. One can construct a quotient topology on the Klein bottle K^2, where a subset U of K^2 is open in K^2 if and only if its preimage $r^{-1}(U)$ is open in the square $[0,1] \times [0,1]$.

5.10 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 $q: X \to Y$ 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.

Example Let S^1 denote the unit circle $\{(x, y) \in \mathbb{R}^2 : x^2+y^2 = 1\}$ in $\mathbb{R}^2$, and let $q: [0, 1] \to S^1$ be the continuous map defined by $q(t) = (\cos 2\pi t, \sin 2\pi t)$ for all $t \in [0, 1]$. We show that $q: [0, 1] \to S^1$ is an identification map. This map is continuous and surjective. It remains to show that if V is a subset of S^1 with the property that $q^{-1}(V)$ is open in $[0, 1]$ then V is open in S^1.

Note that $|q(s) - q(t)| = 2|\sin \pi(s - t)|$ for all $s, t \in [0, 1]$ satisfying $|s - t| \leq \frac{1}{2}$. Let V be a subset of S^1 with the property that $q^{-1}(V)$ is open in $[0, 1]$, and let $\mathbf{v}$ be an element of V. We show that there exists $\varepsilon > 0$ such that all points $\mathbf{u}$ of S^1 satisfying $|\mathbf{u} - \mathbf{v}| < \varepsilon$ belong to V. We consider separately the cases when $\mathbf{v} = (1, 0)$ and when $\mathbf{v} \neq (1, 0)$.

Suppose that $\mathbf{v} = (1, 0)$. Then $(1, 0) \in V$, and hence $0 \in q^{-1}(V)$ and $1 \in q^{-1}(V)$. But $q^{-1}(V)$ is open in $[0, 1]$. It follows that there exists a real number δ satisfying $0 < \delta < \frac{1}{2}$ such that $[0, \delta) \subset q^{-1}(V)$ and $(1 - \delta, 1] \in q^{-1}(V)$. Let $\varepsilon = 2 \sin \pi\delta$. Now if $-\pi \leq \theta \leq \pi$ then the Euclidean distance between the points $(1, 0)$ and $(\cos \theta, \sin \theta)$ is $2 \sin \frac{1}{2}|\theta|$. Moreover, this distance increases monotonically as $|\theta|$ increases from 0 to π. Thus any point on the unit circle S^1 whose distance from $(1, 0)$ is less than ε must be of the form $(\cos \theta, \sin \theta)$, where $|\theta| < 2\pi\delta$. Thus if $\mathbf{u} \in S^1$ satisfies $|\mathbf{u} - \mathbf{v}| < \varepsilon$ then $\mathbf{u} = q(s)$ for some $s \in [0, 1]$ satisfying either $0 \leq s < \delta$ or $1 - \delta < s \leq 1$. But then $s \in q^{-1}(V)$, and hence $\mathbf{u} \in V$.

Next suppose that $\mathbf{v} \neq (1, 0)$. Then $\mathbf{v} = q(t)$ for some real number t satisfying $0 < t < 1$. But $q^{-1}(V)$ is open in $[0, 1]$, and $t \in q^{-1}(V)$. It follows that $(t - \delta, t + \delta) \subset q^{-1}(V)$ for some real number δ satisfying $\delta > 0$. Let $\varepsilon = 2 \sin \pi\delta$. If $\mathbf{u} \in S^1$ satisfies $|\mathbf{u} - \mathbf{v}| < \varepsilon$ then $\mathbf{u} = q(s)$ for some $s \in (t - \delta, t + \delta)$. But then $s \in q^{-1}(V)$, and hence $\mathbf{u} \in V$.

We have thus shown that if V is a subset of S^1 with the property that $q^{-1}(V)$ is open in $[0, 1]$ then there exists $\varepsilon > 0$ such that $\mathbf{u} \in V$ for all elements $\mathbf{u}$ of S^1 satisfying $|\mathbf{u} - \mathbf{v}| < \varepsilon$. It follows from this that V is open in S^1. Thus the continuous surjection $q: [0, 1] \to S^1$ is an identification map.

Lemma 5.18 *Let X be a topological space, let Y be a set, and let $q: X \to Y$ be a surjection. Then there is a unique topology on Y for which the function $q: X \to Y$ is an identification map.*

Proof Let τ be the collection consisting of all subsets U of Y for which $q^{-1}(U)$ is open in X. Now $q^{-1}(\emptyset) = \emptyset$, and $q^{-1}(Y) = X$, so that $\emptyset \in \tau$ and

$Y \in \tau$. 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 τ must themselves belong to τ. Thus τ is a topology on Y, and the function $q: X \to Y$ is an identification map with respect to the topology τ. Clearly τ is the unique topology on Y for which the function $q: X \to Y$ 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.

Let $\sim$ be an equivalence relation on a topological space X. If Y is the corresponding set of equivalence classes of elements of X then there is a surjection $q: X \to Y$ that sends each element of X to its equivalence class. Lemma 5.18 ensures that there is a well-defined quotient topology on Y, where a subset U of Y is open in Y if and only if $q^{-1}(U)$ is open in X. (Appropriate equivalence relations on the square yield the torus and the Klein bottle, as discussed above.)

Lemma 5.19 *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 The Klein bottle K^2 is the identification space obtained from the square $[0,1] \times [0,1]$ by identifying $(0,t)$ with $(1,1-t)$ for all $t \in [0,1]$ and identifying $(s,0)$ with $(s,1)$ for all $s \in [0,1]$. Let $q\colon [0,1] \times [0,1] \to K^2$ be the identification map determined by these identifications. Let Z be a topological space. A function $g\colon [0,1] \times [0,1] \to Z$ mapping the square into Z which satisfies $g(0,t) = g(1,1-t)$ for all $t \in [0,1]$ and $g(s,0) = g(s,1)$ for all $s \in [0,1]$, determines a corresponding function $f\colon K^2 \to Z$, where $g = f \circ q$. It follows from Lemma 5.19 that the function $f\colon K^2 \to Z$ is continuous if and only if $g\colon [0,1] \times [0,1] \to Z$ 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 $q\colon S^n \to \mathbb{R}P^n$ 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 $q\colon S^n \to \mathbb{R}P^n$ 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 5.19 that a function $f\colon \mathbb{R}P^n \to Z$ from $\mathbb{R}P^n$ to any topological space Z is continuous if and only if the composition function $f \circ q\colon S^n \to Z$ is continuous.

5.11 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 5.20 *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 $f: X \to \mathbb{Z}$ 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 5.21 *A topological space X is connected if and only if every continuous function $f: X \to \mathbb{Z}$ from X to the set $\mathbb{Z}$ of integers is constant.*

Proof Suppose that X is connected. Let $f: X \to \mathbb{Z}$ 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 $f: X \to \mathbb{Z}$ is constant, with value n.

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

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

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

Proof Let $f: [a, b] \to \mathbb{Z}$ 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 5.21 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 5.23 *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 5.22), therefore $f\circ\gamma$ is constant (Proposition 5.21). 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 5.21. ∎

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 5.20 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 5.24 *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 5.25 *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 5.21 that $g \circ f$ is constant on A. Therefore g is constant on $f(A)$. We deduce from Proposition 5.21 that $f(A)$ is connected. ∎

Lemma 5.26 *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 5.21 that $X \times Y$ is connected. ∎

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

Proposition 5.27 *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 $f\colon S_x \to \mathbb{Z}$ 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 5.24. 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 $f\colon S_x \cup S_y \to \mathbb{Z}$ 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 5.27 are referred to as the *connected components* of X. We see from Proposition 5.27, 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 $J_n = (n - \frac{1}{2}, n + \frac{1}{2})$.

6 Compact Spaces

6.1 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 6.1 *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 6.2 (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 6.3 *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 6.1 that A is compact, as required. ∎

Lemma 6.4 *Let $f: X \to Y$ 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 $V_1, V_2, \ldots, V_k$ 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 6.5 *Let $f: X \to \mathbb{R}$ 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 $I_1, I_2, \ldots, I_k$ of open intervals of the form $(-m, m)$, where $m \in \mathbb{N}$, since $f(X)$ is compact (Lemma 6.4) 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 $I_1, I_2, \ldots, I_k$. Thus the function f is bounded above and below on X, as required. ∎

Proposition 6.6 *Let $f: X \to \mathbb{R}$ 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 6.5. 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 6.5. But then $f(u) \leq f(x) \leq f(v)$ for all $x \in X$, as required. ∎

Proposition 6.7 *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 6.6 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 δ satisfy $0 < \delta \leq f(q)$. Then the open ball of radius δ 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 6.8 *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 6.9 *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 6.8 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 6.10 *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 6.8 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 6.11 *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 6.3), and therefore $f(K)$ is compact (Lemma 6.4). But any compact subset of a Hausdorff space is closed (Corollary 6.9). Thus $f(K)$ is closed in Y, as required. ∎

Remark If the Hausdorff space Y in Lemma 6.11 is a metric space, then Proposition 6.7 may be used in place of Corollary 6.9 in the proof of the lemma.

Theorem 6.12 *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 6.11. 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 6.13 *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 6.11. It follows that U is open in Y. Conversely if U is open in Y then $f^{-1}(Y)$ 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 $q: [0, 1] \to S^1$ 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 $q: [0, 1] \to S^1$ 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 6.14 *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 6.15 *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 6.4). 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 6.14 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 6.16 *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 6.9). 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 6.2), and C is the Cartesian product of n copies of the compact set $[-L, L]$. It follows from Theorem 6.15 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 6.3. Thus K is compact, as required. ∎

6.2 Compact Metric Spaces

We recall that a metric or topological space is said to be *compact* if every open cover of the space has a finite subcover. We shall obtain some equivalent characterizations of compactness for *metric spaces* (Theorem 6.22); these characterizations do not generalize to arbitrary topological spaces.

Proposition 6.17 *Every sequence of points in a compact metric space has a convergent subsequence.*

Proof Let X be a compact metric space, and let $x_1, x_2, x_3, \ldots$ be a sequence of points of X. We must show that this sequence has a convergent subsequence. Let F_n denote the closure of $\{x_n, x_{n+1}, x_{n+2}, \ldots\}$. We claim that the intersection of the sets $F_1, F_2, F_3, \ldots$ is non-empty. For suppose that this intersection were the empty set. Then X would be the union of the sets $V_1, V_2, V_3, \ldots$, where $V_n = X \setminus F_n$ for all n. But $V_1 \subset V_2 \subset V_3 \subset \cdots$, and each set V_n is open. It would therefore follow from the compactness of X that X would be covered by finitely many of the sets $V_1, V_2, V_3, \ldots$, and therefore $X = V_n$ for some sufficiently large n. But this is impossible, since F_n is non-empty for all natural numbers n. Thus the intersection of the sets $F_1, F_2, F_3, \ldots$ is non-empty, as claimed, and therefore there exists a point p of X which belongs to F_n for all natural numbers n.

We now obtain, by induction on n, a subsequence $x_{n_1}, x_{n_2}, x_{n_3}, \ldots$ which satisfies $d(x_{n_j}, p) < 1/j$ for all natural numbers j. Now p belongs to the closure F_1 of the set $\{x_1, x_2, x_3, \ldots\}$. Therefore there exists some natural number n_1 such that $d(x_{n_1}, p) < 1$. Suppose that x_{n_j} has been chosen so that $d(x_{n_j}, p) < 1/j$. The point p belongs to the closure F_{n_j+1} of the set $\{x_n : n > n_j\}$. Therefore there exists some natural number n_{j+1} such that $n_{j+1} > n_j$ and $d(x_{n_{j+1}}, p) < 1/(j+1)$. The subsequence $x_{n_1}, x_{n_2}, x_{n_3}, \ldots$ constructed in this manner converges to the point p, as required. ∎

We shall also prove the converse of Proposition 6.17: if X is a metric space, and if every sequence of points of X has a convergent subsequence, then X is compact (see Theorem 6.22 below).

Let X be a metric space with distance function d. A *Cauchy sequence* in X is a sequence $x_1, x_2, x_3, \ldots$ of points of X with the property that, given any $\varepsilon > 0$, there exists some natural number N such that $d(x_j, x_k) < \varepsilon$ for all j and k satisfying $j \geq N$ and $k \geq N$.

A metric space (X, d) is said to be *complete* if every Cauchy sequence in X converges to some point of X.

Proposition 6.18 *Let X be a metric space with the property that every sequence of points of X has a convergent subsequence. Then X is complete.*

Proof Let $x_1, x_2, x_3, \ldots$ be a Cauchy sequence in X. This sequence then has a subsequence $x_{n_1}, x_{n_2}, x_{n_3}, \ldots$ which converges to some point p of X. We claim that the given Cauchy sequence also converges to p.

Let $\varepsilon > 0$ be given. Then there exists some natural number N such that $d(x_m, x_n) < \frac{1}{2}\varepsilon$ whenever $m \geq N$ and $n \geq N$, since $x_1, x_2, x_3, \ldots$ is a Cauchy sequence. Moreover n_j can be chosen large enough to ensure that $n_j \geq N$ and $d(x_{n_j}, p) < \frac{1}{2}\varepsilon$. If $n \geq N$ then

$$d(x_n, p) \leq d(x_n, x_{n_j}) + d(x_{n_j}, p) < \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon.$$

This shows that the Cauchy sequence $x_1, x_2, x_3, \ldots$ converges to the point p. Thus X is complete, as required. ∎

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.)

Let X be a metric space with distance function d, and let A be a subset of X. The *closure* $\overline{A}$ of A is the intersection of all closed sets in X that contain the set A: it can be regarded as the smallest closed set in X containing A. Let x be a point of the closure $\overline{A}$ of A. Given any $\varepsilon > 0$, there exists some point x' of A such that $d(x, x') < \varepsilon$. (Indeed the open ball in X of radius ε about the point x must intersect the set A, since otherwise the complement of this open ball would be a closed set in X containing the set A but not including the point x, which is not possible if x belongs to the closure of A.)

Lemma 6.19 *Let X be a metric space, and let A be a subset of X. Then* $\operatorname{diam} A = \operatorname{diam} \overline{A}$, *where $\overline{A}$ is the closure of A.*

Proof Clearly $\operatorname{diam} A \leq \operatorname{diam} \overline{A}$. Let x and y be points of $\overline{A}$. Then, given any $\varepsilon > 0$, there exist points x' and y' of A satisfying $d(x, x') < \varepsilon$ and $d(y, y') < \varepsilon$. It follows from the Triangle Inequality that

$$d(x, y) \leq d(x, x') + d(x', y') + d(y', y) < \operatorname{diam} A + 2\varepsilon.$$

Thus $d(x, y) < \operatorname{diam} A + 2\varepsilon$ for all $\varepsilon > 0$, and hence $d(x, y) \leq \operatorname{diam} A$. This shows that $\operatorname{diam} \overline{A} \leq \operatorname{diam} A$, as required. ∎

Definition A metric space X is said to be *totally bounded* if, given any $\varepsilon > 0$, the set X can be expressed as a finite union of subsets of X, each of which has diameter less than ε.

A subset A of a totally bounded metric space X is itself totally bounded. For if X is the union of the subsets $B_1, B_2, \ldots, B_k$, where $\operatorname{diam} B_n < \varepsilon$ for $n = 1, 2, \ldots, k$, then A is the union of $A \cap B_n$ for $n = 1, 2, \ldots, k$, and $\operatorname{diam} A \cap B_n < \varepsilon$.

Proposition 6.20 *Let X be a metric space. Suppose that every sequence of points of X has a convergent subsequence. Then X is totally bounded.*

Proof Suppose that X were not totally bounded. Then there would exist some $\varepsilon > 0$ with the property that no finite collection of subsets of X of diameter less than 3ε covers the set X. There would then exist an infinite sequence $x_1, x_2, x_3, \ldots$ of points of X with the property that $d(x_m, x_n) \geq \varepsilon$ whenever $m \neq n$. Indeed suppose that points $x_1, x_2, \ldots, x_{k-1}$ of X have already been chosen satisfying $d(x_m, x_n) \geq \varepsilon$ whenever $m < k$, $n < k$ and $m \neq n$. The diameter of each open ball $B_X(x_m, \varepsilon)$ is less than or equal to 2ε. Therefore X could not be covered by the sets $B_X(x_m, \varepsilon)$ for $m < k$, and thus there would exist a point x_k of X which does not belong to $B(x_m, \varepsilon)$ for any $m < k$. Then $d(x_m, x_k) \geq \varepsilon$ for all $m < k$. In this way we can successively choose points $x_1, x_2, x_3, \ldots$ to form an infinite sequence with the required property. However such an infinite sequence would have no convergent subsequence, which is impossible. This shows that X must be totally bounded, as required. ∎

Proposition 6.21 *Every complete totally bounded metric space is compact.*

Proof Let X be some totally bounded metric space. Suppose that there exists an open cover $\mathcal{V}$ of X which has no finite subcover. We shall prove the existence of a Cauchy sequence $x_1, x_2, x_3, \ldots$ in X which cannot converge to any point of X. (Thus if X is not compact, then X cannot be complete.)

Let $\varepsilon > 0$ be given. Then X can be covered by finitely many closed sets whose diameter is less than ε, since X is totally bounded and every subset of X has the same diameter as its closure (Lemma 6.19). At least one of these closed sets cannot be covered by a finite collection of open sets belonging to $\mathcal{V}$ (since if every one of these closed sets could be covered by a such a finite collection of open sets, then we could combine these collections to obtain a finite subcover of $\mathcal{V}$). We conclude that, given any $\varepsilon > 0$, there exists a closed subset of X of diameter less than ε which cannot be covered by any finite collection of open sets belonging to $\mathcal{V}$.

We claim that there exists a sequence $F_1, F_2, F_3, \ldots$ of closed sets in X satisfying $F_1 \supset F_2 \supset F_3 \supset \cdots$ such that each closed set F_n has the following properties: $\operatorname{diam} F_n < 1/2^n$, and no finite collection of open sets belonging to $\mathcal{V}$ covers F_n. For if F_n is a closed set with these properties then F_n is itself totally bounded, and thus the above remarks (applied with F_n in place of X) guarantee the existence of a closed subset F_{n+1} of F_n with the required properties. Thus the existence of the required sequence of closed sets follows by induction on n.

Choose $x_n \in F_n$ for each natural number n. Then $d(x_m, x_n) < 1/2^n$ for any $m > n$, since x_m and x_n belong to F_n and $\operatorname{diam} F_n < 1/2^n$. Therefore the sequence $x_1, x_2, x_3, \ldots$ is a Cauchy sequence. Suppose that this Cauchy sequence were to converge to some point p of X. Then $p \in F_n$ for each natural number n, since F_n is closed and $x_m \in F_n$ for all $m \geq n$. (If a sequence of points belonging to a closed subset of a metric or topological space is convergent then the limit of that sequence belongs to the closed set.) Moreover $p \in V$ for some open set V belonging to $\mathcal{V}$, since $\mathcal{V}$ is an open cover of X. But then there would exist $\delta > 0$ such that $B_X(p, \delta) \subset V$, where $B_X(p, \delta)$ denotes the open ball of radius δ in X centred on p. Thus if n were large enough to ensure that $1/2^n < \delta$, then $p \in F_n$ and $\operatorname{diam} F_n < \delta$, and hence $F_n \subset B_X(p, \delta) \subset V$, contradicting the fact that no finite collection of open sets belonging to $\mathcal{V}$ covers the set F_n. This contradiction shows that the Cauchy sequence $x_1, x_2, x_3, \ldots$ is not convergent.

We have thus shown that if X is a totally bounded metric space which is not compact then X is not complete. Thus every complete totally bounded metric space must be compact, as required. ∎

Theorem 6.22 *Let X be a metric space with distance function d. The following are equivalent:—*

(i) *X is compact,*

(ii) *every sequence of points of X has a convergent subsequence,*

(iii) *X is complete and totally bounded,*

Proof Propositions 6.17, 6.18 6.20 and 6.21 show that (i) implies (ii), (ii) implies (iii), and (iii) implies (i). It follows that (i), (ii) and (iii) are all equivalent to one another. ∎

Remark A subset K of $\mathbb{R}^n$ is complete if and only if it is closed in $\mathbb{R}^n$. Also it is easy to see that K is totally bounded if and only if K is a bounded subset of $\mathbb{R}^n$. Thus Theorem 6.22 is a generalization of the theorem which states that a subset K of $\mathbb{R}^n$ is compact if and only if it is both closed and bounded (Theorem 6.16).

6.3 The Lebesgue Lemma and Uniform Continuity

Lemma 6.23 (Lebesgue Lemma) *Let (X, d) be a compact metric space. Let $\mathcal{U}$ be an open cover of X. Then there exists a positive real number δ such that every subset of X whose diameter is less than δ is contained wholly within one of the open sets belonging to the open cover $\mathcal{U}$.*

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

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

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

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

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

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

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

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

Let X and Y be metric spaces with distance functions d_X and d_Y respectively, and let $f\colon X \to Y$ 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 δ should be independent of both x and x'.)

Theorem 6.24 *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 $f\colon X \to Y$ 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 6.23) that there exists some $\delta > 0$ such that every subset of X whose diameter is less than δ 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 δ. 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 6.16 and Theorem 6.24 that any continuous function $f\colon K \to \mathbb{R}^k$ is uniformly continuous.

6.4 The Equivalence of Norms on a Finite-Dimensional Vector Space

Let $\|.\|$ and $\|.\|_*$ be norms on a real or complex vector space X. The norms $\|.\|$ and $\|.\|_*$ are said to be *equivalent* if and only if there exist constants c and C, where $0 < c \leq C$, such that

$$c\|x\| \leq \|x\|_* \leq C\|x\|$$

for all $x \in X$.

Lemma 6.25 *Two norms $\|.\|$ and $\|.\|_*$ on a real or complex vector space X are equivalent if and only if they induce the same topology on X.*

Proof Suppose that the norms $\|.\|$ and $\|.\|_*$ induce the same topology on X. Then there exists some $\delta > 0$ such that

$$\{x \in X : \|x\| < \delta\} \subset \{x \in X : \|x\|_* < 1\},$$

since the set $\{x \in X : \|x\|_* < 1\}$ is open with respect to the topology on X induced by both $\|.\|_*$ and $\|.\|$. Let C be any positive real number satisfying $C\delta > 1$. Then

$$\left\| \frac{1}{C\|x\|} x \right\| = \frac{1}{C} < \delta,$$

and hence

$$\|x\|_* = C\|x\| \left\| \frac{1}{C\|x\|} x \right\|_* < C\|x\|.$$

for all non-zero elements x of X, and thus $\|x\|_* \leq C\|x\|$ for all $x \in X$. On interchanging the roles of the two norms, we deduce also that there exists a positive real number c such that $\|x\| \leq (1/c)\|x\|_*$ for all $x \in X$. But then $c\|x\| \leq \|x\|_* \leq C\|x\|$ for all $x \in X$. We conclude that the norms $\|.\|$ and $\|.\|_*$ are equivalent.

Conversely suppose that the norms $\|.\|$ and $\|.\|_*$ are equivalent. Then there exist constants c and C, where $0 < c \leq C$, such that $c\|x\| \leq \|x\|_* \leq C\|x\|$ for all $x \in X$. Let U be a subset of X that is open with respect to the topology on X induced by the norm $\|.\|_*$, and let $u \in U$. Then there exists some $\delta > 0$ such that

$$\{x \in X : \|x - u\|_* < C\delta\} \subset U.$$

But then

$$\{x \in X : \|x - u\| < \delta\} \subset \{x \in X : \|x - u\|_* < C\delta\} \subset U,$$

showing that U is open with respect to the topology induced by the norm $\|.\|$. Similarly any subset of X that is open with respect to the topology induced by the norm $\|.\|$ must also be open with respect to the topology induced by $\|.\|_*$. Thus equivalent norms induce the same topology on X. ∎

It follows immediately from Lemma 6.25 that if $\|.\|$, $\|.\|_*$ and $\|.\|_\sharp$ are norms on a real (or complex) vector space X, if the norms $\|.\|$ and $\|.\|_*$ are equivalent, and if the norms $\|.\|_*$ and $\|.\|_\sharp$ are equivalent, then the norms $\|.\|$ and $\|.\|_\sharp$ are also equivalent. This fact can easily be verified directly from the definition of equivalence of norms.

We recall that the usual topology on $\mathbb{R}^n$ is that generated by the Euclidean norm on $\mathbb{R}^n$.

Lemma 6.26 *Let $\|.\|$ be a norm on $\mathbb{R}^n$. Then the function $\mathbf{x} \mapsto \|\mathbf{x}\|$ is continuous with respect to the usual topology on on $\mathbb{R}^n$.*

Proof Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ denote the basis of $\mathbb{R}^n$ given by

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \cdots, \quad \mathbf{e}_n = (0, 0, 0, \ldots, 1).$$

Let $\mathbf{x}$ and $\mathbf{y}$ be points of $\mathbb{R}^n$, given by

$$\mathbf{x} = (x_1, x_2, \ldots, x_n), \qquad \mathbf{y} = (y_1, y_2, \ldots, y_n).$$

Using Schwarz' Inequality, we see that

$$\begin{aligned} \|\mathbf{x} - \mathbf{y}\| &= \left\| \sum_{j=1}^{n} (x_j - y_j)\mathbf{e}_j \right\| \leq \sum_{j=1}^{n} |x_j - y_j| \, \|\mathbf{e}_j\| \\ &\leq \left(\sum_{j=1}^{n} (x_j - y_j)^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^{n} \|\mathbf{e}_j\|^2 \right)^{\frac{1}{2}} = C\|\mathbf{x} - \mathbf{y}\|_2, \end{aligned}$$

where

$$C^2 = \|\mathbf{e}_1\|^2 + \|\mathbf{e}_2\|^2 + \cdots + \|\mathbf{e}_n\|^2$$

and $\|\mathbf{x} - \mathbf{y}\|_2$ denotes the Euclidean norm of $\mathbf{x} - \mathbf{y}$, defined by

$$\|\mathbf{x} - \mathbf{y}\|_2 = \left(\sum_{j=1}^{n} (x_j - y_j)^2 \right)^{\frac{1}{2}}.$$

Also $|\|\mathbf{x}\| - \|\mathbf{y}\|| \leq \|\mathbf{x} - \mathbf{y}\|$, since

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

We conclude therefore that

$$|\|\mathbf{x}\| - \|\mathbf{y}\|| \leq C\|\mathbf{x} - \mathbf{y}\|_2,$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, and thus the function $\mathbf{x} \mapsto \|\mathbf{x}\|$ is continuous on $\mathbb{R}^n$ (with respect to the usual topology on $\mathbb{R}^n$). ∎

Theorem 6.27 *Any two norms on $\mathbb{R}^n$ are equivalent, and induce the usual topology on $\mathbb{R}^n$.*

Proof Let $\|.\|$ be any norm on $\mathbb{R}^n$. We show that $\|.\|$ is equivalent to the Euclidean norm $\|.\|_2$. Let S^{n-1} denote the unit sphere in $\mathbb{R}^n$, defined by

$$S^{n-1} = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = 1\},$$

and let $f\colon S^{n-1} \to \mathbb{R}$ be the real-valued function on S^{n-1} defined such that $f(\mathbf{x}) = \|\mathbf{x}\|$ for all $\mathbf{x} \in S^{n-1}$. Now the function f is a continuous function on S^{n-1} (Lemma 6.26). Also the function f is non-zero at each point of S^{n-1}, and therefore the function sending $\mathbf{x} \in S^{n-1}$ to $1/f(\mathbf{x})$ is continuous. Now any closed bounded set in $\mathbb{R}^n$ is compact (Theorem 6.16), and any continuous real-valued function on a compact topological space is bounded (Lemma 6.5). It follows that there exist positive real numbers C and D such that $f(\mathbf{x}) \leq C$ and $1/f(\mathbf{x}) \leq D$ for all $\mathbf{x} \in S^{n-1}$. Let $c = D^{-1}$. Then $c \leq \|\mathbf{x}\| \leq C$ for all $\mathbf{x} \in S^{n-1}$.

Now

$$\frac{\|\mathbf{x}\|}{\|\mathbf{x}\|_2} = f\left(\|\mathbf{x}\|_2^{-1}\mathbf{x}\right)$$

for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. (This is an immediate consequence of the fact that $\|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$.) It follows that $c\|\mathbf{x}\|_2 \leq \|\mathbf{x}\| \leq C\|\mathbf{x}\|_2$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. These inequalities also hold when $\mathbf{x} = \mathbf{0}$. The result follows. ∎