

APPENDIX A

Metric Spaces

For those readers not already familiar with the elementary properties of metric spaces and the notion of compactness, this appendix presents a sufficiently detailed treatment for a reasonable understanding of this subject matter. However, for those who have already had some exposure to elementary point set topology (or even a solid introduction to real analysis), then the material in this appendix should serve as a useful review of some basic concepts. Besides, any mathematics or physics student should become thoroughly familiar with all of this material.

Let S be any set. Then a function $d: S \times S \rightarrow \mathbb{R}$ is said to be a **metric** on S if it has the following properties for all $x, y, z \in S$:

- (M1) $d(x, y) \geq 0$;
- (M2) $d(x, y) = 0$ if and only if $x = y$;
- (M3) $d(x, y) = d(y, x)$;
- (M4) $d(x, y) + d(y, z) \geq d(x, z)$.

The real number $d(x, y)$ is called the **distance** between x and y , and the set S together with a metric d is called a **metric space** (S, d) .

As a simple example, let $S = \mathbb{R}$ and let $d(x, y) = |x - y|$ for all $x, y \in \mathbb{R}$. From the properties of the absolute value, conditions (M1) – (M3) should be obvious, and (M4) follows by simply noting that

$$|x - z| = |x - y + y - z| \leq |x - y| + |y - z| .$$

For our purposes, we point out that given any normed vector space $(V, \|\cdot\|)$ we may treat V as a metric space by defining

$$d(x, y) = \|x - y\|$$

for every $x, y \in V$. Using Theorem 2.17, the reader should have no trouble showing that this does indeed define a metric space (V, d) . In fact, it is easy to see that $\mathbb{R}^n$ forms a metric space relative to the standard inner product and its associated norm.

Given a metric space (X, d) and any real number $r > 0$, the **open ball** of **radius** r and **center** x_0 is the set $B_d(x_0, r) \subset X$ defined by

$$B_d(x_0, r) = \{x \in X: d(x, x_0) < r\} .$$

Since the metric d is usually understood, we will generally leave off the subscript d and simply write $B(x_0, r)$. Such a set is frequently referred to as an **r -ball**. We say that a subset U of X is **open** if, given any point $x \in U$, there exists $r > 0$ and an open ball $B(x, r)$ such that $B(x, r) \subset U$.

Probably the most common example of an open set is the open unit disk D_1 in $\mathbb{R}^2$ defined by

$$D_1 = \{(x, y) \in \mathbb{R}^2: x^2 + y^2 < 1\} .$$

We see that given any point $x_0 \in D_1$, we can find an open ball $B(x_0, r) \subset D_1$ by choosing $r = 1 - d(x_0, 0)$. The set

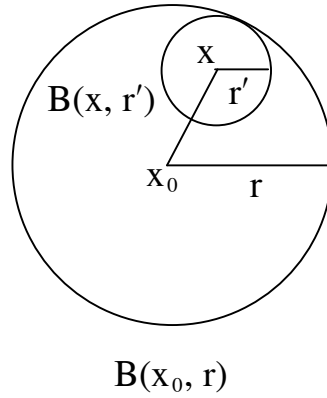
$$D_2 = \{(x, y) \in \mathbb{R}^2: x^2 + y^2 \leq 1\}$$

is not open because there is no open ball centered on any of the boundary points $x^2 + y^2 = 1$ that is contained entirely within D_2 .

The fundamental characterizations of open sets are contained in the following three theorems.

Theorem A1 Let (X, d) be a metric space. Then any open ball is an open set.

Proof Let $B(x_0, r)$ be an open ball in X and let x be any point in $B(x_0, r)$. We must find a $B(x, r')$ contained in $B(x_0, r)$.



Since $d(x, x_0) < r$, we define $r' = r - d(x, x_0)$. Then for any $y \in B(x, r')$ we have $d(y, x) < r'$, and hence

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

which shows that $y \in B(x_0, r)$. Therefore $B(x, r') \subset B(x_0, r)$. ■

Theorem A2 Let (X, d) be a metric space. Then

- (a) Both X and $\emptyset$ are open sets.
- (b) The intersection of a finite number of open sets is open.
- (c) The union of an arbitrary number of open sets is open.

Proof (a) X is clearly open since for any $x \in X$ and $r > 0$ we have $B(x, r) \subset X$. The statement that $\emptyset$ is open is also automatically satisfied since for any $x \in \emptyset$ (there are none) and $r > 0$, we again have $B(x, r) \subset \emptyset$.

(b) Let $\{U_i\}$, $i \in I$, be a finite collection of open sets in X . Suppose $\{U_i\}$ is empty. Then $\cap U_i = X$ because a point is in the intersection of a collection of sets if it belongs to each set in the collection, so if there are no sets in the collection, then every point of X satisfies this requirement. Hence $\cap U_i = X$ is open by (a). Now assume that $\{U_i\}$ is not empty, and let $U = \cap U_i$. If $U = \emptyset$ then it is open by (a), so assume that $U \neq \emptyset$. Suppose $x \in U$ so that $x \in U_i$ for every $i \in I$. Therefore there exists $B(x, r_i) \subset U_i$ for each i , and since there are only a *finite* number of the r_i we may let $r = \min\{r_i\}$. It follows that

$$B(x, r) \subset B(x, r_i) \subset U_i$$

for every i , and hence $B(x, r) \subset \cap U_i = U$. In other words, we have found an open ball centered at each point of U and contained in U , thus proving that U is open.

(c) Let $\{U_i\}$ be an arbitrary collection of open sets. If $\{U_i\}$ is empty, then $U = \bigcup U_i = \emptyset$ is open by (a). Now suppose that $\{U_i\}$ is not empty and $x \in \bigcup U_i$. Then $x \in U_i$ for some i , and hence there exists $B(x, r_i) \subset U_i \subset \bigcup U_i$ so that $\bigcup U_i$ is open. ■

Notice that part (b) of this theorem requires that the collection be finite. To see the necessity of this, consider the infinite collection of intervals in $\mathbb{R}$ given by $(-1/n, 1/n)$ for $1 \leq n < \infty$. The intersection of these sets is the point $\{0\}$ which is not open in $\mathbb{R}$.

In an arbitrary metric space the structure of the open sets can be very complicated. However, the most general description of an open set is contained in the following.

Theorem A3 A subset U of a metric space (X, d) is open if and only if it is the union of open balls.

Proof Assume U is the union of open balls. By Theorem A1 each open ball is an open set, and hence U is open by Theorem A2(c). Conversely, let U be an open subset of X . For each $x \in U$ there exists at least one $B(x, r) \subset U$, so that $\bigcup_{x \in U} B(x, r) \subset U$. On the other hand each $x \in U$ is contained in at least $B(x, r)$ so that $U \subset \bigcup_{x \in U} B(x, r)$. Therefore $U = \bigcup B(x, r)$. ■

As a passing remark, note that a set is never open in and of itself. Rather, a set is open only with respect to a specific metric space containing it. For example, the set of numbers $[0, 1)$ is not open when considered as a subset of the real line because any open interval about the point 0 contains points not in $[0, 1)$. However, if $[0, 1)$ is considered to be the entire space X , then it is open by Theorem A2(a).

If U is an open subset of a metric space (X, d) , then its complement $U^c = X - U$ is said to be **closed**. In other words, a set is closed if and only if its complement is open. For example, a moments thought should convince you that the subset of $\mathbb{R}^2$ defined by $\{(x, y) \in \mathbb{R}^2: x^2 + y^2 \leq 1\}$ is a closed set. The **closed ball** of radius r centered at x_0 is the set $B[x_0, r]$ defined in the obvious way by

$$B[x_0, r] = \{x \in X: d(x_0, x) \leq r\} .$$

We leave it to the reader (see Exercise A.3) to prove the closed set analogue of Theorem A2. The important difference to realize is that the intersection of an arbitrary number of closed sets is closed, while only the union of a finite number of closed sets is closed.

If (X, d) is a metric space and $Y \subset X$, then Y may be considered a metric space in its own right with the same metric d used on X . In other words, if we

let $d|Y$ denote the metric d restricted to points in Y , then the space $(Y, d|Y)$ is said to be a **subspace** of the metric space (X, d) .

Theorem A4 Let (X, d) be a metric space and $(Y, d|Y)$ a metric subspace of X . Then a subset $W \subset Y$ is open in Y (i.e., open with respect to the metric $d|Y$) if and only if $W = Y \cap U$ where U is open in X .

Proof Let $W \subset Y$ be open in Y and suppose $x \in W$. Then there exists $r > 0$ such that the set

$$B_{d|Y}(x, r) = \{y \in Y: (d|Y)(x, y) < r\}$$

is a subset of W . But this is clearly the same as the open set

$$B_d(x, r) = \{y \in X: d(x, y) < r\}$$

restricted to only those points y that are in Y . Another way of saying this is that

$$B_{d|Y}(x, r) = B_d(x, r) \cap Y .$$

Since $W = \bigcup_{x \in W} B_{d|Y}(x, r)$, it follows that (see Exercise 0.1.1(b)) $W = U \cap Y$ where $U = \bigcup_{x \in W} B_d(x, r)$ is open in X (by Theorem A2(c)).

On the other hand, let $W = Y \cap U$ where U is open in X , and suppose $x \in W$. Then $x \in U$ so there exists $r > 0$ with

$$B_d(x, r) = \{y \in X: d(x, y) < r\} \subset U .$$

But $Y \cap B_d(x, r)$ is just $B_{d|Y}(x, r) = \{y \in Y: (d|Y)(x, y) < r\} \subset W$ which shows that W is open in Y . ■

Note that all of our discussion on metric spaces also applies to normed vector spaces where $d(x, y) = \|x - y\|$. Because of this, we can equally well discuss open sets in any normed space V .

Let $f: (X, d_X) \rightarrow (Y, d_Y)$ be a mapping. We say that f is **continuous at** $x_0 \in X$ if, given any real number $\varepsilon > 0$, there exists a real number $\delta > 0$ such that $d_X(f(x), f(x_0)) < \varepsilon$ for every $x \in X$ with $d_Y(x, x_0) < \delta$. Equivalently, f is continuous at x_0 if for each $B(f(x_0), \varepsilon)$ there exists $B(x_0, \delta)$ such that $f(B(x_0, \delta)) \subset B(f(x_0), \varepsilon)$. (Note that these open balls are defined with respect to two different metrics since they are in different spaces. We do not want to clutter the notation by adding subscripts such as d_X and d_Y to B .) In words, “if you tell me how close you wish to get to the number $f(x_0)$, then I will tell you how close x must be to x_0 in order that $f(x)$ be that close.” If f is defined

on a subset $S \subset X$, then f is said to be **continuous on S** if f is continuous at every point of S .

For example, consider the mapping $f: (0, \infty) \subset \mathbb{R} \rightarrow (0, \infty) \subset \mathbb{R}$ defined by $f(x) = 1/x$. For any $x_0 \in (0, \infty)$ we have (using the absolute value as our metric)

$$|f(x) - f(x_0)| = |1/x - 1/x_0| = |x_0 - x|/|x x_0| .$$

If x is such that $|x - x_0| < \delta$, then we see that

$$|f(x) - f(x_0)| < \delta/|x x_0| = \delta/(x x_0) .$$

In particular, choosing $\delta \leq x_0/2$, it follows that $x \geq x_0/2$ (since $|x - x_0| < \delta = x_0/2$), and hence $\delta/(x x_0) < 2\delta/x_0^2$. Therefore, given any $\varepsilon > 0$, if we pick $\delta = \min\{x_0/2, \varepsilon x_0^2/2\}$ then we will have $|f(x) - f(x_0)| < \varepsilon$.

Fortunately one can usually tell by inspection (i.e., by drawing a picture if necessary) whether or not a particular function is continuous without resorting to clever calculations. The general definition is a powerful technique for proving theorems about classes of continuous functions satisfying given properties. Moreover, there is an intrinsic way to characterize continuous mappings that is of the utmost importance.

Theorem A5 Let $f: (X, d_X) \rightarrow (Y, d_Y)$. Then f is continuous if and only if $f^{-1}(U)$ is open in X for all open sets U in Y .

Proof Suppose f is continuous and U is an open subset of Y . If $x \in f^{-1}(U)$, then $f(x) \in U$ so there exists $B(f(x), \varepsilon) \subset U$ (since U is open). But the continuity of f then implies that there exists $B(x, \delta)$ such that

$$f(B(x, \delta)) \subset B(f(x), \varepsilon) \subset U .$$

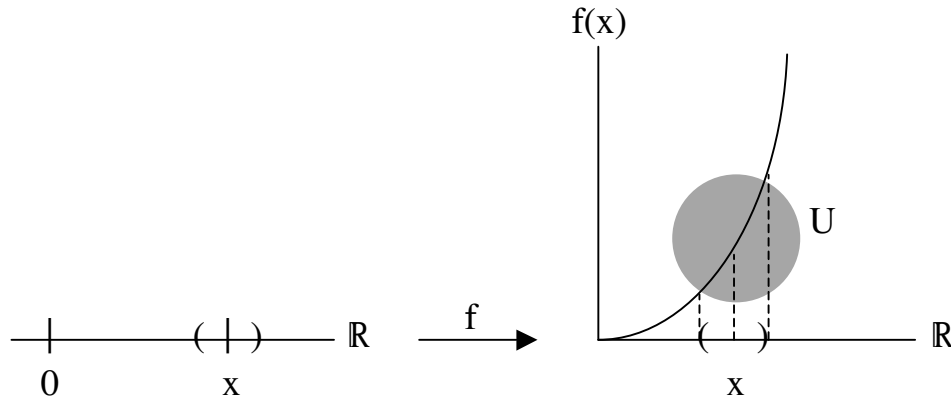
Therefore $B(x, \delta)$ is an r -ball centered on x and contained in $f^{-1}(U)$, and hence $f^{-1}(U)$ is open.

Conversely, assume that $f^{-1}(U)$ is open whenever U is, and let $x \in X$ be arbitrary. Then the open ball $B(f(x), \varepsilon)$ is an open set, so its inverse image is an open set containing x . Therefore there exists an open ball $B(x, \delta)$ contained in this inverse image, and it clearly has the property that $f(B(x, \delta)) \subset B(f(x), \varepsilon)$, hence proving that f is continuous. ■

Corollary If $f: (X, d_X) \rightarrow (Y, d_Y)$, then f is continuous if and only if $f^{-1}(F)$ is closed in X whenever F is closed in Y .

Proof It was shown in Exercise 0.2.1 that if $A \subset Y$, then $f^{-1}(A^c) = f^{-1}(A)^c$. Therefore if $F \subset Y$ is closed, then $F = U^c$ for some open set $U \subset Y$ and so by Theorem A5, $f^{-1}(F) = f^{-1}(U^c) = f^{-1}(U)^c$ must be closed if and only if f is continuous. ■

Note that if $f: X \rightarrow Y$ is continuous and $U \subset Y$ is open, then $f^{-1}(U)$ is open, but if $A \subset X$ is open, then it is not necessarily true that $f(A)$ is open. As a simple example, consider the function $f: \mathbb{R} \rightarrow \mathbb{R}^2$ defined by $f(x) = (x, x^2)$. It should be clear that the open ball $U \subset \mathbb{R}^2$ shown below is an open set whose inverse image is an open interval on $\mathbb{R} \cup \{\emptyset\}$ (since some points of U are not the image under f of any point in $\mathbb{R}$), but that the image under f of an open interval is part of the parabola $y = x^2$ which is not open as a subset of $\mathbb{R}^2$.



Now suppose that (X, d) is a metric space, and let $\{U_i\}$ be a collection of open subsets of X such that $\bigcup U_i = X$. Such a collection of subsets is called an **open cover** of X . A subcollection $\{V_j\}$ of the collection $\{U_i\}$ is said to be an **open subcover** of X if $\bigcup V_j = X$. A space (X, d) is said to be **compact** if *every* open cover has a finite subcover. Similarly, given a subset $A \subset X$, a collection $\{U_i\}$ of open subsets of X with the property that $A \subset \bigcup U_i$ is said to be an **open cover** of A . Equivalently, the collection $\{U_i\}$ of open subsets of X is an open cover of A in X if the collection $\{U_i \cap A\}$ is an open cover of the subset A in the metric $d|_A$ (i.e., in the subspace A). We then say that A is **compact** if every open cover of A has a finite subcover, or equivalently, A is compact if the subspace A is compact. While this is not a particularly easy concept to thoroughly understand and appreciate without detailed study, its importance to us is based on the following two examples.

Example A1 Consider the subset $A = (0, 1)$ of the real line $\mathbb{R}$. We define the collection $\{U_1, U_2, \dots\}$ of open sets by

$$U_n = (1/2^{n+1}, 1 - 1/2^{n+1}) .$$

Thus $U_1 = (1/4, 3/4)$, $U_2 = (1/8, 7/8)$ etc. The collection $\{U_n\}$ clearly covers A since for any $x \in (0, 1)$ we can always find some U_n such that $x \in U_n$. However, A is not compact since given any finite number of the U_n there exists $\varepsilon > 0$ (so that $\varepsilon \in (0, 1)$) which is not in any of the U_n . //

Example A2 Let us show that the subspace $[0, 1]$ of the real line is compact. This is sometimes called the **Heine-Borel theorem**, although we shall prove a more general version below.

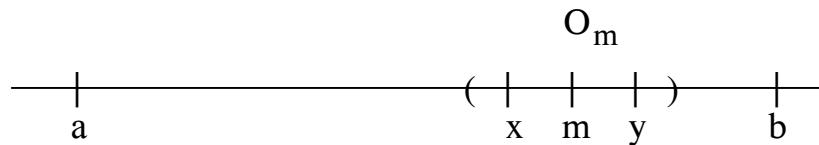
First note that the points 0 and 1 which are included in the subspace $[0, 1]$ are not in the set $(0, 1)$ discussed in the previous example. However, if we have positive real numbers a and b with $a \leq b < 1$, then the collection $\{U_n\}$ defined above together with the sets $[0, a)$ and $(b, 1]$ does indeed form an open cover for $[0, 1]$ (the sets $[0, a)$ and $(b, 1]$ are open by Theorem A4). It should be clear that given the sets $[0, a)$ and $(b, 1]$ we can now choose a finite cover of $[0, 1]$ by including these sets along with a finite number of the U_n . To prove that $[0, 1]$ is compact however, we must show that *any* open cover has a finite subcover.

Somewhat more generally, let $\{O_n\}$ be any open cover of the interval $[a, b]$ in $\mathbb{R}$. Define

$$A = \{x \in [a, b]: [a, x] \text{ is covered by a finite number of the } O_n\} .$$

We see that $A \neq \emptyset$ since clearly $a \in A$, and furthermore A is bounded above by b . Therefore (by the Archimedean axiom) A must have a least upper bound $m = \sup A \leq b$. If A is to be compact, then we must have $b \in A$. We will show that this is true by first proving that $m \in A$, and then that $m = b$.

Since $\{O_n\}$ covers $[a, b]$ and $m \in [a, b]$, it follows that $m \in O_m$ for some $O_m \in \{O_n\}$. Now, O_m is an open subset of $[a, b]$, and hence there are points in O_m that are less than m , and points in O_m that are greater than m .



Since $m = \sup A$, there is an $x < m$ with $x \in O_m$ such that the interval $[a, x]$ is covered by a finite number of the O_n , while $[x, m]$ is covered by the single set O_m . Therefore $[a, m]$ is covered by a finite number of open sets so that $m \in A$.

Now suppose that $m < b$. Then there is a point y with $m < y < b$ such that $[m, y] \subset O_m$. But we just showed that $m \in A$, so the interval $[a, m]$ is covered by finitely many O_n while $[m, y]$ is covered by O_m . Therefore $y \in A$ which contradicts the definition of m , and hence we must have $m = b$. //

An important property of metric spaces is the following. Given two distinct points $x, y \in X$, there exist disjoint open sets U and V in X such that $x \in U$ and $y \in V$. That this does indeed hold for metric spaces is easy to prove by considering open balls of radius $d(x, y)/2$ centered on each of the points x and y . This property is called the **Hausdorff property**. We sometimes refer to a metric space as a “Hausdorff space” if we wish to emphasize this property.

The following theorems describe some of the most fundamental properties of compact spaces.

Theorem A6 Any closed subset of a compact space is compact.

Proof Let $F \subset X$ be a closed subset of a compact space X . If $\{U_i\}$ is any open cover of F , then $(\cup U_i) \cup F^c$ is an open cover of X . Since X is compact, we may select a finite subcover by choosing F^c along with a finite number of the U_i . But then F is covered by this finite subcollection of the U_i , and hence F is compact. ■

Theorem A7 Any compact subset of a metric space is closed.

Proof Let F be a compact subset of a metric space X . We will show that F^c is open. Fix any $x \in F^c$ and suppose $y \in F$. Since X is Hausdorff, there exist open sets U_y and V_y such that $x \in U_y$, $y \in V_y$ and $U_y \cap V_y = \emptyset$. As the point y varies over F , we see that $\{V_y: y \in F\}$ is an open cover for F . Since F is compact, a finite number, say $V_{y_1}, \dots, V_{y_n}$, will cover F . Corresponding to each V_{y_i} there is a U_{y_i} , and we let $U = \cap_i U_{y_i}$ and $V = \cup_i V_{y_i}$. By construction $x \in U$, $F \subset V$ and $U \cap V = \emptyset$. But then U is an open set containing x such that $U \cap F = \emptyset$, and hence F^c is open. ■

Theorem A8 Let (X, d_X) be a compact space and let f be a continuous function from X onto a space (Y, d_Y) . Then Y is compact.

Proof Let $\{U_i\}$ be any open cover of Y . Since f is continuous, each $f^{-1}(U_i)$ is open in X , and hence $\{f^{-1}(U_i)\}$ is an open cover for X . But X is compact, so that a finite number of the $f^{-1}(U_i)$, say $\{f^{-1}(U_{i_1}), \dots, f^{-1}(U_{i_n})\}$ cover X . Therefore $\{U_{i_1}, \dots, U_{i_n}\}$ form a finite subcover for Y , and hence Y is compact. ■

Theorem A9 Let $\{K_i\}$ be a collection of compact subsets of a metric space X , such that the intersection of every finite subcollection of $\{K_i\}$ is nonempty. Then $\bigcap K_i$ is nonempty.

Proof Fix any $K_1 \in \{K_i\}$ and assume that $K_1 \cap (\bigcap_{i \neq 1} K_i) = \emptyset$. We will show that this leads to a contradiction. First note that by our assumption we have $(\bigcap_{i \neq 1} K_i) \subset K_1^c$, and hence from Example 0.1 and Theorem 0.1 we see that

$$K_1 \subset \bigcap_{i \neq 1} K_i^c = \bigcup_{i \neq 1} K_i^c.$$

Thus $\{K_i^c\}$, $i \neq 1$, is an open cover of K_1 . But K_1 is compact so that a finite number of these sets, say $K_{i_1}^c, \dots, K_{i_n}^c$ cover K_1 . Then

$$K_1 \subset (\bigcup_{\alpha=1}^n K_{i_\alpha}^c) = (\bigcap_{\alpha=1}^n K_{i_\alpha})^c$$

which implies

$$K_1 \cap (\bigcap_{\alpha=1}^n K_{i_\alpha}) = \emptyset.$$

However, this contradicts the hypothesis of the theorem. ■

Corollary If $\{K_n\}$ is a *sequence* of nonempty compact sets such that $K_n \supset K_{n+1}$, then $\bigcap K_n \neq \emptyset$.

Proof This is an obvious special case of Theorem A9. ■

As a particular application of this corollary we see that if $\{I_n\}$ is a nonempty sequence of intervals $[a_n, b_n] \subset \mathbb{R}$ such that $I_n \supset I_{n+1}$, then $\bigcap I_n \neq \emptyset$. While this result is based on the fact that each I_n is compact (Example A2), we may also prove this directly as follows. If $I_n = [a_n, b_n]$, we let $S = \{a_n\}$. Then $S \neq \emptyset$ and is bounded above by b_1 . By the Archimedean axiom, we let $x = \sup S$. For any $m, n \in \mathbb{Z}^+$ we have $a_n \leq a_{m+n} \leq b_{m+n} \leq b_m$ so that $x \leq b_m$ for all m . Since $a_m \leq x$ for all m , we must have $x \in [a_m, b_m] = I_m$ for each $m = 1, 2, \dots$ so that $\bigcap I_n \neq \emptyset$. We now show that this result holds in $\mathbb{R}^n$ as well.

Suppose $a, b \in \mathbb{R}^n$ where $a^i < b^i$ for each $i = 1, \dots, n$. By an **n-cell** we mean the set of all points $x \in \mathbb{R}^n$ such that $a^i \leq x^i \leq b^i$ for every i . In other words, an n-cell is just an n-dimensional rectangle.

Theorem A10 Let $\{I_k\}$ be a sequence of n-cells such that $I_k \supset I_{k+1}$. Then $\bigcap I_k \neq \emptyset$.

Proof For each $k = 1, 2, \dots$ the n -cell I_k consists of all points $x \in \mathbb{R}^n$ with the property that $a_k^i \leq x^i \leq b_k^i$ for every $1 \leq i \leq n$, so we let $I_k^i = [a_k^i, b_k^i]$. Now, for each $i = 1, \dots, n$ the sequence $\{I_k^i\}$ satisfies the hypotheses of the corollary to Theorem A9. Hence for each $i = 1, \dots, n$ there exists $z^i \in [a_k^i, b_k^i]$ for every $k = 1, 2, \dots$. If we define $z = (z^1, \dots, z^n) \in \mathbb{R}^n$, we see that $z \in I_k$ for every $k = 1, 2, \dots$. ■

Theorem A11 Every n -cell is compact.

Proof Let I be an n -cell as defined above, and set $\delta = [\sum_{i=1}^n (b_i - a_i)^2]^{1/2}$. Then if $x, y \in I$ we have $\|x - y\| \leq \delta$ (see Example 2.9). Let $\{U_i\}$ be any open cover of I and assume that it contains no finite subcover. We will show that this leads to a contradiction.

Let $c^j = (a^j + b^j)/2$ for each $j = 1, \dots, n$. Then we have 2^n n -cells Q_i defined by the intervals $[a^j, c^j]$ and $[c^j, b^j]$ such that $\cup Q_i = I$. Since I has no finite subcover, at least one of the Q_i , which we call I_1 , can not be covered by any finite number of the U_i . Next we subdivide I_1 into another 2^n n -cells and continue in the same manner. We thus obtain a sequence $\{I_\alpha\}$ of n -cells with the following properties:

- (a) $I \supset I_1 \supset I_2 \supset \dots$;
- (b) I_α is not covered by any finite subcollection of the U_k ;
- (c) $x, y \in I_\alpha$ implies $\|x - y\| \leq 2^{-\alpha} \delta$.

By (a) and Theorem A10, there exists $z \in \cap I_\alpha$, and since $\{U_i\}$ covers I , we must have $z \in U_k$ for some k . Now, U_k is an open set in the metric space $\mathbb{R}^n$, so there exists $\varepsilon > 0$ such that $\|z - y\| < \varepsilon$ implies that $y \in U_k$. If we choose α sufficiently large that $2^{-\alpha} \delta < \varepsilon$ (that this can be done follows from Theorem 0.3), then (c) implies that $I_\alpha \subset U_k$ which contradicts (b). ■

We are now in a position to prove the generalized Heine-Borel theorem. Before doing so however, we first prove a simple result which is sometimes taken as the definition of a compact set. By way of terminology, any open set U containing a point x is said to be a **neighborhood** of x , and the set $U - \{x\}$ is called a **deleted neighborhood** of x . We say that a point $x \in (X, d)$ is an **accumulation point** of $A \subset X$ if every *deleted* neighborhood of x intersects A .

Theorem A12 Any infinite subset A of a compact set K has a point of accumulation in K .

Proof Suppose every point $x \in K$ is not an accumulation point of A . Then there exists a neighborhood U_x of x such that U_x contains at most a single point of A , namely x itself if $x \in A$. Then clearly no finite subcollection of $\{U_x\}$ covers $A \subset K$ so that K can not possibly be compact. ■

Theorem A13 A subset A of a metric space (X, d) is closed if and only if A contains all of its accumulation points.

Proof First suppose that A is closed. Let $x \in X$ be an accumulation point of A and assume that $x \notin A$. Then $x \in A^c$ which is an open set containing x that does not intersect A , and hence contradicts the fact that x is an accumulation point of A . Therefore x must be an element of A .

Conversely, suppose A contains all of its accumulation points. We show that A^c is open. If $x \in A^c$ and hence is not an accumulation point of A , then there exists an open set U containing x such that $A \cap U = \emptyset$. But then $x \in U \subset A^c$ which implies that A^c is open. ■

We say that a subset $A \subset \mathbb{R}^n$ is **bounded** if it can be enclosed in some n -cell. The equivalence of (a) and (b) in the next theorem is called the **(generalized) Heine-Borel theorem**, while the equivalence of (a) and (c) is a general version of the **Bolzano-Weierstrass theorem**.

Theorem A14 Let A be a subset of $\mathbb{R}^n$. Then the following three properties are equivalent:

- (a) A is closed and bounded.
- (b) A is compact.
- (c) Every infinite subset of A has a point of accumulation in A .

Proof (a) $\Rightarrow$ (b): If (a) holds, then A can be enclosed by some n -cell which is compact by Theorem A11. But then A is compact by Theorem A6.

(b) $\Rightarrow$ (c): This follows from Theorem A12.

(c) $\Rightarrow$ (a): We assume that every infinite subset of A has an accumulation point in A . Let us first show that A must be bounded. If A is not bounded, then for each positive integer $k = 1, 2, \dots$ we can find an $x_k \in A$ such that $\|x_k\| > k$. Then the set $\{x_k\}$ is clearly infinite but contains no point of accumulation in $\mathbb{R}^n$, so it certainly contains none in A . Hence A must be bounded.

We now show that A must be closed. Again assume the contrary. Then there exists $x_0 \in \mathbb{R}^n$ which is an accumulation point of A but which does not belong to A (Theorem A13). This means that for each $k = 1, 2, \dots$ there exists $x_k \in A$ such that $\|x_k - x_0\| < 1/k$. The set $S = \{x_k\}$ is then an infinite subset of

A with x_0 as an accumulation point. Since $x_0 \notin A$, we will be finished if we can show that S has no accumulation point in A (because the assumption that A is not closed then leads to a contradiction with the property described in (c)).

First note that if $a, b \in \mathbb{R}^n$, then Example 2.11 shows us that

$$\|a + b\| = \|a - (-b)\| \geq \|a\| - \|b\|.$$

Using this result, if y is any point of $\mathbb{R}^n$ other than x_0 we have

$$\begin{aligned} \|x_k - y\| &= \|x_k - x_0 + x_0 - y\| \\ &\geq \|x_0 - y\| - \|x_k - x_0\| \\ &> \|x_0 - y\| - 1/k. \end{aligned}$$

No matter how large (or small) $\|x_0 - y\|$ is, we can always find a $k_0 \in \mathbb{Z}^+$ such that $1/k \leq (1/2)\|x_0 - y\|$ for every $k \geq k_0$ (this is just Theorem 0.3). Hence

$$\|x_k - y\| > (1/2)\|x_0 - y\|$$

for every $k \geq k_0$. This shows that y can not possibly be an accumulation point of $\{x_k\} = S$ (because the open ball of radius $(1/2)\|x_0 - y\|$ centered at y can contain at most a finite number of elements of S). ■

We remark that the implication “(a) implies (b)” in this theorem is not true in an arbitrary metric space (see Exercise A.5).

Let f be a mapping from a set A into $\mathbb{R}^n$. Then f is said to be **bounded** if there exists a real number M such that $\|f(x)\| \leq M$ for all $x \in A$. If f is a continuous mapping from a compact space X into $\mathbb{R}^n$, then $f(X)$ is compact (Theorem A8) and hence closed and bounded (Theorem A14). Thus we see that any continuous function from a compact set into $\mathbb{R}^n$ is bounded. On the other hand, note that the function $f: \mathbb{R} \rightarrow \mathbb{R}$ defined by $f(x) = 1/x$ is not bounded on the interval $(0, 1)$. We also see that the function $g: \mathbb{R} \rightarrow \mathbb{R}$ defined by $g(x) = x$ for $x \in [0, 1)$ never attains a maximum value, although it gets arbitrarily close to 1. Note that both f and g are defined on non-compact sets.

We now show that a continuous function defined on a compact space takes on its maximum and minimum values at some point of the space.

Theorem A15 Let f be a continuous real-valued function defined on a compact space X , and let $M = \sup_{x \in X} f(x)$ and $m = \inf_{x \in X} f(x)$. Then there exist points $p, q \in X$ such that $f(p) = M$ and $f(q) = m$.

Proof The above discussion showed that $f(X)$ is a closed and bounded subset of $\mathbb{R}$. Hence by the Archimedean axiom, $f(X)$ must have a sup and an inf. Let $M = \sup f(x)$. This means that given $\varepsilon > 0$, there exists $x \in X$ such that

$$M - \varepsilon < f(x) < M$$

(or else M would not be the least upper bound of $f(X)$). This just says that any open ball centered on M intersects $f(X)$, and hence M is an accumulation point of $f(X)$. But $f(X)$ is closed so that Theorem A13 tells us that $M \in f(X)$. In other words, there exists $p \in X$ such that $M = f(p)$. The proof for the minimum is identical. ■

As an application of these ideas, we now prove the Fundamental Theorem of Algebra.

Theorem A16 (Fundamental Theorem of Algebra) The complex number field $\mathbb{C}$ is algebraically closed.

Proof Consider the non-constant polynomial

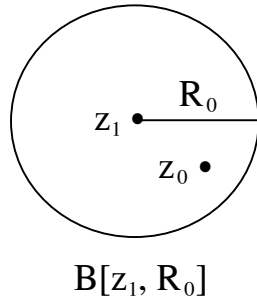
$$f(z) = a_0 + a_1z + \cdots + a_nz^n \in \mathbb{C}[z]$$

where $a_n \neq 0$. Recall that we view $\mathbb{C}$ as the set $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$, and let R be any (finite) real number. Then the absolute value function $|f|: \mathbb{C} \rightarrow \mathbb{R}$ that takes any $z \in \mathbb{C}$ to the real number $|f(z)|$ is continuous on the closed ball $B[0, R]$ of radius R centered at the origin. But $B[0, R]$ is compact (Theorem A14) so that $|f(z)|$ takes its minimum value at some point on the ball (Theorem A15). On the other hand, if we write $f(z)$ in the form

$$f(z) = a_nz^n(a_0/a_nz^n + a_1/a_nz^{n-1} + \cdots + a_{n-1}/a_nz + 1)$$

we see that $|f(z)|$ becomes arbitrarily large as $|z|$ becomes large. To be precise, given any real $C > 0$ there exists $R > 0$ such that $|z| > R$ implies $|f(z)| > C$.

We now combine these two facts as follows. Let z_1 be arbitrary, and define $C = |f(z_1)|$. Then there exists $R_0 > 0$ such that $|f(z)| > |f(z_1)|$ for all $z \in \mathbb{C}$ such that $|z - z_1| > R_0$ (i.e., for all z outside $B[z_1, R_0]$). Since $B[z_1, R_0]$ is compact, there exists a point $z_0 \in B[z_1, R_0]$ such that $|f(z_0)| \leq |f(z)|$ for all $z \in B[z_1, R_0]$.



In particular, $|f(z_0)| \leq |f(z_1)|$ and hence we see that $|f(z_0)| < |f(z)|$ for all $z \in \mathbb{C}$. In other words, z_0 is an absolute minimum of $|f|$. We claim that $f(z_0) = 0$.

To show that $f(z_0) = 0$, we assume that $f(z_0) \neq 0$ and arrive at a contradiction. By a suitable choice of constants c_i , we may write f in the form

$$f(z) = c_0 + c_1(z - z_0) + \cdots + c_n(z - z_0)^n .$$

If $f(z_0) \neq 0$ then $c_0 = f(z_0) \neq 0$. By assumption, $\deg f \geq 1$, so we let m be the smallest integer greater than 0 such that $c_m \neq 0$. Defining the new variable $w = z - z_0$, we may define the polynomial function g by

$$f(z) = g(w) = c_0 + c_m w^m + w^{m+1} h(w)$$

for some polynomial h .

Now let w_1 be a complex number such that $w_1^m = -c_0/c_m$ and consider all values of $w = \lambda w_1$ for real λ with $0 \leq \lambda \leq 1$. Then

$$c_m w^m = c_m \lambda^m w_1^m = -c_0 \lambda^m$$

and hence

$$\begin{aligned} f(z) &= g(\lambda w_1) = c_0 - \lambda^m c_0 + \lambda^{m+1} w_1^{m+1} h(\lambda w_1) \\ &= c_0 [1 - \lambda^m + \lambda^{m+1} w_1^{m+1} c_0^{-1} h(\lambda w_1)] . \end{aligned}$$

But $\lambda \in [0, 1]$ which is compact, and hence $|w_1^{m+1} c_0^{-1} h(\lambda w_1)|$ is a continuous function defined on a compact set. Then the image of this function is a compact subset of $\mathbb{R}$ (Theorem A8) and so is closed and bounded (Theorem A14). This means there exists a number $B > 0$ such that

$$|w_1^{m+1} c_0^{-1} h(\lambda w_1)| \leq B$$

for all $\lambda \in [0, 1]$, and therefore (since $0 \leq \lambda \leq 1$ implies that $0 \leq \lambda^m \leq 1$)

$$\begin{aligned}
|g(\lambda w_1)| &= |c_0| |1 - \lambda^m + \lambda^{m+1} w_1^{m+1} c_0^{-1} h(\lambda w_1)| \\
&\leq |c_0| \{ |1 - \lambda^m| + \lambda^{m+1} |w_1^{m+1} c_0^{-1} h(\lambda w_1)| \} \\
&\leq |c_0| (1 - \lambda^m + \lambda^{m+1} B) .
\end{aligned}$$

Now recall that $|c_0| = |f(z_0)| \leq |f(z)|$ for all $z \in \mathbb{C}$. If we can show that

$$0 < 1 - \lambda^m + \lambda^{m+1} B < 1$$

for sufficiently small λ with $0 < \lambda < 1$, then we will have shown that $|f(z)| = |g(\lambda w_1)| < |c_0|$, a contradiction. But it is obvious that λ can be chosen so that $0 < 1 - \lambda^m + \lambda^{m+1} B$. And to require that $1 - \lambda^m + \lambda^{m+1} B < 1$ is the same as requiring that $\lambda B < 1$ which can certainly be satisfied for small enough λ . ■

Exercises

1. Show the absolute value function is continuous on $\mathbb{R}$.
2. Show that the norm on a vector space V defines a metric on V .
3. Let (X, d) be a metric space. Prove:
 - (a) Both X and $\emptyset$ are closed sets.
 - (b) The intersection of an arbitrary number of closed sets is closed.
 - (c) The union of a finite number of closed sets is closed.
4. Let A be the subset of $[0, 1]$ consisting of all $x \in [0, 1]$ whose decimal expansion contains only the digits 4 and 7. Explain whether or not A is countable, dense in $[0, 1]$, or compact.
5. Show that $\{x: \|x\|_2 \leq 1\}$ is closed and bounded but not compact in the space l_2 (see Example 12.8).
6. A metric space is said to be **separable** if it contains a countable dense subset. Prove that $\mathbb{R}^n$ is separable. [*Hint*: Consider the set of all points in $\mathbb{R}^n$ with rational coordinates.]

APPENDIX B

Sequences and Series

In this appendix we briefly go through all of the theory necessary for an understanding of Section 10.6 and Chapter 12. Furthermore, as we mentioned at the beginning of Appendix A, rather than being simply an overview, we want the reader to understand this material even if it has not been studied before. We do assume however, that the reader has studied Appendix A, and part of this appendix is a direct extension of that material.

A **sequence** $\{x_n\} = \{x_1, x_2, \dots\}$ in a set S is any function from the set $\mathbb{Z}^+$ of positive integers into S . If (X, d) is a metric space, then a sequence $\{x_n\}$ in X is said to **converge** to the **limit** x if for each $B(x, \epsilon)$ there exists $N \in \mathbb{Z}^+$ such that $x_n \in B(x, \epsilon)$ for every $n \geq N$. In other words, given $\epsilon > 0$, there exists a positive integer N such that $n \geq N$ implies $d(x_n, x) < \epsilon$. This is usually written as $\lim x_n = x$ or $x_n \rightarrow x$. If a sequence $\{x_n\}$ does not converge, then it is said to **diverge**. Furthermore, if for every real number M there exists an integer N such that $n \geq N$ implies $x_n \geq M$, then we write $x_n \rightarrow +\infty$. Similarly, if for every real number M there exists an integer N such that $n \geq N$ implies $x_n \leq M$, then we write $x_n \rightarrow -\infty$.

(We remark that the small, *always positive* number ϵ will be used extensively in many proofs, and it is important to realize that for proofs of convergence there is no real difference between the number ϵ and certain simple functions of ϵ such as 2ϵ . For example, suppose we can show that given ϵ , there exists N such that $n \geq N$ implies $d(x_n, x) < 2\epsilon$. We claim this also proves that $x_n \rightarrow x$. Indeed, let $\epsilon' = \epsilon/2$. Then, by assumption, there exists

N' such that $n \geq N'$ implies $d(x_n, x) < 2\varepsilon' = \varepsilon$ which shows that x is the limit of the sequence. It should be clear that the statement “given $\varepsilon > 0$ ” is equivalent to saying “given $C\varepsilon > 0$ ” for any finite $C > 0$.)

The set of all points x_n for $n = 1, 2, \dots$ is called the **range** of the sequence $\{x_n\}$. This may be either a finite or an infinite set of points. A set $A \subset (X, d)$ is said to be **bounded** if there exists a real number M and a point $x_0 \in X$ such that $d(x, x_0) \leq M$ for all $x \in A$. (The point x_0 is almost always taken to be the origin of any given coordinate system in X .) The sequence $\{x_n\}$ is said to be **bounded** if its range is a bounded set. It is easy to show that any convergent sequence is bounded. Indeed, if $x_n \rightarrow x$ then, given 1, there exists N such that $n \geq N$ implies $d(x_n, x) < 1$. This shows that $\{x_n: n \geq N\}$ is bounded. To show that $\{x_n: n = 1, \dots, N-1\}$ is bounded, let

$$r = \max\{1, d(x_1, x), \dots, d(x_{N-1}, x)\} .$$

Since $x_n \in X$, it must be true that each $d(x_n, x)$ is finite, and hence $d(x_n, x) \leq r$ for each $n = 1, \dots, N-1$.

We now prove several elementary properties of sequences, starting with the uniqueness of the limit.

Theorem B1 If $\{x_n\}$ is a sequence in a metric space (X, d) such that $x_n \rightarrow x$ and $x_n \rightarrow y$, then $x = y$.

Proof Given $\varepsilon > 0$, there exists N such that $n \geq N$ implies $d(x_n, x) < \varepsilon$ and $d(x_n, y) < \varepsilon$. But then $d(x, y) \leq d(x_n, x) + d(x_n, y) < 2\varepsilon$. Since this holds for all $\varepsilon > 0$, we must have $x = y$ (see Appendix A, definition (M2)). ■

Theorem B2 Let $s_n \rightarrow s$ and $t_n \rightarrow t$ be convergent sequences of complex numbers. Then

- (a) $\lim (s_n + t_n) = s + t$.
- (b) $\lim cs_n = cs$ and $\lim (c + s_n) = c + s$ for any $c \in \mathbb{C}$.
- (c) $\lim s_n t_n = st$.
- (d) $\lim 1/s_n = 1/s$ if $s \neq 0$ and $s_n \neq 0$ for all n .

Proof (a) Given $\varepsilon > 0$, there exists N_1 and N_2 such that $n \geq N_1$ implies that $|s_n - s| < \varepsilon/2$ and $n \geq N_2$ implies that $|t_n - t| < \varepsilon/2$. Let $N = \max\{N_1, N_2\}$. Then $n \geq N$ implies

$$|(s_n - s) + (t_n - t)| \leq |s_n - s| + |t_n - t| < \varepsilon .$$

- (b) This is Exercise B.1.

(c) Note the algebraic identity

$$s_n t_n - st = (s_n - s)(t_n - t) + s(t_n - t) + t(s_n - s) .$$

By parts (a) and (b) we see that $\lim s(t_n - t) = 0 = \lim t(s_n - s)$. Now, given $\varepsilon > 0$, there exists N_1 and N_2 such that $n \geq N_1$ implies

$$|s_n - s| < \sqrt{\varepsilon}$$

and $n \geq N_2$ implies

$$|t_n - t| < \sqrt{\varepsilon} .$$

If $N = \max\{N_1, N_2\}$, then $n \geq N$ implies that

$$|(s_n - s)(t_n - t)| < \varepsilon$$

and hence $\lim (s_n t_n - st) = \lim (s_n - s)(t_n - t) = 0$.

(d) At the end of Section 0.4 we showed that $|a| - |b| \leq |a + b|$ for any $a, b \in \mathbb{R}$. Reviewing the proof shows that this applies to complex numbers as well. Letting $b \rightarrow -b$ then shows that $|a| - |b| \leq |a - b|$.

Given $|s|/2$, there exists m such that $n \geq m$ implies $|s_n - s| < |s|/2$ (this follows from the fact that $s_n \rightarrow s$). But $|s| - |s_n| \leq |s_n - s| < |s|/2$ implies $|s_n| > |s|/2$.

Alternatively, given $\varepsilon > 0$, there exists N (which we can always choose greater than m) such that $n \geq N$ implies $|s_n - s| < |s|^2 \varepsilon / 2$. Combining these results, we see that for all $n \geq N$ we have

$$\left| \frac{1}{s_n} - \frac{1}{s} \right| = \left| \frac{s_n - s}{s s_n} \right| = \frac{|s_n - s|}{|s| |s_n|} < \frac{2|s_n - s|}{|s|^2} < \varepsilon . \blacksquare$$

Intuitively, we expect that a sequence $\{x_k\}$ of points in $\mathbb{R}^n$ converges to a point in $\mathbb{R}^n$ if and only if each of the n coordinates converges on its own. That this is indeed true is shown in our next theorem.

Theorem B3 Suppose $x_k = (x_k^1, \dots, x_k^n) \in \mathbb{R}^n$. Then $\{x_k\}$ converges to $x = (x^1, \dots, x^n) \in \mathbb{R}^n$ if and only if $x_k^i \rightarrow x^i$ for each $i = 1, \dots, n$.

Proof First note that for any $j = 1, \dots, n$ we have

$$|x - y|^2 = \sum_{i=1}^n (x^i - y^i)^2 \geq (x^j - y^j)^2$$

which implies $|x^j - y^j| \leq |x - y|$. Now assume that $x_k \rightarrow x$. Then given $\varepsilon > 0$, there exists N such that $k \geq N$ implies $|x_k^j - y^j| \leq |x_k - x| < \varepsilon$. This shows that $x_k \rightarrow x$ implies $x_k^j \rightarrow x^j$.

Conversely, assume $x_k^j \rightarrow x^j$ for every $j = 1, \dots, n$. Then given $\varepsilon > 0$, there exists N such that $k \geq N$ implies $|x_k^j - x^j| < \varepsilon/\sqrt{n}$. Hence $k \geq N$ implies

$$|x_k - x| = \left(\sum_{j=1}^n (x_k^j - x^j)^2 \right)^{1/2} < (n\varepsilon^2/n)^{1/2} = \varepsilon$$

so that $x_k \rightarrow x$. ■

A sequence $\{x_k\}$ in a metric space (X, d) is said to be a **Cauchy sequence** if given $\varepsilon > 0$, there exists N such that $n, m \geq N$ implies $d(x_n, x_m) < \varepsilon$. It is easy to see that every convergent sequence is in fact a Cauchy sequence. Indeed, simply note that if $x_k \rightarrow x$, then given $\varepsilon > 0$, there exists N such that $n \geq N$ implies $d(x_n, x) < \varepsilon/2$. Hence if $n, m \geq N$ we have

$$d(x_n, x_m) \leq d(x_n, x) + d(x_m, x) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

However, it is not true that a Cauchy sequence need converge. For example, suppose $X = (0, 1] \subset \mathbb{R}$ and let $\{x_k\} = \{1/k\}$ for $k = 1, 2, \dots$. This is a Cauchy sequence that wants to converge to the point 0 (choose $N = 1/\varepsilon$ so that $|1/n - 1/m| \leq |1/n| + |1/m| < 2\varepsilon$ for all $m, n \geq N$). But $0 \notin (0, 1]$ so that the limit of the sequence is not in the space. This example shows that convergence is not an intrinsic property of sequences, but rather depends on the space in which the sequence lies. A metric space in which every Cauchy sequence converges is said to be **complete** (see Appendix A).

We have shown that any convergent sequence is necessarily a Cauchy sequence, but that the converse is not true in general. However, in the case of $\mathbb{R}^n$, it is in fact true that every Cauchy sequence does indeed converge, i.e., $\mathbb{R}^n$ is a complete metric space. This is easy to prove using the fact that any n -cell in $\mathbb{R}^n$ is compact (Theorem A11), and we outline the proof in Exercise B.10. However, it is worth proving that $\mathbb{R}^n$ is complete without using this result. We begin by proving several other facts dealing with the real number line $\mathbb{R}$. By way of terminology, a sequence $\{x_n\}$ of real numbers with the property that $x_n \leq x_{n+1}$ is said to be **increasing**. Similarly, if $x_n \geq x_{n+1}$ then the sequence is said to be **decreasing**. We will sometimes use the term **monotonic** to refer to a sequence that is either increasing or decreasing.

Theorem B4 Let $\{x_k\}$ be an increasing sequence (i.e., $x_k \leq x_{k+1}$) of real numbers that is bounded above. Then the least upper bound b of the set $\{x_k\}$ is the limit of the sequence.

Proof It should be remarked that the existence of the least upper bound is guaranteed by the Archimedean axiom. Given $\varepsilon > 0$, the number $b - \varepsilon/2$ is not an upper bound for $\{x_k\}$ since b is by definition the least upper bound. Therefore there exists N such that $b - \varepsilon/2 \leq x_N \leq b$ (for otherwise $b - \varepsilon/2$ would be the least upper bound). Since $\{x_k\}$ is increasing, we have

$$b - \varepsilon/2 \leq x_N \leq x_n \leq b$$

for every $n \geq N$. Rearranging, this is just $b - x_n \leq \varepsilon/2 < \varepsilon$ which is the same as $|x_n - b| < \varepsilon$. ■

Since the Archimedean axiom also refers to the greatest lower bound of a set of real numbers, it is clear that Theorem B4 may be applied equally well to the greatest lower bound of a decreasing sequence.

Let (X, d) be a metric space, and suppose A is a subset of X . Recall from Appendix A that a point $x \in X$ is said to be an **accumulation point** of A if every deleted neighborhood of x contains a point of A . The analogous term for sequences is the following. A number x is said to be a **cluster point** (or **limit point**) of a sequence $\{x_n\}$ if given $\varepsilon > 0$, there exist infinitely many integers n such that $|x_n - x| < \varepsilon$. Equivalently, x is a cluster point if given $\varepsilon > 0$ and given N , there exists *some* $n \geq N$ such that $|x_n - x| < \varepsilon$. Note that this does not say that there are infinitely many *distinct* x_n such that $|x_n - x| < \varepsilon$. In fact, all the x_n could be identical. It is important to distinguish between the indices n and the actual elements x_n of the sequence. It is also important to realize that a limit point of a sequence is not the same as the limit of a sequence (why?). Note also that a sequence in X may be considered to be a subset of X , and in this context we may also refer to the accumulation points of a sequence.

Our next result is known as the Bolzano-Weierstrass Theorem.

Theorem B5 (Bolzano-Weierstrass) Let $\{x_k\}$ be a sequence of real numbers, and let $a, b \in \mathbb{R}$ be such that $a \leq x_k \leq b$ for all positive integers k . Then there exists a cluster point c of the sequence with $a \leq c \leq b$.

Proof For each n , the sequence $\{x_n, x_{n+1}, \dots\}$ is bounded below (by a), and hence has a greatest lower bound (whose existence is guaranteed by the Archimedean axiom) which we denote by c_n . Then $\{c_n\}$ forms an increasing sequence $c_n \leq c_{n+1} \leq \dots$ which is bounded above by b . Theorem B4 now

shows that the sequence $\{c_n\}$ has a least upper bound c (with $a \leq c \leq b$) which is in fact the limit of the sequence $\{c_n\}$. We must show that c is a cluster point of the sequence $\{x_k\}$.

To say that c is the limit of the sequence $\{c_n\}$ means that given $\varepsilon > 0$ and given any N , there exists some $m \geq N$ such that

$$|c_m - c| < \varepsilon/2 .$$

By definition, c_m is the greatest lower bound of the set $\{x_m, x_{m+1}, \dots\}$ which means there exists $k \geq m$ such that $c_m \leq x_k < c_m + \varepsilon/2$ or

$$|x_k - c_m| < \varepsilon/2 .$$

Therefore $k \geq m \geq N$ and

$$|x_k - c| = |x_k - c_m + c_m - c| \leq |x_k - c_m| + |c_m - c| < \varepsilon$$

which shows that c is a cluster point of the sequence $\{x_k\}$. ■

Note that Theorem B5 also follows from Theorem A12.

Theorem B6 If $\{x_n\}$ is a Cauchy sequence of numbers, then it is bounded.

Proof By definition of a Cauchy sequence, given $\varepsilon = 1$ there exists N such that $n \geq N$ implies $|x_n - x_N| < 1$. Hence $|x_n| - |x_N| \leq |x_n - x_N| < 1$ implies $|x_n| \leq |x_N| + 1$ for every $n \geq N$. Define $B = \max\{|x_1|, \dots, |x_N|, |x_N| + 1\}$. Then B is clearly a bound for $\{x_n\}$. ■

Theorem B7 Any Cauchy sequence $\{x_n\}$ of numbers converges.

Proof From Theorem B6 the sequence $\{x_n\}$ has a bound B , and hence we have $-B \leq x_n \leq B$ for all n . Hence by Theorem B5, the sequence $\{x_n\}$ has a cluster point c . We claim that c is the limit of the sequence. Since the sequence is Cauchy, given $\varepsilon > 0$ there exists N such that $m, n \geq N$ implies $|x_m - x_n| < \varepsilon/2$. Using this ε , we see that because c is a cluster point, there exists $m \geq N$ such that $|x_m - c| < \varepsilon/2$. Combining these last two results shows that for all $n \geq N$

$$|x_n - c| \leq |x_n - x_m| + |x_m - c| < \varepsilon . \quad \blacksquare$$

We are now in a position to prove our principal assertion.

Theorem B8 $\mathbb{R}^n$ is a complete metric space. In other words, every Cauchy sequence in $\mathbb{R}^n$ converges to a point in $\mathbb{R}^n$.

Proof Let $\{x_k\}$ be a Cauchy sequence in $\mathbb{R}^n$. Then $|x_m^j - x_n^j| \leq |x_m - x_n|$ (see the proof of Theorem B3) so that $\{x_k^j\}$ is also a Cauchy sequence in $\mathbb{R}$ for each $j = 1, \dots, n$. Hence by Theorem B7 each of the sequences $\{x_k^j\}$ also converges in $\mathbb{R}$. Therefore (by Theorem B3) the sequence $\{x_k\}$ must converge in $\mathbb{R}^n$. ■

We have seen that any convergent sequence is a Cauchy sequence and hence bounded. However, the converse is not generally true. (For example, the sequence $\{1, 2, 1, 2, \dots\}$ is clearly bounded but does not converge to either 1 or 2.) There is, however, a special case in which the converse is true that will be of use to us.

Theorem B9 A monotonic sequence $\{x_n\}$ converges if and only if it is bounded.

Proof We consider increasing sequences. The proof for decreasing sequences is similar. It was shown above that any convergent sequence is bounded, so we need only consider the converse. But this was proved in Theorem B4. ■

Finally, accumulation points are useful in determining whether or not a set is closed. The principal result relating these concepts is the following.

Theorem B10 A subset A of a metric space (X, d) is closed if and only if A contains all of its accumulation points.

Proof This is also Theorem A13. ■

Before continuing, we must make a digression to discuss some more basic properties of metric spaces. If the reader already knows that a point x in a subset A of a metric space X is in the closure of A if and only if every neighborhood of x intersects A , then he/she may skip to Theorem B16 below.

Let (X, d) be a metric space, and suppose $A \subset X$. We define

- (a) The **closure** of A , denoted by $\text{Cl } A$, to be the intersection of all closed supersets of A ;
- (b) The **interior** of A , denoted by $\text{Int } A$ (or A°), to be the union of all open subsets of A ;

- (c) The **boundary** of A , denoted by $\text{Bd } A$, to be the set of all $x \in X$ such that every open set containing x contains both points of A and points of $A^c = X - A$;
- (d) The **exterior** of A , denoted by $\text{Ext } A$, to be $(\text{Cl } A)^c = X - \text{Cl } A$;
- (e) The **derived set** of A , denoted by A' , to be the set of all accumulation points of A .

Example B1 Let $X = \mathbb{R}^2$ with the Pythagorean metric. Let A be the open unit ball defined by $A = \{(x, y): x^2 + y^2 < 1\}$. Then the following sets should be intuitively clear to the reader:

$$\begin{aligned}\text{Cl } A &= \{(x, y): x^2 + y^2 \leq 1\}; \\ \text{Int } A &= A; \\ \text{Bd } A &= \{(x, y): x^2 + y^2 = 1\}; \\ \text{Ext } A &= \{(x, y): x^2 + y^2 > 1\}; \\ A' &= \text{Cl } A. \quad //\end{aligned}$$

Theorem B11 Let (X, d) be a metric space and suppose $A, B \subset X$. Then

- (a) $A \subset \text{Cl } A$.
- (b) $\text{Cl}(\text{Cl } A) = \text{Cl } A$.
- (c) $\text{Cl}(A \cup B) = (\text{Cl } A) \cup (\text{Cl } B)$.
- (d) $\text{Cl } \emptyset = \emptyset$.
- (e) A is closed if and only if $A = \text{Cl } A$.
- (f) $\text{Cl}(A \cap B) \subset (\text{Cl } A) \cap (\text{Cl } B)$.

Proof Parts (a), (b), (d) and (e) are essentially obvious from the definition of $\text{Cl } A$, the fact that the intersection of an arbitrary number of closed sets is closed (see Exercise A.3), the fact that the empty set is a subset of every set, and the fact that if A is closed, then A is one of the closed sets which contains A .

(c) First note that if $S \subset T$, then any closed superset of T is also a closed superset of S , and therefore $\text{Cl } S \subset \text{Cl } T$. Next, observe that $A \subset A \cup B$ and $B \subset A \cup B$, so that taking the closure of both sides of each of these relations yields

$$\text{Cl } A \subset \text{Cl}(A \cup B)$$

and

$$\text{Cl } B \subset \text{Cl}(A \cup B) .$$

Together these show that $(\text{Cl } A) \cup (\text{Cl } B) \subset \text{Cl}(A \cup B)$. Since $\text{Cl } A$ and $\text{Cl } B$ are both closed, $(\text{Cl } A) \cup (\text{Cl } B)$ must also be closed and contain $A \cup B$. Hence we also have

$$\text{Cl}(A \cup B) \subset (\text{Cl } A) \cup (\text{Cl } B) .$$

This shows that $\text{Cl}(A \cup B) = (\text{Cl } A) \cup (\text{Cl } B)$.

(f) By (a) we have $A \cap B \subset A \subset \text{Cl } A$ and $A \cap B \subset B \subset \text{Cl } B$, and hence $A \cap B$ is a subset of the *closed* set $(\text{Cl } A) \cap (\text{Cl } B)$. But by definition of closure this means that

$$\text{Cl}(A \cap B) \subset (\text{Cl } A) \cap (\text{Cl } B) . \blacksquare$$

Theorem B12 Let (X, d) be a metric space and suppose $A \subset X$. Then

- (a) $\text{Cl } A = A \cup A'$.
- (b) $\text{Cl } A = \text{Int } A \cup \text{Bd } A$.
- (c) $\text{Bd } A = \text{Bd } A^c$.
- (d) $\text{Int } A = \text{Cl } A - \text{Bd } A$.

Proof (a) Assume $x \in \text{Cl } A$ but that $x \notin A$. We first show that $x \in A'$. Let U be any open set containing x . If $U \cap A = \emptyset$, then U^c is a closed superset of A which implies $\text{Cl } A \subset U^c$. But this contradicts the assumption that $x \in U$ since it was assumed that $x \in \text{Cl } A$. Therefore, since $x \notin A$, we must have

$$(U - \{x\}) \cap A \neq \emptyset$$

so that $x \in A'$. This shows that $\text{Cl } A \subset A \cup A'$.

Now assume that $x \in A \cup A'$. If $x \in A$, then obviously $x \in \text{Cl } A$ (since $A \subset \text{Cl } A$), so suppose $x \in A'$. We will show that x is contained in any closed superset of A . Let F be any closed superset of A not containing x . Then F^c is an open set containing x and such that $(F^c - \{x\}) \cap A = \emptyset$ which says that $x \notin A'$, a contradiction. Thus x is contained in any closed superset of A so that $x \in \text{Cl } A$. Since this shows that $A \cup A' \subset \text{Cl } A$, it follows that $A \cup A' = \text{Cl } A$.

(b) We first suppose that $x \in \text{Cl } A$ but $x \notin \text{Bd } A$. Since $x \notin \text{Bd } A$, there exists an open set U containing x such that either $U \subset A$ or $U \subset A^c$. If it were true that $U \subset A^c$, then U^c would be a closed superset of A (see Example 0.1) so that $\text{Cl } A \subset U^c$ which contradicts the assumption that $x \in \text{Cl } A$. We must therefore have $U \subset A$, and hence $x \in \text{Int } A$. Since the assumption that $x \in \text{Cl } A$ but $x \notin \text{Bd } A$ led to the requirement that $x \in \text{Int } A$, we must have

$$\text{Cl } A \subset \text{Int } A \cup \text{Bd } A .$$

Now assume $x \in \text{Int } A \cup \text{Bd } A$, but that $x \notin \text{Cl } A$. Note $\text{Int } A \subset A \subset \text{Cl } A$ so that $x \notin \text{Int } A$, and hence it must be true that $x \in \text{Bd } A$. However, since $x \notin \text{Cl } A$, $(\text{Cl } A)^c$ is an open set containing x with the property that $(\text{Cl } A)^c \cap A = \emptyset$. But this says that $x \notin \text{Bd } A$ which contradicts our original assumption. In other words, we must have $\text{Int } A \cup \text{Bd } A \subset \text{Cl } A$, so that

$$\text{Cl } A = \text{Int } A \cup \text{Bd } A .$$

(c) If $x \in \text{Bd } A$ and U is any open set containing x , then $U \cap A \neq \emptyset$ and $U \cap A^c \neq \emptyset$. But $A = (A^c)^c$ so that we also have $U \cap (A^c)^c \neq \emptyset$. Together with $U \cap A^c \neq \emptyset$, this shows that $x \in \text{Bd } A^c$. Reversing the argument shows that if $x \in \text{Bd } A^c$, then $x \in \text{Bd } A$. Hence $\text{Bd } A = \text{Bd } A^c$.

(d) This will follow from part (b) if we can show that $\text{Int } A \cap \text{Bd } A = \emptyset$. Now suppose that $x \in \text{Int } A \cap \text{Bd } A$. Then since $x \in \text{Bd } A$, it must be true that every open set containing x intersects A^c . But this contradicts the assumption that $x \in \text{Int } A$ (since by definition there must exist an open set U containing x such that $U \subset A$), and hence we must have $\text{Int } A \cap \text{Bd } A = \emptyset$. ■

It should be remarked that some authors *define* $\text{Bd } A$ as $\text{Cl } A - \text{Int } A$ so that our definition of $\text{Bd } A$ follows as a theorem. This fact, along with some additional insight, is contained in the following theorem.

Theorem B13 Let A be a subset of a metric space (X, d) , and suppose $x \in X$. Then

- (a) $x \in \text{Int } A$ if and only if *some* neighborhood of x is a subset of A .
- (b) $x \in \text{Cl } A$ if and only if *every* neighborhood of x intersects A .
- (c) $\text{Bd } A = \text{Cl } A - \text{Int } A$.

Proof (a) If $x \in \text{Int } A$, then by definition of $\text{Int } A$ there exists a neighborhood U of x such that $U \subset A$. On the other hand, if there exists an open set U such that $x \in U \subset A$, then $x \in \text{Int } A$.

(b) By Theorem B12(a), $\text{Cl } A = A \cup A'$. If $x \in \text{Cl } A$ and $x \in A$, then clearly every neighborhood of x intersects A . If $x \in \text{Cl } A$ and $x \in A'$, then every neighborhood of x also intersects A . Conversely, suppose that every neighborhood of x intersects A . Then either $x \in A$ or $x \in A'$, and hence $x \in A \cup A' = \text{Cl } A$.

(c) By definition, $x \in \text{Bd } A$ if and only if every open set containing x contains both points of A and points of A^c . In other words, $x \in \text{Bd } A$ if and only if every open set containing x contains points of A but is not a subset of A . By parts (a) and (b), this is just $\text{Bd } A = \text{Cl } A - \text{Int } A$. ■

Example B2 An elementary fact that will be referred to again is the following. Let A be a nonempty set of real numbers that is bounded above. Then by the Archimedean axiom, A has a least upper bound $b = \sup A$. Given $\varepsilon > 0$, there must exist $x \in A$ such that $b - \varepsilon < x < b$, for otherwise $b - \varepsilon$ would be an upper bound of A . But this means that any neighborhood of b intersects A , and hence $b \in \text{Cl } A$. Thus $b \in A$ if A is closed. //

Our next example yields an important basic result.

Example B3 Let $X = \mathbb{R}$ with the standard (absolute value) metric, and let $\mathbb{Q} \subset \mathbb{R}$ be the subset of all rational numbers. In Theorem 0.4 it was shown that given any two distinct real numbers there always exists a rational number between them. This may also be expressed by stating that any neighborhood of any real number always contains a rational number. In other words, we have $\text{Cl } \mathbb{Q} = \mathbb{R}$. //

From Theorems B10 and B12(a), we might guess that there is a relationship between sequences and closed sets. This is indeed the case, and our next theorem provides a very useful description of the closure of a set.

Theorem B14 (a) A set $A \subset (X, d)$ is closed if and only if for every sequence $\{x_n\}$ in A that converges, the limit is an element of A .

(b) If $A \subset (X, d)$, then $x \in \text{Cl } A$ if and only if there is a sequence $\{x_n\}$ in A such that $x_n \rightarrow x$.

Proof (a) Suppose that A is closed, and let $x_n \rightarrow x$. Since any neighborhood of x must contain all x_n for n sufficiently large, it follows from Theorem B10 that $x \in A$.

Conversely, assume that any sequence in A converges to an element of A , and let x be any accumulation point of A . We will construct a sequence in A that converges to x . To construct such a sequence, choose $x_n \in B(x, 1/n) \cap A$. This is possible since x is an accumulation point of A . Then given $\varepsilon > 0$, choose $N \geq 1/\varepsilon$ so that $x_n \in B(x, \varepsilon)$ for every $n \geq N$. Hence $x_n \rightarrow x$ so that $x \in A$. Theorem B10 then shows that A is closed.

(b) This is Exercise B.4. ■

If (X, d) is a metric space, then $A \subset X$ is said to be **somewhere dense** if $\text{Int}(\text{Cl } A) \neq \emptyset$. The set A is said to be **nowhere dense** if it is not somewhere dense. If $\text{Cl } A = X$, then A is said to be **dense** in X .

Example B4 Let $X = \mathbb{R}$ with the usual metric. The set $A = [a, b)$ has closure $\text{Cl } A = [a, b]$, and therefore $\text{Int}(\text{Cl } A) = (a, b) \neq \emptyset$. Hence A is somewhere dense. Example B3 showed that the set $\mathbb{Q}$ is dense in $\mathbb{R}$. Now let $A = \mathbb{Z}$, the set of all integers. $\mathbb{Z}^c = \mathbb{R} - \mathbb{Z}$ is the union of open sets of the form $(n, n + 1)$ where n is an integer, and hence $\mathbb{Z}$ is closed. It should also be clear that $\mathbb{Z}' = \emptyset$ since there clearly exist deleted neighborhoods of any integer that do not contain any other integers. By Theorem B13(a), we also see that $\text{Int}(\text{Cl } \mathbb{Z}) = \text{Int } \mathbb{Z} = \emptyset$ so that $\mathbb{Z}$ is nowhere dense. //

Theorem B15 A subset A of a metric space (X, d) is dense if and only if every open subset U of X contains some point of A .

Proof Suppose A is dense so that $\text{Cl } A = X$. If $U \subset X$ is open, then the fact that any $x \in U \subset \text{Cl } A$ implies that every neighborhood of x intersects A (Theorem B13). In particular, U is a neighborhood of x so that $U \cap A \neq \emptyset$. On the other hand, suppose that every open set U intersects A . If $x \in X$, then every neighborhood U of x must intersect A so that $x \in \text{Cl } A$. Since x was arbitrary, it must be true that $\text{Cl } A = X$. ■

After this topological digression, let us return to sequences of numbers. Given a sequence $\{x_n\}$, we may consider a sequence $\{n_k\}$ of positive integers that forms a subset of $\mathbb{Z}^+$ such that $n_1 < n_2 < \dots$. The corresponding subset $\{x_{n_k}\}$ of $\{x_n\}$ is called a **subsequence** of $\{x_n\}$. If $\{x_{n_k}\}$ converges, then its limit is called a **subsequential limit** of $\{x_n\}$. From the definitions, it should be clear that any cluster point of $\{x_n\}$ is the limit of a convergent subsequence. It should also be clear that a sequence $\{x_n\}$ converges to x if and only if every subsequence also converges to x (see Exercise B.5).

Theorem B16 The set of all subsequential limits of a sequence $\{x_n\}$ in a metric space (X, d) forms a closed subset of X .

Proof Let S be the set of all subsequential limits of $\{x_n\}$. If y is an accumulation point of S , we must show that $y \in S$ (Theorem B10), and hence that some subsequence of $\{x_n\}$ converges to y . Choose n_1 such that $x_{n_1} \neq y$ (why can this be done?), and let $\delta = d(x_{n_1}, y)$. Now suppose that $n_1, n_2, \dots, n_{k-1}$ have been chosen. Since y is an accumulation point of S , there exists $z \in S$, $z \neq y$, such that $d(z, y) < 2^{-k}\delta$. But $z \in S$ implies that z is a subsequential limit, and hence there exists $n_k > n_{k-1}$ such that $d(z, x_{n_k}) < 2^{-k}\delta$. Therefore, for each $k = 1, 2, \dots$ we have

$$d(x_{n_k}, y) \leq d(x_{n_k}, z) + d(z, y) < 2^{1-k}\delta$$

so that the subsequence $\{x_{n_k}\}$ converges to y . ■

Given a sequence $\{a_n\}$ of complex numbers (which can be regarded as points in $\mathbb{R}^2$), we may define the **infinite series** (generally called simply a **series**)

$$\sum_{i=1}^{\infty} a_n = a_1 + a_2 + \cdots$$

as the sequence $\{s_n\}$ where

$$s_n = \sum_{k=1}^n a_k .$$

If the *sequence* $\{s_n\}$ converges to a number s , we say that the series **converges** and we write this as $\sum_{n=1}^{\infty} a_n = s$. Note that the number s is the limit of a sequence of partial sums. For notational convenience, the lower limit in the sum may be taken to be 0 instead of 1, and we will frequently write just $\sum a_n$ when there is no danger of ambiguity and the proper limits are understood.

The reader should note that $\sum_{n=1}^{\infty} a_n$ stands for two very different things. On the one hand it is used to stand for the sequence $\{s_n\}$ of partial sums, and on the other hand it stands for $\lim_{n \rightarrow \infty} s_n$. This is a common abuse of notation, and the context usually makes it clear which meaning is being used.

We have seen that any convergent sequence is Cauchy, and in Theorem B8 we showed that any Cauchy sequence in $\mathbb{R}^n$ converges. Thus, a sequence in $\mathbb{R}^n$ converges if and only if it is Cauchy. This is called the **Cauchy criterion**. Since the convergence of a series is defined in terms of the sequence of partial sums, we see that the Cauchy criterion may be restated as follows.

Theorem B17 A series of numbers $\sum a_n$ converges if and only if given $\varepsilon > 0$, there exists an integer N such that $m \geq n \geq N$ implies

$$\left| \sum_{k=n}^m a_k \right| < \varepsilon .$$

Proof If the series $\sum a_n$ converges, then the sequence $\{s_k\}$ of partial sums $s_k = \sum_{n=1}^k a_n$ converges, and hence $\{s_k\}$ is Cauchy. Conversely, if the sequence of partial sums s_k is Cauchy, then $\{s_k\}$ converges (Theorem B8). In either case, this means that given $\varepsilon > 0$, there exists N such that $p \geq q \geq N$ implies

$$|s_p - s_q| = \left| \sum_{k=1}^p a_k - \sum_{k=1}^q a_k \right| = \left| \sum_{k=q+1}^p a_k \right| < \varepsilon .$$

The result now follows by choosing $m = p$ and $n = q + 1$. ■

Another useful way of stating Theorem B17 is to say that a series $\sum a_n$ converges if and only if given $\varepsilon > 0$, there exists N such that $k \geq N$ implies that $|a_k + \cdots + a_{k+p}| < \varepsilon$ for all positive integers $p = 0, 1, 2, \dots$.

Corollary If $\sum a_n$ converges, then given $\varepsilon > 0$ there exists N such that $|a_n| < \varepsilon$ for all $n \geq N$. In other words, if $\sum a_n$ converges, then $\lim_{n \rightarrow \infty} a_n = 0$.

Proof This follows from Theorem B17 by letting $m = n$. ■

While this corollary says that a *necessary* condition for $\sum a_n$ to converge is that $a_n \rightarrow 0$, this is not a *sufficient* condition (see Example B5 below).

If we have a series of nonnegative real numbers, then each partial sum is clearly nonnegative, and the sequence of partial sums forms a non-decreasing sequence. Thus, directly from Theorem B9 we have the following.

Theorem B18 A series of nonnegative real numbers converges if and only if its partial sums form a bounded sequence.

One consequence of this theorem is the next result.

Theorem B19 Suppose $\sum a_n$ is a series such that $a_1 \geq a_2 \geq \cdots \geq 0$. Then $\sum a_n$ converges if and only if

$$\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + \cdots$$

converges.

Proof Let $s_n = a_1 + a_2 + \cdots + a_n$ and let $t_k = a_1 + 2a_2 + \cdots + 2^k a_{2^k}$. Since all terms in the series $\sum a_n$ are nonnegative we may write for $n < 2^k$

$$s_n \leq a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + \cdots + (a_{2^k} + \cdots + a_{2^{k+1}-1})$$

since this is just adding the nonnegative term $a_{2^k+1} + \cdots + a_{2^{k+1}-1}$ to $s_{2^k} \geq s_n$. But $\{a_n\}$ is a decreasing sequence, so noting that the last term in parentheses consists of 2^k terms, we have

$$s_n \leq a_1 + 2a_2 + 4a_4 + \cdots + 2^k a_{2^k} = t_k .$$

Similarly, if $n > 2^k$ we have

$$\begin{aligned} s_n &\geq a_1 + a_2 + (a_3 + a_4) + \cdots + (a_{2^{k-1}+1} + \cdots + a_{2^k}) \\ &\geq (1/2)a_1 + a_2 + 2a_4 + \cdots + 2^{k-1}a_{2^k} \\ &= (1/2)t_k . \end{aligned}$$

We have now shown that $n < 2^k$ implies $s_n \leq t_k$, and $n > 2^k$ implies $2s_n \geq t_k$. Thus the sequences $\{s_n\}$ and $\{t_k\}$ are either both bounded or both unbounded. Together with Theorem B18, this completes the proof. ■

The interesting aspect of this theorem is that the convergence of $\sum a_n$ is determined by the convergence of a rather “sparse” subsequence.

Example B5 Let us show that the series $\sum n^{-p}$ converges if $p > 1$ and diverges if $p \leq 1$. Indeed, suppose $p > 0$. Then by Theorem B19 we consider the series

$$\sum_{k=0}^{\infty} 2^k \cdot 2^{-kp} = \sum_{k=0}^{\infty} 2^{k(1-p)} .$$

By the corollary to Theorem B17, we must have $1 - p < 0$ so that $p > 1$. In this case, $\sum 2^{k(1-p)} = \sum (2^{1-p})^k$ is a geometric series which converges as in Example B1, and hence Theorem B19 shows that $\sum n^{-p}$ converges for $p > 1$. If $p \leq 0$, then $\sum n^{-p}$ diverges by the corollary to Theorem B17. //

If we are given a series $\sum a_n$, we could rearrange the terms in this series to obtain a new series $\sum a'_n$. Formally, we define this rearrangement by letting $\{k_n\}$ be a sequence in which every positive integer appears exactly once. In other words, $\{k_n\}$ is a one-to-one mapping from $\mathbb{Z}^+$ onto $\mathbb{Z}^+$. If we now define $a'_n = a_{k_n}$ for $n = 1, 2, \dots$, then the corresponding series $\sum a'_n$ is called a **rearrangement** of the series $\sum a_n$.

For each of the series $\sum a_n$ and $\sum a'_n$, we form the respective sequences of partial sums $\{s_k\}$ and $\{s'_k\}$. Since these sequences are clearly different in general, it is not generally the case that they both converge to the same limit. While we will not treat this problem in any detail, there is one special case that we will need. This will be given as a corollary to the following theorem.

A series $\sum a_n$ is said to **converge absolutely** if the series $\sum |a_n|$ converges.

Theorem B20 If $\sum a_n$ converges absolutely, then $\sum a_n$ converges.

Proof Note $|\sum_{k=n}^m a_k| \leq \sum_{k=n}^m |a_k|$ and apply Theorem B17. ■

Corollary If $\sum a_n$ converges absolutely, then every rearrangement of $\sum a_n$ converges to the same sum.

Proof Let $\sum a'_n$ be a rearrangement with partial sums s'_k . Since $\sum a_n$ converges absolutely, we may apply Theorem B17 to conclude that for every $\varepsilon > 0$ there exists N such that $m \geq n \geq N$ implies

$$\sum_{i=n}^m |a_i| < \varepsilon . \quad (*)$$

Using the notation of the discussion preceding the theorem, we let $p \in \mathbb{Z}^+$ be such that the integers $1, \dots, N$ are contained in the collection $k_1, \dots, k_p$ (note that we must have $p \geq N$). If for any $n > p$ we now form the difference $s_n - s'_n$, then the numbers $a_1, \dots, a_N$ will cancel out (as may some other numbers if $p > N$) and hence, since (*) applies to all $m \geq n \geq N$, we are left with $|s_n - s'_n| < \varepsilon$. This shows that $\{s_n\}$ and $\{s'_n\}$ both converge to the same sum (since if $s_n \rightarrow s$, then $|s'_n - s| \leq |s'_n - s_n| + |s_n - s| < 2\varepsilon$ which implies that $s'_n \rightarrow s$ also). ■

We remark that Theorems B17 and B20 apply equally well to any complete normed space if we replace the absolute value by the appropriate norm.

Before presenting any examples of series, we first compute the limits of some commonly occurring sequences of real numbers.

Theorem B21 (a) If $p > 0$, then $\lim_{n \rightarrow \infty} 1/n^p = 0$.

(b) If $p > 0$, then $\lim_{n \rightarrow \infty} p^{1/n} = 1$.

(c) $\lim_{n \rightarrow \infty} n^{1/n} = 1$.

(d) If $p > 0$ and r is real, then $\lim_{n \rightarrow \infty} n^r/(1+p)^n = 0$.

(e) If $|x| < 1$, then $\lim_{n \rightarrow \infty} x^n = 0$.

Proof (a) Given $\varepsilon > 0$, we seek an N such that $n \geq N$ implies $1/n^p < \varepsilon$. Then choose $N \geq (1/\varepsilon)^{1/p}$.

(b) If $p = 1$, there is nothing to prove. For $p > 1$, define $x_n = p^{1/n} - 1 > 0$ so that by the binomial theorem (Example 0.7) we have

$$p = (1 + x_n)^n \geq 1 + nx_n .$$

Thus $0 < x_n \leq (p-1)/n$ so that $\lim x_n = 0$, and hence $\lim p^{1/n} = 1$. If $0 < p < 1$, we define $y_n = (1/p)^{1/n} - 1 > 0$. Then

$$p = (1 + y_n)^n \geq 1 + ny_n$$

so that $y_n \rightarrow 0$, and hence we again have $\lim p^{1/n} = 1$.

(c) Let $x_n = n^{1/n} - 1 \geq 0$, so that using only the quadratic term in the binomial expansion yields

$$n = (1 + x_n)^n \geq \binom{n}{2} x_n^2 = \frac{n(n-1)}{2} x_n^2 .$$

Thus (for $n \geq 2$) we have $0 \leq x_n \leq [2/(n-1)]^{1/2}$ so that $x_n \rightarrow 0$. Therefore $\lim n^{1/n} = 1$.

(d) Let k be any integer > 0 such that $k > r$. Choosing the k th term in the binomial expansion we have (since $p > 0$)

$$(1 + p)^n \geq \binom{n}{k} p^k = \frac{n(n-1) \cdots (n-(k-1))}{k!} p^k .$$

If we let $n > 2k$, then $k < n/2$ so that $n > n/2$, $n-1 > n/2$, $\dots$, $n-(k-1) > n/2$ and hence $(1 + p)^n > (n/2)^k p^k / k!$. Thus (for $n > 2k$)

$$0 < \frac{n^r}{(1 + p)^n} < \frac{2^k k!}{p^k} n^{r-k} .$$

Since $r - k < 0$, it follows that $n^{r-k} \rightarrow 0$ by (a).

(e) Choose $r = 0$ in (d). ■

Corollary If $N > 0$ is any finite integer, then $\lim_{n \rightarrow \infty} (n^N)^{1/n} = 1$.

Proof $(n^N)^{1/n} = n^{N/n} = (n^{1/n})^N$ so that (by Theorem B2(c))

$$\lim (n^{1/n})^N = (\lim n^{1/n})^N = 1^N = 1 . \blacksquare$$

Example B6 The geometric series $\sum_{n=0}^{\infty} x^n$ converges for $|x| < 1$ and diverges for $|x| \geq 1$. Indeed, from elementary algebra we see that (for $x \neq 1$)

$$1 + x + x^2 + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x} .$$

If $|x| < 1$, we clearly have $\lim x^{n+1} = 0$, and hence

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} .$$

If $|x| > 1$, then $|x|^{n+1} \rightarrow \infty$ and the series diverges. In the case that $|x| = 1$, we see that $x^n \not\rightarrow 0$ so the series diverges. //

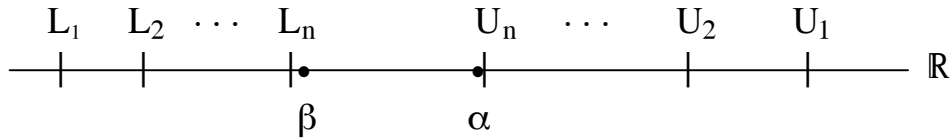
Let $\{x_n\}$ be a sequence of real numbers. In general this sequence may have many cluster points, in which case it will not converge to a definite limit. Now define the sequences

$$U_n = \sup_{k \geq n} x_k$$

and

$$L_n = \inf_{k \geq n} x_k .$$

Note that U_n is a decreasing sequence and L_n is an increasing sequence. If α is the largest cluster point of $\{x_n\}$, then clearly the U_n will approach α as n increases. Furthermore, no matter how large n gets, U_n will always remain $\geq \alpha$. Similarly, if β is the smallest cluster point of $\{x_n\}$, then all the L_n must be $\leq \beta$. This situation is represented schematically in the figure below.



Let $U = \inf_n U_n$. By Theorem B4 and the remarks following it we see that U_n converges to U . The limit U is called the **upper limit** (or **limit superior**) of the sequence $\{x_n\}$ and will be denoted by $\bar{x}$. In other words,

$$\bar{x} = \inf_n \sup_{k \geq n} x_k = \lim_{n \rightarrow \infty} \sup_{k \geq n} x_k .$$

The upper limit is frequently also written as $\limsup x_n$.

Similarly, L_n converges to $L = \sup_n L_n$. The limit L is called the **lower limit** (or **limit inferior**) of the sequence $\{x_n\}$, and is denoted by $\underline{x}$. Thus

$$\underline{x} = \sup_n \inf_{k \geq n} x_k = \lim_{n \rightarrow \infty} \inf_{k \geq n} x_k$$

which is also written as $\liminf x_n$. Note that either or both $\bar{x}$ and $\underline{x}$ could be $\pm\infty$.

Theorem B22 If $x_n \leq y_n$ for all n greater than or equal to some fixed N , then $\limsup x_n \leq \limsup y_n$ and $\liminf x_n \leq \liminf y_n$.

Proof This is Exercise B.6. ■

We have already remarked that in general a sequence may have many (or no) cluster points, and hence will not converge. However, suppose $\{x_n\}$ converges to x , and let $\lim U_n = U$. We claim that $x = U$.

To see this, we simply use the definitions involved. Given $\varepsilon > 0$, we may choose N such that for all $n \geq N$ we have both $|x - x_n| < \varepsilon$ and $|U - U_n| < \varepsilon$. Since $U_N = \sup_{k \geq N} x_k$, we see that given this ε , there exists $k \geq N$ such that $U_N - \varepsilon < x_k$ or $U_N - x_k < \varepsilon$. But then we have

$$|U - x| \leq |U - U_N| + |U_N - x_k| + |x_k - x| < 3\varepsilon$$

which proves that $U = x$. In an exactly analogous way, it is easy to prove that $L = \lim L_n = x$ (see Exercise B.7). We have therefore shown that $x_n \rightarrow x$ implies $\limsup x_n = \liminf x_n = x$. That the converse of this statement is also true is given in the next theorem. It should be clear however, that all but a finite number of terms in the sequence $\{x_n\}$ will be caught between U and L , and hence if $U = L$ it must be true that $x_n \rightarrow x = U = L$.

Theorem B23 A real-valued sequence $\{x_n\}$ converges to the number x if and only if $\limsup x_n = \liminf x_n = x$.

Proof Let $U_n = \sup_{k \geq n} x_k$ and $L_n = \inf_{k \geq n} x_k$, and first suppose that $\lim U_n = \lim L_n = x$. Given $\varepsilon > 0$, there exists N such that $|U_n - x| < \varepsilon$ for all $n \geq N$, and there exists M such that $|L_n - x| < \varepsilon$ for all $n \geq M$. These may be written as (see Example 0.6) $x - \varepsilon < U_n < x + \varepsilon$ for all $n \geq N$, and $x - \varepsilon < L_n < x + \varepsilon$ for all $n \geq M$. But from the definitions of U_n and L_n we know that $x_n \leq U_n$ and $L_n \leq x_n$. Hence $x_n < x + \varepsilon$ for all $n \geq N$ and $x - \varepsilon < x_n$ for all $n \geq M$. Therefore $|x_n - x| < \varepsilon$ for all $n \geq \max\{N, M\}$ so that $x_n \rightarrow x$.

The converse was shown in the discussion preceding the theorem. ■

Define S to be the set of all cluster points of $\{x_n\}$. Since any cluster point is the limit of some subsequence, it follows that S is just the set of all subsequential limits of $\{x_n\}$. From the figure above, we suspect that $\sup S = \bar{x}$ and $\inf S = \underline{x}$. It is not hard to prove that this is indeed the case.

Theorem B24 Let $\{x_n\}$ be a sequence of real numbers and let S , $\bar{x}$ and $\underline{x}$ be defined as above. Then $\sup S = \bar{x}$ and $\inf S = \underline{x}$.

Proof This is Exercise B.8. ■

Example B7 Let $x_n = (-1)^n/(1 + 1/n)$. Then it should be clear that we have $\limsup x_n = 1$ and $\liminf x_n = -1$. //

Our next theorem will be very useful in proving several tests for the convergence of series.

Theorem B25 Let $\{x_n\}$ be a sequence of real numbers, let S be the set of all subsequential limits of $\{x_n\}$, and let $\bar{x} = \limsup x_n$ and $\underline{x} = \liminf x_n$. Then

- (a) $\bar{x} \in S$.
- (b) If $r > \bar{x}$, then there exists N such that $n \geq N$ implies $x_n < r$.
- (c) $\bar{x}$ is unique.

Of course, the analogous results hold for $\underline{x}$ as well.

Proof We will show only the results for $\bar{x}$, and leave the case of $\underline{x}$ to the reader.

(a) Since S (the set of all subsequential limits) lies in the extended number system, we must consider three possibilities. If $-\infty < \bar{x} < +\infty$, then S is bounded above so that at least one subsequential limit exists. Then the set S is closed (by Theorem B16), and hence $\bar{x} = \sup S \in S$ (see Example B2).

If $\bar{x} = +\infty$, then S is not bounded above so that $\{x_n\}$ is not bounded above. Thus there exists a subsequence $\{x_{n_k}\}$ such that $x_{n_k} \rightarrow +\infty$. But then $+\infty \in S$ so that $\bar{x} \in S$.

If $\bar{x} = -\infty$, then there is no finite subsequential limit (since $\bar{x}$ is the least upper bound of the set of such limits), and hence S consists solely of the element $-\infty$. This means that given any real M , $x_n > M$ for at most a finite number of indices n so that $x_n \rightarrow -\infty$, and hence $\bar{x} = -\infty \in S$.

(b) If there existed an $r > \bar{x}$ such that $x_n \geq r$ for an infinite number of n , then there would be a subsequential limit x' of $\{x_n\}$ such that $x' \geq r > \bar{x}$. This contradicts the definition of $\bar{x}$.

(c) Let $\bar{x}$ and $\bar{y}$ be distinct numbers that satisfy (a) and (b), and suppose $\bar{x} < \bar{y}$. Let r be any number such that $\bar{x} < r < \bar{y}$ (that such an r exists was shown in Theorem 0.4). Since $\bar{x}$ satisfies (b), there exists N such that $x_n < r$ for all $n \geq N$. But then $\bar{y}$ can not possibly satisfy (a). ■

We now have the background to prove three basic tests for the convergence of series.

Theorem B26 (a) (**Comparison test**) If $\sum b_n$ converges, and if $|a_n| \leq b_n$ for $n \geq N_0$ (N_0 fixed), then $\sum a_n$ converges. If $\sum c_n$ diverges, and if $a_n \geq c_n \geq 0$ for $n \geq N_0$, then $\sum a_n$ diverges.

(b) (**Root test**) Given the series $\sum a_n$, let $\bar{a} = \limsup |a_n|^{1/n}$. If $\bar{a} < 1$, then $\sum a_n$ converges, and if $\bar{a} > 1$, then $\sum a_n$ diverges.

(c) (**Ratio test**) The series $\sum a_n$ converges if $\limsup |a_{n+1}/a_n| < 1$, and diverges if $|a_{n+1}/a_n| \geq 1$ for $n \geq N$ (N fixed).

Proof (a) Given $\varepsilon > 0$, there exists $N \geq N_0$ such that $m \geq n \geq N$ implies that $|\sum_{k=n}^m b_k| < \varepsilon$ (Theorem B17). Hence $\sum a_n$ converges since

$$\left| \sum_{k=n}^m a_k \right| \leq \sum_{k=n}^m |a_k| \leq \sum_{k=n}^m b_k = \left| \sum_{k=n}^m b_k \right| < \varepsilon .$$

By what has just been shown, we see that if $0 \leq c_n \leq a_n$ and $\sum a_n$ converges, then $\sum c_n$ must also converge. But the contrapositive of this statement is then that if $\sum c_n$ diverges, then so must $\sum a_n$.

(b) First note that $\bar{a} \geq 0$ since $|a_n|^{1/n} \geq 0$. Now suppose that $\bar{a} < 1$. By Theorem B25(b), for any r such that $\bar{a} < r < 1$, there exists N such that $n \geq N$ implies $|a_n|^{1/n} < r$, and thus $|a_n| < r^n$. But $\sum r^n$ converges (Example B5) so that $\sum a_n$ must also converge by the comparison test.

If $\bar{a} > 1$, then by Theorems B22(a) and B14(b), there must exist a sequence $\{n_k\}$ such that $|a_{n_k}|^{1/n_k} \rightarrow \bar{a}$. But this means that $|a_n| > 1$ for infinitely many n so that $a_n \not\rightarrow 0$ and $\sum a_n$ does not converge (corollary to Theorem B17).

(c) If $\limsup |a_{n+1}/a_n| < 1$ then, by Theorem B25(b), we can find a number $r < 1$ and an integer N such that $n \geq N$ implies $|a_{n+1}/a_n| < r$. We then see that

$$\begin{aligned} |a_{N+1}| &< r |a_N| \\ |a_{N+2}| &< r |a_{N+1}| < r^2 |a_N| \\ &\vdots \\ |a_{N+p}| &< r^p |a_N| . \end{aligned}$$

Therefore, letting $n = N + p$ we have

$$|a_n| < r^{n-N} |a_N| = r^{-N} |a_N| r^n$$

for $n \geq N$, and hence $\sum a_n$ converges by the comparison test and Example B6.

If $|a_{n+1}| \geq |a_n|$ for $n \geq N$ (N fixed), then clearly $a_n \not\rightarrow 0$ so that $\sum a_n$ can not converge (corollary to Theorem B17). ■

Note that if $\bar{a} = 1$ when applying the root test we get no information since, for example, $\sum 1/n$ and $\sum 1/n^2$ both have $\bar{a} = 1$ (corollary to Theorem B21), but the first diverges whereas the second converges (see Example B5).

Example B8 Consider the function e^x ($= \exp x$) defined by

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

where x may be real or complex. To test for convergence, we observe that $|a_{n+1}/a_n| = |x/(n+1)|$ so that

$$\begin{aligned} \limsup \left| \frac{a_{n+1}}{a_n} \right| &= \lim_{n \rightarrow \infty} \sup_{k \geq n} \left| \frac{x}{k+1} \right| \\ &= \lim_{n \rightarrow \infty} \left| \frac{x}{n+1} \right| \\ &= 0 \end{aligned}$$

and hence the series converges by the ratio test.

It is of great use in both mathematics and physics to prove that

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n} \right)^n .$$

While this can be proved by taking the logarithm of both sides and using l'Hospital's rule, we shall follow a direct approach. Let $x_n = (1 + x/n)^n$. Expanding x_n by the binomial theorem we have (for $n > 2$)

$$\begin{aligned} x_n &= \sum_{k=0}^n \frac{n!}{k!(n-k)!} \left(\frac{x}{n} \right)^k \\ &= 1 + x + \frac{n(n-1)}{2!} \frac{x^2}{n^2} + \frac{n(n-1)(n-2)}{3!} \frac{x^3}{n^3} + \cdots + \frac{x^n}{n^n} . \end{aligned}$$

If we write

$$\begin{aligned} \frac{1}{n^n} &= \frac{n!}{n!} \frac{1}{n^n} = \frac{1}{n!} \frac{n(n-1)(n-2) \cdots (n-(n-1))}{n^n} \\ &= \frac{1}{n!} \left(1 - \frac{1}{n} \right) \left(1 - \frac{2}{n} \right) \cdots \left(1 - \frac{n-1}{n} \right) \end{aligned}$$

then

$$\begin{aligned}
x_n = 1 + x + \frac{1}{2!} \left(1 - \frac{1}{n}\right) x^2 + \frac{1}{3!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) x^3 \\
+ \cdots + \frac{1}{k!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) x^k \\
+ \cdots + \frac{1}{n!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{n-1}{n}\right) x^n .
\end{aligned}$$

We now treat each x_n as an infinite series by defining all terms with $k > n$ to be zero, and we consider the difference

$$e^x - x_n = \sum_{k=2}^{\infty} \frac{1}{k!} \left[1 - \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \right] x^k . \quad (*)$$

Applying Theorem B17 to the convergent series $e^x = \sum |x|^n/n!$, we see that for fixed x and $\varepsilon > 0$, we can choose an integer m sufficiently large that

$$\sum_{k=m+1}^{\infty} \frac{|x|^k}{k!} < \varepsilon/2 .$$

Writing $(*)$ in the form $\sum_{k=2}^{\infty} = \sum_{k=2}^m + \sum_{k=m+1}^{\infty}$ and noting that the coefficient of x^k in the (second) sum is ≥ 0 but $\leq 1/k!$, we obtain (for $n > m$)

$$|e^x - x_n| \leq \sum_{k=2}^m \frac{1}{k!} \left[1 - \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \right] |x^k| + \varepsilon/2 .$$

Since the sum in this expression consists of a finite number of terms, each of which approaches 0 as $n \rightarrow \infty$, we may choose an $N > 0$ such that the sum is less than $\varepsilon/2$ for $n > N$. Therefore, for $n > N$ we have $|e^x - x_n| < \varepsilon$ which proves that $x_n \rightarrow e^x$. //

Exercises

1. Prove Theorem B2(b).
2. Let $\{x_n\}$ and $\{y_n\}$ be sequences of real numbers such that $x_n \leq y_n$ for all $n \geq N$ where N is fixed. If $x_n \rightarrow x$ and $y_n \rightarrow y$, prove that $x \leq y$.
3. If A is a subset of a metric space X and $x \in X$, prove that $x \in \text{Ext } A$ if and only if x has *some* neighborhood disjoint from A .

4. Prove Theorem B14(b).
5. Prove that any cluster point is the limit of a subsequence.
6. Prove Theorem B22.
7. If $\{x_n\}$ is a sequence of real numbers converging to x , and $L_n = \inf_{k \geq n} x_k$ converges to L , show that $x = L$.
8. Prove Theorem B24.
9. Let $\bar{\mathbb{R}}$ denote the extended real number system, and let $f: (a, b) \subset \mathbb{R} \rightarrow \bar{\mathbb{R}}$. Define

$$\lim_{x \rightarrow y} \sup f(x) = \inf_{\delta > 0} \sup_{0 < |x-y| < \delta} f(x)$$

and suppose that $\lim_{x \rightarrow y} f(x) = L$ (i.e., $\lim_{x \rightarrow y} f(x)$ exists). Show

$$\lim_{x \rightarrow y} \sup f(x) = L.$$

[Hint: Let $S_\delta = \sup_{|x-y| < \delta} f(x)$ and define $S = \inf_\delta S_\delta$. Then note that

$$|S - L| \leq |S - S_\delta| + |S_\delta - f(x)| + |f(x) - L|.$$

10. (a) Let $\{x_n\}$ be a Cauchy sequence in $\mathbb{R}^n$, and assume that $\{x_n\}$ has a cluster point c . Prove that $\{x_n\}$ converges to c .
 (b) Using this result, prove that any Cauchy sequence in $\mathbb{R}^n$ converges to a point of $\mathbb{R}^n$.
11. Show that $\text{Bd } A = \text{Cl } A \cap \text{Cl } A^c$.
12. If $U \subset (X, d)$ is open and $A \subset X$ is dense, show that $\text{Cl } U = \text{Cl}(U \cap A)$.
13. If $\{x_n\}$ is a Cauchy sequence with a convergent subsequence $\{x_{n_k}\}$, show that $\{x_n\}$ converges.

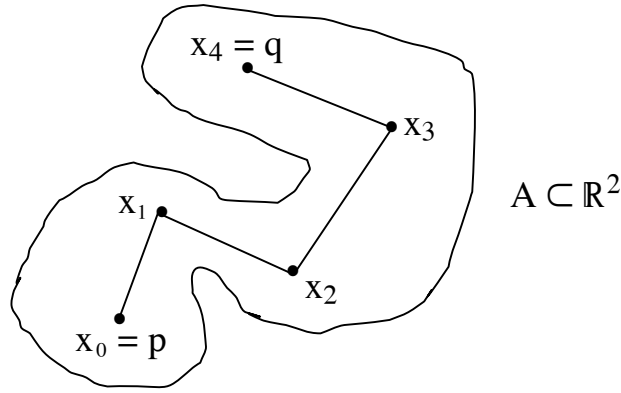
Path Connectedness

In order to avoid having to define a general topological space, we shall phrase this appendix in terms of metric spaces. However, the reader should be aware that this material is far more general than we are presenting it. We assume that the reader has studied Appendix A.

In elementary analysis and geometry, one thinks of a curve as a collection of points whose coordinates are continuous functions of a real variable t . For example, a curve in the plane $\mathbb{R}^2$ may be specified by giving its coordinates $(x = f(t), y = g(t))$ where f and g are continuous functions of the parameter t . If we require that the curve join two points p and q , then the parameter can always be adjusted so that $t = 0$ at p and $t = 1$ at q . Thus we see that the curve is described by a continuous mapping from the unit interval $I = [0, 1]$ into the plane.

Let X be a metric space, and let $I = [0, 1]$ be a subspace of $\mathbb{R}$ with the usual metric. We define a **path** in X , joining two points p and q of X , to be a continuous mapping $f: I \rightarrow X$ such that $f(0) = p$ and $f(1) = q$. This path will be said to lie in a subset $A \subset X$ if $f(I) \subset A$. It is important to realize that *the path is the mapping* f , and not the set of image points $f(I)$. The space X is said to be **path connected** if for every $p, q \in X$ there exists a path in X joining p and q . If $A \subset X$, then A is path connected if every pair of points of A can be joined by a path in A . (We should note that what we have called path connected is sometimes called **arcwise connected**.)

Let us consider for a moment the space $\mathbb{R}^n$. If $x_i, x_j \in \mathbb{R}^n$, then we let $\overline{x_i x_j}$ denote the closed line segment joining x_i and x_j . A subset $A \subset \mathbb{R}^n$ is said to be **polygonally connected** if given any two points $p, q \in A$ there are points $x_0 = p, x_1, x_2, \dots, x_m = q$ in A such that $\bigcup_{i=1}^m \overline{x_{i-1} x_i} \subset A$.



Just because a subset of $\mathbb{R}^n$ is path connected does not mean that it is polygonally connected. For example, the unit circle in $\mathbb{R}^2$ is path connected since it is actually a path itself, but it is not polygonally connected.

Example C1 The space $\mathbb{R}^n$ is path connected. Indeed, if $p \in \mathbb{R}^n$ has coordinates $(x^1, \dots, x^n)$ and $q \in \mathbb{R}^n$ has coordinates $(y^1, \dots, y^n)$, then we define the mapping $f: I \rightarrow \mathbb{R}^n$ by $f(t) = (f^1(t), \dots, f^n(t))$ where $f^i(t) = (1 - t)x^i + ty^i$. This mapping is clearly continuous and satisfies $f(0) = p$ and $f(1) = q$. Thus f is a path joining the arbitrary points p and q of $\mathbb{R}^n$, and hence $\mathbb{R}^n$ is path connected. //

The following is a simple consequence of Theorem A5 that we shall need for our main result (i.e., Theorem C2).

Theorem C1 Let $f: (X_1, d_1) \rightarrow (X_2, d_2)$ and $g: (X_2, d_2) \rightarrow (X_3, d_3)$ both be continuous functions. Then $g \circ f: (X_1, d_1) \rightarrow (X_3, d_3)$ is a continuous function.

Proof If $U \subset X_3$ is open, then the continuity of g shows that $g^{-1}(U) \subset X_2$ is open. Therefore $(g \circ f)^{-1}(U) = (f^{-1} \circ g^{-1})(U) = f^{-1}(g^{-1}(U))$ is open by the continuity of f . ■

Theorem C2 Let f be a continuous mapping from a metric space X onto a metric space Y . Then Y is path connected if X is.

Proof Let x', y' be any two points of Y . Then (since f is surjective) there exist $x, y \in X$ such that $f(x) = x'$ and $f(y) = y'$. Since X is path connected, there exists a path g joining x and y such that $g(0) = x$ and $g(1) = y$. But then $f \circ g$ is a continuous function (Theorem C1) from I into Y such that $(f \circ g)(0) = x'$ and $(f \circ g)(1) = y'$. In other words, $f \circ g$ is a path joining x' and y' , and hence Y is path connected. ■

It is an obvious corollary of Theorem C2 that if f is a continuous mapping from the path connected space X into Y , then $f(X)$ is path connected in Y since f maps X onto the subspace $f(X)$.

Bibliography

This bibliography lists those books referred to in the text as well as many of the books that we found useful in our writing.

Abraham, R., Marsden, J. E. and Ratiu, T., *Manifolds, Tensor Analysis, and Applications*, Addison-Wesley, Reading, MA, 1983.

Adler, R., Bazin, M. and Schiffer, M., *Introduction to General Relativity*, McGraw-Hill, New York, 1965.

Arnold, V. I., *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York, 1978.

Biggs, N. L., *Discrete Mathematics*, Clarendon Press, Oxford, 1985.

Bishop, R. L. and Goldberg, S. I., *Tensor Analysis on Manifolds*, Macmillan, New York, 1968.

Boothby, W. M., *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press, New York, 1975.

Byron, F. W. Jr., and Fuller, R. W., *Mathematics of Classical and Quantum Physics*, Addison-Wesley, Reading, MA, 1969.

Curtis, C. W., *Linear Algebra*, Springer-Verlag, New York, 1984.

Curtis, W. D. and Miller, F. R., *Differential Manifolds and Theoretical Physics*, Academic Press, Orlando, FL, 1985.

Dennery, P. and Krzywicki, A., *Mathematics for Physicists*, Harper & Row, New York, 1967.

Durbin, J. R., *Modern Algebra*, John Wiley & Sons, New York, 1985.

Flanders, H., *Differential Forms*, Academic Press, New York, 1963.

Frankel, T., *The Geometry of Physics*, 2nd edition, Cambridge University Press, New York, NY, 2004.

Frankel, T., *Linear Algebra*, unpublished lecture notes, University of California, San Diego, 1980.

Friedberg, S. H. and Insel, A. J., *Introduction to Linear Algebra with Applications*, Prentice-Hall, Englewood Cliffs, NJ, 1986.

Friedberg, S. H., Insel, A. J. and Spence, L. E., *Linear Algebra*, Prentice-Hall, Englewood Cliffs, NJ, 1979.

Gemignani, M. C., *Elementary Topology*, 2nd edition, Addison-Wesley, Reading, MA, 1972.

Geroch, R., *Mathematical Physics*, University of Chicago Press, Chicago, IL, 1985.

Goldstein, H., *Classical Mechanics*, 2nd edition, Addison-Wesley, Reading, MA, 1980.

Halmos, P. R., *Finite-Dimensional Vector Spaces*, Springer-Verlag, New York, 1974.

Herstein, I. N., *Topics in Algebra*, Xerox College Publishing, Lexington, MA, 1964.

Hoffman, K. and Kunze, R., *Linear Algebra*, 2nd edition, Prentice-Hall, Englewood Cliffs, NJ, 1971.

Jackson, J. D., *Classical Electrodynamics*, 2nd edition, John Wiley & Sons, New York, 1975.

Johnson, R. E., *Linear Algebra*, Prindle, Weber & Schmidt, Boston, MA, 1967.

Knopp, K., *Infinite Sequences and Series*, Dover Publications, New York, 1956.

Kolmogorov, A. N. and Fomin, S. V., *Functional Analysis*, Graylock Press, Rochester, NY, 1957.

Lang, S., *Analysis I*, Addison-Wesley, Reading, MA, 1968.

Lang, S., *Linear Algebra*, 2nd edition, Addison-Wesley, Reading, MA, 1971.

Lang, S., *Real Analysis*, 2nd edition, Addison-Wesley, Reading, MA, 1983.

Lipschutz, S., *Linear Algebra*, Schaum's Outline Series, McGraw-Hill, New York, 1968.

Marcus, M. and Minc, H., *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Boston, MA, 1964.

Marcus, M., *Finite-Dimensional Multilinear Algebra, Parts I and II*, Marcel Dekker, New York, 1973.

Marcus, M., *Introduction to Linear Algebra*, Dover Publications, New York, 1988.

Marion, J. B., *Classical Dynamics of Particles and Systems*, 2nd edition, Academic Press, Orlando, FL, 1970.

Marsden, J. E., *Elementary Classical Analysis*, W. H. Freeman, San Francisco, CA, 1974.

Misner, C. W., Thorne, K. S. and Wheeler, J. A., *Gravitation*, W. H. Freeman, San Francisco, CA, 1973.

Munkres, J. R., *Topology*, Prentice-Hall, Englewood Cliffs, NJ, 1975.

- Murdoch, D. C., *Linear Algebra*, John Wiley & Sons, New York, 1970.
- Prugovecki, E., *Quantum Mechanics in Hilbert Space*, 2nd edition, Academic Press, New York, 1981.
- Reed, M. and Simon, B., *Functional Analysis*, Revised and Enlarged edition, Academic Press, Orlando, FL, 1980.
- Roach, G. F., *Green's Functions*, 2nd edition, Cambridge University Press, Cambridge, 1982.
- Royden, H. L., *Real Analysis*, 2nd edition, Macmillan, New York, 1968.
- Rudin, W., *Functional Analysis*, 2nd edition, McGraw-Hill, New York, 1974.
- Rudin, W., *Principles of Mathematical Analysis*, 3rd edition, McGraw-Hill, New York, 1976.
- Ryder, L. H., *Quantum Field Theory*, Cambridge University Press, Cambridge, 1985.
- Schutz, B., *Geometrical Methods of Mathematical Physics*, Cambridge University Press, Cambridge, 1980.
- Shankar, R., *Principles of Quantum Mechanics*, Plenum Press, New York, 1980.
- Simmons, G. F., *Introduction to Topology and Modern Analysis*, McGraw-Hill, New York, 1963.
- Talman, J. D., *Special Functions*, W. A. Benjamin, New York, 1968.
- Taylor, J. R., *Scattering Theory*, John Wiley & Sons, New York, 1972.
- Tung, W., *Group Theory in Physics*, World Scientific, Philadelphia, PA, 1985.