Metric Spaces, Topological Spaces, and Compactness

A metric space is a set $X$, together with a distance function $d : X \times X \to [0, \infty)$, having the properties that

\[
\begin{align*}
    d(x, y) &= 0 \iff x = y, \\
    d(x, y) &= d(y, x), \\
    d(x, y) &\leq d(x, z) + d(y, z).
\end{align*}
\]

The third of these properties is called the triangle inequality. An example of a metric space is the set of rational numbers $\mathbb{Q}$, with $d(x, y) = |x - y|$. Another example is $X = \mathbb{R}^n$, with

\[
    d(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.
\]

If $(x_\nu)$ is a sequence in $X$, indexed by $\nu = 1, 2, 3, \ldots$, i.e., by $\nu \in \mathbb{Z}^+$, one says $x_\nu \to y$ if $d(x_\nu, y) \to 0$, as $\nu \to \infty$. One says $(x_\nu)$ is a Cauchy sequence if $d(x_\nu, x_\mu) \to 0$ as $\mu, \nu \to \infty$. One says $X$ is a complete metric space if every Cauchy sequence converges to a limit in $X$. Some metric spaces are not complete; for example, $\mathbb{Q}$ is not complete. You can take a sequence $(x_\nu)$ of rational numbers such that $x_\nu \to \sqrt{2}$, which is not rational. Then $(x_\nu)$ is Cauchy in $\mathbb{Q}$, but it has no limit in $\mathbb{Q}$.

If a metric space $X$ is not complete, one can construct its completion $\hat{X}$ as follows. Let an element $\xi$ of $\hat{X}$ consist of an equivalence class of Cauchy
sequences in \( X \), where we say \((x_\nu) \sim (y_\nu)\) provided \(d(x_\nu, y_\nu) \to 0\). We write the equivalence class containing \((x_\nu)\) as \([x_\nu]\). If \(\xi = [x_\nu]\) and \(\eta = [y_\nu]\), we can set \(d(\xi, \eta) = \lim_{\nu \to \infty} d(x_\nu, y_\nu)\) and verify that this is well defined and that it makes \(\hat{X}\) a complete metric space.

If the completion of \(\mathbb{Q}\) is constructed by this process, we get \(\mathbb{R}\), the set of real numbers. This construction provides a good way to develop the basic theory of the real numbers.

There are a number of useful concepts related to the notion of closeness. We define some of them here. First, if \(p\) is a point in a metric space \(X\) and \(r \in (0, \infty)\), the set

\[
(B_r(p)) = \{x \in X : d(x, p) < r\}
\]

is called the open ball about \(p\) of radius \(r\). Generally, a neighborhood of \(p \in X\) is a set containing such a ball, for some \(r > 0\).

A set \(U \subset X\) is called open if it contains a neighborhood of each of its points. The complement of an open set is said to be closed. The following result characterizes closed sets.

**Proposition A.1.** A subset \(K \subset X\) of a metric space \(X\) is closed if and only if

\[
x_j \in K, \ x_j \to p \in X \implies p \in K.
\]

**Proof.** Assume \(K\) is closed, \(x_j \in K, \ x_j \to p\). If \(p \notin K\), then \(p \in X \setminus K\), which is open, so some \(B_\varepsilon(p) \subset X \setminus K\), and \(d(x_j, p) \geq \varepsilon\) for all \(j\). This contradiction implies \(p \in K\).

Conversely, assume (A.3) holds, and let \(q \in U = X \setminus K\). If \(B_{1/n}(q)\) is not contained in \(U\) for any \(n\), then there exists \(x_n \in K \cap B_{1/n}(q)\); hence \(x_n \to q\), contradicting (A.3). This completes the proof.

The following is straightforward.

**Proposition A.2.** If \(U_\alpha\) is a family of open sets in \(X\), then \(\bigcup_\alpha U_\alpha\) is open. If \(K_\alpha\) is a family of closed subsets of \(X\), then \(\bigcap_\alpha K_\alpha\) is closed.

Given \(S \subset X\), we denote by \(\overline{S}\) (the closure of \(S\)) the smallest closed subset of \(X\) containing \(S\), i.e., the intersection of all the closed sets \(K_\alpha \subset X\) containing \(S\). The following result is also straightforward.

**Proposition A.3.** Given \(S \subset X\), \(p \in \overline{S}\) if and only if there exist \(x_j \in S\) such that \(x_j \to p\).
Given $S \subset X$, $p \in X$, we say $p$ is an accumulation point of $S$ if and only if, for each $\varepsilon > 0$, there exists $q \in S \cap B_\varepsilon(p)$, $q \neq p$. It follows that $p$ is an accumulation point of $S$ if and only if each $B_\varepsilon(p)$, $\varepsilon > 0$, contains infinitely many points of $S$. One straightforward observation is that all points of $\overline{S} \setminus S$ are accumulation points of $S$.

The interior of a set $S \subset X$ is the largest open set contained in $S$, i.e., the union of all the open sets contained in $S$. Note that the complement of the interior of $S$ is equal to the closure of $X \setminus S$.

We now turn to the notion of compactness. We say a metric space $X$ is compact provided the following property holds:

(A) Each sequence $(x_k)$ in $X$ has a convergent subsequence.

We will establish various properties of compact metric spaces and provide various equivalent characterizations. For example, it is easily seen that (A) is equivalent to the following:

(B) Each infinite subset $S \subset X$ has an accumulation point.

The following property is known as total boundedness:

Proposition A.4. If $X$ is a compact metric space, then

(C) given $\varepsilon > 0$, $\exists$ a finite set $\{x_1, \ldots, x_N\}$ such that $B_\varepsilon(x_1), \ldots, B_\varepsilon(x_N)$ covers $X$.

Proof. Take $\varepsilon > 0$ and pick $x_1 \in X$. If $B_\varepsilon(x_1) = X$, we are done. If not, pick $x_2 \in X \setminus B_\varepsilon(x_1)$. If $B_\varepsilon(x_1) \cup B_\varepsilon(x_2) = X$, we are done. If not, pick $x_3 \in X \setminus [B_\varepsilon(x_1) \cup B_\varepsilon(x_2)]$. Continue, taking $x_{k+1} \in X \setminus [B_\varepsilon(x_1) \cup \cdots \cup B_\varepsilon(x_k)]$, if $B_\varepsilon(x_1) \cup \cdots \cup B_\varepsilon(x_k) \neq X$. Note that, for $1 \leq i, j \leq k$,

$$i \neq j \implies d(x_i, x_j) \geq \varepsilon.$$ 

If one never covers $X$ this way, consider $S = \{x_j : j \in \mathbb{N}\}$. This is an infinite set with no accumulation point, so property (B) is contradicted.

Corollary A.5. If $X$ is a compact metric space, it has a countable dense subset.

Proof. Given $\varepsilon = 2^{-n}$, let $S_n$ be a finite set of points $x_j$ such that $\{B_\varepsilon(x_j)\}$ covers $X$. Then $\mathcal{C} = \bigcup_n S_n$ is a countable dense subset of $X$.

Here is another useful property of compact metric spaces, which will eventually be generalized even further, in (E) below.
Proposition A.6. Let $X$ be a compact metric space. Assume $K_1 \supset K_2 \supset K_3 \supset \cdots$ form a decreasing sequence of closed subsets of $X$. If each $K_n \neq \emptyset$, then $\bigcap_n K_n \neq \emptyset$.

Proof. Pick $x_n \in K_n$. If (A) holds, $(x_n)$ has a convergent subsequence, $x_{n_k} \to y$. Since $\{x_{n_k} : k \geq \ell\} \subset K_{n_{\ell}}$, which is closed, we have $y \in \bigcap_n K_n$.

Corollary A.7. Let $X$ be a compact metric space. Assume $U_1 \subset U_2 \subset U_3 \subset \cdots$ form an increasing sequence of open subsets of $X$. If $\bigcup_n U_n = X$, then $U_N = X$ for some $N$.

Proof. Consider $K_n = X \setminus U_n$.

The following is an important extension of Corollary A.7.

Proposition A.8. If $X$ is a compact metric space, then it has the property:

(D) every open cover $\{U_\alpha : \alpha \in \mathcal{A}\}$ of $X$ has a finite subcover.

Proof. Each $U_\alpha$ is a union of open balls, so it suffices to show that (A) implies the following:

(D') Every cover $\{B_\alpha : \alpha \in \mathcal{A}\}$ of $X$ by open balls has a finite subcover.

Let $\mathcal{C} = \{z_j : j \in \mathbb{N}\} \subset X$ be a countable dense subset of $X$, as in Corollary A.5. Each $B_\alpha$ is a union of balls $B_{r_j}(z_j)$, with $z_j \in \mathcal{C} \cap B_\alpha$, $r_j$ rational. Thus it suffices to show that

(D'') every countable cover $\{B_j : j \in \mathbb{N}\}$ of $X$ by open balls has a finite subcover.

For this, we set

$$U_n = B_1 \cup \cdots \cup B_n$$

and apply Corollary A.7.

The following is a convenient alternative to property (D):

(E) If $K_\alpha \subset X$ are closed and $\bigcap_\alpha K_\alpha = \emptyset$ then some finite intersection is empty.

Considering $U_\alpha = X \setminus K_\alpha$, we see that

$$(D) \iff (E).$$

The following result completes Proposition A.8.
Theorem A.9. For a metric space $X$,

$$(A) \iff (D).$$

**Proof.** By Proposition A.8, $(A) \Rightarrow (D)$. To prove the converse, it will suffice to show that $(E) \Rightarrow (B)$. So let $S \subset X$ and assume $S$ has no accumulation point. We claim

such $S$ must be closed.

Indeed, if $z \in \overline{S}$ and $z \notin S$, then $z$ would have to be an accumulation point. Say $S = \{x_{\alpha} : \alpha \in \mathcal{A}\}$. Set $K_{\alpha} = S \setminus \{x_{\alpha}\}$. Then each $K_{\alpha}$ has no accumulation point hence $K_{\alpha} \subset X$ is closed. Also $\bigcap_{\alpha} K_{\alpha} = \emptyset$. Hence there exists a finite set $\mathcal{F} \subset \mathcal{A}$ such that $\bigcap_{\alpha \in \mathcal{F}} K_{\alpha} = \emptyset$ if $(E)$ holds. Hence $S = \bigcup_{\alpha \in \mathcal{F}} \{x_{\alpha}\}$ is finite, so indeed $(E) \Rightarrow (B)$.

**Remark.** So far we have that for any metric space $X$,

$$(A) \iff (B) \iff (D) \iff (E) \iff (C).$$

We claim that $(C)$ implies the other conditions if $X$ is *complete*. Of course, compactness implies completeness, but $(C)$ may hold for incomplete $X$, e.g., $X = (0,1) \subset \mathbb{R}$.

**Proposition A.10.** If $X$ is a complete metric space with property $(C)$, then $X$ is compact.

**Proof.** It suffices to show that $(C) \Rightarrow (B)$ if $X$ is a complete metric space. So let $S \subset X$ be an infinite set. Cover $X$ by balls $B_{1/2}(x_1), \ldots, B_{1/2}(x_N)$. One of these balls contains infinitely many points of $S$, and so does its closure, say $X_1 = \overline{B_{1/2}(y_1)}$. Now cover $X$ by finitely many balls of radius 1/4; their intersection with $X_1$ provides a cover of $X_1$. One such set contains infinitely many points of $S$, and so does its closure $X_2 = \overline{B_{1/4}(y_2)} \cap X_1$. Continue in this fashion, obtaining

$$X_1 \supset X_2 \supset X_3 \supset \cdots \supset X_k \supset X_{k+1} \supset \cdots,$$

$$X_j \subset \overline{B_{2^{-j}}(y_j)},$$

each containing infinitely many points of $S$. One sees that $(y_j)$ forms a Cauchy sequence. If $X$ is complete, it has a limit, $y_j \to z$, and $z$ is seen to be an accumulation point of $S$. 
If \( X_j \), \( 1 \leq j \leq m \), is a finite collection of metric spaces, with metrics \( d_j \), we can define a Cartesian product metric space

\[
X = \prod_{j=1}^{m} X_j, \quad d(x,y) = d_1(x_1,y_1) + \cdots + d_m(x_m,y_m).
\]

Another choice of metric is \( \delta(x,y) = \sqrt{d_1(x_1,y_1)^2 + \cdots + d_m(x_m,y_m)^2} \). The metrics \( d \) and \( \delta \) are equivalent; i.e., there exist constants \( C_0, C_1 \in (0,\infty) \) such that

\[
C_0 \delta(x,y) \leq d(x,y) \leq C_1 \delta(x,y), \quad \forall x, y \in X.
\]

A key example is \( \mathbb{R}^m \), the Cartesian product of \( m \) copies of the real line \( \mathbb{R} \).

We describe some important classes of compact spaces.

**Proposition A.11.** If \( X_j \) are compact metric spaces, \( 1 \leq j \leq m \), so is \( X = \prod_{j=1}^{m} X_j \).

**Proof.** If \( (x_\nu) \) is an infinite sequence of points in \( X \), say \( x_\nu = (x_{1\nu}, \ldots, x_{m\nu}) \), pick a convergent subsequence of \( (x_{1\nu}) \) in \( X_1 \) and consider the corresponding subsequence of \( (x_\nu) \), which we relabel \( (x_\nu) \). Using this, pick a convergent subsequence of \( (x_{2\nu}) \) in \( X_2 \). Continue. Having a subsequence such that \( x_{j\nu} \to y_j \) in \( X_j \) for each \( j = 1, \ldots, m \), we then have a convergent subsequence in \( X \).

The following result is useful for calculus on \( \mathbb{R}^n \).

**Proposition A.12.** If \( K \) is a closed bounded subset of \( \mathbb{R}^n \), then \( K \) is compact.

**Proof.** The discussion above reduces the problem to showing that any closed interval \( I = [a,b] \) in \( \mathbb{R} \) is compact. This compactness is a corollary of Proposition A.10. For pedagogical purposes, we redo the argument here, since in this concrete case it can be streamlined.

Suppose \( S \) is a subset of \( I \) with infinitely many elements. Divide \( I \) into two equal subintervals, \( I_1 = [a,b_1] \), \( I_2 = [b_1,b] \), \( b_1 = (a+b)/2 \). Then either \( I_1 \) or \( I_2 \) must contain infinitely many elements of \( S \). Say \( I_j \) does. Let \( x_1 \) be any element of \( S \) lying in \( I_j \). Now divide \( I_j \) into two equal pieces, \( I_j = I_{j1} \cup I_{j2} \). One of these intervals (say \( I_{jk} \)) contains infinitely many points of \( S \). Pick \( x_2 \in I_{jk} \) to be one such point (different from \( x_1 \)). Then subdivide \( I_{jk} \) into two equal subintervals, and continue. We get an infinite sequence of distinct points \( x_\nu \in S \), and \( |x_\nu - x_{\nu+k}| \leq 2^{-\nu}(b-a) \), for \( k \geq 1 \). Since \( \mathbb{R} \) is complete, \( (x_\nu) \) converges, say to \( y \in I \). Any neighborhood of \( y \) contains infinitely many points in \( S \), so we are done.
Appendix A. Metric Spaces, Topological Spaces, and Compactness

If $X$ and $Y$ are metric spaces, a function $f : X \to Y$ is said to be continuous provided $x_\nu \to x$ in $X$ implies $f(x_\nu) \to f(x)$ in $Y$. An equivalent condition, which the reader can verify, is

$$U \text{ open in } Y \implies f^{-1}(U) \text{ open in } X.$$

**Proposition A.13.** If $X$ and $Y$ are metric spaces, $f : X \to Y$ continuous, and $K \subset X$ compact, then $f(K)$ is a compact subset of $Y$.

**Proof.** If $(y_\nu)$ is an infinite sequence of points in $f(K)$, pick $x_\nu \in K$ such that $f(x_\nu) = y_\nu$. If $K$ is compact, we have a subsequence $x_{\nu_j} \to p$ in $X$, and then $y_{\nu_j} \to f(p)$ in $Y$.

If $F : X \to \mathbb{R}$ is continuous, we say $f \in C(X)$. A useful corollary of Proposition A.13 is

**Proposition A.14.** If $X$ is a compact metric space and $f \in C(X)$, then $f$ assumes a maximum and a minimum value on $X$.

**Proof.** We know from Proposition A.13 that $f(X)$ is a compact subset of $\mathbb{R}$. Hence $f(X)$ is bounded, say $f(X) \subset I = [a, b]$. Repeatedly subdividing $I$ into equal halves, as in the proof of Proposition A.12, at each stage throwing out intervals that do not intersect $f(X)$ and keeping only the leftmost and rightmost interval amongst those remaining, we obtain points $\alpha \in f(X)$ and $\beta \in f(X)$ such that $\alpha = f(x_0)$ for some $x_0 \in X$ is the minimum and $\beta = f(x_1)$ for some $x_1 \in X$ is the maximum.

At this point, the reader might take a look at the proof of the Mean Value Theorem, given in Chapter 1, which applies this result.

A function $f \in C(X)$ is said to be uniformly continuous provided that, for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$(A.6) \quad x, y \in X, \; d(x, y) \leq \delta \implies |f(x) - f(y)| \leq \varepsilon.$$ 

An equivalent condition is that $f$ have a modulus of continuity, i.e., a monotonic function $\omega : [0, 1) \to [0, \infty)$ such that $\delta \searrow 0 \Rightarrow \omega(\delta) \searrow 0$ and such that

$$(A.7) \quad x, y \in X, \; d(x, y) \leq \delta \leq 1 \implies |f(x) - f(y)| \leq \omega(\delta).$$

Not all continuous functions are uniformly continuous. For example, if $X = (0, 1) \subset \mathbb{R}$, then $f(x) = \sin 1/x$ is continuous, but not uniformly continuous, on $X$. The following result is useful, for example, in the development of the Riemann integral in Chapter 1.
Appendix A. Metric Spaces, Topological Spaces, and Compactness

Proposition A.15. If $X$ is a compact metric space and $f \in C(X)$, then $f$ is uniformly continuous.

Proof. If not, there exist $x_\nu, y_\nu \in X$ and $\varepsilon > 0$ such that $d(x_\nu, y_\nu) \leq 2^{-\nu}$ but
\begin{equation}
|f(x_\nu) - f(y_\nu)| \geq \varepsilon. \tag{A.8}
\end{equation}

Taking a convergent subsequence $x_{\nu_j} \to p$, we also have $y_{\nu_j} \to p$. Now continuity of $f$ at $p$ implies $f(x_{\nu_j}) \to f(p)$ and $f(y_{\nu_j}) \to f(p)$, contradicting (A.8).

If $X$ and $Y$ are metric spaces, the space $C(X, Y)$ of continuous maps $f : X \to Y$ has a natural metric structure, under some additional hypotheses. We use
\begin{equation}
D(f, g) = \sup_{x \in X} d(f(x), g(x)). \tag{A.9}
\end{equation}

This sup exists provided $f(X)$ and $g(X)$ are bounded subsets of $Y$, where to say $B \subset Y$ is bounded is to say $d : B \times B \to [0, \infty)$ has bounded image. In particular, this supremum exists if $X$ is compact. The following result is frequently useful.

Proposition A.16. If $X$ is a compact metric space and $Y$ is a complete metric space, then $C(X, Y)$, with the metric (A.9), is complete.

Proof. That $D(f, g)$ satisfies the conditions to define a metric on $C(X, Y)$ is straightforward. We check completeness. Suppose $(f_\nu)$ is a Cauchy sequence in $C(X, Y)$, so, as $\nu \to \infty$,
\[
\sup_{k \geq 0} \sup_{x \in X} d(f_{\nu+k}(x), f_\nu(x)) \leq \varepsilon_\nu \to 0.
\]

Then in particular $(f_\nu(x))$ is a Cauchy sequence in $Y$ for each $x \in X$, so it converges, say to $g(x) \in Y$. It remains to show that $g \in C(X, Y)$ and that $f_\nu \to g$ in the metric (A.9).

In fact, taking $k \to \infty$ in the estimate above, we have
\[
\sup_{x \in X} d(g(x), f_\nu(x)) \leq \varepsilon_\nu \to 0,
\]
i.e., $f_\nu \to g$ uniformly. It remains only to show that $g$ is continuous. For this, let $x_j \to x$ in $X$ and fix $\varepsilon > 0$. Pick $N$ so that $\varepsilon_N < \varepsilon$. Since $f_N$ is continuous, there exists $J$ such that $j \geq J \Rightarrow d(f_N(x_j), f_N(x)) < \varepsilon$. Hence
\[
j \geq J \Rightarrow d(g(x_j), g(x)) \leq d(g(x_j), f_N(x_j)) + d(f_N(x_j), f_N(x)) + d(f_N(x), g(x)) + d(f_N(x), g(x)) < 3\varepsilon.
\]
This completes the proof.

We next give a couple of slightly more sophisticated results on compactness. The following extension of Proposition A.11 is a special case of Tychonov’s Theorem.

**Proposition A.17.** If \( \{X_j : j \in \mathbb{Z}^+\} \) are compact metric spaces, so is \( X = \prod_{j=1}^{\infty} X_j \).

Here, we can make \( X \) a metric space by setting

\[
d(x, y) = \sum_{j=1}^{\infty} \frac{2^{-j} d_j(x, y)}{1 + d_j(x, y)}.
\]

It is easy to verify that, if \( x_\nu \in X \), then \( x_\nu \to y \) in \( X \), as \( \nu \to \infty \), if and only if, for each \( j \), \( p_j(x_\nu) \to p_j(y) \) in \( X_j \), where \( p_j : X \to X_j \) is the projection onto the \( j \)th factor.

**Proof.** Following the argument in Proposition A.11, if \( (x_\nu) \) is an infinite sequence of points in \( X \), we obtain a nested family of subsequences

\[
(A.11) \quad (x_\nu) \supset (x_1^\nu) \supset (x_2^\nu) \supset \cdots \supset (x_j^\nu) \supset \cdots
\]

such that \( p_\ell(x_j^\nu) \) converges in \( X_\ell \), for \( 1 \leq \ell \leq j \). The next step is a diagonal construction. We set

\[
(A.12) \quad \xi_\nu = x_j^\nu \in X.
\]

Then, for each \( j \), after throwing away a finite number \( N(j) \) of elements, one obtains from \( (\xi_\nu) \) a subsequence of the sequence \( (x_j^\nu) \) in (A.11), so \( p_\ell(\xi_\nu) \) converges in \( X_\ell \) for all \( \ell \). Hence \( (\xi_\nu) \) is a convergent subsequence of \( (x_\nu) \).

The next result is a special case of Ascoli’s Theorem.

**Proposition A.18.** Let \( X \) and \( Y \) be compact metric spaces, and fix a modulus of continuity \( \omega(\delta) \). Then

\[
(A.13) \quad \mathcal{C}_\omega = \{ f \in C(X, Y) : d(f(x_1), f(x_2)) \leq \omega(d(x_1, x_2)), \forall x_1, x_2 \in X \}
\]

is a compact subset of \( C(X, Y) \).

**Proof.** Let \( (f_\nu) \) be a sequence in \( \mathcal{C}_\omega \). Let \( \Sigma \) be a countable dense subset of \( X \), as in Corollary A.5. For each \( x \in \Sigma \), \( (f_\nu(x)) \) is a sequence in \( Y \), which hence has a convergent subsequence. Using a diagonal construction similar
to that in the proof of Proposition A.17, we obtain a subsequence \((\varphi_\nu)\) of 
\((f_\nu)\) with the property that \(\varphi_\nu(x)\) converges in \(Y\), for each \(x \in \Sigma\), say

\[(A.14) \quad x \in \Sigma \implies \varphi_\nu(x) \to \psi(x),\]

where \(\psi : \Sigma \to Y\).

So far, we have not used (A.13), but this hypothesis readily yields

\[(A.15) \quad d(\psi(x), \psi(y)) \leq \omega(d(x, y)),\]

for all \(x, y \in \Sigma\). Using the denseness of \(\Sigma \subset X\), we can extend \(\psi\) uniquely to a continuous map of \(X \to Y\), which we continue to denote by \(\psi\). Also, (A.15) holds for all \(x, y \in X\), i.e., \(\psi \in C_\omega\).

It remains to show that \(\varphi_\nu \to \psi\) uniformly on \(X\). Pick \(\varepsilon > 0\). Then pick \(\delta > 0\) such that \(\omega(\delta) < \varepsilon/3\). Since \(X\) is compact, we can cover \(X\) by finitely many balls \(B_\delta(x_j), 1 \leq j \leq N, x_j \in \Sigma\). Pick \(M\) so large that \(\varphi_\nu(x_j)\) is within \(\varepsilon/3\) of its limit for all \(\nu \geq M\) (when \(1 \leq j \leq N\)). Now, for any \(x \in X\), picking \(\ell \in \{1, \ldots, N\}\) such that \(d(x, x_\ell) \leq \delta\), we have, for \(k \geq 0, \nu \geq M\),

\[(A.16) \quad d(\varphi_{\nu+k}(x), \varphi_\nu(x)) \leq d(\varphi_{\nu+k}(x), \varphi_{\nu+k}(x_\ell)) + d(\varphi_{\nu+k}(x_\ell), \varphi_\nu(x_\ell)) + d(\varphi_\nu(x_\ell), \varphi_\nu(x)) \leq \varepsilon/3 + \varepsilon/3 + \varepsilon/3,

proving the proposition.

We next define the notion of a connected space. A metric space \(X\) is said to be connected provided that it cannot be written as the union of two disjoint open subsets. The following is a basic class of examples.

**Proposition A.19.** An interval \(I\) in \(\mathbb{R}\) is connected.

**Proof.** Suppose \(A \subset I\) is nonempty, with nonempty complement \(B \subset I\), and both sets are open. Take \(a \in A, b \in B\); we can assume \(a < b\). Let \(\xi = \sup\{x \in [a, b] : x \in A\}\). This exists as a consequence of the basic fact that \(\mathbb{R}\) is complete.

Now we obtain a contradiction, as follows. Since \(A\) is closed, \(\xi \in A\). But then, since \(A\) is open, there must be a neighborhood \((\xi - \varepsilon, \xi + \varepsilon)\) contained in \(A\); this is not possible.

We say \(X\) is path-connected if, given any \(p, q \in X\), there is a continuous map \(\gamma : [0, 1] \to X\) such that \(\gamma(0) = p\) and \(\gamma(1) = q\). It is an easy consequence of Proposition A.19 that \(X\) is connected whenever it is path-connected.

The next result, known as the Intermediate Value Theorem, is frequently useful.
Appendix A. Metric Spaces, Topological Spaces, and Compactness

Proposition A.20. Let $X$ be a connected metric space and $f : X \to \mathbb{R}$ continuous. Assume $p, q \in X$ and $f(p) = a < f(q) = b$. Then, given any $c \in (a, b)$, there exists $z \in X$ such that $f(z) = c$.

Proof. Under the hypotheses, $A = \{ x \in X : f(x) < c \}$ is open and contains $p$, while $B = \{ x \in X : f(x) > c \}$ is open and contains $q$. If $X$ is connected, then $A \cup B$ cannot be all of $X$; so any point in its complement has the desired property.

We turn now to the notion of a topological space. This is a set $X$, together with a family $\mathcal{O}$ of subsets, called open, satisfying the following conditions:

$$\begin{align*}
X, \emptyset &\text{ open,} \\
U_j \text{ open, } 1 \leq j \leq N &\Rightarrow \bigcap_{j=1}^{N} U_j \text{ open,} \\
U_\alpha \text{ open, } \alpha \in A &\Rightarrow \bigcup_{\alpha \in A} U_\alpha \text{ open,}
\end{align*}$$

(A.17)

where $A$ is any index set. It is obvious that the collection of open subsets of a metric space, defined above, satisfies these conditions. As before, a set $S \subset X$ is closed provided $X \setminus S$ is open. Also, we say a subset $N \subset X$ containing $p$ is a neighborhood of $p$ provided $N$ contains an open set $U$ which in turn contains $p$.

If $X$ is a topological space and $S$ is a subset, $S$ gets a topology as follows. For each $U$ open in $X$, $U \cap S$ is declared to be open in $S$. This is called the induced topology.

A topological space $X$ is said to be Hausdorff provided that any distinct $p, q \in X$ have disjoint neighborhoods. Clearly any metric space is Hausdorff. Most important topological spaces are Hausdorff.

A Hausdorff topological space is said to be compact provided the following condition holds. If $\{ U_\alpha : \alpha \in A \}$ is any family of open subsets of $X$, covering $X$, i.e., $X = \bigcup_{\alpha \in A} U_\alpha$, then there is a finite subcover, i.e., a finite subset $\{ U_{\alpha_1}, \ldots, U_{\alpha_N} : \alpha_j \in A \}$ such that $X = U_{\alpha_1} \cup \cdots \cup U_{\alpha_N}$. An equivalent formulation is the following, known as the finite intersection property. Let $\{ S_\alpha : \alpha \in A \}$ be any collection of closed subsets of $X$. If each finite collection of these closed sets has nonempty intersection, then the complete intersection $\bigcap_{\alpha \in A} S_\alpha$ is nonempty. In the first part of this chapter, we have shown that any compact metric space satisfies this condition.

Any closed subset of a compact space is compact. Furthermore, any compact subset of a Hausdorff space is necessarily closed.
Most of the propositions stated above for compact metric spaces have extensions to compact Hausdorff spaces. We mention one nontrivial result, which is the general form of Tychonov’s Theorem. One can find a proof in many topology texts, such as [Dug] or [Mun].

**Theorem A.21.** If $S$ is any nonempty set (possibly uncountable) and, for any $\alpha \in S$, $X_\alpha$ is a compact Hausdorff space, then so is $X = \prod_{\alpha \in S} X_\alpha$.

A Hausdorff space $X$ is said to be **locally compact** provided every $p \in X$ has a neighborhood $N$ which is compact (with the induced topology).

A Hausdorff space is said to be **paracompact** provided every open cover $\{U_\alpha : \alpha \in A\}$ has a locally finite refinement, i.e., an open cover $\{V_\beta : \beta \in B\}$ such that each $V_\beta$ is contained in some $U_\alpha$ and each $p \in X$ has a neighborhood $N_p$ such that $N_p \cap V_\beta$ is nonempty for only finitely many $\beta \in B$. A typical example of a paracompact space is a locally compact Hausdorff space $X$ which is also $\sigma$-compact, i.e., $X = \bigcup_{n=1}^{\infty} X_n$ with $X_n$ compact. Paracompactness is a natural condition under which to construct partitions of unity.

A map $F : X \to Y$ between two topological spaces is said to be **continuous** provided $F^{-1}(U)$ is open in $X$ whenever $U$ is open in $Y$. If $F : X \to Y$ is one-to-one and onto and both $F$ and $F^{-1}$ are continuous, $F$ is said to be a **homeomorphism**. For a bijective map $F : X \to Y$, the continuity of $F^{-1}$ is equivalent to the statement that $F(V)$ is open in $Y$ whenever $V$ is open in $X$; another equivalent statement is that $F(S)$ is closed in $Y$ whenever $S$ is closed in $X$.

If $X$ and $Y$ are Hausdorff, $F : X \to Y$ continuous, then $F(K)$ is compact in $Y$ whenever $K$ is compact in $X$. In view of the discussion above, there arises the following useful sufficient condition for a continuous map $F : X \to Y$ to be a homeomorphism. Namely, if $X$ is compact, $Y$ Hausdorff, and $F$ one-to-one and onto, then $F$ is a homeomorphism.

We turn to a discussion of the Stone-Weierstrass Theorem, which gives a sufficient condition for a family of functions on a compact Hausdorff space $X$ to be dense in the space $C(X)$ of continuous functions on $X$, with the sup norm. This result is used several times in the text. It is an extension of the Weierstrass Approximation Theorem, which we state first.

**Theorem A.22.** If $I = [a, b]$ is an interval in $\mathbb{R}$, the space $\mathcal{P}$ of polynomials in one variable is dense in $C(I)$.

There are many proofs of this. One close to Weierstrass’ original (and my favorite) goes as follows. Given $f \in C(I)$, extend it to be continuous and compactly supported on $\mathbb{R}$, convolve this with a highly peaked Gaussian, and approximate the result by power series.
Such an argument involves a little use of complex variable theory. We sketch another proof of Theorem A.22, which avoids complex variable theory. Our starting point will be the result that the power series for \((1 - x)^a\) converges uniformly on \([-1, 1]\), for any \(a > 0\). This is a simple consequence of Taylor's formula with remainder. We will use it with \(a = 1/2\).

From the identity \(x^{1/2} = (1 - (1 - x))^{1/2}\), we have \(x^{1/2} \in \mathcal{P}([0, 2])\). More to the point, from the identity (A.18)

\[|x| = (1 - (1 - x^2))^{1/2},\]
we have \(|x| \in \mathcal{P}([-\sqrt{2}, \sqrt{2}])\). Using \(|x| = b^{-1}|bx|\), for any \(b > 0\), we see that \(|x| \in \mathcal{P}(I)\) for any interval \(I = [-c, c]\) and also for any closed subinterval, hence for any compact interval \(I\). By translation, we have

(A.19) \(|x - a| \in \mathcal{P}(I)|\)
for any compact interval \(I\). Using the identities

(A.20) \(\max(x, y) = \frac{1}{2}(x + y) + \frac{1}{2}|x - y|, \quad \min(x, y) = \frac{1}{2}(x + y) - \frac{1}{2}|x - y|,\)
we see that for any \(a \in \mathbb{R}\) and any compact \(I\),

(A.21) \(\max(x, a), \min(x, a) \in \mathcal{P}(I).\)

We next note that \(\mathcal{P}(I)\) is an algebra of functions, i.e.,

(A.22) \(f, g \in \mathcal{P}(I), c \in \mathbb{R} \implies f + g, fg, cf \in \mathcal{P}(I).\)

Using this, one sees that, given \(f \in \mathcal{P}(I)\), with range in a compact interval \(J\), one has \(h \circ f \in \mathcal{P}(I)\) for all \(h \in \mathcal{P}(J)\). Hence \(f \in \mathcal{P}(I) \implies |f| \in \mathcal{P}(I)\), and, via (A.20), we deduce that

(A.23) \(f, g \in \mathcal{P}(I) \implies \max(f, g), \min(f, g) \in \mathcal{P}(I).\)

Suppose now that \(I' = [a', b']\) is a subinterval of \(I = [a, b]\). With the notation \(x_+ = \max(x, 0)\), we have

(A.24) \(f_{I'}(x) = \min((x - a')_+, (b' - x)_+) \in \mathcal{P}(I).\)
This is a piecewise linear function, equal to zero off \(I \setminus I'\), with slope 1 from \(a'\) to the midpoint \(m'\) of \(I'\) and slope \(-1\) from \(m'\) to \(b'\).

Now if \(I\) is divided into \(N\) equal subintervals, any continuous function on \(I\) that is linear on each such subinterval can be written as a linear combination of such “tent functions,” so it belongs to \(\mathcal{P}(I)\). Finally, any \(f \in C(I)\) can be uniformly approximated by such piecewise linear functions, so we have \(f \in \mathcal{P}(I)\), proving the theorem.

We now state the Stone-Weierstrass Theorem.
Appendix A. Metric Spaces, Topological Spaces, and Compactness

**Theorem A.23.** Let $X$ be a compact Hausdorff space, $\mathcal{A}$ a subalgebra of $C_\mathbb{R}(X)$, the algebra of real-valued continuous functions on $X$. Suppose $1 \in \mathcal{A}$ and that $\mathcal{A}$ separates points of $X$, i.e., for distinct $p, q \in X$, there exists $h_{pq} \in \mathcal{A}$ with $h_{pq}(p) \neq h_{pq}(q)$. Then the closure $\overline{\mathcal{A}}$ is equal to $C_\mathbb{R}(X)$.

We give a proof of Theorem A.23, making use of the argument above, which implies that, if $f \in \overline{\mathcal{A}}$ and $\varphi : \mathbb{R} \to \mathbb{R}$ is continuous, then $\varphi \circ f \in \overline{\mathcal{A}}$. Consequently, if $f \in \overline{\mathcal{A}}$, then $\sup(f_1, f_2) = (1/2)|f_1 - f_2| + (1/2)(f_1 + f_2) \in \overline{\mathcal{A}}$. Similarly $\inf(f_1, f_2) \in \overline{\mathcal{A}}$.

The hypothesis of separating points implies that, for distinct $p, q \in X$, there exists $f_{pq} \in \overline{\mathcal{A}}$, equal to 1 at $p$, 0 at $q$. Applying appropriate $\varphi$, we can arrange also that $0 \leq f_{pq}(x) \leq 1$ on $X$ and that $f_{pq}$ is 1 near $p$ and 0 near $q$. Taking infima, we can obtain $f_{pU} \in \overline{\mathcal{A}}$, equal to 1 on a neighborhood of $p$ and equal to 0 off a given neighborhood $U$ of $p$. Applying sups to these, we obtain for each compact $K \subset X$ and open $U \supset K$ a function $g_{KU} \in \overline{\mathcal{A}}$ such that $g_{KU}$ is 1 on $K$, 0 off $U$, and $0 \leq g_{KU}(x) \leq 1$ on $X$. Once we have gotten this far, it is easy to approximate any continuous $u \in \overline{\mathcal{A}}$ on $X$ by a sup of (positive constants times) such $g_{KU}$’s and from there to prove the theorem.

Theorem A.23 has a complex analogue. In that case, we add the assumption that $f \in \mathcal{A} \Rightarrow \overline{f} \in \mathcal{A}$ and conclude that $\overline{\mathcal{A}} = C(X)$. This is easily reduced to the real case.

The next result is known as the Contraction Mapping Principle, and it has many uses in analysis. In particular, we will use it in the proof of the Inverse Function Theorem, in Appendix B.

**Theorem A.24.** Let $X$ be a complete metric space, and let $T : X \to X$ satisfy

\[(A.25) \quad \text{dist}(Tx, Ty) \leq r \text{dist}(x, y),\]

for some $r < 1$. (We say $T$ is a contraction.) Then $T$ has a unique fixed point $x$. For any $y_0 \in X$, $T^k y_0 \to x$ as $k \to \infty$.

**Proof.** Pick $y_0 \in X$ and let $y_k = T^k y_0$. Then $\text{dist}(y_{k+1}, y_k) \leq r^k \text{dist}(y_1, y_0)$, so

\[(A.26) \quad \text{dist}(y_{k+m}, y_k) \leq \text{dist}(y_{k+m}, y_{k+m-1}) + \cdots + \text{dist}(y_{k+1}, y_k) \leq (r^k + \cdots + r^{k+m-1}) \text{dist}(y_1, y_0) \leq r^k(1 - r)^{-1} \text{dist}(y_1, y_0).\]

It follows that $(y_k)$ is a Cauchy sequence, so it converges; $y_k \to x$. Since $T y_k = y_{k+1}$ and $T$ is continuous, it follows that $Tx = x$, i.e., $x$ is a fixed
point. Uniqueness of the fixed point is clear from the estimate $\text{dist}(Tx, Tx') \leq r \text{ dist}(x, x')$, which implies $\text{dist}(x, x') = 0$ if $x$ and $x'$ are fixed points. This proves Theorem A.24.