Appendix
A Primer on Topological Spaces

This appendix aims to provide a succinct, and yet self-contained, discussion of some basic elements of general topology and metric analysis. We concentrate here only on those aspects of topological notions and metric space concepts that are relevant for the present coverage of order theory. Furthermore, for brevity, we leave the proofs of most of the (easy) propositions as exercises. All major theorems that are essential to the present discussion are proved, however.

1 Topological Spaces

1.1 Basic Definitions

We begin with the axiomatic definition of the notion of “topological space.”

**Definition.** Let $X$ be a nonempty set. A **topology** on $X$ is a collection $\mathcal{O}_X$ of subsets of $X$ such that

- $\emptyset \in \mathcal{O}_X$ and $X \in \mathcal{O}_X$;
- $\bigcup U \in \mathcal{O}_X$ for any nonempty subset $U$ of $\mathcal{O}_X$; and
- $\bigcap U \in \mathcal{O}_X$ for any nonempty finite subset $U$ of $\mathcal{O}_X$.

Formally speaking, a **topological space** is a list $(X, \mathcal{O}_X)$ where $X$ is a nonempty set and $\mathcal{O}_X$ is a topology on $X$. However, in practice, we omit giving specific reference to the involved topology, and refer to $X$ itself as a topological space.

We shall denote the topology of an arbitrarily given topological space $X$ as $\mathcal{O}_X$ in what follows. In addition, we refer to every member of $\mathcal{O}_X$ as **open** (in $X$). An open...
subset of $X$ that contains an element $x$ of $X$ is, in turn, called an **open neighborhood** of $x$. More generally, a **neighborhood** of $x$ is a subset of $X$ that contains an **open neighborhood** of $x$. Finally, we refer to any subset $S$ of $X$ such that $X \setminus S$ is open as **closed**. The collection of all closed subsets of $X$ is denoted as $\mathcal{C}_X$. Finally, any subset of $X$ that is both open and closed, such as $\emptyset$ and $X$, is called **clopen**.

**Note.** $\mathcal{C}_X$ is nonempty for any topological space $X$, for it surely contains $\emptyset$ and $X$. This collection is closed under arbitrary intersections (that is, the intersection of any collection of closed sets is closed), and it is closed under finite unions (that is, the union of finitely many closed sets is closed).

**Example 1.1.1.** Let $X$ be a nonempty set. Then, $2^X$ and $\{\emptyset, X\}$ are topologies on $X$. We say that $X$ is a **discrete space** if it is endowed with the topology $2^X$, and that it is an **indiscrete space** if it is endowed with the topology $\{\emptyset, X\}$. In a discrete space, every set is open. (In particular, in this case, $\{x\}$ is an open neighborhood of $x$ for any $x \in X$.) This, in turn, implies that every subset of a discrete space is closed as well. By contrast, the only open (and hence closed) subsets of an indiscrete space $X$ is $\emptyset$ and $X$.

Let $X$ be a topological space, and $S$ a subset of $X$. The **interior** of $S$ (in $X$) – denoted as $\text{int}(S)$ – is the $\supseteq$-maximum open set in $X$ that is contained in $S$. It is easy to see that this means

$$\text{int}(S) = \bigcup \{O \in \mathcal{O}_X : O \subseteq S\},$$

which, in particular, shows that $S$ is open iff $\text{int}(S) = S$. Similarly, the **closure** of $S$ (in $X$) – denoted as $\text{cl}(S)$ – is the $\subseteq$-minimum closed set in $X$ that contains $S$. It is easy to see that this means

$$\text{cl}(S) = \bigcap \{C \in \mathcal{C}_X : S \subseteq C\},$$

which, in particular, shows that $S$ is closed iff $\text{cl}(S) = S$. (The set $\text{cl}(S) \setminus \text{int}(S)$ is called the **boundary** of $S$.) Finally, we say that $S$ is **dense** in $X$, if $\text{cl}(S) = X$. (Dense sets are thought of as “large” from the topological point of view.)

The following proposition, whose proof is easy, gives an alternative characterization of the closure of a set.

**Proposition 1.1.1.** Let $X$ be a topological space, and $S$ a subset of $X$. Then, an element $x$ of $X$ belongs to $\text{cl}(S)$ if, and only if, every open neighborhood of $x$ intersects $S$.

As an immediate corollary we get:
Corollary 1.1.2. A subset $S$ of a topological space $X$ is dense if, and only if, every open set in $X$ intersects $S$.

We introduce next a very important class of topological spaces.

Definition. We say that a topological space $X$ (or the topology of that space) is Hausdorff, if for any distinct elements $x$ and $y$ of $X$, there exist disjoint open subsets $O$ and $U$ of $X$ such that $x \in O$ and $y \in U$.

It is easily checked that every singleton, and hence every finite set, in a Hausdorff topological space is closed. (We shall prove something stronger than this shortly.) It is thus fortunate that most (although not all) topological spaces that one encounters in practice are Hausdorff.

Example 1.1.2. A discrete space is Hausdorff, because every singleton is open in that space. By contrast, an indiscrete space is not Hausdorff, unless it contains a single element.

1.2 Basis for a Topology

In practice, one rarely specifies all contents of a topology. Indeed, most topologies are defined first by choosing a particular collection of sets to be called “open” in the space, and then constructing the other “open” sets by taking unions of these prespecified sets. To be able to do this, however, that collection of sets must have some special properties.

Definition. Let $X$ be a nonempty set. A collection $\mathcal{B}$ of subsets of $X$ is said to be a basis for a topology on $X$ if, for each $x$ in $X$,

- there is an element $B$ of $\mathcal{B}$ with $x \in B$; and
- if $x \in A \cap B$ for some $A$ and $B$ in $\mathcal{B}$, then there is an element $C$ of $\mathcal{B}$ with $x \in C \subseteq A \cap B$.

If $\mathcal{B}$ is a basis for a topology on $X$, then the topology generated by $\mathcal{B}$ is the collection of all subsets $O$ of $X$ such that

$$x \in O \quad \text{implies} \quad x \in B \subseteq O \quad \text{for some} \quad B \in \mathcal{B}.$$
Put differently, if $\mathcal{B}$ is a basis for a topology on $X$, then a subset $O$ of $X$ is open relative to this topology iff for every $x$ in $O$ there is a $B$ in $\mathcal{B}$ such that

$$x \in B \subseteq O.$$ 

Equivalently, $O$ is open relative to the topology generated by $\mathcal{B}$ iff

$$O = \bigcup\{B : B \in \mathcal{B} \text{ and } B \subseteq O\}.$$ 

Of course, for any of this to make sense, we need to check that the topology generated by a basis is indeed a topology. We leave this as an (easy) exercise.

**Example 1.2.1.** Let $\mathcal{B}$ stand for the collection of all open and bounded intervals. Then, $\mathcal{B}$ is a basis for a topology, and the topology generated by $\mathcal{B}$ on $\mathbb{R}$ is one relative to which a subset $O$ of $\mathbb{R}$ is open iff for every $x$ in $O$ there is a real number $\varepsilon > 0$ such that $(x - \varepsilon, x + \varepsilon) \subseteq O$. (This is called the **standard topology** of $\mathbb{R}$.) This topology is, obviously, Hausdorff.

**Example 1.2.2.** Let $(X, \succeq)$ be a loset with $|X| \geq 2$. Let $\mathcal{B}$ be the collection of all sets of the form $\uparrow x$, or $\downarrow x$, or $\uparrow x \cap \downarrow x$, where $x$ and $y$ are arbitrary elements of $X$. Then, $\mathcal{B}$ is a basis for a topology on $X$. The topology generated by $\mathcal{B}$ is called the $\succeq$-**order topology** on $X$. This topology is Hausdorff. Indeed, if $x$ and $y$ are two distinct elements of $X$, and say, $x \succeq y$, then either $x \succ z \succ y$ for some $z \in X$, or there is no such $z$ in $X$. Letting $O := \uparrow z$ and $U := \downarrow z$ in the former case, and $O := \uparrow y$ and $U := \downarrow x$ in the latter, we see that there exist two disjoint open subsets $O$ and $U$ of $X$ such that $x \in O$ and $y \in U$.

Topological spaces that have countable bases, that is, those spaces whose topologies are generated by bases that contain countably many elements, appear rather frequently in analysis. Such spaces are called **second-countable**, and enjoy a number of useful properties. Even more general than these are the **first-countable** spaces.

**Definition.** Let $X$ be a topological space. If, for any $x$ in $X$, there exists a countable collection $\mathcal{B}(x)$ of open neighborhoods of $x$ such that

$$x \in O \in \mathcal{O}_x \implies x \in B \in O \text{ for some } B \in \mathcal{B}(x),$$

we say that $X$ is **first-countable** (or that “$X$ has a countable basis at each of its points.”) More specifically, if the topology of $X$ is generated by a countable basis, we say that $X$ is **second-countable** (or that “$X$ has a countable basis.”)
Note. Every second-countable topological space is first-countable, but, as we shall see later, the converse of this is not true.

**Example 1.2.3.** Every discrete (or indiscrete) space is second-countable.

**Example 1.2.4.** $\mathbb{R}$ is second-countable, because the collection of all open intervals with rational endpoints constitute a basis for the topology of $\mathbb{R}$.

One reason why first-countable topological spaces is useful in analysis is that in such spaces the notions of being “open” and “closed” can be captured by means of convergent sequences. We turn to this matter next.

**Definition.** Let $X$ be a topological space and $(x_m)$ a sequence in $X$. We say that $(x_m)$ converges to some element $x$ of $X$ – in this case we refer to $(x_m)$ as convergent and $x$ as a limit of $(x_m)$ – if for every open neighborhood of $O$, there is a positive integer $M$ such that $x_m \in O$ for each $m \geq M$. We denote this situation by writing

$$x_m \rightarrow x \quad \text{or} \quad \lim x_m = x.$$

**Note.** If the space $X$ has a basis, say $B$, then $x_m \rightarrow x$ iff for every element $B$ of $B$ that contains $x$, we have $x_m \in B$ for all but finitely many $m$.

Put differently, a sequence $(x_m)$ in a topological space $X$ converges to a point $x$ in this space, if every open neighborhood of $x$ contains all but finitely many terms of this sequence. So, for instance, the real sequence $(1, \frac{1}{2}, \frac{1}{3}, \ldots)$ converges to zero (with respect to the standard topology of $\mathbb{R}$), because, for any basis open set that contains 0, say, $(−\varepsilon, \varepsilon)$, there exists a positive integer $M$ with $1/m \in (−\varepsilon, \varepsilon)$ for each $m \geq M$.

**Note.** A convergent sequence in a topological space $X$ need not have a unique limit. (In an indiscrete space, for instance, every sequence converges to every point in the space.) But, if $X$ is Hausdorff, then a sequence in $X$ can have at most one limit. We leave the (easy) proof of this fact as an exercise.

The following theorem brings sequences into play in the case of first-countable topological spaces.
Proposition 1.2.1. Let $X$ be a first-countable topological space. Then, a subset $S$ of $X$ is closed if, and only if, the limit of any convergent sequence in $S$ belongs to $S$.

Proof. Let $S$ be a closed subset of $X$, and take any sequence $(x_m)$ in $S$ such that $x_m \to x$ for some $x$ in $X$. Suppose $x$ lies in $X \setminus S$. As $x_m \to x$, and $X \setminus S$ is open in $X$, all but finitely many terms of $(x_m)$ belong to $X \setminus S$, which is a contradiction.

Conversely, suppose $S$ is not a closed subset of $X$. To complete our proof, it is enough to find a sequence in $S$ that converges to a point outside $S$.

To this end, observe that there is an element $x$ in $X \setminus S$ every open neighborhood of which intersects $S$. (If there was no such $x$, then $\text{cl}(S) = S$, that is, $S$ would be closed.) As $X$ is first-countable, there is a countable collection $\mathcal{B}(x)$ of open neighborhoods of $x$ such that (1) holds. We enumerate $\mathcal{B}(x)$ as $\{B_1, B_2, \ldots\}$, and define $O_1 := B_1$ and

$$O_m := B_1 \cap \cdots \cap B_m \quad \text{for each } m = 1, 2, \ldots$$

Then, each $O_m$ is an open neighborhood of $x$, and $O_1 \supseteq O_2 \supseteq \cdots$. As $O_m \cap S \neq \emptyset$, we can use the Axiom of Choice to pick an element $x_m$ from $O_m \cap S$, for each $m$. Clearly, we will be done if we can show that $x_m \to x$. But, if $O$ is an open neighborhood of $x$, by (1), there is a positive integer $M$ with

$$x \in O_M \subseteq B_M \subseteq O.$$ 

Consequently, $x_m \in O$ for each $m \geq M$, and we may conclude that $x_m \to x$.

Second-countable topological spaces possess even more desirable properties than first-countable ones. Here is one major example:

Proposition 1.2.2. Let $X$ be a topological space with a countable basis. Then, there exists a countable dense subset of $X$.

Proof. Let $\mathcal{B}$ be a countable basis for the topology of $X$. By the Axiom of Choice, there is a function $f : \mathcal{B} \to \bigcup \mathcal{B}$ such that $f(B) \in B$ for every $B \in \mathcal{B}$. Then, $f(\mathcal{B})$ is a countable subset of $X$ which is dense in $X$.

We conclude this section by reviewing a method of “topologizing” a given nonempty set $X$. A collection $\mathcal{S}$ of subsets of $X$ is said to be a subbasis for a topology on $X$ if the union of its members equal $X$. One can easily check that the collection of all intersections of finitely many elements of $\mathcal{S}$ is a basis for a topology on $X$. We refer to this topology
as “the topology of $X$ generated by using $S$ as a subbasis.” In practice, many topologies are introduced by this method. We give two examples.

**Example 1.2.5.** Let $(X, \succ)$ be a poset with $|X| \geq 2$. Let $S$ be the collection of all sets of the form $\uparrow x$, or $\downarrow x$, where $x$ and $y$ are arbitrary elements of $X$. The topology on $X$ generated by using $S$ as a subbasis is called \textbf{$\succ$-interval topology} on $X$, and reduces to the $\succ$-order topology on $X$ when $\succ$ is a linear order. Relative to this topology, for instance, a subset of $X$ is closed iff it is the intersection of a collection of subsets of $X$ each of which is a finite union of sets of the form $x\uparrow$ or $x\downarrow$. (Verify!)

The interval topology is not Hausdorff in general. For instance, consider the partial order $\succ$ on the set $\{x, y, a, b\}$ such that $a \succ \{x, y\}$ and $\{x, y\} \succ b$ with no other comparabilities. We leave it as an exercise to check that the $\succ$-interval topology on $\{x, y, a, b\}$ is not Hausdorff.

**Note.** There are many other interesting topologies one may induce by using the order structure of a poset (or a lattice). See Gierz et al. (2003) for an extensive discussion of this matter.

**Example 1.2.6.** Let $X$ be an arbitrary nonempty set and $Y$ a topological space. Recall that $Y^X$ stands for the collection of all maps from $X$ into $Y$. We next describe a standard way of topologizing $Y^X$. For any $x$ in $X$ and any open subset $O$ of $Y$, define

$$U(x, O) := \{ f \in Y^X : f(x) \in O \},$$

and let $S := \{ U(x, O) : x \in X \text{ and } O \in \mathcal{O}_Y \}$. The topology on $Y^X$ generated by using $S$ as a subbasis is called \textbf{topology of pointwise convergence} on $Y^X$. The name for this topology owes to the following fact: A sequence $(f_m)$ in $Y^X$ converges to a map $f : X \to Y$ under the topology of pointwise convergence iff $f_m(x) \to f(x)$ for each $x \in X$ (relative to the topology of $Y$). The proof is not difficult, and is left as an exercise.

We note that, for any positive integer $n$, $\mathbb{R}^n$ is none other than $\mathbb{R}^n$. In this case, endowing $\mathbb{R}$ with its standard topology, and $\mathbb{R}^n$ with the topology of pointwise convergence, yields the \textbf{standard topology} on $\mathbb{R}^n$.

### 1.3 The Subspace Topology

There is a natural way of regarding a subset of a topological space itself as a topological space.
Definition. Let $X$ be a topological space and $Y$ a nonempty subset of $X$. Then, the collection

$$\{ Y \cap O : O \in \mathcal{O}_X \},$$

which is easily checked to be a topology on $Y$, is called the **subspace** (or **relative**) topology on $Y$. We refer to $Y$ as a subspace of $X$ when it is endowed with the subspace topology.

Note. If $\mathcal{B}$ is a basis for the topology of $X$, then $\{ Y \cap B : B \in \mathcal{B} \}$ is a basis for the subspace topology on $Y$.

Let $X$ be a topological space, and $Y$ a subspace of $X$. Then, a set $U$ in $Y$ is open (closed) with respect to the subspace topology on $Y$ – we refer to this situation by saying “$U$ is open (closed) in $Y$” – if $U$ is the intersection of an open (closed) subset of $X$ with $Y$. It follows that if $Y$ is open (closed) in $X$, and $U$ is open (closed) in $Y$, then $U$ is open (closed) in $X$ as well.

A property that is passed from any topological space to any of its subspaces is called **hereditary**. It is plain that the properties of being Hausdorff, first-countable and second-countable are hereditary. That is, $Y$ is Hausdorff if so is $X$, and it has a countable basis (at each of its points), if so does $X$. (Check!)

From now on, and throughout this text, when we refer to a nonempty subset of a topological space itself as a topological space, this will be with respect to the subspace topology (unless otherwise is explicitly mentioned).

### 1.4 The (Finite) Product Topology

Let $X$ and $Y$ be two topological spaces. It is easy to check that the products of the open subsets of these spaces, that is,

$$\mathcal{B}_{X \times Y} := \{ O \times U : O \in \mathcal{O}_X \text{ and } U \in \mathcal{O}_Y \},$$

is basis for a topology on $X \times Y$.

**Definition.** Let $X$ and $Y$ be two topological spaces. The **product topology** on $X \times Y$ is the topology generated by $\mathcal{B}_{X \times Y}$.

**Note.** If $\mathcal{B}_X$ is a basis for the topology of $X$, and $\mathcal{B}_Y$ for that of $Y$, then $\mathcal{B}_X \times \mathcal{B}_Y$ is a basis for the product topology on $X \times Y$.  

8
Needless to say, this definition may be extended, by induction, to the case of any (finite) number of topological spaces.

**Example 1.4.1.** The standard topology on $\mathbb{R}^2$ is the product topology on $\mathbb{R} \times \mathbb{R}$. With respect to this topology, a subset $O$ of $\mathbb{R}^2$ is open iff for every $x$ in $O$ there is a real number $\varepsilon > 0$ such that

$$(x_1 - \varepsilon, x_1 + \varepsilon) \times (x_2 - \varepsilon, x_2 + \varepsilon) \subseteq O.$$ 

Similarly, the standard topology on $\mathbb{R}^n$ is defined as the product topology on the $n$-fold product of $\mathbb{R}$. This is identical to the way we defined this topology at the end of Example 1.2.6.

It is easy to check that the product topology inherits the properties of being Hausdorff, first-countable and second-countable from its mother topologies. That is, $X \times Y$ is Hausdorff if $X$ and $Y$ are Hausdorff topological spaces, and it has a countable basis (at each of its points), if so do $X$ and $Y$. (Check!)

**Example 1.4.2.** $\mathbb{R}^n$ (relative to its standard topology) is a Hausdorff space with countable basis. This follows from the previous remark and the fact that $\mathbb{R}$ is a second-countable Hausdorff space.

From now on, and throughout this text, when we refer to the product of two (or more) topological spaces itself as a topological space, this will be with respect to the product topology (unless otherwise is explicitly mentioned). In Section 4 of Chapter 4 we extend this topology to the context of the product of an arbitrary collection of topological spaces.

### 1.5 Compactness

Compactness is a property for subsets of a topological space that exhibits a number of characteristics of finiteness, and plays an essential role in various branches of mathematical analysis.

**Definition.** Let $X$ be a topological space and $S$ subset of $X$. A collection $\mathcal{O}$ of subsets of $X$ is said to be a cover of $S$ if $S \subseteq \bigcup \mathcal{O}$. If, in addition, every member of $\mathcal{O}$ is open in $X$ — that is, $\mathcal{O} \subseteq \mathcal{O}_X$ — we say that $\mathcal{O}$ is an open cover of $S$. Finally, if every open
cover $O$ of $S$ has a finite subset $U$ that covers $S$, we say that $S$ is compact. (Such a set $U$ is said to be a finite cover of $S$ extracted from $O$.)

**Definition.** Let $X$ be a topological space and $S$ subset of $X$. If the closure of $S$ (in $X$) is compact, we say that $S$ is relatively compact in $X$.

**Note.** $S$ is a compact subset of a topological space $X$ iff it is compact with respect to the subspace topology. (Thus, there is no ambiguity in the phrase “compact subspace.”) The proof is easy, and hence omitted.

**Note.** Let $Y$ be a subspace of a topological space $X$. If $S$ is a subset of $Y$ which is compact in $X$, then $S$ is compact in $Y$. Again, the proof is easy.

**Example 1.5.1.** Any finite subset of a topological space is compact.

**Example 1.5.2.** A nonempty compact subset $S$ of $\mathbb{R}$ contains its supremum (and infimum). For, if $\sup S$ belongs to $\mathbb{R}\setminus S$, then letting $O_m := (-\infty, \sup S - \frac{1}{m})$, we see that $O := \{O_1, O_2, \ldots\}$ is an open cover of $S$, but there is no finite cover of $S$ that can be extracted from $O$. Similarly, if $\sup S = \infty$, examining the open cover $O := \{(-\infty, m) : m = 1, 2, \ldots\}$ of $S$, we see that $S$ is not compact.

For instance, neither $[0, 1)$ nor $\mathbb{R}$ are compact. By contrast, every interval of the form $[a, b]$ is compact if in particular, $[0, 1)$, but not $\mathbb{R}$, is relatively compact in $\mathbb{R}$ – but this is not a trivial fact.

**Example 1.5.3.** We say that a subset $S$ of $\mathbb{R}^n$ is bounded if

$$\sup\{\|x - y\| : x, y \in S\} < \infty,$$

where $\|\cdot\|$ is the Euclidean norm. A famous result of real analysis states that $S$ is compact iff it is closed and bounded. This is the Heine-Borel Theorem. (Thus, every bounded set in $\mathbb{R}^n$ is relatively compact in $\mathbb{R}^n$.) The “if” part of this theorem is a special property that need not hold in topological spaces other than $\mathbb{R}^n$, and is far from being a triviality. We shall prove this theorem later as a consequence of a deeper result in metric space theory.

**Example 1.5.4.** Let $X$ be a discrete space with infinitely many elements. Then $X$ is closed and bounded, but it is not compact.
**Example 1.5.5.** Consider $(0, 1)$ as a topological space space under the auspices of the subspace topology. This space is closed (in itself) and bounded, but it is not compact.

**Example 1.5.6.** Let $(X, \succ)$ be a poset with $|X| \geq 2$. Then, $X$ is compact with respect to the $\succ$-interval topology iff $(X, \succ)$ is a complete lattice. This is the Birkhoff-Frink Compactness Theorem. We outline a proof for this result in Section 4 of Chapter 4.

We now go through some basic properties of compact sets. First, we note that every closed set contained within a compact set is compact itself. More precisely:

**Proposition 1.5.1.** Every closed subset $S$ of a compact topological space $X$ is compact.

**Proof.** If $\mathcal{O}$ is an open cover of $S$ (with sets open in $X$), then $\mathcal{O} \cup \{X \setminus S\}$ is an open cover of $X$. Since $X$ is compact, we can extract a finite cover of $X$ from $\mathcal{O} \cup \{X \setminus S\}$, say $\mathcal{U}$. Then, $\mathcal{U} \setminus \{X \setminus S\}$ is a finite cover of $S$ extracted from $\mathcal{O}$.

Conversely, when the underlying topology is Hausdorff, every compact set is closed. To prove this, we need the following useful observation:

**Lemma 1.5.2.** Let $X$ be a Hausdorff topological space, $S$ a compact subset of $X$, and $x$ an element of $X \setminus S$. Then, there exist two disjoint open subsets $O$ and $U$ of $X$ such that $x \in O$ and $S \subseteq U$.

**Proof.** For each $y$ in $S$, we choose two disjoint open neighborhoods $O_y$ and $U_y$ of $x$ and $y$, respectively. (As $X$ is Hausdorff, we can do this.) Obviously, $\{U_y : y \in S\}$ is an open cover of $S$. As $S$ is compact, we can extract a finite cover of $S$ from this open cover. That is, there exist finitely many elements $y_1, \ldots, y_k$ of $S$ such that

$$S \subseteq U_{y_1} \cup \cdots \cup U_{y_k}.$$  

Setting $U := U_{y_1} \cup \cdots \cup U_{y_k}$ and $O := O_{y_1} \cap \cdots \cap O_{y_k}$, we are done.

**Proposition 1.5.3.** Every compact subset $S$ of a Hausdorff topological space $X$ is closed.

**Proof.** For any $x$ in $X \setminus S$, we use Lemma 1.5.2 to find an open neighborhood of $x$ which is contained in $X \setminus S$. This means that $\operatorname{int}(X \setminus S) = X \setminus S$, and hence $X \setminus S$ is open in $X$.  

11
Note. A compact subset of a non-Hausdorff topological space need not be closed. For instance, every subset of an indiscrete space is compact while this space has only two closed sets.

Another important observation about compactness is the fact that this property is inherited through the product topology. That is:

**Proposition 1.5.4.** If $X$ and $Y$ are two compact topological spaces, then $X \times Y$ is a compact topological space.

Leaving the proof of this result as an exercise, we turn to an alternative characterization of the compactness property. In optimization theory, for instance, compactness is often utilized through this formulation.

**Definition.** We say that a nonempty collection $\mathcal{X}$ of sets has the **finite intersection property** if $\bigcap S \neq \emptyset$ for every nonempty finite subset $S$ of $\mathcal{X}$.

**Proposition 1.5.5.** Let $X$ be a topological space. Then, $X$ is compact if, and only if, for every nonempty collection $\mathcal{C}$ of closed subsets of $X$ with the finite intersection property, we have $\bigcap \mathcal{C} \neq \emptyset$.

Note. An easy corollary of this result is that $S_1 \cap S_2 \cap \cdots \neq \emptyset$ for any sequence $(S_m)$ of compact subsets of a Hausdorff topological space with $S_1 \supseteq S_2 \supseteq \cdots$. (Such a sequence is called **nested**.)

This result is easily proved by focusing on the contrapositive of its statement and working with the complements of the involved closed sets. We leave working out the details as an exercise.

We conclude this section with a local formulation of the notion of compactness. As we shall have limited use for this concept in this text, however, our discussion will be brief.

**Definition.** We say that a topological space $X$ is **locally compact** if for every $x$ in $X$, and every open neighborhood $O$ of $x$, there exist an open neighborhood $U$ of $x$ and a compact subset $S$ of $X$ such that $x \in U \subseteq S \subseteq O$. 

12
For Hausdorff topological spaces, local compactness is easier to formulate:

**Proposition 1.5.6.** Let $X$ be a Hausdorff topological space. Then, $X$ is locally compact if, and only if, for every $x$ in $X$, and every open neighborhood $O$ of $x$, there exists a compact subset $S$ of $X$ such that $x \in S \subseteq O$.

A proof of this fact can be obtained by using Lemma 1.5.2. Leaving the details as an exercise, we point to a few examples that illustrate that local compactness is significantly weaker than compactness.

**Example 1.5.7.** $\mathbb{R}$ is locally compact (but $\mathbb{Q}$ is not).

**Example 1.5.8.** $\mathbb{R}^n$ is locally compact.

Local compactness property is, in general, not hereditary. There are, however, two special cases in which no difficulty arises in this regard.

**Example 1.5.9.** If $Y$ is a subspace of a locally compact topological space $X$, then $Y$ is locally compact itself, *provided that $Y$ is closed in $X$*. We leave the (easy) proof as an exercise.

**Example 1.5.8.** If $Y$ is a subspace of a locally compact Hausdorff topological space $X$, then $Y$ is locally compact itself, *provided that $Y$ is open in $X$*. To see this, take any $x$ in $Y$ and an open subset $O$ of $Y$ with $x \in O$. As $Y$ is Hausdorff, we will be done if we can find a compact subset of $Y$ that contains $x$ and that is contained in $O$. But, as $Y$ is open, $O$ is open in $X$ as well. Then, by Proposition 1.5.6, there exists a compact subset $S$ of $X$ such that $x \in S \subseteq O$. As $S$ is compact in $Y$ as well, we are done.

### 1.6 Connectedness

Intuitively speaking, a connected subset of a metric space is one that cannot be partitioned into two (or more) separate pieces, it is rather in one whole piece. In $\mathbb{R}$, for instance, we like to think of the interval $(0, 1)$ as connected and $[0, 1] \cup [2, 3)$ as disconnected. The definition below formalizes this simple geometric intuition.
**Definition.** We say that a topological space $X$ is **connected** if there do not exist two nonempty and disjoint open subsets $O$ and $U$ of $X$ such that $O \cup U = X$. In turn, a subset $S$ of $X$ is said to be **connected in** $X$ (or a **connected subset of** $X$) if $S$ is a connected subspace of $X$. (So, $S \subseteq X$ is connected in $X$ iff it cannot be written as a disjoint union of two nonempty sets that are open in $S$.)

The following simple result provides an interesting characterization of connected topological spaces.

**Proposition 1.6.1.** Let $X$ be a topological space. Then, $X$ is connected if, and only if, $\emptyset$ and $X$ are the only subsets of $X$ that are both open and closed.

**Proof.** If $S \notin \{\emptyset, X\}$ is a subset of $X$ that is both open and closed, then $X$ cannot be connected since $X = S \cup (X \setminus S)$. Conversely, assume that $X$ is not connected. In this case we can find nonempty and disjoint open subsets $O$ and $U$ of $X$ such that $O \cup U = X$. But then $U = X \setminus O$ so that $O$ must be both open and closed. Since $O \notin \{\emptyset, X\}$, this proves our assertion.

**Example 1.6.1.** A discrete space is not connected unless it contains only one element, because any subset of a discrete space is both open and closed.

**Example 1.6.2.** $\mathbb{Q}$ is not connected in $\mathbb{R}$, because $\mathbb{Q} = (-\infty, \pi) \cup (\pi, \infty)$ and $\pi$ is not a rational number. As $\mathbb{R}$ is connected, this shows that the connectedness property is not inherited by subspaces of a topological space.

**Example 1.6.3.** A nonempty subset of $\mathbb{R}$ is connected iff it is an interval. We leave the proof of this fact as an exercise here.

The connectedness property is inherited under the product topology. That is:

**Proposition 1.6.2.** If $X$ and $Y$ are two connected topological spaces, then $X \times Y$ is a connected topological space.

We leave the (easy) proof of this result as an exercise.
1.7 Continuity

The following definition extends the familiar notion of continuity of a function from a Euclidean space into another to the context of functions that map a topological space into another.

**Definition.** Let $X$ and $Y$ be two topological spaces, and $f : X \to Y$ a function. For any $x$ in $X$, we say that $f$ is **continuous at** $x$ if for every open neighborhood $O$ of $f(x)$, there exists an open neighborhood $U$ of $x$ such that $f(U) \subseteq O$. We say that $f$ is **continuous** if it is continuous at every $x$ in $X$.

As we shall see later, this definition reduces to the usual formulation of continuity in the context of metric spaces. For now, we offer some elementary examples.

**Example 1.7.1.** For every topological space $X$, and any $x_*$ in $X$, the (constant) self-map $f$ on $X$, defined by $f(x) := x_*$, is, obviously, continuous. Also obvious is the fact that the identity function on $X$ — we denote this function by $\text{id}_X$ — that map each $x$ in $X$ to itself is a continuous self-map on $X$.

**Example 1.7.2.** By Example 1.2.1, a real function $f$ on $\mathbb{R}$ is continuous iff for every real numbers $x$ and $\varepsilon$ with $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$f(y) \in (f(x) - \varepsilon, f(x) + \varepsilon) \quad \text{for every } y \in (x - \delta, x + \delta),$$

or equivalently,

$$|f(y) - f(x)| < \varepsilon \quad \text{whenever} \quad |y - x| < \delta,$$

which is the familiar formulation of continuity in one-variable calculus.

**Example 1.7.3.** A real function on $\mathbb{R}^n$ is continuous iff for every $n$-vector $x$ and real number $\varepsilon > 0$, there exists a real number $\delta > 0$ such that

$$f(y) \in (f(x) - \varepsilon, f(x) + \varepsilon) \quad \text{whenever } y_i \in (x_i - \delta, x_i + \delta), \ i = 1, \ldots, n$$

A moment’s reflection shows that this is the same thing as saying that for every $n$-vector $x$ and $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$|f(y) - f(x)| < \varepsilon \quad \text{whenever} \quad \sum_{i=1}^{n} |y_i - x_i| < \delta.$$
For instance, \( f : \mathbb{R}^n \to \mathbb{R} \), defined by

\[
  f(x) := \sum_{i=1}^{n} x_i,
\]

is continuous, because for any \( n \)-vector \( x \) and \( \varepsilon > 0 \), setting \( \delta := \varepsilon / n \), and using the triangle inequality,

\[
  |f(y) - f(x)| \leq \sum_{i=1}^{n} |y_i - x_i| < \varepsilon
\]

whenever \( |y_1 - x_1| + \cdots + |y_n - x_n| < \delta \).

Before presenting some more examples, we would like to obtain another way of looking at the notion of continuity. This requires us to consider the following auxiliary definition.

**Definition.** Let \( X \) and \( Y \) be two nonempty sets and \( f : X \to Y \) a function. For any subset \( S \) of \( Y \), the **inverse image** of \( S \) under \( f \) is

\[
  f^{-1}(S) := \{ x \in X : f(x) \in S \}.
\]

**Note.** If \( f : X \to Y \) is an invertible function, then, obviously, \( f^{-1}(f(S)) = S \) for every subset \( S \) of \( Y \). In general, however, all we have is \( f^{-1}(f(S)) \supseteq S \).

It is a useful fact that inverse images preserve unions, that is,

\[
  f^{-1}(\bigcup S) = \bigcup\{ f^{-1}(S) : S \in S \}
\]

for any collection \( S \) of subsets of \( Y \). (Exactly the same is true for intersections as well.)

Here is how the inverse images of a function can be used to characterize the continuity of that function.

**Proposition 1.7.1.** Let \( X \) and \( Y \) be two topological spaces and \( f : X \to Y \) a function. Then, the following are equivalent:

(a) \( f \) is continuous;

(b) \( f^{-1}(O) \) is open in \( X \) for every open subset \( O \) of \( Y \);

(c) \( f^{-1}(C) \) is closed in \( X \) for every closed subset \( C \) of \( Y \).

**Proof.** Assume that \( f \) is continuous, and take any open subset \( O \) of \( Y \). We wish to show that \( f^{-1}(O) \) is open in \( X \). To this end, note that, for any \( x \) in \( f^{-1}(O) \), the set \( O \) is an
open neighborhood of $f(x)$, so by continuity of $f$, there is an open neighborhood $U_x$ of $x$ such that $f(U_x) \subseteq O$, that is, $U_x \subseteq f^{-1}(O)$. It follows that

$$f^{-1}(O) = \bigcup \{U_x : x \in f^{-1}(O)\}.$$  

As the union of open sets is open, we find that $f^{-1}(O)$ is open. Conversely, to see that (b) implies (a), take any $x$ in $X$, and let $O$ be an open neighborhood of $f(x)$. Then, by (b), $U := f^{-1}(O)$ is an open neighborhood of $x$ such that $f(U) \subseteq O$.

To see that (b) implies (c), let $C$ be a closed subset of $Y$. Then, $Y \setminus C$ is open in $Y$, and hence, by (b), $f^{-1}(Y \setminus C)$ is open in $X$. Since $X \setminus f^{-1}(C) = f^{-1}(Y \setminus C)$, this means that $f^{-1}(C)$ is closed in $X$. That (c) implies (b) is proved analogously.

Note. If $Y$ has a basis, say $B_Y$, then $f$ is continuous if and only if the inverse image of each element of $B_Y$ under $f$ is open in $X$. To see this, suppose $f^{-1}(B) \in \mathcal{O}_X$ for every $B \in B_Y$. Then, if $O$ is an open subset of $Y$, we have $O = \bigcup B$ for some $B \subseteq B_Y$, and hence

$$f^{-1}(O) = f^{-1}(\bigcup B) = \bigcup \{f^{-1}(B) : B \in B\},$$

which means that $f^{-1}(O)$ is open in $X$. The “only if” part of our assertion is immediate from Proposition 1.7.1.

The two main preservation properties for the notion of continuity are that the restriction of a continuous function is continuous, and that the composition of continuous functions is, again, continuous:

**Proposition 1.7.2.** Let $X$, $Y$, and $Z$ be topological spaces and $f : X \to Y$ and $g : Y \to Z$ two continuous functions. Then, $f|_S$ is continuous for any subspace $S$ of $X$, and $g \circ f$ is continuous.

This is proved easily by using Proposition 1.7.1; we leave the details an exercise. Instead, we use this result to deduce the continuity of certain types of functions.

**Example 1.7.4.** Let $X$ and $Y$ be two topological spaces. The projection map $\rho_X$ from $X \times Y$ into $X$ is defined by

$$\rho_X(x, y) := x,$$

while $\rho_Y$ is defined dually. An immediate application of Proposition 1.7.1 shows that both of these maps are continuous.
**Example 1.7.5.** Let $X$, $Y$, and $Z$ be topological spaces and $f : Z \to X \times Y$ a function. Suppose for the functions $f_1 : Z \to X$ and $f_2 : Z \to Y$, we have

$$f(z) = (f_1(z), f_2(z))$$

for every $z \in Z$.

Then, $f$ is continuous if and only if both $f_1$ and $f_2$ are continuous.

To see this, observe that $f_1 = \rho_X \circ f$ and $f_2 = \rho_Y \circ f$, so the “only if” part of our assertion follows from Example 1.7.4 and Proposition 1.7.2. To prove the converse, it is enough to show that the inverse image of every basis element of the product topology on $X \times Y$ is open in $Z$. That is, we wish to prove that, for any open sets $O$ and $U$ in $X$ and $Y$, respectively, $f^{-1}(O \times U)$ is open in $Z$. As

$$f^{-1}(O \times U) = f_1^{-1}(O) \cap f_2^{-1}(U),$$

this is an immediate consequence of Proposition 1.7.1 and the continuity of $f_1$ and $f_2$.

**Example 1.7.6.** Fix any positive integer $n$ and any topological space $X$. Let $f_i : X \to \mathbb{R}$ be a continuous map, $i = 1, \ldots, n$, and pick any continuous function $g : \mathbb{R}^n \to \mathbb{R}$. We wish to show that the map $F : X \to \mathbb{R}$, defined by

$$F(x) := g(f_1(x), \ldots, f_n(x)),$$

is continuous. To this end, define $f : X \to \mathbb{R}^n$ by

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

and observe that $F = g \circ f$. Applying the results of Proposition 1.7.2 and (the inductive generalization of) Example 1.7.5, we may conclude that $F$ is continuous.

**Example 1.7.7.** Fix any positive integer $n$ and any topological space $X$. Then, given any continuous real maps $f_1, \ldots, f_n$ on $X$,

$$x \mapsto \sum_{i=1}^{n} f_i(x)$$

is a continuous real function on $X$. This is an obvious consequence of the observations we made in Examples 1.7.3 and 1.7.6. One can similarly show that the real maps

$$\sum_{i=1}^{n} |f_i|, \quad \prod_{i=1}^{n} \varphi_i \quad \text{and} \quad \max\{f_1, \ldots, f_n\}$$

are continuous on $X$ as well.
In the case of first-countable topological spaces, there is a very useful way of characterizing the continuity of a function in terms of convergent sequences.

**Proposition 1.7.3.** Let $X$ and $Y$ be two first-countable topological spaces, and $f : X \to Y$ a function. Then, $f$ is continuous if, and only if, for every sequence $(x_m)$ in $X$ and every element $x$ of $X$,

$$x_m \to x \quad \text{implies} \quad f(x_m) \to f(x).$$

**Proof.** Assume that $f$ is continuous, and take any elements $x, x_1, x_2, \ldots$ in $X$ such that $x_m \to x$. Let $O$ be an arbitrary open neighborhood of $f(x)$. Then, by Proposition 1.7.1, $f^{-1}(O)$ is an open neighborhood of $x$, and hence $x_m \to x$ implies that $x_m \in f^{-1}(O)$, and thus $f(x_m) \in O$, for all but finitely many $m$. Conclusion: $f(x_m) \to f(x)$.

Conversely, assume that (2) holds for every $x, x_1, x_2, \ldots$ in $X$. Take an arbitrary closed subset $C$ of $Y$. In view of Proposition 1.7.1, we can conclude that $f$ is continuous if we can show that $f^{-1}(C)$ is closed in $X$. By Proposition 1.2.1, therefore, it is enough to show that the limit of any convergent sequence in $f^{-1}(C)$ belongs to $f^{-1}(C)$. To this end, take any sequence $(x_m)$ in $f^{-1}(C)$, and assume that $x_m \to x$ for some $x$ in $X$. Then, $f(x_m) \in C$ for each $m$, and we have $f(x_m) \to f(x)$ by hypothesis. As $S$ is closed, we then have $f(x) \in C$ by Proposition 1.2.1, and hence $x \in f^{-1}(C)$, as we sought.

The images of continuous functions inherit some (but not all) important topological properties from their preimages. In particular, this is the case for connectedness and compactness (but not, for instance, openness or closedness). Let us first deal with the case of connectedness.

**Proposition 1.7.4.** Let $X$ and $Y$ be two topological spaces, and $f : X \to Y$ a continuous function. If $X$ is connected, then $f(X)$ is a connected subset of $Y$.

**Proof.** If $f(X)$ is not connected in $Y$, then we can find two nonempty and disjoint open subsets $O$ and $U$ of $f(X)$ such that $O \cup U = f(X)$. But then, by Proposition 1.7.1, $f^{-1}(O)$ and $f^{-1}(U)$ are nonempty open subsets of $X$. Moreover, these sets are disjoint, and we have

$$f^{-1}(O) \cup f^{-1}(U) = f^{-1}(O \cup U) = X,$$

that is, $X$ is not connected.

The analogous result holds for the compactness property as well.
**Proposition 1.7.5.** Let $X$ and $Y$ be two topological spaces, and $f : X \to Y$ a continuous function. If $X$ is compact, then $f(X)$ is a compact subset of $Y$.

We omit the proof of this fact at present, as we shall shortly prove something more general than this. However, let us note that Proposition 1.7.5 has a number of important consequences. First of all, it gives us a useful sufficient condition for the inverse of an invertible continuous function to be continuous.

**The Homeomorphism Theorem.** Let $X$ be a compact Hausdorff topological space and $f : X \to Y$ a continuous and invertible function. Then, $f^{-1}$ is a continuous function.

**Proof.** Take any closed subset $C$ of $X$. As $X$ is compact, $C$ is a compact subspace of $X$ (Proposition 1.5.1) and hence $f(C)$ is a compact, and hence closed, subset of $Y$ by Propositions 1.7.5 and 1.5.3. That is: $f(C) \subseteq C_Y$ for every $C \subseteq C_X$. By Proposition 1.7.1, this means that $f^{-1}$ is continuous.

**Remark.** A map $f$ from a topological space $X$ onto another, say, $Y$, is said to be a **homeomorphism** if it is continuous, bijective and has a continuous inverse. Clearly, if such a map exists, in which case we say that $X$ and $Y$ are **homeomorphic**, then the open sets of $Y$ can be obtained from those of $X$ by mere relabelling of the elements of $X$, and vice versa. It follows that the topological structure of $X$ and $Y$ are identical – for instance, a subset $S$ of $X$ is closed/connected/compact iff $f(S)$ is closed/connected/compact in $Y$, $X$ is first/second countable iff $Y$ is first/second countable, and so on – provided that $X$ and $Y$ are homeomorphic. From this point of view, the Homeomorphism Theorem tells us that the continuous image of a compact topological space $X$ under an injection is topologically equivalent to $X$.

As a final application, we deduce the following fundamental theorem of optimization theory from Proposition 1.7.5.

**The Weierstrass Extremum Value Theorem.** Let $X$ be a compact topological space and $f$ a continuous real function on $X$. Then, there exists an $x$ in $X$ such that

$$f(x) = \sup f(X).$$

**Proof.** By Proposition 1.7.5, $f(X)$ is a compact subset of $\mathbb{R}$, and hence it contains its supremum (Example 1.5.2).
Corollary 1.7.6. Every continuous real function on a compact topological space is bounded.

Proof. Let $f$ be such a function and apply the Weierstrass Maximum Value Theorem to $f$ and $-f$.

Let us now turn to assimilating the notion of continuity for correspondences. Recall that a correspondence $\Gamma$ from a nonempty set $X$ into another nonempty set $Y$ is a map from $X$ into $2^Y \setminus \{\emptyset\}$. Thus, for each $x \in X$, the image $\Gamma(x)$ of $x$ under $\Gamma$ is a nonempty subset of $Y$. We write

$$\Gamma : X \rightrightarrows Y$$

to denote that $\Gamma$ is a correspondence from $X$ into $Y$. For any nonempty subset $S$ of $X$, we let

$$\Gamma(S) := \bigcup \{\Gamma(x) : x \in S\},$$

and by convention, set $\Gamma(\emptyset) := \emptyset$.

There are several ways of devising a notion of continuity for correspondences. For brevity, we shall consider here only the most commonly used notion of continuity for such set-valued maps, namely, upper hemicontinuity. (This is the only continuity notion used for correspondences in the body of the text.)

Definition. Let $X$ and $Y$ be two topological spaces, and $\Gamma : X \rightrightarrows Y$ a correspondence. For any $x$ in $X$, we say that $\Gamma$ is \textbf{upper hemicontinuous at } $x$ if for every open subset $O$ of $Y$ with $\Gamma(x) \subseteq O$, there exists an open neighborhood $U$ of $x$ such that $\Gamma(U) \subseteq O$. We say that $\Gamma$ is \textbf{upper hemicontinuous} if it is upper hemicontinuous at each $x \in X$. Finally, $\Gamma$ is said to be \textbf{compact-valued} if $\Gamma(x)$ is compact for each $x \in X$.

Clearly, this definition mimics that of continuity of a function, and it reduces to the definition of the latter when $\Gamma$ is single-valued. That is to say, every upper hemicontinuous single-valued correspondence “is” a continuous function, and conversely.

However, while a continuous function maps compact sets to compact sets, this is not the case with upper hemicontinuous correspondences. For instance, while $\Gamma : [0,1] \rightrightarrows \mathbb{R}_+$ defined by $\Gamma(x) := \mathbb{R}_+$ is obviously upper hemicontinuous, it doesn’t map even a singleton set to a compact set. This is not a big difficulty in most applications, because the correspondences that one deals with usually have some additional structure that might circumvent this problem. In particular, as we show next, every compact-valued upper hemicontinuous correspondence maps compact sets to compact sets. (This observation generalizes Proposition 1.7.5.)
Proposition 1.7.7. Let $X$ and $Y$ be two topological spaces, and $\Gamma : X \rightrightarrows Y$ a compact-valued upper hemicontinuous correspondence. If $X$ is compact, then $\Gamma(X)$ is a compact subset of $Y$.

Proof. Assume that $X$ is compact, and let $\mathcal{O}$ be an open cover of $\Gamma(X)$. We wish to find a finite subset of $\mathcal{O}$ that would also cover $\Gamma(X)$. Note that, for each $x \in X$, $\mathcal{O}$ is also an open cover of $\Gamma(x)$, so, since $\Gamma(x)$ is compact, there exist finitely many open sets $O_1(x), \ldots, O_{m_x}(x)$ in $\mathcal{O}$ such that

$$\Gamma(x) \subseteq \bigcup\{O_i(x) : i = 1, \ldots, m_x\}.$$ 

For any $x \in X$, let us denote the open set on the right-hand side of this statement by $O(x)$, and define

$$\Gamma^{-1}(O(x)) := \{\omega \in X : \Gamma(\omega) \subseteq O(x)\}.$$ 

It is easy to check that $\Gamma^{-1}(O(x))$ is an open subset of $X$ (for each $x \in X$) because $\Gamma$ is upper hemicontinuous. Moreover,

$$\Gamma(X) \subseteq \bigcup\{O(x) : x \in X\},$$

which ensures that $\{\Gamma^{-1}(O(x)) : x \in X\}$ is an open cover of $X$. By compactness of $X$, therefore, there exist finitely many points $x^1, \ldots, x^m \in X$ such that $\{\Gamma^{-1}(O(x^1)), \ldots, \Gamma^{-1}(O(x^m))\}$ covers $X$. But then $\{O(x^1), \ldots, O(x^m)\}$ must cover $\Gamma(X)$. Therefore, $\{O_j(x^i) : j = 1, \ldots, m_x, i = 1, \ldots, m\}$ is a finite subset of $\mathcal{O}$ that covers $\Gamma(X)$. Conclusion: $\Gamma(X)$ is a compact subset of $Y$.

1.8 Semicontinuity

In the case of real functions, there is a weakening of the continuity property that is particularly useful for optimization theory.

Definition. Let $X$ be a topological space, and $f : X \to \mathbb{R}$ a function. For any $x$ in $X$, we say that $f$ is upper semicontinuous at $x$ if for every $\varepsilon > 0$, there exists an open neighborhood $U$ of $x$ such that

$$f(y) \leq f(x) + \delta \quad \text{for every } y \in U.$$ 

We say that $f$ is upper semicontinuous if it is upper semicontinuous at every $x$ in $X$. Finally, we say that $f$ is lower semicontinuous if $-f$ is upper semicontinuous.
Note. It follows easily from this definition that a continuous transformation (such as sum or product) of two upper semicontinuous functions is upper semicontinuous. Similarly, the pointwise maximum of finitely many upper semicontinuous functions is upper semicontinuous, while the pointwise infimum of any nonempty collection of upper semicontinuous functions is upper semicontinuous, provided that it is well-defined. Also, the composition of two upper semicontinuous functions is, again, upper semicontinuous.

Let $f$ be a real map on a topological space $X$, and assume that $f$ is upper semicontinuous at some element $x$ of $X$. Then, for any $\varepsilon > 0$, there exists an open neighborhood $U_\varepsilon$ of $x$ such that
\[
\sup\{f(y) : y \in U_\varepsilon\} \leq f(x) + \varepsilon.
\]
But then, for every $\varepsilon > 0$,
\[
\inf_{U \in \mathcal{U}(x)} \sup\{f(y) : y \in U\} \leq f(x) + \varepsilon,
\]
where $\mathcal{U}(x)$ is the collection of all open neighborhoods of $x$. It follows that
\[
\inf_{U \in \mathcal{U}(x)} \sup\{f(y) : y \in U\} \leq f(x).
\]
As, obviously, $x \in U$ for each $U \in \mathcal{U}(x)$, the converse inequality holds as well, and we have
\[
\inf_{U \in \mathcal{U}(x)} \sup\{f(y) : y \in U\} = f(x). \tag{3}
\]
As we can reverse this reasoning as well, we may conclude that $f$ is upper semicontinuous at $x$ iff (3) holds.

Definition. Let $X$ be a topological space, and $f : X \to \mathbb{R}$ a function. The limsup of $f$ is defined as the map
\[
x \mapsto \inf_{U \in \mathcal{U}(x)} \sup\{f(y) : y \in U\}
\]
on $X$. (This map is $\mathbb{R}$-valued.)

Consequently, what we have shown above can be restated as:

**Proposition 1.8.1.** Let $X$ be a topological space, and $f : X \to \mathbb{R}$ a function. Then, $f$ is upper semicontinuous (at $x$) iff it equals to its limsup (at $x$).

It is easy to check that the limsup of the limsup of a real function is equal to itself. Hence:
Corollary 1.8.2. Let $X$ be a topological space, and $f : X \to \mathbb{R}$ a function. Then, the limsup of $f$ is upper semicontinuous.

The following result, whose proof we leave as an exercise, catalogues some other ways of formulating the notion of semicontinuity.

Proposition 1.8.3. Let $X$ be a topological space, and $f : X \to \mathbb{R}$ a function. Then, the following are equivalent:

(a) $f$ is upper semicontinuous;
(b) $\{\omega \in X : f(\omega) < \alpha\}$ is open in $X$ for every real number $\alpha$;
(c) $\{\omega \in X : f(\omega) \geq \alpha\}$ is closed in $X$ for every real number $\alpha$.

Example 1.8.1. A real function is continuous iff it is both upper and lower semicontinuous.

Example 1.8.2. Let $X$ be a topological space and $C$ a closed subset of $X$. Then, the indicator function of $C$ on $X$, that is, the map $1_C : X \to \{0, 1\}$, defined by

$$1_C(x) := \begin{cases} 1, & \text{if } x \in C \\ 0, & \text{otherwise} \end{cases}$$

is upper semicontinuous.

Example 1.8.3. $1_{\mathbb{Q}}$ (on $\mathbb{R}$) is upper semicontinuous at each rational number, and lower semicontinuous at each irrational number. This function is continuous at no point in $\mathbb{R}$.

Finally, we show that continuity can be relaxed to upper semicontinuity in the Weierstrass Maximum Value Theorem.

The Baire Maximum Value Theorem. Let $X$ be a compact topological space and $f$ an upper semicontinuous real function on $X$. Then,

$$\{x \in X : f(x) = \sup f(X)\}$$

is a nonempty compact subset of $X$.

Proof. Let $\Lambda_m := (-\infty, m)$ for each positive integer $m$. Then, by Proposition 1.8.3, $\{f^{-1}(\Lambda_m) : m = 1, 2, ...\}$ is an open cover of $X$. Since $X$ is compact, we can extract a
finite cover of $X$ from this cover. It follows that there is a positive integer $M$ such that $X \subseteq f^{-1}(\Lambda_M)$, and this means that $s := \sup f(X) < \infty$. But then, by definition of $s$ and Proposition 1.8.3,

$$\mathcal{A} := \{f^{-1}([s - \frac{1}{m}, \infty)) : m = 1, 2, \ldots\}$$

is a collection of closed subsets of $X$ that has the finite intersection property. Thus, by Proposition 1.5.5, there exists an $x$ in $\bigcap \mathcal{A}$. Clearly, $f(x) = s$, which proves that $f^{-1}(s)$ is a nonempty subset of $X$. Besides, $f^{-1}(s) = \{x \in X : f(x) \geq s\}$ is a closed, and hence compact, subset of $X$ by Propositions 1.8.3 and Proposition 1.5.1.

### 1.9 The Quotient Topology

Our final task in this section is to introduce a method of topologizing a given partition of a topological space. (This subsection assumes familiarity with Section 2 of Chapter 1.)

Let $X$ be a topological space, and $\sim$ an equivalence relation on $X$. The **quotient map** on $X$ induced by $\sim$ is the map

$$x \mapsto [x]_{\sim},$$

which we shall denote by $\rho_{\sim}$. We would like to topologize $X_{/\sim}$ such that this map is continuous. Of course, the more open sets this topology has, the harder it becomes for $\rho_{\sim}$ to be continuous, and more stringent this requirement becomes. Fortunately, there is a largest (finest) topology on $X_{/\sim}$ that meets this requirement.

**Definition.** Let $X$ be a topological space, and $\sim$ an equivalence relation on $X$. The **quotient topology** on the quotient set $X_{/\sim}$ of $X$ relative to $\sim$ is the collection

$$\mathcal{O}_{X_{/\sim}} := \{O \in 2^{X_{/\sim}} : \rho_{\sim}^{-1}(O) \in \mathcal{O}_X\}.$$ 

As inverse images preserve intersections and unions, $\mathcal{O}_{X_{/\sim}}$ is indeed a topology on $\mathcal{O}_{X_{/\sim}}$. Furthermore, it is plain from its definition that the quotient topology on $X_{/\sim}$ is the largest (finest) topology on $X_{/\sim}$ that renders the quotient map $\rho_{\sim}$ continuous.

This is pretty much all one needs to know about the quotient topology for the purposes of this text. In passing, however, let us note that the Hausdorff property is not hereditary in the case of this topology. For instance, if $\sim$ is the equivalence relation on $\mathbb{R}$ defined by $x \sim y$ iff $x - y \in \mathbb{Q}$, then the quotient topology on $\mathbb{R}_{/\sim}$ is $\{\emptyset, \mathbb{Q}, \mathbb{R} \setminus \mathbb{Q}, \mathbb{R}\}$, so it is not Hausdorff. Nevertheless, we have the following result:
Proposition 1.9.1. Let $X$ be a topological space, and $\sim$ an equivalence relation on $X$. If $X/\sim$ is Hausdorff, then $\sim$ is a closed subset of $X \times X$. Conversely, if $\rho_\sim$ maps open sets to open sets, and $\sim$ is a closed subset of $X \times X$, then $X/\sim$ is Hausdorff.

Proof. Suppose $X/\sim$ is Hausdorff. We wish to show that $\Lambda := (X \times X) \setminus \sim$ is open in $X \times X$. To this end, take any $(x, y) \in \Lambda$. Then, $\rho_\sim(x)$ and $\rho_\sim(y)$ are distinct elements of $X/\sim$. As $X/\sim$ is Hausdorff, there are disjoint open neighborhoods $O$ and $U$ of $\rho_\sim(x)$ and $\rho_\sim(y)$, respectively. As $\rho_\sim$ is continuous, $\rho_\sim^{-1}(O)$ and $\rho_\sim^{-1}(U)$ are open in $X$, and we have $(a, b) \in \Lambda$ for any $a \in \rho_\sim^{-1}(O)$ and $b \in \rho_\sim^{-1}(U)$, because $O \cap U = \emptyset$. Therefore, $\rho_\sim^{-1}(O) \times \rho_\sim^{-1}(U)$ is an open neighborhood of $(x, y)$ that is contained in $\Lambda$. It follows that $\Lambda$ is open in $X \times X$.

Conversely, assume that $\rho_\sim$ maps open sets to open sets and $\Lambda$ is open in $X \times X$. Take any $x$ and $y$ of $X$ such that $[x]_\sim$ and $[y]_\sim$ are distinct elements of $X/\sim$. Then, $(x, y)$ belongs to $\Lambda$, so, because $\sim$ is a closed subset of $X \times X$, there exists a basis element of the product topology on $X \times X$ that contains $(x, y)$ and that is contained in $\Lambda$. That is, there exist open neighborhoods $O$ and $U$ of $x$ and $y$, respectively, such that $O \times U \subseteq \Lambda$, which means that $(a, b) \in \Lambda$ for any $a \in O$ and $b \in U$. But then, as $\rho_\sim$ maps open sets to open sets, $\rho_\sim(O)$ and $\rho_\sim(U)$ are disjoint open neighborhoods of $[x]_\sim$ and $[y]_\sim$, respectively.

2 Metric Spaces

The majority (but not all) of topological spaces one encounters in practice are metric spaces. While you are likely to be familiar with at least the rudiments of metric space theory, we shall nevertheless give a brief review here for the sake of completeness, and demonstrate how some of this theory is readily inherited from general topology.

2.1 Definitions and Examples

Definition. Let $X$ be a nonempty set. A nonnegative function $d : X \times X \to \mathbb{R}_+$ that satisfies the following properties is called a distance function (or a metric) on $X$: For any $x, y, z \in X$,

- $d(x, y) = 0$ iff $x = y$,
- (Symmetry) $d(x, y) = d(y, x)$,
- (Triangle Inequality) $d(x, y) \leq d(x, z) + d(z, y)$.
If $d$ is a distance function on $X$, we say that $(X, d)$ is a **metric space**.

Recall that we think of the distance between two points $x$ and $y$ on the real line as $|x - y|$. Thus the map $(x, y) \mapsto |x - y|$ serves as a function that tells us how much apart are any two elements of $\mathbb{R}$ from each other. Among others, this function is nonnegative and satisfies the three properties of the definition above. By way of abstraction, the notion of distance function is built *only* on these properties.

**Example 2.1.1.** Let $X$ be a nonempty set, and define $d : X \times X \to \mathbb{R}_+$ by

$$d(x, y) := \begin{cases} 1, & \text{if } x \neq y \\ 0, & \text{if } x = y \end{cases}.$$ 

Then, $(X, d)$ is a metric space. (It is said to be a **discrete metric space**.)

**Example 2.1.2.** Let $X$ be a linear space. A nonnegative function $\|\cdot\| : X \to \mathbb{R}_+$ that satisfies the following properties is called a **norm** on $X$: For any $x, y \in X$, and real number $\lambda$,

- $\|x\| = 0$ iff $x = 0$,
- (Positive Homogeneity) $\|\lambda x\| = |\lambda| \|x\|$,
- (Subadditivity) $\|x + y\| \leq \|x\| + \|y\|$.

If $\|\cdot\|$ is a norm on $X$, we say that $(X, \|\cdot\|)$ is a **normed linear space**.

Every norm $\|\cdot\|$ on $X$ induces a metric $d_{\|\cdot\|}$ on $X$ in the following manner:

$$d_{\|\cdot\|}(x, y) := \|x - y\|.$$ 

This is how most metric spaces arise in practice.

**Example 2.1.3.** For every real number $\rho$ in the interval $[1, \infty)$, the $\rho$-norm $\|\cdot\|_\rho$ on $\mathbb{R}^n$ is defined by

$$\|x\|_\rho := \left( \sum_{i=1}^n |x_i|^{\rho} \right)^{1/\rho}.$$ 

It is not trivial that this function satisfies subadditivity (when $\rho > 1$), which is a fact called the **Minkowski Inequality**. For brevity, we omit giving a proof for this, which would be provided in any text on real analysis.
In what follows, we shall write $\mathbb{R}^n_\rho$ for the normed linear space $(\mathbb{R}^n, \|\cdot\|_\rho)$, and view this space as a metric space whose distance function is $d_\rho := d_{\|\cdot\|_\rho}$. When we talk of $\mathbb{R}^n$ as a normed linear space what we have in mind is $\mathbb{R}^n_2$, and when we talk about it as a metric space what we have in mind is $(\mathbb{R}^n, d_2)$. ($d_2$ is called the Euclidean metric.)

**Example 2.1.4.** For any nonempty set $X$, we denote by $B(X)$ the linear space of all bounded real functions on $X$. We define the real function $\|\cdot\|_{\infty}$ on $B(X)$ by

$$\|f\|_{\infty} := \sup\{|f(x)| : x \in X\}.$$ 

It is easy to see that this function satisfies the first two properties of being a norm. It is also subadditive, because for any $f$ and $g$ in $B(X)$, we have

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_{\infty} + \|g\|_{\infty}$$

for every $x \in X$. When we talk of $B(X)$ as a normed linear space, what we have in mind is $(B(X), \|\cdot\|_{\infty})$, and when we talk about it as a metric space what we have in mind is $(B(X), d_{\|\cdot\|_{\infty}})$.

If $(X, d)$ is a metric space and $Y$ a nonempty subset of $X$, then, obviously, $(Y, d|_{Y \times Y})$ is also a metric space – it is said to be a metric subspace of $(X, d)$. Naturally, we denote this metric space simply as $(Y, d)$.

**Example 2.1.5.** For any topological space $X$, we denote the collection of all continuous, and continuous and bounded real functions on $X$ by $C(X)$ and $C_b(X)$, respectively. Unless otherwise is explicitly mentioned, the latter set is viewed as a metric subspace of $B(X)$.

**Note.** As a continuous real function on a compact topological space is bounded by the Weierstrass Theorem, we have $C_b(X) = C(X)$ when $X$ is compact.

**Example 2.1.6.** Let $(X_i, d_i)$ be metric spaces, $i = 1, ..., n$. The product of these spaces is defined as the metric space $(X, d_X)$, where $X := X_1 \times \cdots \times X_n$ and

$$d_X((x_1, ..., x_n), (y_1, ..., y_n)) := \sum_{i=1}^n d_i(x_i, y_i).$$

It is easy to check that $(X, \rho)$ is indeed a metric space. (In particular, we see that $\mathbb{R}^{n,1}$ is the $n$-fold product of $\mathbb{R}^n$.)
2.2 Topology of Metric Spaces

There is a natural way in which one can view a metric space \((X, d)\) as a topological space. To see this, let us define

\[ N_\varepsilon(x) := \{ \omega \in X : d(x, y) < \varepsilon \} , \]

which is called the \(\varepsilon\)-neighborhood of \(x\) relative to the metric space \((X, d)\). It is an easy exercise to check that

\[ N := \{ N_\varepsilon(x) : \varepsilon > 0 \text{ and } x \in X \} \]

is a basis for a topology on \(X\). The resulting topology is called the **metric topology** on \(X\) (induced by the metric \(d\)). Obviously, relative to this topology, a subset \(O\) of \(X\) is open iff for every \(x\) in \(O\) there is a number \(\varepsilon > 0\) such that \(N_\varepsilon(x) \subseteq O\). We refer to such a set \(O\) as open in \((X, d)\), and denote the collection of all such sets as \(\mathcal{O}_{(X,d)}\). The notions of closedness, interiority, closure, separability, compactness, etc., in \((X, d)\) are then deduced from this characterization in the straightforward way.

In what follows, and throughout this text, whenever we talk about a topological concept in the context of a metric space, it is the metric topology that we have in mind. For instance, for any two metric spaces \((X, d_X)\) and \((Y, d_Y)\), when we say that a function \(f : X \to Y\) is continuous at the point \(x\) of \(X\), what we mean is that for every \(\varepsilon > 0\), there exists a \(\delta > 0\) such that

\[ d_Y(f(x), f(y)) < \varepsilon \quad \text{for every } y \in X \text{ with } d_X(x, y) < \delta. \]

Similarly, given that the ambient topological space is a metric space \((X, d)\), when we say that a sequence \((x_m)\) in \(X\) converges to an element \(x\) of \(X\) in \((X, d)\), what we mean is that for every \(\varepsilon > 0\), there is a positive integer \(M\) such that

\[ d(x_m, x) < \varepsilon \quad \text{for every } m \geq M, \]

or, equivalently, \(d(x_m, x) \to 0\). (As usual, we write \(x_m \to x\) when this is the case.)

**Example 2.2.1.** Let \((X, d)\) be a metric space and \(x\) a point in \(X\). Every \(\varepsilon\)-neighborhood of \(x\) is open in \((X, d)\). Indeed, if \(y \in N_\varepsilon(x)\), then, by using the triangle inequality for \(d\), it is easily verified that \(N_{\varepsilon-d(x,y)}(x) \subseteq N_\varepsilon(x)\).

**Example 2.2.2.** Let \((X, d)\) be a metric space. Then, the map \(d : X \times X \to \mathbb{R}\) is continuous, where we think of the domain of \(d\) as the product metric space \((X \times X, d_{X \times X})\) as in Example 2.1.6. Then, by the triangle inequality for \(d\), we have

\[ d(x, z) - d(y, z) \leq d(x, y) \quad \text{and} \quad d(y, z) - d(x, z) \leq d(y, x), \]
and hence by the symmetry of $d$,

$$|d(x, z) - d(y, z)| \leq d(x, y) \quad \text{for every } x, y, z \in X.$$ 

It follows that

$$|d(x, z) - d(y, w)| \leq |d(x, z) - d(y, z)| + |d(y, z) - d(y, w)|$$
$$\leq d(x, y) + d(z, w)$$

for every $x, y, z$ and $w$ in $X$. Continuity of $d$ follows easily from this inequality.

**Note.** A major problem in general topology is to identify those topological spaces $X$ for which there exists a metric $d$ on $X$ such that the metric topology on $X$ induced by $d$ is identical to the topology of $X$. (Such topological spaces are called **metrizable**.) We shall not discuss this matter here, however.

**Note.** It is easily checked that the metric topology of a metric subspace of a metric space is the same as the subspace topology induced by the metric topology of the mother metric space.

The metric topologies of apparently distinct metric spaces may turn out to be identical. For instance, there is no difference between $\mathbb{R}^{n,p}$ and $\mathbb{R}^{n,q}$ from the topological viewpoint for any $p, q \geq 1$, as we show next.

**Example 2.2.3.** For any $n$-vectors $x$ and $y$, we define

$$d_\infty(x, y) := \max\{|x_1 - y_1|, ..., |x_n - y_n|\}.$$ 

Then, $(\mathbb{R}^n, d_\infty)$, which we shall denote simply by $\mathbb{R}^{n,\infty}$, is a metric space. (In fact, this metric space is none other than $B(\{1, ..., n\})$.) Obviously, we have

$$d_\infty(x, y) \leq d_p(x, y) \leq nd_\infty(x, y)$$

for any $n$-vectors $x$ and $y$. It follows that every $\varepsilon$-neighborhood of $x$ relative to $\mathbb{R}^{n,p}$ is contained in the $\varepsilon$-neighborhood of $x$ relative to $\mathbb{R}^{n,\infty}$, and conversely, every $\frac{\varepsilon}{n}$-neighborhood of $x$ relative to $\mathbb{R}^{n,\infty}$ is contained in the $\varepsilon$-neighborhood of $x$ relative to $\mathbb{R}^{n,p}$. Thus: The metric topologies on $\mathbb{R}^n$ induced by $d_\infty$ and $d_p$ are the same, for any $p \geq 1$.

In passing, we note that every metric space is Hausdorff and has a countable basis at each of its points. That is:
**Proposition 2.2.1.** Let \((X, d)\) be a metric space. Then, \(X\) is a first-countable Hausdorff topological space (relative to its metric topology).

**Proof.** For any distinct \(x\) and \(y\) in \(X\), we have \(N_\epsilon(x) \cap N_\epsilon(y) = \emptyset\) for \(\epsilon = \frac{1}{2}d(x, y)\). As \(N_\epsilon(x)\) and \(N_\epsilon(y)\) are open neighborhoods of \(x\) and \(y\), respectively (Example 2.2.2), this shows that \(X\) is Hausdorff relative to its metric topology.

For any \(x\) in \(X\), and any \(O \in O(X, d)\) with \(x \in O\), it is obvious that we have \(N_{1/m}(x) \subseteq O\) for a positive integer \(m\) large enough. Thus, for every \(x\) in \(X\),

\[
\mathcal{B}(x) := \{N_{1/m}(x) : m = 1, 2, \ldots\}
\]

is a countable collection of open neighborhoods of \(x\) such that

\[
x \in O \in O(X, d) \quad \text{implies} \quad x \in B \in O \text{ for some } B \in \mathcal{B}(x),
\]

which means that \(X\) is first-countable (relative to its metric topology).

Combining this result with Propositions 1.2.1 and 1.7.3 shows that we can characterize the notions of closedness and continuity in terms of convergent sequences in the context of metric spaces.

**Corollary 2.2.2.** Let \((X, d)\) be a metric space. Then, a subset \(S\) of \(X\) is closed in \((X, d)\) if, and only if, the limit of any convergent sequence in \(S\) belongs to \(S\).

**Corollary 2.2.3.** Let \((X, d_X)\) and \((Y, d_Y)\) be two metric spaces, and \(f : X \to Y\) a function. Then, \(f\) is continuous if, and only if, for every sequence \((x_m)\) in \(X\) and every element \(x\) of \(X\),

\[
d_X(x_m, x) \to 0 \quad \text{implies} \quad d_Y(f(x_m), f(x)) \to 0.
\]

### 2.3 Separable Metric Spaces

A metric space is said to be **separable** if it contains a countable dense subset. For instance, \(\mathbb{R}\) is separable, because \(\mathbb{Q}\) is a countable dense set in \(\mathbb{R}\). Similarly:

**Example 2.3.1.** \(\mathbb{R}^n\) is separable, because \(\mathbb{Q}^n\) is a countable dense set in \(\mathbb{R}^n\).

**Example 2.3.2.** A discrete space is separable iff it is countable.
Example 2.3.3. $C[a,b]$ is separable for any $-\infty < a \leq b < \infty$. An elementary proof of this claim could be given by using the piecewise linear functions with kinks occurring at rational points with rational values. Since the construction is a bit tedious, we omit the details here.

As these examples illustrate, separable metric spaces are encountered routinely in practice. It is thus useful to note that the separability property for a metric space is none other than the second countability of that metric space relative to the metric topology.

Proposition 2.3.1. Let $(X, d)$ be a metric space. Then, $X$ has a countable basis if, and only if, $(X, d)$ is separable.

Proof. The “if” part of this assertion is a special case of Proposition 1.2.2. To see its “only if” part, suppose that $(X, d)$ is a separable metric space. Let $S := \{x^1, x^2, \ldots\}$ be a dense set in $X$, and define

$$B := \left\{ N_{1/m}(x^i) : i, m = 1, 2, \ldots \right\}.$$ 

It is an easy exercise to check that $B$ is a basis for the (metric) topology of $X$. As $B$ is countable, this completes our proof.

This proposition shows that, while a metric space is always first-countable, it need not be second-countable: Any non-separable metric space is an example of a topological space that is first-countable but not second-countable. Next, we give a concrete example of such a space.

Example 2.3.4. The set of all bounded real sequences, that is, all real sequences $(x_m)$ with $\sup\{|x_m| : m = 1, 2, \ldots\} < \infty$, is denoted as $\ell_\infty$. Then, $(\ell_\infty, d_\infty)$, which we shall denote simply by $\ell_\infty$, is a metric space. (In fact, this metric space is none other than $B(\mathbb{N})$.) We shall show next that $\ell_\infty$ is not separable.

Note first that $\{0, 1\}^\infty$ is an uncountable subset of $\ell_\infty$. As $d_\infty((x_m), (y_m)) = 1$ for any two distinct members of this set,

$$O := \{ N_{1/2}((x_m)) : (x_m) \in \{0, 1\}^\infty \}$$

is a collection of uncountably many open sets in $\ell_\infty$ such that $O \cap U = \emptyset$ for any two distinct $O, U \in O$. Then, a subset of $\ell_\infty$ that intersects every member of $O$ must be
uncountable. As a dense subset of a topological space intersects every open subset of that space, therefore, we must conclude that any dense subset of $\ell_\infty$ is uncountable. Conclusion: $\ell_\infty$ is not separable.

In passing, we note that separability property is inherited by metric subspaces.

**Proposition 2.3.2.** Any metric subspace of a separable metric space is separable.

**Proof.** Let $(X, d)$ be a metric space that has a countable dense subset, say, $Y$. Take any metric subspace $(Z, d)$ of $(X, d)$. Define

$$Y_m := \{y \in Y : N_{1/m}(y) \cap Z \neq \emptyset\}.$$  

(As $Y$ is dense in $X$, each $Y_m$ is nonempty.) Now pick an arbitrary point $z^m(y)$ in $N_{1/m}(y) \cap Z$ for each positive integer $m$ and $y \in Y_m$, and define

$$W := \{z^m(y) : y \in Y_m \text{ and } m = 1, 2, \ldots\}.$$  

Clearly, $W$ is a countable subset of $Z$. (Yes?) Now take any $z$ in $Z$. By denseness of $Y$, for each positive integer $m$ we can find a $y^m \in Y$ with $d(z, y^m) < 1/m$. So $z \in N_{1/m}(y^m) \cap Z$, and hence $y^m \in Y_m$, $m = 1, 2, \ldots$ Therefore,

$$d(z, z^m(y^m)) \leq d(z, y^m) + d(y^m, z^m(y^m))$$

$$< \frac{1}{m} + \frac{1}{m}$$

$$= \frac{2}{m}.$$  

It follows that any element of $Z$ is in fact a limit of some sequence in $W$, that is, $W$ is dense in $Z$.

**Note.** The separability property is not hereditary in general, that is, a subspace of a topological space with a countable dense set need not possess a countable dense set. Here is an example: Let $B$ be the collection of all half-open rectangles of the form $[a, b) \times [c, d)$ – $B$ is easily checked to be a basis for a topology – and let $\mathcal{O}$ be the topology generated by $B$. Then, there is a countable dense set in $(\mathbb{R}^2, \mathcal{O})$, but $\{(x, y) \in \mathbb{R}^2 : x + y = 0\}$ does not possess a countable subset which is dense relative to the subspace topology.

### 2.4 Complete Metric Spaces

Complete metric spaces play a significant role in metric analysis. These are the metric spaces in which any sequence whose terms eventually get arbitrarily close to one another is, per force, convergent. Such sequences are called **Cauchy sequences**.
**Definition.** Let \((X, d)\) be a metric space. A sequence \((x_m)\) in \(X\) is said to be a **Cauchy sequence** in \((X, d)\) if for any \(\varepsilon > 0\), there exists a positive integer \(M\) such that

\[
d(x_k, x_l) < \varepsilon \quad \text{for every } k, l \geq M.
\]

For instance, \((1/m)\) is a Cauchy sequence in \(\mathbb{R}\), for the terms of this sequence get closer and closer towards its tail. More formally, this sequence is Cauchy, because for any \(\varepsilon > 0\), there exists a positive integer \(M\) such that \(|1/m| < \varepsilon/2\) for every \(m \geq M\). But then, by the triangle inequality, \(|1/k - 1/l| < \varepsilon\) whenever \(k, l \geq M\). On the other hand, \(((−1)^m)\) is not a Cauchy sequence in \(\mathbb{R}\), for \(|(−1)^m − (−1)^{m+1}| = 2\) for every positive integer \(m\).

The following result, whose (easy) proof is omitted provides an alternative characterization of Cauchy sequences.

**Proposition 2.4.1.** Let \((X, d)\) be a metric space and \((x_m)\) a sequence in \(X\). This sequence is Cauchy in \((X, d)\) if, and only if,

\[
\sum_{i=1}^{\infty} d(x_i, x_{i+1}) < \infty.
\]

**Note.** For any sequence \((x_m)\) in a metric space \((X, d)\) the condition that consecutive terms of the sequence get closer and closer, that is, \(d(x_m, x_{m+1}) \to 0\), is not a sufficient condition for \((x_m)\) to be Cauchy. For instance, \((\ln m)\) is a divergent real sequence but \(\ln(m+1) − \ln m = \ln(1 + 1/m) \to 0\).

While a Cauchy sequence need not converge – consider the sequence \((1/m)\) in the metric space \((0, 1]\) – it is plain that a Cauchy sequence is necessarily bounded, that is, the distance between the terms of any such sequence cannot be arbitrarily large. Furthermore, a Cauchy sequence with a convergent subsequence is itself convergent. We summarize these observations in the following proposition, whose (easy) proof is left as an exercise.

**Proposition 2.4.2.** Let \((X, d)\) be a metric space, and \((x_m)\) a sequence in \(X\).

(a) If \((x_m)\) is converges in \((X, d)\), then it is Cauchy in \((X, d)\).

(b) If \((x_m)\) is Cauchy in \((X, d)\), then \(\{x^1, x^2, \ldots\}\) is bounded in \((X, d)\).

(c) If \((x_m)\) is Cauchy in \((X, d)\), and has a subsequence that converges in \((X, d)\), then it converges in \((X, d)\) as well.
We are now ready for the following fundamental definition.

**Definition.** A metric space \((X, d)\) is said to be **complete** if every Cauchy sequence in \((X, d)\) converges in \((X, d)\). In turn, a normed linear space that is complete as a metric space is called referred to as a **Banach space**.

A useful example of a metric space that is not complete is supplied by \(\mathbb{Q}\) (viewed as a metric subspace of \(\mathbb{R}\)). Indeed, since \(\mathbb{Q}\) is dense in \(\mathbb{R}\), for any irrational number \(x\) we can find a sequence \((x_m)\) of rational numbers with \(x_m \to x\). Then, \((x_m)\) is Cauchy, but it does not converge in \(\mathbb{Q}\).

Here are some less trivial examples.

**Example 2.4.1.** Every discrete metric space is complete. For, a Cauchy sequence \((x_m)\) in such a space \((X, d)\) must be eventually constant, and this implies that \((x_m)\) converges in \((X, d)\).

**Example 2.4.2.** \(\mathbb{R}\) is complete. Indeed, if \((x_m)\) is a Cauchy sequence in \(\mathbb{R}\), it is bounded, and hence, by the Bolzano-Weierstrass Theorem and Proposition 2.4.2, it converges.

**Example 2.4.3.** \(\mathbb{R}^n\) is complete. This is obtained by using Example 2.4.2 and Proposition 2.4.2 in tandem.

**Example 2.4.4.** \(B(X)\) is complete for any nonempty set \(X\) – as a normed linear space, therefore, we may refer to \(B(X)\) as a Banach space (Example 2.1.4). To see this, let \((f_m)\) be a Cauchy sequence in \(B(X)\). For each \(x \in X\), it is clear that \((f_m(x))\) is a Cauchy sequence in \(\mathbb{R}\), so we have \(f_m(x) \to f(x)\) for some real number \(f(x)\). Let us denote the map \(x \mapsto f(x)\) obtained this way by \(f\). We wish to show first that \(f \in B(X)\). To this end, fix any \(\varepsilon > 0\), and notice that, as \((f_m)\) is Cauchy in \(B(X)\), there is a positive integer \(M\) such that \(\|f_k - f_l\|_{\infty} < \varepsilon\) for every \(k, l \geq M\). Then, for any \(m \geq M\),

\[
|f(x) - f_m(x)| = \lim_{k \to \infty} |f_k(x) - f_m(x)| \leq \lim_{k \to \infty} \|f_k - f_m\|_{\infty} \leq \varepsilon
\]

for each \(x \in X\). It follows that \(|f(x)| \leq |f_m(x)| + \varepsilon\) for each \(x \in X\), and hence, since \(f_m \in B(X)\), we have \(f \in B(X)\). Moreover, (4) gives us also that \(\|f - f_m\|_{\infty} \leq \varepsilon\) for every integer \(m \geq M\). Since \(\varepsilon\) is arbitrary here, we may thus conclude that \(\|f - f_m\|_{\infty} \to 0\), that is, \((f_m)\) converges in \(B(X)\), as we sought.
There is a tight connection between the closedness of a set and its completeness as a metric subspace. Indeed, a complete subspace of a metric space is necessarily closed. We even have a partial converse of this fact.

**Proposition 2.4.3.** Let \((X, d)\) be a metric space, and \(Y\) a nonempty subset of \(X\). If \((Y, d)\) is complete, then \(Y\) is closed in \((X, d)\). Conversely, if \(Y\) is closed in \((X, d)\) and \((X, d)\) is complete, then \((Y, d)\) is complete.

**Proof.** Let \((Y, d)\) be complete, and take any sequence \((x_m)\) in \(Y\) that converges in \((X, d)\). Then \((x_m)\) is Cauchy in \((X, d)\), and thus \(\lim x_m \in Y\). It follows from Corollary 2.2.2 that \(Y\) is closed in \((X, d)\).

To prove our second assertion, assume that \((X, d)\) is complete, and \(Y\) is closed in \((X, d)\). If \((x_m)\) is a Cauchy sequence in \((Y, d)\), then by completeness of \((X, d)\), it must converge in \((X, d)\). But since \(Y\) is closed in \((X, d)\), \(\lim x_m\) must belong to \(Y\) by Corollary 2.2.2. It follows that \((Y, d)\) is complete.

**Corollary 2.4.4.** A metric subspace \((Y, d)\) of a complete metric space \((X, d)\) is complete if, and only if, \(Y\) is closed in \((X, d)\).

Here is an important application of this fact.

**Example 2.4.5.** Let \((X, d)\) be a metric space. We wish to prove that \(C_b(X)\) is a complete metric space. In view of Example 2.4.4 and Corollary 2.4.4, it is enough to show that \(C_b(X)\) is closed in \(B(X)\).

Let \((f_m)\) be a sequence in \(C_b(X)\) and let \(\|f_m - f\|_\infty \to 0\) for some \(f \in B(X)\). We wish to show that \(f\) is continuous. Take any \(\varepsilon > 0\), and observe that there exists a positive integer \(M\) such that \(\sup\{|f(\omega) - f_M(\omega)| : \omega \in X\} < \varepsilon/3\). Thus,

\[
|f(\omega) - f_M(\omega)| < \frac{\varepsilon}{3} \quad \text{for every } \omega \in X. \tag{5}
\]

Now take an arbitrary \(x \in X\). By continuity of \(f_M\), there exists a \(\delta > 0\) such that \(f_M(N_\delta(x))\) is contained in \(N_{\varepsilon/3}(f_M(x))\). But then, for any \(y \in N_\delta(x)\),

\[
|f(x) - f(y)| \leq |f(x) - f_M(x)| + |f_M(x) - f_M(y)| + |f_M(y) - f(y)| < \varepsilon
\]

where we used (5) twice. Since \(x\) was chosen arbitrarily in \(X\), we may conclude that \(f\) is continuous.
2.5 Compact Metric Spaces

In this section we shall show that there are various equivalent ways of looking at the notion of compactness in the context of metric spaces. We begin by noting that every compact metric space \((X, d)\) is bounded, that is,

\[
\text{diam}(X) := \sup \{d(x, y) : x, y \in X\} < 1.
\]

(Here \(\text{diam}(X)\) is called the diameter of \(X\).) In fact, more generally, such a metric space is totally bounded, that is, for any \(\varepsilon > 0\), there is a finite subset \(S\) of \(X\) such that \(\{B_\varepsilon(x) : x \in S\}\) covers \(X\). A certain converse of this also true, as we prove next.

**Theorem 2.5.1.** For any metric space \((X, d)\), the following are equivalent:

(a) \((X, d)\) is compact;

(b) \((X, d)\) is sequentially compact, that is, every sequence in \(X\) has a subsequence that converges in \((X, d)\);

(c) \((X, d)\) is complete and totally bounded.

**Proof.** Assume that \((X, d)\) is compact, and take any sequence \((x_m)\) in \(X\). To derive a contradiction, suppose \((x_m)\) does not have a subsequence that converges in \((X, d)\). In that case, \(S := \{x^1, x^2, \ldots\}\) must be a closed subset of \((X, d)\). Since \((X, d)\) is compact, then, \(S\) is compact in \((X, d)\) by Proposition 1.5.1. On the other hand, since \((x_m)\) lacks a convergent subsequence, for any positive integer \(m\) there is an \(\varepsilon_m > 0\) such that \(N_{\varepsilon_m}(x_m) \cap \{x^1, x^2, \ldots\} = \{x_m\}\). But \(\{N_{\varepsilon_m}(x_m) : m = 1, 2, \ldots\}\) is an open cover of \(S\). So, by compactness of \(S\), there is a finite cover of \(S\) that can be extracted from this collection. It follows that \(S\) is a finite set, which means that at least one term of \((x_m)\) must be repeated infinitely often in the sequence, that is, there is a constant subsequence of \((x_m)\), which is trivially convergent in \((X, d)\), a contradiction. Conclusion: (a) implies (b).

Assume that \((X, d)\) is sequentially compact. In view of part (c) of Proposition 2.4.2, it is readily verified that \((X, d)\) is complete. Next, to derive a contradiction, suppose \((X, d)\) is not totally bounded, that is, there exists an \(\varepsilon > 0\) such that \(\{B_\varepsilon(x) : x \in S\} = X\) does not hold for \textit{any} finite subset \(S\) of \(X\). We wish to construct a sequence in \(X\) no subsequence of which converges in \((X, d)\). Begin by picking an \(x^1 \in X\) arbitrarily. By hypothesis, we cannot have \(X \subseteq B_\varepsilon(x^1)\) so there must exist an \(x^2 \in X\) such that \(d(x^1, x^2) \geq \varepsilon\). Again, \(X \subseteq B_\varepsilon(x^1) \cup B_\varepsilon(x^2)\) cannot hold, so we can find an \(x^3 \in X\) such that \(d(x^1, x^3) \geq \varepsilon, i = 1, 2\). Continuing this way, and invoking the Axiom of Choice, we obtain a sequence \((x_m)\) in \(X\) such that \(d(x^i, x^j) \geq \varepsilon\) for any distinct positive integers \(i\)
and \( j \). Since \((X,d)\) is sequentially compact, there exists a subsequence, say \((x^{m_k})\), of \((x_m)\) which converges in \((X,d)\). But, as \((x_m)\) is not even Cauchy in \((X,d)\), this is impossible. Conclusion: (b) implies (c).

Assume that \((X,d)\) is complete and totally bounded. Let \(O\) be an arbitrary open cover of \(X\). To derive a contradiction, assume that no finite subset of \(O\) covers \(X\). Since \((X,d)\) is totally bounded, \(X\) can be covered by finitely many nonempty closed sets of diameter at most 1. Therefore, at least one of these closed sets, say \(S_1\), cannot be covered by finitely many elements of \(O\). Obviously, \((S_1,d)\) itself is totally bounded, as \(S \subseteq X\) and \((X,d)\) is totally bounded. So, \(S_1\) can be covered by finitely many nonempty closed sets of diameter at most \(\frac{1}{2}\), and at least one of these sets, say \(S_2\), cannot be covered by finitely many elements of \(O\). Continuing inductively, and invoking the Axiom of Choice, we obtain a sequence \((S_m)\) of nonempty closed subsets of \(X\) such that \(S_1 \supseteq S_2 \supseteq \cdots\) and \(\text{diam}(S_m) \leq \frac{1}{m}\) for each \(m\), whereas no \(S_m\) can be covered by finitely many sets in \(O\).

Now, again by the Axiom of Choice, pick any point \(x_m\) in \(S_m\) for each positive integer \(m\), and note that \((x_m)\) is a Cauchy sequence in \((X,d)\). So, since \((X,d)\) is complete, we have \(x_m \to x\) for some \(x \in X\). Since each \(S_m\) is closed and \(x_k \in S_m\) for all \(k \geq m\), it follows that \(x \in \bigcap_{i=1}^{\infty} S_i\). Yet, \(x \in O\) for some \(O \in O\) (because \(O\) covers \(X\)). Since \(O\) is open, there exists an \(\varepsilon > 0\) such that \(N_{\varepsilon}(x) \subseteq O\). But for any \(m > 1/\varepsilon\), we have \(\text{diam}(S_m) < \varepsilon\) and \(x \in S_m\), so \(S_m \subseteq N_{\varepsilon}(x)\). Thus \(S_m\) is covered by an element of \(O\), contradiction. Conclusion: (c) implies (a).

In passing, we note that combining this result with Corollary 2.4.4 leads us to the following fact:

**Corollary 2.5.2.** A closed subset of a complete metric space is compact if, and only if, it is totally bounded.

Theorem 2.5.1 simplifies the task of identifying compact subsets of metric spaces. Here is a major case in point.

**The Heine-Borel Theorem.** A subset of \(\mathbb{R}^n\) is compact (as a metric subspace of \(\mathbb{R}^n\)) if, and only if, it is closed and bounded.

**Proof.** As every metric topology is Hausdorff, the “only if” part of the claim follows from Proposition 1.5.3 and Theorem 2.5.1. To see its “if” part, notice that, for any real numbers \(a\) and \(b\) with \(a \leq b\), the compactness of \([a,b]\) is an immediate consequence of
the Bolzano-Weierstrass Theorem and Theorem 2.5.1. To prove the same for \([a, b]^n\) for any positive integer \(n\), take any sequence \((x_m)\) in \([a, b]^n\). Then \((x_{m,1})\) is a real sequence in \([a, b]\). So, by Theorem 2.5.1, \((x_{m,1})\) has a convergent subsequence in \([a, b]\), call it \((x_{m_k,1})\). Now observe that \((x_{m_k,2})\) must have a convergent subsequence in \([a, b]\):

Continuing this way, we can obtain a subsequence of \((x_m)\) that converges in \([a, b]^n\). This proves that \([a, b]^n\) is sequentially compact. Thus, by Theorem 2.5.1, \([a, b]^n\) is compact.

Finally, take any closed and bounded subset \(S\) of \(\mathbb{R}^n\). As \(\text{diam}(S) < \infty\), there are real numbers \(a\) and \(b\) such that \(S \subseteq [a, b]^n\). As \([a, b]^n\) is compact, and \(S\) is a closed subset of \([a, b]^n\), it follows from Proposition 1.5.1 that \(S\) is compact.

**Corollary 2.5.3.** A subset \(S\) of \(\mathbb{R}^n\) is bounded if, and only if, it is totally bounded.

**Proof.** The “if” part of the assertion is straightforward. To see its “only if” part, let \(S\) be a bounded subset of \(\mathbb{R}^n\). Then, \(\text{cl}(S)\) is a closed and bounded, and hence compact, subset of \(\mathbb{R}^n\), and therefore, it is totally bounded.

### 2.6 The Stone-Weierstrass Theorem

(TBW)...for compact Hausdorff spaces only - mention that the result is valid for locally compact Hausdorff spaces as well in a note.

Prove that \(C(X)\) is separable whenever \(X\) is compact Hausdorff.

### 2.7 Lipschitz Continuity

Let \((X, d_X)\) and \((Y, d_Y)\) be two metric spaces and \(f : X \to Y\) a function. The presence of the metrics on the domain and codomain of \(f\) allows us to relate the distances between two points in \(X\) and their images in \(Y\) under \(f\). In particular, we define the **Lipschitz constant** of \(f\) as the extended real number

\[
\text{Lip}(f) := \inf \left\{ \frac{d_Y(f(x), f(y))}{d_X(x, y)} : x, y \in X \text{ and } x \neq y \right\}.
\]

The size of this number provides a useful classification criteria for maps from \(X\) into \(Y\). In particular, we say that \(f\) is **Lipschitz continuous** if \(\text{Lip}(f) < \infty\). (It follows readily from Proposition 1.7.3 that all Lipschitz continuous functions are continuous.) We say that \(f\) is **nonexpansive** if \(\text{Lip}(f) \leq 1\), and that it is a **contraction** if \(\text{Lip}(f) < 1\).

One of the most important results about contractions is:

**The Contraction Mapping Theorem.** Let \((X, d)\) be a complete metric space, and \(f : X \to X\) a contraction. Then, \(f\) has a unique fixed point.
We shall not prove this theorem here, as we prove substantially more general results within the body of the text.

We conclude by making note of the following (easy) consequence of the Stone-Weierstrass Theorem:

**Proposition 2.7.1.** Let $(X, d)$ be a complete metric space, and $\mathcal{F}$ is the collection of all Lipschitz continuous real maps on $X$. Then, $\mathcal{F}$ is dense in $C(X)$.

### 3 The Hausdorff Metric

Let $(X, d)$ be metric space and $\text{cb}(X)$ the class of all closed and bounded subsets of $X$. There are various settings in which one needs to metrize $\text{cb}(X)$ in a way loyal to the metrization of $X$ in the sense of judging the distance between the sets $\{x\}$ and $\{y\}$ exactly as $d(x, y)$, for any $x, y \in X$. There are many different methods of doing this. In this section, we shall introduce the most popular of these methods.

#### 3.1 Distance of a Point from a Set

Let $S$ be any nonempty subset of a metric space $(X, d)$, and $x$ a point in $X$. The **distance of $x$ from $S$** is defined as

$$d(x, S) := \inf\{d(x, \omega) : \omega \in S\}.$$  

Thus, the function $f : X \to \mathbb{R}_+$ defined by $f(x) := d(x, S)$ measures the distance of any given point in $X$ from $S$ in terms of the metric $d$. We note that this function is (Lipschitz) continuous. Indeed, for each elements $x$ and $y$ of $X$, the triangle inequality for $d$ yields

$$f(x) = d(x, S) \leq \inf\{d(x, y) + d(y, z) : z \in S\} = d(x, y) + f(y),$$  

and similarly, $f(y) \leq d(y, x) + f(x)$. Thus

$$|f(x) - f(y)| \leq d(x, y) \quad \text{for all } x, y \in X,$$

and it follows that $f$ is continuous.

Combining this observation with the Weierstrass Extremum Value Theorem, it is evident that $d(x, S) = \min\{d(x, \omega) : \omega \in S\}$ when $S$ is compact. However, in general,
this equation is not true when all we know is that \( S \) is closed and bounded, as we show next.

**Example 3.1.1.** Let \( X \) be the collection of all real sequences \((x_m)\) such that \( x_m = 0\) for all but finitely many \( m \), and defined \( d : X \times X \to \mathbb{R} \) by

\[
d((x_m), (y_m)) = \left( \sum_{i=1}^{\infty} (x_i - y_i)^2 \right)^{1/2}.
\]

It is easy to check that \((X, d)\) is a metric space (by using the Cauchy-Schwarz Inequality). Next, define

\[
S := \left\{ (x_m) \in X : \sum_{i=1}^{\infty} 2^{-i} x_i = 0 \right\}.
\]

It is easy to check that \( S \) is a closed subset of \( X \) (using, again, the Cauchy-Schwarz Inequality). Let \( e := (1, 0, 0, \ldots) \). We leave it as an exercise to prove that \( d((x_m), e) \) is not equal to \( d(e, S) \) for any \((x_m)\) in \( S \). (In fact, in this example, \( e \) can be replaced with any member of \( X \) outside \( S \).)

In passing, we note the following property of the \( d(\cdot, S) \) function.

**Proposition 3.1.1.** Let \((X, d)\) be a metric space and \( S \) a nonempty closed subset of \( X \). Then, \( x \in S \) if, and only if, \( d(x, S) = 0 \).

We leave the (easy) proof of this result as an exercise.

### 3.2 The Hausdorff Metric

Let \((X, d)\) be metric space and \( \text{cb}(X) \) the class of all closed and bounded subsets of \( X \). For any \( A \) and \( B \) in \( \text{cb}(X) \), we define

\[
\Lambda(A, B) := \sup\{d(a, B) : a \in A\},
\]

which is a real number because \( \text{diam}(A \cup B) < \infty \). The **Hausdorff metric** on \( d_H : \text{cb}(X) \times \text{cb}(X) \to \mathbb{R}_+ \) is defined by

\[
d^H(A, B) := \max\{\Lambda(A, B), \Lambda(B, A)\}.
\]

**Proposition 3.2.1.** Let \((X, d)\) be a metric space. Then, \((\text{cb}(X), d^H)\) is a metric space.
**Proof.** The only nontrivial part of the assertion is that $d^H$ satisfies the triangle inequality. To see this, take any $A$, $B$ and $C$ in $\mathbf{cb}(X)$, and note that $d(x, B) \leq d(x, y)$ for each $y$ in $B$. Therefore, by the triangle inequality,

$$d(x, B) \leq d(x, z) + d(z, y) \quad \text{for every } (x, y, z) \in A \times B \times C.$$ 

Taking the infimum over $y$ in the right-hand side of this inequality, we get

$$d(x, B) \leq d(x, z) + d(z, B) \quad \text{for every } (x, z) \in A \times C.$$ 

As $d(z, B) \leq \Lambda(C, B) \leq d^H(C, B)$, therefore,

$$d(x, B) \leq d(x, z) + d^H(C, B) \quad \text{for every } (x, z) \in A \times C.$$ 

Taking the infimum over $z$ in the right-hand side of this inequality, and noting that $d(x, C) \leq \Lambda(A, C) \leq d^H(A, C)$, we find

$$d(x, B) \leq d^H(A, C) + d^H(C, B) \quad \text{for every } x \in A.$$ 

Finally, taking the supremum over $x$ in the left-hand side, we get

$$\Lambda(A, B) \leq d^H(A, C) + d^H(C, B).$$ 

As $d^H$ is symmetric, replacing the roles of $A$ and $B$ in this inequality yields

$$\Lambda(B, A) \leq d^H(A, C) + d^H(C, B)$$ 

as well, and our assertion is established.

There are other characterizations of the Hausdorff metric which are occasionally useful. The following result, whose proof we leave as an exercise, provides two such characterizations:

**Proposition 3.2.2.** Let $(X, d)$ be a metric space, and $A$ and $B$ two nonempty closed and bounded subsets of $(X, d)$. Then,

$$d^H(A, B) = \inf\{\varepsilon > 0 : A \subseteq N_\varepsilon(B) \text{ and } B \subseteq N_\varepsilon(A)\},$$

and

$$d^H(A, B) = \sup_{x \in X} |d(x, A) - d(x, B)|.$$ 

42
Note. The second part of Proposition 3.2.2 shows that convergence relative to the Hausdorff metric can be characterized by means of the uniform convergence of certain distance maps. Put precisely, this result shows that \( A_m \to_H A \) iff \( d(\cdot, A_m) \to d(\cdot, A) \) uniformly. This suggests considering the following pointwise convergence notion: \( A_m \to_W A \) iff \( d(x, A_m) \to d(x, A) \) for each \( x \in X \). This is called Wijsman convergence. When \( X \) is a separable metric space, it is possible to introduce a metric on \( \mathcal{C}(X) \) that would induce precisely this convergence notion. Put precisely, if \( X \) is a complete and separable metric space, and \( \{x_1, x_2, \ldots\} \) a countable dense subset of \( X \), then, where the real map \( d^W \) on \( \mathcal{C}(X) \times \mathcal{C}(X) \) is defined by

\[
d^W(A, B) := \sum_{i=1}^{\infty} 2^{-i} \min\{1, |d(x_i, A) - d(x_i, B)|\},
\]

\( (\mathcal{C}(X), d^W) \) is a complete and separable metric space such that \( A_m \to_W A \) iff \( d^W(A_m, A) \to 0 \), for any \( A, A_1, A_2, \ldots \) in \( \mathcal{C}(X) \). (This is Beer’s Theorem.)

A general theme concerning the theory of Hausdorff metric is to determine those properties of the metric space \( (X, d) \) that are preserved by \( (\mathcal{C}(X), d^H) \). One major property of this nature is completeness.

**Proposition 3.2.3.** Let \( (X, d) \) be a complete metric space. Then, \( (\mathcal{C}(X), d^H) \) is a complete metric space.

**Proof.** Let \( (S_m) \) be a Cauchy sequence in \( (\mathcal{C}(X), d^H) \). Then, clearly, there is a subsequence of \( (S_m) \), say \( (S_{m_k}) \), such that

\[
d^H(S_{m_k}, S_{m_k+1}) < 2^{-m_k} \quad \text{for each } k = 1, 2, \ldots.
\]

We wish to prove that \( (S_{m_k}) \) is convergent. As a Cauchy sequence with a convergent subsequence converges, this is enough to prove our assertion.

Letting \( T_k := S_{m_k} \) for each \( k \), we may denote the subsequence \( (S_{m_k}) \) as \( (T_k) \). Observe that for any sequence \( (x_k) \) in \( X \) with \( d(x_i, x_{i+1}) < 2^{-i} \) for each \( i \), we have

\[
\sum_{i=1}^{\infty} d(x_i, x_{i+1}) < \infty,
\]

which means that \( (x_k) \) is Cauchy (Proposition 2.4.1), and hence convergent (because \( (X, d) \) is complete). It follows that the set

\[
T := \{\lim x_k : x_i \in T_i \text{ and } d(x_i, x_{i+1}) < 2^{-i}, i = 1, 2, \ldots\}.
\]

is well-defined. It is also easily checked to be nonempty. We wish to prove that

\[
T_m \to_H \text{cl}(T).
\]
To this end, we pick an arbitrary $\varepsilon > 0$, and establish the following claims:

**Claim 1.** There exists a positive integer $K$ such that

$$\sup_{x \in T_k} d(x, \text{cl}(T)) \leq \varepsilon$$

whenever $k \geq K$.

**Proof of Claim 1.** Choose $K$ large enough so that

$$\sum_{i=K}^{\infty} 2^{-i} < \frac{\varepsilon}{2},$$

fix any integer $k \geq K$, and pick any $x$ in $T_k$. As $d(T_i, T_{i+1}) < 2^{-i}$ for every integer $i \geq K$, we can use the Axiom of Choice to find a sequence $(x_{k+1}, x_{k+2}, \ldots)$ in $T_{k+1} \times T_{k+2} \times \cdots$ such that $d(x, x_{k+1}) < 2^{-k}$ and $d(x_i, x_{i+1}) < 2^{-i}$ for each $i \geq K + 1$. Then, clearly, $(x, x_{k+1}, x_{k+2}, \ldots)$ is a Cauchy sequence in $(X, d)$. As $(X, d)$ is complete, this sequence converges to some point $x(k)$ in $X$. By definition of $T$, we have $x(k) \in T$. Moreover,

$$d(x, \text{cl}(T)) = d(x, T) \leq d(x, x(k)),$$

whereas, for $l > k + 1$ large enough that $d(x_l, x(k)) < \varepsilon/2$, we have

$$d(x, x(k)) \leq d(x, x_{k+1}) + \cdots + d(x_{l-1}, x_l) + d(x_l, x(k))$$

$$< \sum_{i=K}^{\infty} 2^{-i} + \frac{\varepsilon}{2}$$

$$< \varepsilon,$$

and hence, $d(x, \text{cl}(T)) < \varepsilon$. As $x$ was arbitrarily chosen in $T_k$, Claim 1 follows from this observation.

**Claim 2.** There exists a positive integer $K$ such that

$$\sup_{x \in \text{cl}(T)} d(x, T_k) \leq \varepsilon$$

whenever $k \geq K$.

**Proof of Claim 2.** Choose $K$ large enough so that

$$\sum_{i=K}^{\infty} 2^{-i} < \varepsilon,$$

fix any integer $k \geq K$, and pick any $x$ in $\text{cl}(T)$. Then, we can find a point $y$ in $T$ such that $d(x, y) < \varepsilon$. By definition of $T$, there exists a sequence $(x_1, x_2, \ldots)$ in $T_1 \times T_2 \times \cdots$ such that $d(x_i, x_{i+1}) < 2^{-i}$ for every positive integer $i$, and $x_m \rightarrow y$. Then,

$$d(x, T_k) \leq d(x, y) + d(y, T_k)$$

$$< \varepsilon + d(y, x_k)$$

$$\leq \varepsilon + d(\lim_{m \to \infty} x_m, x_k)$$

$$\leq \varepsilon + d(x_m, x_k)$$
while
\[ \lim_{m \to \infty} d(x_m, x_k) = \lim_{m \to \infty} \sum_{i=m}^{\infty} d(x_i, x_{i+1}) = 0, \]
and hence, letting \( m \to \infty \) in the previous inequality yields \( d(x, T_k) < \varepsilon \). As \( x \) was arbitrarily chosen in \( T \), Claim 2 follows from this observation.

Combining these claims shows that \( T_m \to_H \text{cl}(T) \), and completes our proof.

Another important property that is preserved under the Hausdorff metric is total boundedness. That is:

**Proposition 3.2.4.** Let \((X, d)\) be a totally bounded metric space. Then, \((\text{cb}(X), d^H)\) is a totally bounded metric space.

**Proof.** Take any \( \varepsilon > 0 \). As \((X, d)\) is totally bounded, we can pick an arbitrary \( \delta \) in the interval \((0, \varepsilon)\), and find a finite subset \( S \) of \( X \) such that
\[ X = \bigcup \{ N_\delta(x) : x \in S \}. \]
Let \( \mathcal{S} \) be the collection of all nonempty subsets of \( S \). Obviously, \( |\mathcal{S}| < \infty \). We wish to show that
\[ \text{cb}(X) = \bigcup_{T \in \mathcal{S}} \{ A \in \text{cb}(X) : d^H(A, T) < \varepsilon \}, \]
which would complete our proof. To this end, take any nonempty closed and bounded subset \( B \) of \( X \). Define
\[ T := \{ x \in S : N_\delta(x) \cap B \neq \emptyset \}. \]
Then,
\[ B \subseteq \bigcup \{ N_\delta(x) : x \in T \} = N_\delta(T). \]
It follows that \( d^H(B, T) \leq \alpha < \varepsilon \), and hence \( B \) belongs to the \( \varepsilon \)-neighborhood of \( T \) in the space \((\text{cb}(X), d^H)\) as we sought.

Combining Propositions 3.2.3 and 3.2.4, and applying Theorem 2.5.1, we reach to the main result of this section:

**Proposition 3.5.** Let \((X, d)\) be a compact metric space. Then, \((\text{cb}(X), d^H)\) is a compact metric space.