Chapter 2

Metric Spaces and Topology

From an engineering perspective, the most important way to construct a topology on a set is to define the topology in terms of a metric on the set. This approach underlies our intuitive understanding of open and closed sets on the real line. Generally speaking, a metric captures the notion of a distance between two elements of a set. Topologies that are defined through metrics possess a number of properties that make them suitable for analysis. Identifying these common properties permits the unified treatment of different spaces that are useful in solving engineering problems. To gain better insight into metric spaces, we need to review the notion of a metric and to introduce a definition for topology.

2.1 Metric Spaces

A **metric space** is a set that has a well-defined "distance" between any two elements of the set. Mathematically, the notion of a metric space abstracts a few basic properties of Euclidean space. Formally, a metric space (X, d) is a set X and a function d that is a metric on X.

Definition 2.1.1. *A* ***metric*** *on a set X is a function*

$$d : X \times X \to \mathbb{R}$$

that satisfies the following properties,

1. $d(x, y) \geq 0 \quad \forall x, y \in X$*; equality holds if and only if* $x = y$

2. $d(x, y) = d(y, x) \quad \forall x, y \in X$

3. $d(x, y) + d(y, z) \geq d(x, z) \quad \forall x, y, z \in X.$

Example 2.1.2. *The set of real numbers equipped with the metric of absolute distance* $d(x, y) = |x - y|$ *defines the standard metric space of real numbers* $\mathbb{R}$.

Example 2.1.3. *Given* $\underline{x} = (x_1, \ldots, x_n), \underline{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$*, the* ***Euclidean metric*** d *on* $\mathbb{R}^n$ *is defined by the equation*

$$d\left(\underline{x}, \underline{y}\right) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

As implied by its name, the function d *defined above is a metric.*

P 2.1.4. *Let* $\underline{x} = (x_1, \ldots, x_n), \underline{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ *and consider the function* ρ *given by*

$$\rho\left(\underline{x}, \underline{y}\right) = \max\left\{|x_1 - y_1|, \ldots, |x_n - y_n|\right\}.$$

Show that ρ *is a metric.*

P 2.1.5. *Let* X *be a metric space with metric* d*. Define* $\bar{d} : X \times X \to \mathbb{R}$ *by*

$$\bar{d}(x, y) = \min\left\{d(x, y), 1\right\}.$$

Show that $\bar{d}$ *is also a metric.*

Let (X, d) be a metric space. Then, elements of X are called **points** and the number $d(x, y)$ is called the **distance** between x and y. Let $\epsilon > 0$ and consider the set $B_d(x, \epsilon) = \{y \in X | d(x, y) < \epsilon\}$. This set is called the d**-open ball** of radius ϵ centered at x.

P 2.1.6. *Suppose* $a \in B_d(x, \epsilon)$ *with* $\epsilon > 0$*. Show that there exists a* d*-open ball centered at* a *of radius* δ*, say* $B_d(a, \delta)$*, that is contained in* $B_d(x, \epsilon)$*.*

One of the main benefits of having a metric is that it provides some notion of "closeness" between points in a set. This alows one to discusss limits, convergence, open sets, and closed sets.

Definition 2.1.7. *A* ***sequence*** *of elements from a set* X *is an infinite list* $x_1, x_2, \ldots$ *where* $x_i \in X$ *for all* $i \in \mathbb{N}$*. Formally, a sequence is equivalent to a function* $f : \mathbb{N} \to X$ *where* $x_i = f(i)$ *for all* $i \in \mathbb{N}$*.*

Definition 2.1.8. *Consider a sequence $x_1, x_2, \ldots$ of points in a metric space (X, d). This sequence d-**converges** to $x \in X$ if, for any $\epsilon > 0$, there is natural number N such that $d(x, x_n) < \epsilon$ for all $n > N$.*

Definition 2.1.9. *A sequence $x_1, x_2, \ldots$ in (X, d) is a **Cauchy sequence** if, for any $\epsilon > 0$, there is a natural number N (depending on ϵ) such that, for all $m, n > N$,*

$$d\left(x_m, x_n\right) < \epsilon.$$

Theorem 2.1.10. *Every d-convergent sequence is a Cauchy sequence.*

Proof. Since $x_1, x_2, \ldots$ d-converges to some x, there is an N, for any $\epsilon > 0$, such that $d(x, x_n) < \epsilon/2$ for all $n > N$. The triangle inequality for $d(x_m, x_n)$ shows that, for all $m, n > N$,

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

Therefore, $x_1, x_2, \ldots$ is a Cauchy sequence. □

Definition 2.1.11. *Let W be a subset of a metric space (X, d). The set W is called d-**open** if, for every $w \in W$, there is an $\epsilon > 0$ such that $B_d(w, \epsilon) \subseteq W$.*

Theorem 2.1.12. *For any metric space (X, d),*

1. *$\emptyset$ and X are d-open*
2. *any union of d-open sets is d-open*
3. *any finite intersection of d-open sets if d-open*

Proof. This proof is left as an exercise for the reader. □

One might be curious why only finite intersections are allowed in Theorem 2.1.12. The following example highlights the problem with allowing infinite intersections.

Example 2.1.13. *Let $I_n = \left(-\frac{1}{n}, \frac{1}{n}\right) \subset \mathbb{R}$, for $n \in \mathbb{N}$, be a sequence of open real intervals. The infinite intersection*

$$\bigcap_{n \in \mathbb{N}} I_n = \{x \in \mathbb{R} | \forall n \in \mathbb{N}, x \in I_n\} = \{0\}.$$

But, it is easy to verify that $\{0\}$ is not a d-open set.

Definition 2.1.14. *A subset W of a metric space (X,d) is d-closed if its complement $W^c = X - W$ is d-open.*

Corollary 2.1.15. *For any metric space (X,d),*

1. *$\emptyset$ and X are d-closed*
2. *any intersection of d-closed sets is d-closed*
3. *any finite union of d-closed sets if d-closed*

Sketch of proof. Using the definition of d-closed, one can apply De Morgan's Laws to Theorem 2.1.12 verify this result. □

Definition 2.1.16. *Suppose $f : X \to Y$ is a function from the metric space (X, d_X) to the metric space (Y, d_Y). Then, f is d-**continuous** at x_0 if, for any $\epsilon > 0$, there exists a $\delta > 0$ such that, for all $x \in X$ satisfying $d_X(x_0, x) < \delta$,*

$$d_Y\left(f(x_0), f(x)\right) < \epsilon.$$

In precise mathematical notation, one has

$$\begin{aligned}&(\forall\, \epsilon > 0)(\exists\, \delta > 0),\\ &((x \in X) \wedge (d_X(x_0, x) < \delta)) \Rightarrow d_Y\left(f(x_0), f(x)\right) < \epsilon.\end{aligned}$$

Definition 2.1.17. *A function $f : X \to Y$ is called d-**continuous** if, for all $x_0 \in X$, it is d-continuous at x_0. In precise mathematical notation, one has*

$$\begin{aligned}&(\forall\, x_0 \in X)(\forall\, \epsilon > 0)(\exists\, \delta > 0),\\ &((x \in X) \wedge (d_X(x_0, x) < \delta)) \Rightarrow d_Y\left(f(x_0), f(x)\right) < \epsilon.\end{aligned}$$

Definition 2.1.18. *A metric space (X,d) is said to be **complete** if every Cauchy sequence in X converges to a limit $x \in X$.*

Example 2.1.19. *Any closed subset of $\mathbb{R}^n$ (or $\mathbb{C}^n$) is complete.*

Example 2.1.20. *Consider the sequence $x_n \in \mathbb{Q}$ defined by $x_n = \left(1 + \frac{1}{n}\right)^n$. It is well-known that this sequence converges to $e \in \mathbb{R}$, but this number is not rational. Therefore, the rational numbers $\mathbb{Q}$ are not complete.*

Definition 2.1.21. *An* ***isometry*** *is a mapping* $\phi : X \to Y$ *between two metric spaces* (X, d_X) *and* (Y, d_Y) *that is distance preserving (i.e., it satisfies* $d_X(x, x') = d_Y(\phi(x), \phi(x'))$ *for all* $x, x' \in X$*).*

Definition 2.1.22. *A function* $f : X \to Y$ *is called* ***uniformly*** d***-continuous*** *if it* d*-continuous and, for all* $\epsilon > 0$*, the* $\delta > 0$ *can be chosen independently of* x_0*. In precise mathematical notation, one has*

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall x_0 \in X),$$
$$((x \in X) \wedge (d_X(x_0, x) < \delta)) \Rightarrow d_Y(f(x_0), f(x)) < \epsilon.$$

2.2 General Topology

Definition 2.2.1. *A* ***topology*** *on a set* X *is a collection* $\mathcal{J}$ *of subsets of* X *that satisfies the following properties,*

1. $\emptyset$ *and* X *are in* $\mathcal{J}$
2. *the union of the elements of any subcollection of* $\mathcal{J}$ *is in* $\mathcal{J}$
3. *the intersection of the elements of any finite subcollection of* $\mathcal{J}$ *is in* $\mathcal{J}$*.*

A subset $A \subseteq X$ is called an **open set** of X if $A \in \mathcal{J}$. Using this terminology, a topological space is a set X together with a collection of subsets of X, called *open sets*, such that $\emptyset$ and X are both open and such that arbitrary unions and finite intersections of open sets are open.

Definition 2.2.2. *If* X *is a set, a* ***basis*** *for a topology on* X *is a collection* $\mathcal{B}$ *of subsets of* X *(called* basis elements*) such that:*

1. *for each* $x \in X$*, there exists a basis element* B *containing* x*.*
2. *if* $x \in B_1$ *and* $x \in B_2$ *where* $B_1, B_2 \in \mathcal{B}$*, then there exists a basis element* B_3 *containing* x *such that* $B_3 \subseteq B_1 \cap B_2$*.*
3. *a subset* $A \subseteq X$ *is open in the topology on* X *generated by* $\mathcal{B}$ *if and only if, for every* $x \in A$*, there exists a basis element* $B \in \mathcal{B}$ *such that* $x \in B$ *and* $B \subseteq A$*.*

Probably the most important and frequently used way of imposing a topology on a set is to define the topology in terms of a metric.

Example 2.2.3. *If d is a metric on the set X, then the collection of all ϵ-balls*

$$\{B_d(x, \epsilon) | x \in X, \epsilon > 0\}$$

*is a basis for a topology on X. This topology is called the **metric topology** induced by d.*

Applying the meaning of open set from Definition 2.2.2 to this basis, one finds that a set A is open if and only if, for each $x \in A$, there exists a $\delta > 0$ such that $B_d(x, \delta) \subset A$. Clearly, this condition agrees with the definition of d-open from Definition 2.1.11.

Definition 2.2.4. *Let X be a topological space. This space is said to be **metrizable** if there exists a metric d on the set X that induces the topology of X.*

Example 2.2.5. *While most of the spaces discussed in these notes are metrizable, there is a very common notion of convergence that is not metrizable. The topology on the set of functions $f : [0, 1] \to \mathbb{R}$ where the open sets are defined by pointwise convergence is not metrizable.*

2.2.1 Closed Sets and Limit Points

Definition 2.2.6. *A subset A of a topological space X is **closed** if the set*

$$A^c = X - A = \{x \in X | x \notin A\}$$

is open.

Note that a set can be open, closed, both, or neither! It can be shown that the collection of closed subsets of a space X has properties similar to those satisfied by the collection of open subsets of X.

Fact 2.2.7. *Let X be a topological space. The following conditions hold,*

1. *$\emptyset$ and X are closed*

2. *arbitrary intersections of closed sets are closed*

3. *finite unions of closed sets are closed.*

Definition 2.2.8. *Given a subset A of a topological space X, the **interior** of A is defined as the union of all open sets contained in A. The **closure** of A is defined as the intersection of all closed sets containing A.*

The interior of A is denoted by A° and the closure of A is denoted by $\overline{A}$. We note that A° is open and $\overline{A}$ is closed. Furthermore, $A^\circ \subseteq A \subseteq \overline{A}$.

Theorem 2.2.9. *Let A be a subset of the topological space X. The element x is in $\overline{A}$ if and only if every open set B containing x intersects A.*

Proof. We prove instead the equivalent contrapositive statement: $x \notin \overline{A}$ if and only if there is an open set B containing x that does not intersect A. Clearly, if $x \notin \overline{A}$, then $\overline{A}^c = X - \overline{A}$ is an open set containing x that does not intersect A. Conversely, if there is an open set B containing x that does not intersect A, then $B^c = X - B$ is a closed set containing A. The definition of closure implies that B^c must also contain $\overline{A}$. But $x \notin B^c$, so $x \notin \overline{A}$. □

Definition 2.2.10. *An open set O containing x is called a **neighborhood** of x.*

Definition 2.2.11. *Suppose A is a subset of the topological space X and let x be an element of X. Then x is a **limit point** of A if every neighborhood of x intersects A in some point other than x itself.*

In other words, $x \in X$ is a limit point of $A \subset X$ if $x \in \overline{A - \{x\}}$, the closure of $A - \{x\}$. The point x may or may not be in A.

Theorem 2.2.12. *A subset of a topological space is closed if and only if it contains all its limit points.*

Definition 2.2.13. *A subset A of a topological space X is **dense** in X if every $x \in X$ is a limit point of the set A. This is equivalent to its closure $\overline{A}$ being equal to X.*

Definition 2.2.14. *A topological space X is **separable** if it contains a countable subset that is dense in X.*

Example 2.2.15. *Since every real number is a limit point of rational numbers, it follows that $\mathbb{Q}$ is a dense subset of $\mathbb{R}$. This also implies that $\mathbb{R}$, the standard metric space of real numbers, is separable.*

2.2.2 Continuity

Definition 2.2.16. *Let X and Y be topological spaces. A function $f : X \to Y$ is **continuous** if for each open subset $O \subseteq Y$, the set $f^{-1}(O)$ is an open subset of X.*

Recall that $f^{-1}(B)$ is the set $\{x \in X | f(x) \in B\}$. Continuity of a function depends not only upon the function f itself, but also on the topologies specified for its domain and range!

Theorem 2.2.17. *Let X and Y be topological spaces and consider a function $f : X \to Y$. The following are equivalent:*

1. *f is continuous*
2. *for every subset $A \subseteq X$, $f\left(\overline{A}\right) \subseteq \overline{f(A)}$*
3. *for every closed set $C \subseteq Y$, the set $f^{-1}(C)$ is closed in X.*

Proof. $(1 \Rightarrow 2)$. Assume f is a continuous function. Suppose $x \in \overline{A}$, where A is a subset of X. Let O be a neighborhood of $f(x)$. Since $f^{-1}(O)$ is an open set of X containing $x \in \overline{A}$, it must intersect with A in some point x'. It follows that O intersects $f(A)$ in the point $f(x')$. By Theorem 2.2.9, we find that $f(x) \in \overline{f(A)}$.

$(2 \Rightarrow 3)$. Suppose that for every subset $A \subseteq X$, $f\left(\overline{A}\right) \subseteq \overline{f(A)}$. Let $C \subseteq Y$ be a closed set and let $A = f^{-1}(C)$. By P 1.5.4, we have $f(A) \subseteq C$. If $x \in \overline{A}$, then

$$f(x) \in f\left(\overline{A}\right) \subseteq \overline{f(A)} \subseteq \overline{C} = C.$$

So that $x \in f^{-1}(C) = A$ and, as a consequence, $\overline{A} \subseteq A$. Thus, $A = \overline{A}$ is closed.

$(3 \Rightarrow 1)$. Let O be an open set in Y. Let $O^c = Y - O$; then O^c is closed in Y. By assumption, $f^{-1}(O^c)$ is closed in X. Using elementary set theory, we have

$$X - f^{-1}(O^c) = \{x \in X | f(x) \notin O^c\} = \{x \in X | f(x) \in O\} = f^{-1}(O).$$

That is, $f^{-1}(O)$ is open. □

Theorem 2.2.18. *Suppose X and Y are two metrizable spaces with metrics d_X and d_Y. Consider a function $f : X \to Y$. The function f is continuous if and only if it is d-continuous with respect to these metrics.*

Proof. Suppose that f is continuous. For any $x_1 \in X$ and $\epsilon > 0$, let $O_y = B_{d_Y}(f(x_1), \epsilon)$ and consider the set

$$O_x = f^{-1}(O_y)$$

which is open in X and contains the point x_1. Since O_x is open and $x_1 \in O_x$, there exists a d-open ball $B_{d_X}(x_1, \delta)$ of radius $\delta > 0$ centered at x_1 such that $B_{d_X}(x_1, \delta) \subset O_x$. We also see that $f(x_2) \in O_y$ for any $x_2 \in B_{d_X}(x_1, \delta)$ because $A \subseteq O_x$ implies $f(A) \subseteq O_y$. It follows that $d_Y(f(x_1), f(x_2)) < \epsilon$ for all $x_2 \in B_{d_X}(x_1, \delta)$.

Conversely, let O_y be an open set in Y and suppose that the function f is d-continuous with respect to d_X and d_Y. For any $x \in f^{-1}(O_y)$, there exists a d-open ball $B_{d_Y}(f(x), \epsilon)$ of radius $\epsilon > 0$ centered at $f(x)$ that is entirely contained in O_y. By the definition of d-continuous, there exits a d-open ball $B_{d_X}(x, \delta)$ of radius $\delta > 0$ centered at x such that $f(B_{d_X}(x, \delta)) \subset B_{d_Y}(f(x), \epsilon)$. Therefore, every $x \in f^{-1}(O_y)$ has a neighborhood in the same set, and that implies $f^{-1}(O_y)$ is open. □

Definition 2.2.19. *A sequence $x_1, x_2, \ldots$ of points in X is said to* ***converge*** *to $x \in X$ if for every neighborhood O of x there exists a positive integer N such that $x_i \in O$ for all $i \geq N$.*

A sequence need not converge at all. However, if it converges in a metrizable space, then it converges to only one element.

Theorem 2.2.20. *Suppose that X is a metrizable space, and let $A \subseteq X$. There exists a sequence of points of A converging to x if and only if $x \in \overline{A}$.*

Proof. Suppose $x_n \to x$, where $x_n \in A$. Then, for every open set O containing x, there is an N, such that $x_n \in O$ for all $n > N$. By Theorem 2.2.9, this implies that $x \in \overline{A}$. Let d be a metric for the topology of X and x be a point in $\overline{A}$. For each positive integer n, consider the neighborhood $B_d\left(x, \frac{1}{n}\right)$. Since $x \in \overline{A}$, the set $A \cap B_d\left(x, \frac{1}{n}\right)$ is not empty and we choose x_n to be any point in this set. It follows that the sequence $x_1, x_2, \ldots$ converges to x. Notice that the "only if" proof holds for any topological space, while "if" requires a metric. □

Theorem 2.2.21. *Let $f : X \to Y$ where X is a metrizable space. The function f is continuous if and only if for every convergent sequence $x_n \to x$ in X, the sequence $f(x_n)$ converges to $f(x)$.*

Proof. Suppose that f is continuous. Let O be a neighborhood of $f(x)$. Then $f^{-1}(O)$ is a neighborhood of x, and so there exists an integer N such that $x_n \in f^{-1}(O)$ for $n \geq N$. Thus, $f(x_n) \in O$ for all $n \geq N$ and $f(x_n) \to f(x)$.

To prove the converse, assume that the convergent sequence condition is true. Let $A \subseteq X$. Since X is metrizable, one finds that $x \in \overline{A}$ implies that there exists a sequence $x_1, x_2, \ldots$ of points of A converging to x. By assumption, $f(x_n) \to f(x)$. Since $f(x_n) \in f(A)$, Theorem 2.2.21 implies that $f(x) \in \overline{f(A)}$. Hence $f\left(\overline{A}\right) \subseteq \overline{f(A)}$ and f is continuous. □

2.2.3 Completeness

Suppose X is a metrizable space. From Definition 2.2.19, we know that a sequence $x_1, x_2, \ldots$ of points in X converges to $x \in X$ if for every neighborhood A of x there exists a positive integer N such that $x_i \in A$ for all $i \geq N$.

It is possible for a sequence in a metrizable space X to satisfy the Cauchy criterion, but not to converge in X.

Example 2.2.22. *Let $C[-1,1]$ be the vector space of continuous functions on the interval $[-1,1]$ and consider the L_2 norm*

$$\|f(t)\|_2 = \left(\int_{-1}^{1} |f(t)|^2 dt \right)^{\frac{1}{2}}.$$

Define the sequence of functions $f_n(t)$ given by

$$f_n(t) = \begin{cases} 0 & t \in \left[-1, -\frac{1}{n}\right] \\ \frac{nt}{2} + \frac{1}{2} & t \in \left(-\frac{1}{n}, \frac{1}{n}\right) \\ 1 & t \in \left[\frac{1}{n}, 1\right] \end{cases}.$$

Assuming that $m \geq n$, we get

$$d(f_n, f_m) = \|f_n(t) - f_m(t)\|_2 = \left(\int_{-1}^{1} |f_n(t) - f_m(t)|^2 dt \right)^{\frac{1}{2}} = \frac{(m-n)^2}{6m^2 n}.$$

This sequence satisfies the Cauchy criterion, but it does not converge to a continuous function in $C[-1,1]$.

Definition 2.2.23. *A metrizable space X is said to be* ***complete*** *if every Cauchy sequence in X converges to a limit $x \in X$.*

Theorem 2.2.24. *A closed subset A of a complete metrizable space X is itself a complete metrizable space.*

Definition 2.2.25. *The **completion** of a metrizable space X consists of a complete metric space X' and an isometry $\phi : X \to X'$ such that $\phi(X)$ is a dense subset of X'. Moreover, the completion is unique up to isometry.*

Example 2.2.26. *Consider the metric space $\mathbb{Q}$ of rational numbers equipped with the metric of absolute distance. The completion of this metric space is $\mathbb{R}$ because the isometry is given by the identity mapping and $\mathbb{Q}$ is a dense subset of $\mathbb{R}$.*

Cauchy sequences have many applications in analysis and signal processing. For example, they can be used to construct the real numbers from the rational numbers. In fact, the same approach is used to construct the completion of any metric space.

Definition 2.2.27. *Two Cauchy sequences $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ are equivalent if for every $\epsilon > 0$ there exists an integer N such that $d(x_k, y_k) \leq \epsilon$ for all $k \geq N$.*

Example 2.2.28. *Let $\mathcal{C}(\mathbb{Q})$ denote the set of all Cauchy sequences $q_1, q_2, \ldots$ of rational numbers where $\sim$ represents the equivalence relation on this set defined above. Then, the set of equivalence classes (or quotient set) $\mathcal{C}(\mathbb{Q})\backslash\sim$ is in one-to-one correspondence with the real numbers. This construction is the standard completion of $\mathbb{Q}$. Since every Cauchy sequence of rationals converges to a real number, the isometry is given by mapping each equivalence class to its limit point in $\mathbb{R}$.*

Definition 2.2.29. *Let A be a subset of a metric space (X, d) and $f : X \to X$ be a function such that $f(A) \subseteq A$. The function f is a **contraction** on A if there exists a constant $\gamma < 1$ such that $d(f(x), f(y)) \leq \gamma d(x, y)$ for all $x, y \in A$.*

Consider the following important results in applied mathematics: Picard's theorem for differential equations, the implicit function theorem, and Bellman's principle of optimality for Markov decision processes. What do they have in common? They each establish the existence and uniqueness of a function and have relatively simple proofs based on the contraction mapping theorem.

Theorem 2.2.30 (Contraction Mapping Theorem)**.** *Let (X, d) be a complete metric space and f be contraction on a closed subset $A \subseteq X$. Then, f has a unique fixed point x^* in A such that $f(x^*) = x^*$. Moreover, the sequence $x_{n+1} = f(x_n)$ converges to x^* for any point $x_1 \in A$, with $d(x^*, f^n(x_1)) \leq c\gamma^n$ for a constant c depending on x_1.*

Proof. Suppose f has two fixed points $y, z \in A$. Then, $d(y, z) = d\left(f(y), f(z)\right) \leq \gamma d(y, z)$ and $d(y, z) = 0$ because $\gamma \in [0, 1)$. This shows that $y = z$ and any two fixed points in A must be identical.

Using induction, it is easy to see that $d(x_n, x_{n+1}) \leq \gamma^{n-1} d(x_1, x_2)$. From this, we can bound the distance between x_m and x_n (for $m < n$) with

$$\begin{aligned} d(x_m, x_n) &\leq d(x_m, x_{m+1}) + d(x_{m+1}, x_n) \\ &\leq \sum_{i=m}^{n-1} d(x_i, x_{i+1}) \leq \sum_{i=m}^{n-1} \gamma^{i-1} d(x_1, x_2) \\ &\leq \sum_{i=m}^{\infty} \gamma^{i-1} d(x_1, x_2) \leq \frac{\gamma^{m-1}}{1-\gamma} d(x_1, x_2). \end{aligned}$$

The sequence x_n is Cauchy because $d(x_m, x_n)$ can be made arbitrarily small (for all $n > m$) by increasing m. Since (X, d) is complete, it must converge to a fixed point and x^* is the unique fixed point in A. □

Example 2.2.31. *Consider the cosine function restricted to the subset $[0, 1] \subseteq \mathbb{R}$. Since $\cos(x)$ is decreasing for $0 \leq x < \pi$, we have $\cos([0, 1]) = [\cos(1), 1]$ with $\cos(1) \approx 0.54$. The mean value theorem of calculus also tells us that $\cos(y) - \cos(x) = \cos'(t)(y - x)$ for some $t \in [x, y]$. Since $\cos'(t) = -\sin(t)$ and $\sin(t)$ is increasing on $[0, 1]$, we find that $\sin([0, 1]) = [0, \sin(1)]$ with $\sin(1) \approx 0.84$.*

Taking the absolute value, shows that $|\cos(y) - \cos(x)| \leq 0.85|y - x|$. Therefore, $\cos(t)$ is a contraction on $[0, 1]$ and the sequence $x_{n+1} = \cos(x_n)$ converges to the unique fixed point $x^* = \cos(x^*)$ for all $x_1 \in [0, 1]$.

2.2.4 Compactness

Definition 2.2.32. *A metrizable space X is **totally bounded** if, for any $\epsilon > 0$, there exists a finite set of $B(x, \epsilon)$ balls that cover (i.e., whose union equals) X.*

Definition 2.2.33. *A metrizable space X is* **compact** *if it is complete and totally bounded.*

The closed interval $[0,1] \subset \mathbb{R}$ is compact. In fact, a subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded. On the other hand, the standard metric space of real numbers is not compact because it is not totally bounded.

Theorem 2.2.34. *A closed subset A of a compact space X is itself a compact space.*

The following theorem highlights one of the main reasons that compact spaces are desirable in practice.

Theorem 2.2.35. *Let X be a compact space and $x_1, x_2, \ldots \in X$ be a sequence. Then, there is a subsequence $x_{n_1}, x_{n_2}, \ldots$, defined by some increasing sequence $n_1, n_2, \ldots \in \mathbb{N}$, that converges.*

Proof. Suppose for some $x \in X$, it holds that, for any $\epsilon > 0$, the set $B(x,\epsilon)$ contains infinitely many elements in the sequence $x_1, x_2, \ldots$. In this case, we can choose $n_1, n_2, \ldots$ so that $|x - x_{n_j}| < \frac{1}{j}$. Therefore, the subsequence converges to x.

On the other hand, suppose the sequence has no convergent subsequence. Then, for every $x \in X$, there is an $\epsilon > 0$ such that $B(x,\epsilon)$ does not contain infinitely many elements in the sequence. Since X is totally bounded, we also find that, for any $\epsilon > 0$, X is covered by a finite set of balls of radius ϵ. But, for each $\epsilon > 0$, this gives a contradiction because all the elements in the infinite sequence must fall into a finite number balls each containing finitely many elements. By contradiction, this implies that there is a subsequence that converges. □

Functions from compact sets to the real numbers are very important in practice. To keep the discussion self-contained, we first review the extreme values of sets of real numbers. First, we must define the **extended real numbers** $\overline{\mathbb{R}}$ by augmenting the real numbers to include limit points for unbounded sequences $\overline{\mathbb{R}} \triangleq \mathbb{R} \cup \{\infty, -\infty\}$. Surprisingly, it turns out that $\overline{\mathbb{R}}$ is a compact metrizable space.

Definition 2.2.36. *The* **supremum** *(or least upper bound) of $X \subseteq \mathbb{R}$, denoted $\sup X$, is the smallest extended real number $M \in \overline{\mathbb{R}}$ such that $x \leq M$ for all $x \in X$. It is always well-defined and equals $-\infty$ if $X = \emptyset$.*

Definition 2.2.37. *The* **maximum** *of $X \subseteq \mathbb{R}$, denoted $\max X$, is the largest value achieved by the set. It equals $\sup X$ if $\sup X \in X$ and is undefined otherwise.*

Definition 2.2.38. *The **infimum** (or greatest lower bound) of* $X \subseteq \mathbb{R}$*, denoted* $\inf X$*, is the largest extended real number* $m \in \overline{\mathbb{R}}$ *such that* $x \geq m$ *for all* $x \in X$*. It is always well-defined and equals* ∞ *if* $X = \emptyset$*.*

Definition 2.2.39. *The **minimum** of* $X \subseteq \mathbb{R}$*, denoted* $\min X$*, is the smallest value achieved by the set. It equals* $\inf X$ *if* $\inf X \in X$ *and is undefined otherwise.*

Lemma 2.2.40. *Let* X *be a metrizable space and* $f : X \to \mathbb{R}$ *be a function from* X *to the real numbers. Let* $M = \sup f(A)$ *for some non-empty* $A \subseteq X$*. Then, there exists a sequence* $x_1, x_2, \ldots \in A$ *such that* $\lim_n f(x_n) = M$*.*

Proof. If $M = \infty$, then A has no finite upper bound and, for any $y \in \mathbb{R}$, there exists an $x \in A$ such that $f(x) > y$. In this case, we can let x_1 be any element of A and x_{n+1} be any element of A such that $f(x_{n+1}) > f(x_n) + 1$.

If $M < \infty$, then A has a finite upper bound and, for any $\epsilon > 0$, there is an x such that $M - f(x) < \epsilon$. Otherwise, one arrives at the contradiction $\sup f(A) < M$. Therefore, we can construct the sequence $x_1, x_2, \ldots$ by choosing $x_n \in A$ to be any point that satisfies $M - f(x_n) \leq \frac{1}{n}$. □

Theorem 2.2.41. *Let* X *be a metrizable space and* $f : X \to \mathbb{R}$ *be a continuous function from* X *to the real numbers. If* A *is a compact subset of* X*, then there exists* $x \in A$ *such that* $f(x) = \sup f(A)$ *(i.e.,* f *achieves a maximum on* A*).*

Proof. Using Lemma 2.2.40, one finds that there is a sequence $x_1, x_2, \ldots \in A$ such that $\lim_n f(x_n) = \sup f(A)$. Since A is compact, there must also be a subsequence $x_{n_1}, x_{n_2}, \ldots$ that converges to some $x^* \in A$. Finally, the continuity of f shows that

$$\sup f(A) = \lim_n f(x_n) = \lim_k f(x_{n_k}) = f(\lim_k x_{n_k}) = f(x^*).$$

□

Corollary 2.2.42. *A continuous function from a compact subset* A*, of a metrizable space* X*, to the real numbers achieves a minimum on* A*.*

Theorem 2.2.43. *Any bounded non-decreasing sequence of real numbers converges to its supremum.*

Proof. Let $x_1, x_2, \ldots \in \mathbb{R}$ be a sequence satisfying $x_{n+1} \geq x_n$ and $x_n \leq M < \infty$ for all $n \in \mathbb{N}$. Without loss of generality, we can take the upper bound M to be the supremum $\sup\{x_1, x_2, \ldots\}$. This sequence converges to M if, for any $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that $M - x_n < \epsilon$ for all $n > N$.

Proving by contradiction, we assume that it does not converge to M. Since x_n is non-decreasing, this implies that there is an $\epsilon > 0$ such that $M - x_n \geq \epsilon$ for all $n \in \mathbb{N}$. But, this result contradicts $\sup f(A) = M$. Therefore, the sequence converges to M. □